



UNIVERSIDADE DE ÉVORA

# **TÍTULO | Bibliotecas digitais para as Humanidades**

**Subtítulo | novos desafios e oportunidades**

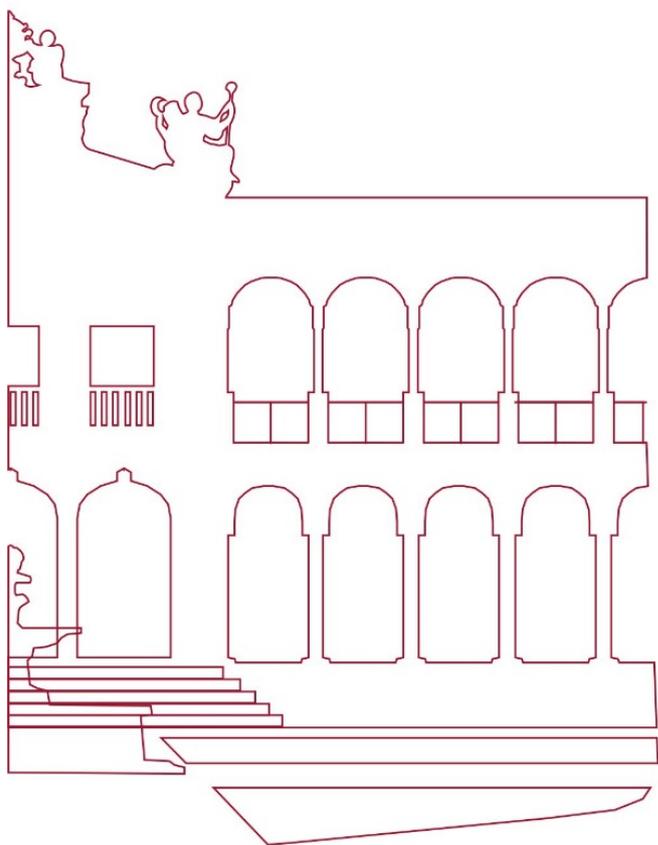
Dália Maria Godinho Guerreiro

Orientador | Doutor José Luís Borbinha

Co-Orientador | Doutor José António Calixto

Tese apresentada à Universidade de Évora para obtenção do Grau de  
Doutor em Ciências da Informação e da Documentação

Évora, 2017







UNIVERSIDADE DE ÉVORA

# **TÍTULO | Bibliotecas digitais para as Humanidades**

**Subtítulo | novos desafios e oportunidades**

Dália Maria Godinho Guerreiro

Orientador | Doutor José Luís Borbinha

Co-Orientador | Doutor José António Calixto

Tese apresentada à Universidade de Évora para obtenção do Grau de  
Doutor em Ciências da Informação e da Documentação

Évora, 2017



## **Agradecimentos**

Aos meus orientadores, Professores Doutor José Luís Borbinha e Doutor José António Calixto, pelo apoio, conselhos, orientação e sugestões, fundamentais para concretizar esta investigação, o meu mais profundo agradecimento.

Agradeço ao CIDEHUS - Centro Interdisciplinar de História, Culturas e Sociedades da Universidade de Évora, pelo acolhimento e motivação e, em particular, à Prof. Doutora Fernanda Olival, Diretora do Centro, e à Prof. Doutora Maria Filomena Gonçalves, Coordenadora da Linha Literacias e Património Textual, pelo apoio constante e permissões concedidas na aquisição de bibliografia e nas deslocações em missão ou a congressos internacionais. Também agradeço à Dr.<sup>a</sup> Carla Malheiro, pela eficiência e ajuda em todas as situações.

Agradeço ao Professor Doutor Francisco Vaz, Diretor do Programa de Doutoramento em Ciências da Informação e da Documentação, por todo o apoio e incentivo.

Agradeço às Profs. Doutora Laura Pires, Doutora Maria Isabel Roque, Doutora Maria Clara Paixão Sousa e Doutora Fernanda Campos, ao Prof. Doutor Daniel Alves e à Doutora Manuela Domingues pela discussão, pelos comentários e pelas sugestões com que contribuíram para este trabalho. Agradeço, também, à Doutora Maria Joaquina Barrulas, pelos esclarecimentos e revisão do guião do grupo focal.

Agradeço ao Dr. Duarte Meireles, Diretor da Biblioteca Francisco Pereira de Moura, do Instituto Superior de Economia e Gestão, por todas as facilidades concedidas para a conclusão deste trabalho.

Aos meus colegas do Programa de Doutoramento em Ciências da Informação e da Documentação, Luísa Alvim, Margarida Vargues, Nelson Vaquinhas, Teresa Costa, Zélia Parreira, agradeço a amizade, a partilha e o incentivo com que nos juntámos neste desafio.

Aos membros do grupo focal e do inquérito por questionário, os meus penhorados agradecimentos pela disponibilidade e pelos argumentos apresentados na discussão que, por isso, resultou tão profícua. Agradeço à Prof. Maria João Forte, pela ajuda nas transcrições do grupo focal e à Dr.<sup>a</sup> Paula Gonçalves, pela revisão do texto.

Agradeço, de forma especial, aos meus amigos António Baptista, Encarnação Silva, Luís Guilherme, Luís Neves, Manuela Domingues, Maria Gilda Terra, Maria Isabel Roque, Maria João Forte, Maria Natália Rothes, Mário Rita, Mauro Oliveira, Óscar Cavaleiro, Paula Gonçalves e Paulo Pinto, à minha afilhada Beatriz e família, pela amizade, pelo apoio constante e pelo incentivo com que acompanharam este trabalho. Agradeço, também, à Prof. Doutora Natália Correia Guedes, pela amizade com que me desvendou novos horizontes.

E, por fim, deixo o maior agradecimento à minha família, por todo o apoio e afeto.

## Financiamento

Esta tese foi financiada pela Fundação para a Ciência e a Tecnologia (FCT) através da concessão da bolsa de doutoramento SFRH/BD/82229/2011 no âmbito do QREN-POPH (Programa Operacional Potencial Humano), comparticipado pelo Fundo Social Europeu e por fundos nacionais do MEC.





## Resumo

As bibliotecas digitais são cruciais à investigação em Humanidades, exigindo maior eficácia e personalização no acesso à informação. No caso dos incunábulos e livro antigo disponibilizados em linha, as exigências da investigação apresentam requisitos específicos de pesquisa e acesso ao conteúdo, de transcrição e marcação do texto, criação de pontos de ligação, nomeadamente, através de sumários com o conteúdo textual e iconográfico das obras e ainda de ferramentas para a tradução do texto. A revisão da literatura teve como referência o domínio emergente das Humanidades Digitais e o das bibliotecas digitais, permitindo caracterizar o estado da investigação, complementada com a investigação em estudos do livro para a descrição de incunábulos e livro antigo. A observação direta de bibliotecas digitais e de coleções de incunábulos e livro antigo disponibilizadas permitiu coletar informação sobre modelos de edição digital e de pesquisa e recuperação da informação, bem como identificar e avaliar as particularidades inerentes a essas obras e a forma como estas afetam a sua disponibilização em linha. O método de investigação incluiu ainda a interação com um grupo focal com investigadores em Humanidades com o objetivo de recolher dados acerca da forma como utilizam as bibliotecas digitais, das vantagens ou desvantagens que reconhecem na utilização das bibliotecas digitais face às congéneres físicas e das expectativas que têm para a sua otimização e maior adequação à investigação, e a validação dos resultados por um painel de investigadores em Humanidades.

**Palavras-chave:** biblioteca digital; humanidades digitais; incunábulo; livro antigo; pesquisa e acesso à informação



# Digital library for Humanities: new challenges and opportunities

## Abstract

Digital libraries are crucial to the research in Humanities, requiring more efficiency and flexibility in tailoring the access to the information. The body of incunabula and antique books available on line is expressive and tends to correspond to the Humanities research demands, stressing the need to improve instruments to the research and access to the works content: text transcription and text markers; creation of connection points, such as summaries with the works textual and iconographic content; implementation of translation tools. Literature review about digital libraries was articulated with the Digital Humanities issues, consenting to characterise the current state research and to identify new projects in this domain; it was complemented with the research in book studies to describe incunabula and antique books concepts. The direct observation of digital libraries websites and incunabula and antique books collections available in digital libraries allowed to collect data about digital edition and research and retrieval data models, as well as to identify the incunabula and antique books specific features and the way how these interferes in making them available online. The research method included the interaction with a focal group with Humanities researchers with the goal of collect data about the way how they use digital libraries, the advantages and disadvantages towards their physical congeners and their expectations for improvement and better correspondence to the research. The results were validated by a Humanities researcher's panel.

**Keywords:** data research and access; digital libraries; Digital Humanities; incunabula; old book



## Sumário

TABELAS.....	XV
FIGURAS.....	XVII
GRÁFICOS.....	XXI
ABREVIATURAS.....	XXIII
INTRODUÇÃO.....	1
1.    Enquadramento da investigação.....	4
2.    Propósito e objetivos da investigação.....	9
3.    Estrutura do trabalho.....	12
1.    MÉTODO.....	17
1.1.    Métodos de análise.....	17
1.2.    Técnicas de recolha de dados.....	22
1.2.1.    Pesquisa documental.....	23
1.2.2.    Estudo de caso.....	30
1.2.3.    Grupo focal.....	38
1.2.4.    Inquérito por questionário.....	44
1.3.    Instrumentos da investigação.....	47
1.3.1.    Grelha de observação das BDLA.....	48
1.3.2.    Grelha de observação dos incunábulos e livro antigo.....	49
1.3.3.    Guião do grupo focal.....	49
1.3.4.    Questionário.....	51
1.4.    Amostra.....	53
1.4.1.    Bibliotecas digitais.....	54
1.4.2.    Incunábulos e livro antigo.....	54
1.4.3.    Edição digital das obras de Aldo Manuzio.....	54
1.4.4.    Caracterização dos grupos focal e do inquérito por questionário.....	55

2.	HUMANIDADES DIGITAIS.....	57
2.1.	Desenvolvimento das Humanidades Digitais .....	62
2.2.	Infraestruturas Europeias.....	68
2.3.	Comunidades de interesses.....	71
2.4.	Linhas de investigação.....	74
2.5.	Projetos.....	80
2.6.	Publicações .....	86
3.	BIBLIOTECAS DIGITAIS .....	91
3.1.	Digitalização e preservação .....	104
3.2.	Metadados.....	106
3.3.	Conversão de imagens em texto .....	112
3.4.	Aplicativos para a construção de bibliotecas digitais.....	117
3.5.	Pesquisa e recuperação de informação .....	119
3.6.	Adequação das bibliotecas digitais às Humanidades.....	126
3.7.	Bibliotecas digitais para as Humanidades .....	134
4.	BIBLIOTECAS DIGITAIS: ESTUDO CRÍTICO COMPARATIVO .....	147
4.1.	Biblioteca Nacional Digital.....	149
4.2.	Breve descrição de outras bibliotecas digitais de referência.....	165
4.2.1.	Alma Mater .....	165
4.2.2.	Gallica.....	168
4.2.3.	Internet Archive .....	176
4.2.4.	Bibliothèques Virtuelles Humanistes .....	185
4.2.5.	Perseus Digital Library.....	191
4.2.6.	Estudo comparativo .....	196
5.	INCUNÁBULOS E LIVRO ANTIGO.....	207
5.1.	Materialidade dos Incunábulos e livro antigo.....	208

5.2.	Os incunábulos e livro antigo na Biblioteca Nacional Digital .....	217
5.3.	As edições digitais de Aldo Manuzio .....	236
5.3.1.	Aldo Manuzio .....	237
5.3.2.	Representação nas várias coleções digitais .....	242
6.	FATORES DE ACESSIBILIDADE DA INFORMAÇÃO: LEVANTAMENTO DAS NECESSIDADES DA INVESTIGAÇÃO EM CIÊNCIAS SOCIAIS E HUMANAS.....	255
6.1.	Grupo focal .....	255
6.2.	Inquérito por questionário .....	265
7.	CONCLUSÕES .....	289
7.1.	Contexto da investigação .....	290
7.2.	Objetivos e conclusões .....	291
7.3.	Limitações do estudo.....	307
7.4.	Linhas futuras de investigação .....	307
	REFERÊNCIAS BIBLIOGRÁFICAS.....	309
	SÍTIOS INSTITUCIONAIS.....	329
	GLOSSÁRIO .....	333
	ANEXOS.....	355
I.	Grelha observação livro antigo na BND.....	357
II.	Grelha de observação das bibliotecas digitais com livro antigo .....	359
III.	Sondagem grupo focal.....	361
IV.	Guião do grupo focal .....	363
V.	Transcrição do grupo focal .....	365
VI.	Grupo focal resultados estruturados nível 2.....	377
VII.	Inquérito por questionário ronda 1.....	383
VIII.	Inquérito por questionário ronda 2.....	389

IX.	Glossário do Inquérito por questionário .....	395
X.	Convite à participação no inquérito por questionário, ronda 1 e 2.....	397
XI.	Normas ISO.....	399
XII.	Lista do Livro antigo na BND.....	407
XIII.	Dimensões das obras na BNP .....	417
XIV.	Lista das obras de Aldo Manuzio na Gallica .....	418
XV.	Lista das obras de Aldo Manuziu no Internet Archive .....	422

## Tabelas

Tabela 1 – Premissas dos estudos quantitativos e qualitativos .....	18
Tabela 2 – Termos associados a bibliotecas com acervos digitais .....	97
Tabela 3 – BND: acesso às coleções .....	156
Tabela 4 – BND: número de títulos digitalizados por coleção .....	163
Tabela 5 – Quadro resumo bibliotecas digitais .....	197
Tabela 6 – Símbolos e funções .....	198
Tabela 7 – Divulgação nas bibliotecas digitais .....	203
Tabela 8 - Bibliotecas Digitais .....	206
Tabela 9 – Número de obras livro antigo na BND .....	217
Tabela 10 – BND: livro antigo, por monografias e folhetos .....	220
Tabela 11 – BND: livro antigo, por língua .....	220
Tabela 12 – BND: livro antigo, por local de edição das obras .....	221
Tabela 13 – BND: livro antigo, por outros locais de edição .....	222
Tabela 14 – BND: livro antigo, obras com e sem transcrição textual .....	224
Tabela 15 – BND: livro antigo, transcrição, por idioma.....	224
Tabela 16 – BND: livro antigo, obras com e sem sumário.....	224
Tabela 17 – BND: livro antigo, obras com sumário, por língua.....	227
Tabela 18 – BND: livro antigo, por características das obras .....	228
Tabela 19 – BND: livro antigo, obras com transcrição, por tipo de letra .....	229
Tabela 20 – BND: Livro antigo, obras por ilustração e mapas.....	230
Tabela 21 – Obras Aldo Manuzio nas bibliotecas digitais .....	242

Tabela 22 – Características das obras de Aldo Manasio, por biblioteca digital.....	252
Tabela 23 – Fator de ponderação das respostas.....	286

## Figuras

Figura 1 – Robert Busa na consola de controle do IBM 705 .....	58
Figura 2 – Distribuição global dos centros de estudos em HD.....	63
Figura 3 – Símbolo ADHO .....	66
Figura 4 – The Spatial History Project: Holocaust Geographies.....	81
Figura 5 – The Spatial History Project: Chinese Canadian Immigrant Pipeline.....	83
Figura 6 – Rome Reborn 1.1 .....	84
Figura 7 – Esquema da constituição da biblioteca digital .....	92
Figura 8 – Humanidades Digitais e bibliotecas.....	138
Figura 9 – BN: biblioteca digital, 2000.....	150
Figura 10 – BND: página de acolhimento, 2002 .....	151
Figura 11 – BND: página de acolhimento .....	153
Figura 12 – BND: página de acesso à obra .....	157
Figura 13 – BND: obra em versão flash com sumário .....	158
Figura 14 – BND: barra de ferramentas, obra em flash .....	158
Figura 15 – BND: barra de ferramentas, obra em JPG .....	159
Figura 16 – BND: secções dos conteúdos na página da obra.....	160
Figura 17 – Alma Mater: página de acolhimento .....	166
Figura 18 – Gallica: página de acolhimento .....	170
Figura 19 – Gallica: opções de pesquisa.....	171
Figura 20 – Gallica: página de consulta .....	175
Figura 21 – IA: página de acolhimento.....	178
Figura 22 – IA: visualização ecrã inteiro .....	181

Figura 23 – IA: Open Library, página de acolhimento .....	183
Figura 24 – Open Library, acesso à obra .....	184
Figura 25 – BVH: página de acolhimento .....	186
Figura 26 – BVH: pesquisa avançada .....	188
Figura 27 – Perseus: página de acolhimento.....	193
Figura 28 – Perseus: pesquisa .....	194
Figura 29 – Bíblia, dita das 42 linhas .....	207
Figura 30 – A Regra & diffinções [sic]: sumário de nível 1.....	225
Figura 31 – Aqui come[n]ça o liuro chamado [...]: sumário de nível 2.....	225
Figura 32 – Instituciones o doctrinas, sumário de nível 3.....	226
Figura 33 – Reg[ra] e statut[us], [...]: capitais e caldeirões.....	228
Figura 34 – De medica matéria: obra sem indicação relativa às ilustrações .....	232
Figura 35 – Visualização da imagem em <i>zoom</i> , sem barras de navegação.....	232
Figura 36 – Exercicios sp[irit]uais & diuinos: páginas com o fundo alterado.....	234
Figura 37 – Obra com a margem superior aparada, truncando o texto. ....	235
Figura 38 – Registo com nota relativa à digitalização .....	236
Figura 39 – Epistole[...] <i>sancta Catharina da Siena</i> .....	239
Figura 40 – Vergilius. [ <i>Bucolica...</i> ].....	240
Figura 41 – Marcas do impressor Aldo Manuzio.....	241
Figura 42 – BNP: resultado da pesquisa para “Manuzio, Aldo” .....	243
Figura 43 – BND: <i>Institutiones graecae grammaticae</i> .....	243
Figura 44 – BND: <i>Institutiones graecae grammaticae</i> .....	244
Figura 45 – Gallica: pesquisa avançada, por autor e data.....	245
Figura 46 – Gallica / e rara: <i>P. Ovidii Metamorphoseon libri XV</i> .....	247

Figura 47 – BVH: Institutionum grammaticarum libri quatuor .....	248
Figura 48 – IA: Petri Bembi de Aetna dialogus. ....	249
Figura 49 – IA: Hypnerotomachia Poliphili .....	250



## **Gráficos**

Gráfico 1 – BND: livro antigo, por distribuição cronológica da data de edição .....	223
Gráfico 2 – Respondentes, por área de investigação (ronda 1) .....	266
Gráfico 3 – Perceção do utilizador face às BD (ronda 1) .....	267
Gráfico 4 – Vantagens da biblioteca digital (ronda 1) .....	269
Gráfico 5 – Desvantagens da biblioteca digital (ronda 1) .....	272
Gráfico 6 – Expetativas de melhoria para as bibliotecas digitais (ronda 1) .....	274
Gráfico 7 – Respondentes, por área de investigação.....	277
Gráfico 8 – Perceção do utilizador face às BD (ronda 2) .....	278
Gráfico 9 – Vantagens da biblioteca digital (ronda 2) .....	280
Gráfico 10 – Desvantagens da biblioteca digital (ronda 2) .....	282
Gráfico 11 – Expetativas de melhoria para as bibliotecas digitais (ronda 2) .....	284



## **Abreviaturas**

aaDH – Australasian Association for Digital Humanities

a.C. – antes de Cristo

ACH – Association for Computers and the Humanities

ACM – Association for Computing Machinery

ACRL – Association of College and Research Libraries

ADHO – Alliance of Digital Humanities Organizations

AFED – Annotated Facsimile Edition

AHDig – Associação das Humanidades Digitais [de língua portuguesa]

ALA – American Library Association

ALTO – Analyzed Layout and Text Object

APA – American Psychological Association

ARK – Archival Resource Key [Identificador persistente]

BD – Biblioteca Digital

BDLA – Biblioteca digital de livro antigo

BDPI – Biblioteca Digital do Património Ibero-Americano

Bibframe – Bibliographic Framework Initiative

BIBO – Bibliographic Ontology Specification

BND – Biblioteca Nacional Digital [Portugal]

BnF – Bibliothèque Nationale de France

BNP – Biblioteca Nacional de Portugal

BVH – Bibliothèques Virtuelles Humanistes

CBRS – Centre national de la recherche scientifique

CERS – Centre d'Études Supérieures de la Renaissance

CESSDA – Council of European Social Science Data Archives

CiC – Concept-in-Context

Clarín – Common Language Resources and Technology Infrastructure

CLÉO – Centre pour l'édition électronique ouverte

CNRS – Centre national de la recherche scientifique

CSDH/SCHN – Canadian Society for Digital Humanities/Société canadienne des humanités numériques

CSN – Computer Says Yes

Dariah – Digital Research Infrastructure for the Arts and Humanities

DLF – Digital Library Federation

DLib – Dlib Magazine

DPLA – Digital Public Library of America

DSH – Digital Scholarship in the Humanities

EADH – European Association for Digital Humanities

EEBO – Early English Books

EHESS – École des hautes études en sciences sociales

ePub – Electronic Publication

ESF – European Science Foundation

ÉU – European Union [União Europeia]

F.A.Q. – Frequently Asked Questions

FaBiO – FRBR aligned Bibliographic Ontology

Fedora – Flexible Extensible Digital Object Repository Architecture

GEDE – Gabinete de Edições Eletrónicas

GIF – Graphics Interchange Format

GO::DH – Global Outlook::Digital Humanities

HASTAC – Humanities, Arts, Science, and Technology Alliance and Collaboratory

HMM – Hidden Markov Model

HP – Hewlett-Packard

HTML – HyperText Markup Language

IA – Internet Archive

ICC – International Color Consortium

IFLA – International Federation of Library Associations and Institutions

IPSA – Imaginum Patavinae Scientiae Archivum - Archive of the Patavinian Scientific Images

IRHT – Institut de recherche et d’histoire des textes

JADH – Japanese Association for Digital Humanites

JCDL – Joint Conference on Digital Libraries

JPG ou JPEG – Joint Photographic Experts Group

JSON – JavaScript Object Notation

LLC – Literary and Linguistic Computing

LOD – Linked Open Data

MADSRDF – Metadata Authority Description Schema ontology

MARC – Machine Readable Cataloging

METRO – Metropolitan New York Library Council's

METS – Metadata Encoding and Transmission Standard

MIT – Massachusetts Institute of technology

MODS – Metadata Object Description Schema

MODSRDF – Metadata Object Description Schema ontology

NeDiMAH – Network for Digital Methods in the Arts and Humanities

NLS – National Library Service for the Blind and Physically Handicapped

OAI – Open Archive Initiative

OAI-PMH – Open Archives Initiative Protocol for Metadata Harvesting

OCA – Open Content Alliance

OCR – Optical character recognition [reconhecimento ótico de caracteres]

ONG – Organização não-governamental [Non-governmental organizations]

OPAC – Online Public Access Catalog [catálogo em linha]

OPDS – Open Publication Distribution System

PDF – Portable Document Format

PNG – Portable Network Graphics

ppp – pixel por polegada

PURL – Persistent Uniform Resource Locator

R – Respondentes

RDF – Resource Description Framework

RDF – Resource Description Framework

RNOD – Registo Nacional de Objectos Digitais

RSS – Rich Site Summary ou Really Simple Syndication

s.d. – sem data

S.J. – Societas Iesu [Companhia de Jesus]

s.l. – sem local

Samtla – Search and Mining Tools with Linguistic Analysis

SEB – Structure Extraction from Books

SECO – SErial Converter

SIBUC – Serviço Integrado das Bibliotecas da Universidade de Coimbra

SLM – statistical language model

sRGB – standard RG B

TBLE – Text-Image-Link-Editor

TEI – Text Encoding Initiative

TEL – The European Library

TIC – Tecnologia de Informação e Comunicação

TIFF [ou TIF] – Tagged Image File Format

UK – United Kingdom [Reino Unido]

UNESCO – United Nations Educational, Scientific and Cultural Organization

W3C – World Wide Web Consortium

WDL – World Digital Library

Web – World Wide Web

XML – Extensible Markup Language



## Introdução

**Biblioteca Digital** – Biblioteca que disponibiliza em linha obras digitalizadas ou nadas digitais, com um sistema que permite a pesquisa e recuperação da informação. Com os serviços inerentes a uma biblioteca como a organização, gestão da coleção, preservação e disseminação.

**Biblioteca Digital Livro Antigo (BDLA)** - Bibliotecas digitais que disponibilizam, em formato digital, os incunábulo e o livro antigo.

**Humanidades Digitais** - Área de atividade que conjuga a informática com as disciplinas das humanidades, que inclui o uso sistemático de recursos digitais no trabalho académico, bem como a reflexão sobre sua aplicação.

**Incunábulo** - Primeiras obras impressas com caracteres móveis, no ocidente, entre 1450 e 1500.

**Livro Antigo** - Primeiras obras impressas com caracteres móveis, no ocidente, entre 1501 e 1800.

A crescente disponibilização em linha de acervos digitalizados traz óbvios contributos à investigação, no sentido em que facilita o acesso à informação e ao conhecimento. No entanto, apresenta também alguns desafios aos investigadores, nomeadamente, aos investigadores em Humanidades que, perante o imenso acervo documental digitalizado e disponível, têm dificuldades em identificar, localizar, utilizar e reutilizar as obras pretendidas. A web é um universo caótico, onde nem sempre é perceptível a lógica que preside à organização dos dados e à forma como são apresentados os resultados da pesquisa. Apesar da navegação ser cada vez mais amigável, tornando-se fácil de utilizar e recolher dados pertinentes, a informação disponibilizada é hierarquizada pelos motores de busca e a ordem em que é apresentada deriva dos algoritmos predefinidos que podem não corresponder, exatamente, à formalização da pesquisa. Por outro lado, as bibliotecas digitais disponibilizam informação que é credível e está organizada, mas a pesquisa, a recuperação e a utilização da informação

podem ser muito dificultadas. Reside aqui a questão de partida que sustenta este estudo: de que forma se pode adequar a disponibilização de obras digitalizadas nas bibliotecas digitais, em particular os incunábulos e o livro antigo, às necessidades dos utilizadores, nomeadamente, aos investigadores em Humanidades?

A proposta é analisar como os incunábulos e o livro antigo são disponibilizados nas bibliotecas digitais, e se essa forma é ou não adequada à investigação no domínio das Humanidades, mas também avaliar como os investigadores gostariam de os encontrar, quais os aspetos a melhorar para uma maior eficácia no acesso e, ainda, quais as suas expectativas em relação ao desenvolvimento de novos modelos e funcionalidades, justificando o título “Bibliotecas Digitais para as Humanidades: novos desafios e oportunidades”.

Este trabalho é motivado pelo crescimento das Humanidades Digitais, facto que implica uma necessidade crescente na utilização das obras digitalizadas e disponibilizadas nas bibliotecas digitais pelos investigadores em Humanidades.

A teorização em torno do conceito de Humanidades Digitais corresponde, essencialmente, à reflexão conjunta elaborada pelos seus intervenientes – investigadores nos diversos campos das Humanidades, como as Ciências Sociais e Humanas, informáticos e especialistas em multimédia – a qual tem vindo a ser promovida, ao longo das últimas décadas, em encontros, seminários ou colóquios, e cujos resultados são divulgados em publicações de cariz científico, impressas ou em formato digital. A maioria destas publicações faz uma resenha dos projetos efetuados ou em curso, apontando algumas linhas de desenvolvimento. Em geral, formula mais questões do que apresenta soluções ou indicam estratégias. As Humanidades Digitais

têm vindo a afirmar-se como um campo fértil de investigação, como se constata através das inúmeras solicitações para participação e apresentação de resultados em eventos, pelo que, atualmente, se considera que o principal desafio é a sistematização da informação reunida (Gold, 2012).

Estão abertas várias linhas de investigação que estudam a otimização dos acervos digitais nas áreas da visualização da informação, da construção de semânticas coerentes, do tratamento e indexação de coleções de textos, da representação do espaço e do tempo e da edição digital para utilização académica e escolar (Schreibman, Siemens, & Unsworth, 2004, 2016; Terras, Nyhan, & Vanhoutte, 2013; Vinopal & McCormick, 2013; Warwick et al., 2009).

Considera-se que, para analisar a forma como as bibliotecas devem evoluir para corresponder às exigências da investigação, é essencial conhecer as fontes que utilizam e a forma como estas são disponibilizadas em linha. Constatando que, na tipologia de livro antigo, se encontram fontes de informação relevantes para as Humanidades, pretende-se definir as particularidades dos livros impressos entre 1450 e 1800, incluindo, portanto, os incunábulos (Faria & Pericão, 2008, pp. 652–653) e o livro antigo, nomeadamente, no que respeita às características que, de alguma forma, possam interferir no processo entre a digitalização e a disponibilização em linha. Por outro lado, afigura-se essencial analisar um conjunto significativo de bibliotecas digitais, constituídas essencialmente por documentos digitalizados, e, em particular, que possuam incunábulos e livro antigo, que constituem as fontes de referência para a investigação histórica em Humanidades, avaliando o seu contributo para a investigação científica neste domínio. O resultado destas análises será confrontado

com os contributos dos próprios investigadores, permitindo definir um conjunto de procedimentos que otimize a consulta de obras digitalizadas incluídas em bibliotecas digitais.

## **1. Enquadramento da investigação**

A produção intelectual, a *obra*, realiza-se através das obras que se materializam em *item*, o suporte da escrita .

O suporte de escrita e o seu formato foi sendo alterado com o passar do tempo. Pode dizer-se, de forma geral, sem referir formatos mais específicos, que o suporte da escrita foi, primeiro, em tabuinhas ou placas cerâmicas e em papiro, depois, em pergaminho, a seguir, em papel e, agora, em suporte digital. Não houve só variações no suporte, como a própria escrita, ou a grafia das obras, também sofreu alterações ao longo do tempo: as tabuinhas eram incisas com caracteres cuneiformes, obrigando a uma escrita sintética; no papiro, em rolo ou *volumen*, a obra era escrita e lida de forma contínua, obrigando a uma leitura sequencial; o pergaminho, mais resistente, era também utilizado em *volumen* que, ao ser cortado em folhas, para um melhor aproveitamento do material, deu origem ao códice, o que permitia escrever no reto e no verso e facilitava a arrumação; e, por fim, o hipertexto que, através de hiperligações, permite a navegação dentro de uma obra e entre obras. As sucessivas mudanças de suporte provocaram outras alterações, nomeadamente, no modo de leitura. “Até ao século I da nossa era, [a leitura] foi primordialmente oral e pública [...]. Era, de novo, a dificuldade do sistema gráfico que exigia que as pessoas alfabetizadas e cultas fossem capazes de reconstruir a sintaxe, enquanto liam [...]” (Castillo Gómez, 2004, p. 21). Posteriormente, a substituição do *volumen* pelo *codex* produziu uma nova

revolução no modo de ler, dado que este, ao permitir “a libertação de uma das mãos tornou possível compartilhar a escrita e a leitura, isto é fazer anotações à medida que ia lendo” (Castillo Gómez, 2004, p. 23). Assim, passou-se de uma leitura pública para uma leitura individual, ao mesmo tempo que o princípio e o fim da obra eram mais evidenciados e se desenvolvia a divisão interna dos textos, a pontuação e as ilustrações. A invenção da imprensa que se lhe seguiu provocou uma revolução epistemológica, ao aumentar exponencialmente a difusão do conhecimento.

Atualmente, são as tecnologias digitais que sustentam uma nova mudança no suporte da escrita, nas formas de escrita e de leitura e na difusão do conhecimento, com um forte impacto na investigação científica, uma vez que as bibliotecas digitais têm vindo a reabilitar a leitura de códices e impressos antigos a uma comunidade mais alargada de investigadores. “El libro electrónico, las bibliotecas digitales le han dado nueva vida a los códices e impresos del pasado” (Lucía Megías, 2010, p. 400). No entanto, é consensual que uma biblioteca digital não pode cingir-se a uma mera transposição de suportes e disponibilização de documentos em linha (Berry, 2012; Besser, 2004; Borgman, 2009; Lucía Megías, 2010; Vinopal & McCormick, 2013), implicando a criação de novos modelos que se ajustem à atualidade e aos novos parâmetros da investigação. “El reto del hipertexto nos colocará ante nuevos modelos textuales, y por tanto, ante nuevos modelos de lectura y de comprensión de nuestro tiempo. Un reto que los filólogos, los humanistas, los científicos que nos movemos alrededor del texto, debemos liderar” (Lucía Megías, 2010, p. 400). Por conseguinte, a criação de novos modelos para os textos antigos implica um modelo de investigação interdisciplinar, ainda que liderado pelos investigadores em Humanidades, os potenciais utilizadores das bibliotecas digitais de livro antigo.

A biblioteca é, literalmente, o lugar dos livros: o termo Biblioteca (do grego βιβλιον-θήκη, composto de βιβλιον, "livro", e θήκη "depósito") significa o espaço físico em que se guardam livros (Frédéric Barbier, 2013, p. 7). Choudhury e Seaman (2008) salientam que "it is important to remind ourselves of the enduring role of the library as a physical place" (s.p.). Além de permitirem o acesso às obras, as bibliotecas também fornecem um espaço de trabalho.

Em sentido lato, as bibliotecas constituem-se como repositórios da produção intelectual da humanidade, com o objetivo de guardar, preservar, organizar e disponibilizar os acervos, sejam físicos ou digitais. A origem da biblioteca coincide com o início da História. "The oldest [2500 a.C.] known library was found in the lost ruins of Ebla in northern Syria. [...] recovered 20,000 clay tablets with cuneiform writing. [...] documented the economic and cultural life of the city" (Murray, 2009, s.p.). Esta biblioteca, constituída por tabuinhas de argila, continha os documentos de arquivo, mas também as obras que refletiam a vida cultural da cidade. Na Antiguidade, destacam-se, pela sua relevância, a biblioteca de Assurbanipal (668-627 a.C.), em Nínive, com um vastíssimo acervo<sup>1</sup>, onde se incluía a *Epopéia* de Gilgamesh (Charpin, 2008) e a biblioteca de Alexandria (3.º séc. d.C.), considerada a maior biblioteca da antiguidade (Barbier, 2013, p. 21). São elas que definem a missão da biblioteca como repositório intelectual de uma cultura e marcam a ligação entre a biblioteca e a produção do conhecimento, tal como se mantêm até aos nossos dias.

---

<sup>1</sup> É difícil de quantificar dado que, aquando da escavação arqueológica, não houve uma clara separação das várias tabuinhas encontradas, mas o Museu Britânico, em Londres, estima este acervo em cerca de 10000 textos.

As bibliotecas sempre foram importantes para a Academia, mas, em particular, para as Humanidades, tal como Besser afirma: "Historically, libraries have been critical to humanities scholarship. Libraries provide access to original works, to correspondence and other commentary that helps contextualize those works, and to previous humanities commentaries" (Besser, 2004, p. 557). Contudo, atualmente a investigação realiza-se com o suporte da tecnologia digital, obrigando as bibliotecas a um processo de reconversão, alargando a missão e os objetivos que a têm norteado e assumindo um novo conjunto de serviços e funcionalidades. "Indeed, it is becoming more and more evident that research is increasingly being mediated through digital technology. [...] it is rare to find an academic today who has had no access to digital technology as part of their research activity" (Berry, 2012, p. 1). As bibliotecas têm vindo a acompanhar esta tendência, primeiro, informatizando os catálogos e, posteriormente, implementando as bibliotecas digitais.

No entanto, as bibliotecas digitais são, em primeiro lugar e fundamentalmente, bibliotecas. Ou seja, o processo de reconversão que tem vindo a ser pedido às bibliotecas, mesmo incluindo a curadoria digital, não pode desvirtuar a sua missão fulcral: organizar, selecionar, recolher, conservar, preservar e dar acesso à informação. "Librarians tend to take a broad view of the concept of a library. In general terms, they see libraries as organizations that select, collect, organize, conserve, preserve, and provide access to information on behalf of a community of users" (Borgmann, 2000, p. 38).

Da mesma forma que se afirma que uma biblioteca digital não pode cingir-se a uma mera transposição de suportes, também a disponibilização de uma coleção em linha

não configura uma biblioteca digital. “Existing digital libraries invariably manifest themselves in this way. But a Web site that provides a wealth of digital objects, along with appropriate methods of access and retrieval, should not necessarily be considered a ‘library’” (Witten, Bainbridge, & Nichols, 2009, p. 9). Enquanto processo orgânico, a biblioteca digital implica a adição de novos itens e a gestão da coleção em constante atualização. A constituição de uma biblioteca digital implica a definição de uma política de edição digital, ou seja, os princípios que norteiam a seleção das obras a adicionar, as normas a seguir ao longo de todo o processo, o método adotado para a pesquisa e recuperação da informação.

As bibliotecas digitais começaram por ser criadas por bibliotecários, como uma extensão dos serviços prestados aos seus utentes. “De ahí que *Biblioteca digital* se haya convertido en el nombre genérico para designar este modelo de almacenamiento y difusión de documentos, ya sea a partir de la digitalización de objetos anteriores con diversos formatos (texto, audio, video, imágenes...) o de objetos creados en esta tecnología” (Lucía Megías, 2010, p. 373). Porém, ao longo da última década, este conceito de biblioteca tornou-se demasiado redutor e limitado.

Os vários autores (Besser, 2004; Borgmann, 2000; Choudhury & Seaman, 2008; Lawless, Conlan, & Hampson, 2016; Witten et al., 2009), que têm vindo a estudar este tema, são consensuais na definição de biblioteca digital como uma coleção coerente de obras disponibilizadas em linha de forma organizada e estruturada, com mecanismos de pesquisa e recuperação da informação, mantida ao longo do tempo, isto é, cujos conteúdos sejam preservados e incrementados em função de uma política de gestão da coleção definida.

A investigação acerca de temas relacionados com as bibliotecas digitais tem vindo também a ser muito incrementada. “Digital library research builds upon a long history of related work in information retrieval, databases, user interfaces, networks, information seeking, classification and organization, library automation, publishing, and other areas” (Borgmann, 2000, p. 40). A produção científica relativa a bibliotecas digitais com obras digitalizadas remonta à sua origem e tem acompanhado toda a sua evolução. As primeiras obras, mais operativas que analíticas, descrevem o processo da disponibilização em linha de forma relativamente exaustiva (Fox, 1993; Lesk, 2005; Ross et al., 2002). No entanto, cedo se começou a refletir sobre a utilidade das bibliotecas digitais e quais os caminhos a seguir (Besser, 2002; Borges, 2003; Borgman, 2009). Na atualidade, verifica-se uma tendência dominante para articular a disponibilização em linha com as Humanidades Digitais, a crescente utilização e reutilização da informação (Borgman, 2009; Galina Russell, 2011a; Koller, 2012; Lucía Megías, 2010), na perspetiva de que as bibliotecas digitais se tornem úteis para o ensino e para a investigação e aliciem o público em geral para a leitura.

## **2. Propósito e objetivos da investigação**

Esta investigação aborda a problemática do livro antigo nas bibliotecas digitais, constituídas por coleções digitalizadas de documentos, a forma como disponibilizam esse tipo de obras em linha, e a sua utilidade para os investigadores da área de Humanidades, com o intuito de contribuir com novas propostas para a otimização do acesso aos conteúdos nas bibliotecas digitais, a sua utilização e reutilização. Pretende-se, também, avaliar a relação e a interação entre os recursos informáticos disponíveis

e a pesquisa, a divulgação e a recuperação efetuadas no âmbito das bibliotecas digitais, e as necessidades dos investigadores em Ciências Sociais e Humanas.

Os objetivos são:

- Analisar um conjunto representativo de bibliotecas digitais que disponibilizem incunábulos e livro antigo;
- Caracterizar os incunábulos e o livro antigo através das obras disponibilizados em linha na Biblioteca Nacional Digital (BND);
- Avaliar a forma como as características específicas dos incunábulos e do livro antigo afetam a edição digital das obras, desde a digitalização à disponibilização em linha;
- Compreender quais as funcionalidades que os investigadores em Humanidades gostariam de ver implementadas nas bibliotecas digitais;
- Propor um conjunto de diretivas que adequem as bibliotecas digitais aos métodos de investigação das Ciências Sociais e Humanas.

Em função destes objetivos, as questões genéricas que se formulam e que constituem as premissas que orientam este estudo são as seguintes:

1. Quais são os atributos mais relevantes das bibliotecas digitais de livro antigo?
2. Quais as características dos incunábulos e do livro antigo na BND?
3. Quais as particularidades dos incunábulos e do livro antigo?
4. Em que medida as características físicas, materiais e formais, dos incunábulos e do livro antigo afetam a sua edição digital?
5. Que desenvolvimentos nas bibliotecas digitais se percecionam como mais vantajosos para a investigação em Humanidades?

As bibliotecas digitais existentes foram construídas, maioritariamente, em função das perspetivas dos seus produtores (bibliotecários e informáticos), pelo que agora se pretende identificar as expectativas dos seus utilizadores e elaborar um conjunto de diretivas que as adequem aos métodos de investigação das Ciências Sociais e Humanas.

Esta investigação assenta em métodos qualitativos de análise, adequados à compreensão e descrição de um fenómeno, no caso, as bibliotecas digitais, na complexidade das modalidades que o compõem. “Qualitative research [...], sets out to tell you why it is happening. It is all about developing a detailed understanding of individuals' views, attitudes and behavior” (Moore, 2006, p. 141).

A investigação começa pela análise documental, tendo em vista o levantamento da literatura relevante e a recolha de informação sustentada sobre o tema. Este procedimento mantém-se recorrente ao longo de todo o trabalho e cumpre as várias fases da análise documental: pesquisa da informação; localização, seleção e análise crítica dos documentos.

De entre as várias abordagens dos métodos qualitativos, o estudo de caso é o que mais se adequa ao estudo que pretendemos efetuar, sem prejuízo de algumas análises correlativas, nomeadamente, a análise do impacto das bibliotecas digitais na investigação em Humanidades. Este método qualitativo é caracterizado pela descrição de um acontecimento ou evento (caso) de uma forma transversal. “Case study research is a method designed to study the particular within context and has a very specific purpose. That is to say a system that operates within well defined boundaries; the size and nature of that system are not the issue, what dictates the case is the

purpose of the investigation” (Pickard, 2007, p. 85). Este método surgiu da necessidade de estudar situações reais complexas: “The purpose of a case study is to provide a holistic account of the case and in-depth knowledge of the specific thought rich descriptions situated in context. This may lead to an understanding of a particular phenomenon but is understanding the case that should be paramount” (Pickard, 2007, p. 86). Segundo Yin (2003, p. 19), o método de estudo de caso, como estratégia de pesquisa, pode-se aplicar em «estudos organizacionais». Dado que, neste estudo, se pretende responder às questões de «como» e «porquê» das bibliotecas digitais, tendo como contexto acontecimentos contemporâneos e múltiplos, alguma da investigação a efetuar pode ser configurada como um estudo de caso. A realização dos estudos de caso tem como objetivo explorar, descrever e explicar a realidade das bibliotecas digitais e a sua utilização pelos investigadores na área das Humanidades.

Pretendemos, desta forma, criar as bases para a conceção de bibliotecas digitais amigáveis para o utilizador, em especial, os investigadores em Ciências Sociais e Humanas, com a intenção de que este trabalho seja útil sobretudo aos profissionais das Ciências da Informação e da Documentação.

### **3. Estrutura do trabalho**

No âmbito da preparação deste trabalho foram publicados artigos científicos, todos com revisão por pares, respetivamente: Bibliotecas Digitais para as Humanidades: novos desafios e oportunidades [Revisão de literatura]. In *11.º Congresso Nacional de Bibliotecários, Arquivistas e Documentalistas. Lisboa - Fundação Calouste Gulbenkian, 18, 19 e 20 de outubro de 2012*; Humanidades Digitais: novos desafios e oportunidades [Bibliotecas Digitais]. *Revista Internacional del Libro, Digitalización y Bibliotecas*, 2(1),

13–22; Humanidades Digitais: novos desafios e oportunidades [Ciências da Informação]. *Cadernos BAD*, (1), 63–78. E O livro antigo na era digital. In *12.º Congresso Congressos Nacionais de Bibliotecários, Arquivistas e Documentalistas: Ligar. Transformar. Criar Valor. Évora - Universidade de Évora, 21, 22 e 23 de outubro de 2015*. Estes artigos foram incorporados no presente trabalho.

Este trabalho divide-se em sete capítulos, após a Introdução: 1. Método; 2. Humanidades Digitais; 3. Bibliotecas digitais; 4. Bibliotecas digitais estudo crítico comparativo; 5. Incunábulos e livro antigo; 6. Fatores de acessibilidade da informação: levantamento das necessidades da investigação em Ciências Sociais e Humanas e 7. Conclusões.

Nos capítulos 2 e 3, a revisão da literatura corresponde às questões 1 e 5, sobre a definição de Humanidades Digitais e bibliotecas digitais e qual a sua relação, que são essencialmente esclarecidas através da pesquisa documental. Nestes capítulos efetuamos a síntese da literatura relevante, onde incluímos a análise aos periódicos de referência *Dlib Magazine*, *Literary and Linguistic Computing/Digital Scholarship in the Humanities* e às atas da *Joint Conference on Digital Libraries*. Foi ainda elaborada uma síntese da bibliografia que auxiliou a definir os principais conceitos, bem como a identificar as boas práticas em projetos similares.

No capítulo 2, através da revisão da literatura, definiu-se o conceito de Humanidades Digitais, o percurso efetuado, as repercussões nas infraestruturas europeias de apoio à Ciência, as comunidades de interesse, as linhas de investigação existentes, e reconheceu-se alguns projetos relevantes, e a respetiva divulgação através das publicações existentes.

No capítulo 3, procurou elucidar-se os vários conceitos que se podem invocar diretamente relacionados com as bibliotecas digitais. Seguidamente, conferiu-se o estado da arte dos temas nucleares relacionados com as bibliotecas digitais, como seja: a digitalização e preservação; a metacodificação; a conversão das imagens em texto editável; os aplicativos disponíveis para a construção e manutenção das bibliotecas digitais; a pesquisa e recuperação de informação; os projetos de aproximação às necessidades dos investigadores em Humanidades dividido em dois subcapítulos: Adequação das bibliotecas digitais às Humanidades e Bibliotecas digitais para as Humanidades.

Os capítulos 4 e 5 sobre, respetivamente, “Bibliotecas digitais estudo crítico comparativo” e “Incunábulos e livro antigo”, correspondem às questões 1, 2, 3 e 4 relativas ao estudo de caso das bibliotecas digitais e, em particular, da Biblioteca Nacional Digital, da caracterização dos incunábulos e livro antigo e da edição eletrónica das obras de Aldo Manuzio, assente essencialmente na observação dos espólios disponibilizados em linha.

No capítulo 4, fez-se o estudo de seis bibliotecas digitais: Biblioteca Nacional Digital (BND); Alma Mater; Gallica; Internet Archive; Bibliothèques Virtuelles Humanistes e Perseus Digital Library, através da observação dos respetivos sítios eletrónicos, com o auxílio de uma grelha de avaliação. Foi realizada uma descrição mais elaborada da BND. Por fim, efetuou-se a comparação entre as várias bibliotecas digitais.

O capítulo 5, “Incunábulos e livro antigo”, inicia com a definição de incunábulo e de livro antigo e suas principais características. Em seguida, analisa-se estas espécies bibliográficas, através dos exemplos existentes em linha na BND. Para se poder

comparar o tratamento destas espécies bibliográficas nas várias bibliotecas digitais, elegeu-se a obra tipográfica de Aldo Manuzio, considerado o inventor do livro moderno, e estudou-se a forma como a sua obra foi colocada em linha nas bibliotecas digitais estudadas.

O capítulo 6, “Fatores de acessibilidade da informação: levantamento das necessidades das Ciências Sociais e Humanas”, corresponde à questão 5, o tema central das entrevistas feitas no âmbito do grupo focal, cujos resultados são aferidos pelo inquérito estruturado. Foi realizado um grupo focal a um conjunto de investigadores das áreas das Humanidades. Com base na transcrição efetuada das várias intervenções, foi elaborado um inquérito em duas rondas, para validar os resultados anteriormente apurados.

Por fim, o capítulo 7 apresenta as conclusões.

Foi também incluído, no início do trabalho, a lista das abreviaturas, e, no final, o glossário e os anexos.



## **1. Método**

“Os métodos de investigação constituem o caminho a seguir para se chegar ao conhecimento” (Coutinho, 2011, p. 22). Neste capítulo vamos detalhar quais os métodos seguidos, as técnicas de recolha de dados, os instrumentos de investigação utilizados e a amostra selecionada, durante a investigação.

### **1.1. Métodos de análise**

A análise quantitativa solicita a obtenção de uma grande quantidade de dados para que se possa extrair os padrões e relacionar as variáveis. Em contrapartida, na análise qualitativa pode bastar a análise de um evento, desde que este seja suficientemente rico para proporcionar a informação, pois “the qualitative researcher is more likely to be interpretive, tending to begin with evidence and then building theory” (Gorman, & Clayton, 2005, p. 10). Enquanto a análise quantitativa se baseia na verificação de relações ou resultados previstos, a análise qualitativa foca-se na interpretação, pelo que os resultados obtidos são interpretações da realidade, sendo tão plurais quantos os indivíduos que as analisam (Gorman & Clayton, 2005).

As questões, ou premissas, necessárias ao desenvolvimento da investigação podem ser de carácter ontológico (qual a natureza da realidade?), epistemológico (qual a relação investigador-objeto?), axiológico (quais os valores envolvidos?) retórico (qual a linguagem utilizada?) ou metodológico (qual o processo de investigação?).

These differences are the nature of reality (the ontological assumption), the relationship of the researcher to that being researched (the epistemological

assumption), the role of values (the axiological assumption), the use of language and words (the rhetorical assumption), and the overall process of the research study (the methodological assumptions). (Creswell, 1994, p. 15)

Os métodos quantitativos ou qualitativos têm desempenhos diferenciados em função destas premissas (Tabela 1):

Tabela 1 – Premissas dos estudos quantitativos e qualitativos (adaptado de Creswell, 1994, p. 5)

<b>Premissa</b>	<b>Estudo quantitativo</b>	<b>Estudo qualitativo</b>
Ontológica	Realidade objetiva e singular, separada do investigador	Realidade subjetiva e múltipla, percebida pelos participantes no estudo
Epistemológica	Investigador é independente do objeto de estudo	Investigador interage com o objeto de estudo
Axiológica	Independente de juízos de valor e imparcial	Dependente de juízos de valor e parcial
Retórica	Formal Baseada em definições Impessoal Uso de terminologia quantitativa	Informal Definições em evolução Primeira pessoa (pessoal) Uso de terminologia qualitativa
Metodológica	Processo dedutivo Causa-efeito Processo estático (categorias definidas antes do estudo) Descontextualizado Generalizações conducentes à previsão, explicação e compreensão Preciso e fiável através da validação científica	Processo indutivo Multifacetado Processo emergente e alterável ao longo da investigação Contextualizado Teorias conducentes à compreensão Preciso e fiável através da verificação

Pela comparação do desempenho de ambas os métodos face às diversas premissas, o modelo qualitativo, é o que melhor se adapta à investigação em curso e aos respetivos objetivos, por se fundamentar na análise e na interpretação de um fenómeno a partir

das experiências dos seus múltiplos utilizadores. “Qualitative research, [...], sets out to tell you why it is happening. It is all about developing a detailed understanding of individuals' views, attitudes and behavior” (Moore, 2006, p. 141). Assim, a escolha do método qualitativo (Maxwell, 2013) é também justificada pelo facto de se pretender avaliar a relação e a interação entre os recursos informáticos disponíveis e a pesquisa, a divulgação e a recuperação efetuadas no âmbito das BDLA, bem como propor alterações de forma a dar resposta às necessidades sentidas pelos investigadores em Ciências Sociais e Humanas.

Por outro lado, a metodologia qualitativa é adequada à análise das BDLA, no sentido proposto por Bogdan e Biklen:

Utilizamos a expressão investigação qualitativa como um termo genérico que agrupa diversas estratégias de investigação que partilham determinadas características. Os dados recolhidos são designados por qualitativos, o que significa ricos em pormenores descritivos relativamente a pessoas, locais e conversas, e de complexo tratamento estatístico. As questões a investigar não se estabelecem mediante a operacionalização de variáveis, sendo, outrossim, formuladas com o objectivo de investigar os fenómenos em toda a sua complexidade e em contexto natural. (Bogdan & Biklen, 1994, p. 16)

Os métodos qualitativos adequam-se à complexidade das bibliotecas, enquanto ambiente informacional e social. “The outcomes of much qualitative research are the generation of research questions and conjectures, not the verification of predicted relationships or outcomes” (Gorman, & Clayton, 2005, p. 7). Dado que este projeto de investigação se pretende aberto na formulação de pontos de vista e hipóteses, não

tendo como objetivo a obtenção de um formulário estático e definitivo, confirma-se a sua conformidade com os métodos qualitativos.

A investigação acerca da aplicabilidade da tecnologia digital nos domínios das Humanidades justifica a aplicação de método qualitativo, com o objetivo de analisar e interpretar o fenómeno das bibliotecas digitais. “The full understanding of a discourse or piece of writing is a kind of artistic achievement and thus requires an «art doctrine» or technology, which we designate by the term «hermeneutics»” (Schleiermacher, & Tice, 2011, p. 53). De acordo com a disciplina hermenêutica, definida por Friedrich Schleiermacher, o conhecimento baseia-se na interpretação e adquire-se através da compreensão do significado de um fenómeno, isto é, da forma como é apreendido pelos seus intervenientes ou observadores num determinado contexto. Neste sentido, o investigador assume-se como intérprete e construtor do conhecimento: “You need, to a substantial extent, to *construct* and *reconstruct* your research design” (Maxwell, 2013, p. 3). Privilegia-se, por conseguinte, um método analítico, seguindo um modelo em espiral, do todo para a parte, ou do geral para o particular.

Qualitative research claims to describe life-worlds ‘from the inside out’, from the point of view of the people who participate. By so doing it seeks to contribute to a better understanding of social realities and to draw attention to processes, meaning patterns and structural features. (Flick, Kardorff, & Steinke, 2004, p. 3)

No modelo qualitativo pretende-se “substituir as noções científicas de *explicação, previsão, e controlo* do paradigma positivista pela *compreensão, significado e acção*” (Coutinho, 2011, p. 16). A compreensão de um fenómeno implica a capacidade de

identificar e distinguir os factos isolados que constituem um fenómeno, e relacioná-los entre si de forma adequada e coerente. A racionalização, que permite apreender e conhecer o fenómeno, elabora também conceitos que clarificam os factos. O significado, ou o sentido, do fenómeno constitui o fundamento das ações subsequentes, isto é, das práticas, procedimentos ou comportamentos, determinados em função do estudo efetuado. Coincide, portanto, com os objetivos desta investigação: analisar as variáveis da utilização do digital no domínio das Humanidades (*compreensão*), para determinar as respetivas possibilidades e expectativas (*significado* ou sentido) e propor soluções adequadas (*ação*).

A formulação do objetivo da investigação baseia-se na observação prévia do fenómeno (Denscombe, 2014; Gorman & Clayton, 2005). A definição do objetivo é crucial, pois “it is necessary to understand the object of research as preliminary until the end of the research, because the object will present itself in its true colors only at the end” (Flick, 2009, p. 18). Porém, na fase inicial do processo, não se pretende restringir demasiado a investigação, evitando truncar a observação e análise e comprometer a execução das etapas seguintes. “An investigator cannot know in advance what such phenomena mean to those being studied” (Gorman, & Clayton, 2005, p. 3). A aproximação ao objeto de estudo é, por conseguinte, ampla e envolvente, sem prejuízo da objetividade e da sistematização pretendidas ao longo de todo o processo. “The researcher begins with a question, or the need to discover something” (Moore, 2006, p. XI). As questões de investigação permitem centrar o estudo nos limites dos seus objetivos, mas não constituem um elemento fixo e imutável. Tal como o plano é dinâmico em função da investigação em progresso, também há questões que ficam sujeitas a reformulações.

In this model, in contrast to some other views of research design, the research questions are not the starting point or controlling piece of the design, to which all other components must conform. Instead, they are the *center* of the design; they are the heart, the hub, of the model, the component that connects most directly to all of the other components. (Maxwell, 2013, p. 4)

A pesquisa qualitativa é um processo de investigação desenhada a partir dos contextos originais onde os eventos ocorrem. A investigação é construída a partir das hipóteses formuladas, às quais se juntam novos entendimentos, adquiridos durante o processo, e que ajudam a ampliar as perspetivas e conduzem ao conhecimento (Coutinho, 2011). A validação das hipóteses fundamenta-se, em primeiro lugar, na elaboração do estado da arte feito a partir da pesquisa bibliográfica, após a qual, a investigação se orienta para a análise concreta de casos nas suas particularidades temporais e locais. Assim, o conhecimento constrói-se ao longo da investigação, de forma não linear, como um processo interativo e não, apenas, sumativo dos dados recolhidos.

As técnicas utilizadas são, por um lado, a pesquisa documental, para uma delineação prévia do objeto de estudo, convergindo na introdução teórica acerca do tema, e por outro lado, na observação direta de bibliotecas digitais e na realização de um grupo focal, seguido de questionário.

## **1.2. Técnicas de recolha de dados**

Utilizamos como técnica de recolha de dados a análise documental, a observação, o grupo focal e o inquérito. A análise documental permite a fundamentação teórica do trabalho. Utilizamos a observação com grelha para recolher os dados sobre as

bibliotecas digitais e sobre os incunábulos e livro antigo. O grupo focal e o inquérito foram utilizados para recolher e validar as opiniões dos investigadores.

### **1.2.1. Pesquisa documental**

A pesquisa e análise documental, podendo constituir um método de investigação, integra-se, neste caso, como etapa fundamental e propedêutica do estudo de caso (Yin, 2001), mantendo-se ativa ao longo de todo o processo. A análise documental fornece o suporte teórico que fundamenta e consolida as hipóteses e conclusões formuladas ao longo do estudo, pelo que se afirma como um procedimento adequado aos vários parâmetros da investigação (Quivy & Campenhoudt, 2008). Não obstante, no âmbito de um estudo sobre Humanidades Digitais, um tema ainda recente acerca do qual existe alguma matéria relevante, mas cujo corpo teórico pode ser considerado ainda incipiente, a obtenção de dados é essencialmente empírica, procurando realizar uma pesquisa documental sistemática, mas relegando-a para um nível de complementaridade, e não predominante, no decurso da investigação. A pesquisa e análise documental funcionam aqui como uma das estratégias integradas no estudo de caso. O carácter complementar que lhe é atribuído não invalida a exigência de rigor na recolha e no tratamento dos dados. Para isso, requer-se um levantamento de fontes de informação, quer sejam similares, complementares, ou contraditórias entre si, estabelecendo protocolos que permitam organizar e sistematizar os dados obtidos, bem como testar a respetiva validade (Reto & Nunes, 1999).

A investigação inicia com a análise documental, tendo em vista o levantamento da literatura relevante e a recolha de informação sustentada sobre o tema. Em conformidade com o modelo proposto por Hart (1998, p. 4), o processo inicia com a

pesquisa prévia e genérica, a partir da qual é procurada a informação, produzida pela investigação da comunidade científica que, neste caso, é essencialmente difundida através de monografias, revistas científicas e atas de conferências. Este procedimento mantém-se recorrente ao longo de todo o trabalho, cumprindo as várias fases da análise documental: pesquisa da informação; tratamento e organização dos dados recolhidos; processamento e análise crítica da informação; cruzamento dos dados, de forma a criar uma leitura ou interpretação inédita dos factos ou dos fenómenos relacionados com o objeto de estudo. “[A análise documental permite] introduzir algum valor acrescido à produção científica sem correr o risco de estudar o que já está estudado, tomando como original o que já outros descobriram” (Carmo, & Ferreira, 1998, p. 59). Neste sentido, a análise documental é um processo dinâmico de produção de informação, criando unidades de informação a partir das fontes primárias.

A validação das fontes documentais é feita através da crítica: interna, ou hermenêutica; externa, ou da testemunha (Saint-Georges, 2011). Pretende-se, desta forma, “[...] examinar metodicamente os documentos para se esforçar por determinar o seu alcance real e tentar medir o grau de confiança que possa ser-lhes concedido, tanto no que *são* como no que *dizem*” (Saint-Georges, 2011, p. 42).

A crítica das fontes tem em conta os seguintes parâmetros: *quem?*, identificando o autor da informação e avaliando o respetivo reconhecimento científico pelos pares; *o quê?*, analisando a natureza e as características da informação; *quando?*, elaborando a cronologia da informação; *como?*, verificando o respetivo trajeto de difusão, registando as reações que provoca, tais como argumentações, polémicas, contradições

e correções; *porquê?*, identificando as motivações do autor (investigação, manipulação, propaganda, etc.).

A análise documental conduz à revisão da literatura sobre o tema em investigação. Por revisão da literatura, entende-se:

The selection of available documents (both published and unpublished) on the topic, which contain information, ideas, data and evidence written from a particular standpoint to fulfil certain aims or express certain views on the nature of the topic and how it is to be investigated, and the effective evaluation of these documents to the research being proposed. (Hart, 1998, p. 13)

A pesquisa incide em informação escrita, sobretudo, em português, inglês e francês. Dada a natureza do tema e a constante atualização a que é sujeito, dá-se preferência a documentos (livros ou artigos) disponibilizados em linha, embora também seja feita a pesquisa em documentos impressos.

Foi definido um vocabulário temático, com termos e expressões utilizados nas pesquisas efetuadas em linha:

- Digitalização; digitization or digitizing; numérisation
- Meta informação; *metadata*
- Preservação digital; *digital preservation*
- “Biblioteca digital” site:pt
- Digital library; bibliothèque numérique
- Humanidades Digitais; Digital Humanities; Humanités Numériques
- “Digital Humanities” site:pt; “Digital Humanities” site:fr
- Ciências Sociais e Humanas; Human Sciences; Sciences Humaines

Considerando que as temáticas envolvidas começaram a ser apresentadas em meados da última década do século passado, o intervalo temporal de pesquisa situa-se entre 1990 e a atualidade, sendo dada preferência à documentação mais recente.

Para lá dos estudos monográficos, a consulta prossegue nos sítios institucionais de referência e nos jornais em linha, onde se inclui a exaustiva revisão de literatura em duas revistas científicas, a *DLib Magazine* (DLib) e a *Digital Scholarship in the Humanities* (DSH), e nas atas da *Joint Conference on Digital Libraries* (JCDL), abrangendo o arco temporal entre 2001 e 2016. O critério para a data de início prende-se com o facto de se considerar que as bibliotecas digitais se autonomizam e atingem a fase de maturação na viragem do milénio (Besser, 2004, pp. 560–561), sendo por essa altura que, apesar de a maioria continuar a funcionar a partir de bibliotecas físicas, analógicas, e manter os serviços e valências complementado com as mais-valias da tecnologia digital, adquirem um estatuto autónomo. Além disso, a série de conferências JCDL começou em 2001, permitindo abarcar as respetivas atas desde o início. Para uniformizar o critério cronológico, o levantamento das revistas científicas é feito, também, a partir de 2001. A data de término, 2016, prende-se com o limite imposto pela contingência da apresentação deste trabalho em 2017. Todos os artigos publicados nas revistas e nas atas são submetidos aos procedimentos correntes de revisão por pares.

A DLib (DOI: 10.1045/dlib.magazine; ISSN: 1082-9873) é uma revista bimestral, publicada desde 1995, em formato exclusivamente eletrónico. Tem como principal objetivo publicar os resultados da investigação em bibliotecas digitais, não só as

novidades tecnológicas e aplicações, bem como as questões socioeconómicas. É mantida em acesso aberto pela Corporation for National Research Initiatives.

A revista *DSH* (ISSN: 2055-7671; eISSN 2055-768X) foi publicada, entre 1986 e 2014, com o título *Literary and Linguistic Computing* (LLC) (ISSN 0268-1145) e, uma vez que este se revelava demasiado redutor, apontando apenas para estudos de linguística, desde dezembro de 2014, assumiu o título *Digital Scholarship in the Humanities*. A revista é publicada pela Alliance of Digital Humanities Organizations (ADHO) e só se encontra disponível para subscritores. Apresenta artigos com pesquisa teórica, metodológica, experimental e aplicada e inclui resultados de projetos de pesquisa, descrições e avaliações de ferramentas, técnicas e metodologias, relatórios de trabalhos em curso, assim como resenhas críticas a livros e recursos.

A JCDL é uma conferência que tem vindo a realizar-se anualmente, desde 2001. As respetivas atas são publicadas pela ACM [Association for Computing Machinery] Digital Library. Os temas tratados focam-se nas bibliotecas digitais, mas abrangem as suas múltiplas formas. Entre 2001 e 2016, foram publicados, em suporte papel e em linha, 919 artigos. Em 2001, foram publicados 76 artigos. A partir desta data, é feita a distinção entre *full paper* e *short paper*, sendo que 433 são em *full paper* e 364 em *short paper*.

Como procedimento metodológico, a primeira fase de recolha consiste na análise dos índices das publicações, selecionando todos os artigos cujo título contenha um dos seguintes termos ou expressões: “digital library”; “OCR”; “optical character recognition”; “recognition”; “annotation”; “old book”; “texto”; “images”; “retrieval”;

“historical document”. Foram, igualmente, selecionados todos os artigos, cujo título sugerisse uma ligação às bibliotecas digitais, preferencialmente, de livro antigo.

No sítio eletrónico da DLib, a pesquisa com o termo “digital library”, com a opção de correspondência exata dos termos, obteve 2571 correspondências. Alterando o modo de seriação entre ordenação por data, título ou classificação, observou-se que, nesta seleção, estavam incluídas as páginas de metadados dos artigos, as páginas introdutórias de cada número, cartas ao editor, guia para escrita de artigo, etc., pelo que o processo não foi considerado eficaz em função dos objetivos pretendidos. O sistema de pesquisa da revista não permite realizar seleções em campos específicos, como, por exemplo, no título dos artigos.

Dado que os resultados obtidos eram artigos genericamente relacionados com bibliotecas digitais, mas não como tema específico de bibliotecas digitais constituídas por livro antigo digitalizado, optou-se por realizar a navegação nos índices da revista para selecionar os títulos de artigos potencialmente relacionáveis com o objeto de estudo. Nesta fase, foram selecionados 50 artigos, procedendo, em seguida, à análise do respetivo resumo. Este procedimento levou à exclusão de 19 artigos, dado que o resumo apontava para estudos marginais ao tema em estudo. Os restantes 31 artigos foram integralmente lidos para excluir os que apresentassem resultados com menor relação ao tema, com estudos de caso cuja particularidade não permitisse uma generalização consistente com o tema principal. Nesta última fase, foram selecionados os 13 artigos mais relevantes para o objeto de estudo.

Na revista *LLC/DSH*, entre 2001 e 2016, foram publicados 646 artigos, dos quais 60 correspondem à introdução feita a cada um dos números da revista e 104 a resenhas

críticas a livros e a artigos. Foram selecionados 25 artigos que tratam explicitamente assuntos relacionados com as bibliotecas digitais. Destaca-se, igualmente, os artigos cuja temática é o TEI, havendo mesmo um número especial sobre o tema (September 2009, Vol. 24, Is. 3). Os artigos sobre bibliotecas digitais exploram, sobretudo, as filosofias subjacentes à sua construção e o impacto junto dos utilizadores, enquanto que os que se referem ao TEI apresentam uma vertente mais didática, incluindo indicações para a sua utilização.

Na primeira fase de seleção nas atas JCDL, foram assinalados 242 artigos. Destes, foram excluídos 120 *short papers*, com uma a três páginas, a maioria dos quais servindo de introdução a *workshops*, enquanto outros eram pequenos apontamentos sobre produtos ou experiências efetuadas. Os artigos cujo conteúdo se centrava em algoritmos de programação ficaram igualmente fora da análise, por não se inserirem na área em estudo. Foram, em seguida, selecionados 42 artigos lidos na íntegra, dos quais se procedeu à análise dos 31 artigos diretamente relacionados com o objeto de estudo e acerca dos quais foi realizada a revisão da literatura.

A análise documental, revelando as linhas de investigação e os argumentos mais prementes dos autores consultados (Hart, 1998), permite identificar conceitos, teorias, perspetivas, metodologias, técnicas, problemas, resultados e conclusões. Este levantamento contribui para validar os parâmetros de observação e análise das bibliotecas digitais. Assim, a análise documental afigura-se como um dos pilares da avaliação, permitindo estabelecer o confronto entre o acervo teórico, a informação fornecida pelas instituições e a realidade das coleções disponibilizadas em linha.

A bibliografia e as referências bibliográficas no corpo do texto seguem a norma bibliográfica APA<sup>2</sup> (2010). Os autores são identificados pelo apelido e a data é colocada a seguir ao autor, pelo que na bibliografia se faz uma ordenação simultânea pelo autor e data dos respetivos trabalhos. Este modelo torna mais imediata a relação entre a referenciação bibliográfica e a respetiva obra referenciada na bibliografia.

### **1.2.2. Estudo de caso**

A análise dos Incunábulos e livro antigo e dos sítios eletrónicos das diversas bibliotecas digitais, tendo como suporte as grelhas de observação (Anexo I e II) previamente elaboradas e complementadas com a informação disponibilizada pelas próprias instituições, assenta no método descritivo do estudo de caso.

No âmbito da metodologia qualitativa, o estudo de caso é um método de recolha de dados de forma sistemática, de acordo com um plano previamente estruturado e delineado. Este método permite obter um conjunto de variáveis e estabelecer uma relação lógica entre os dados, a partir dos quais se pode traçar a história, descrever o fenómeno em estudo e fazer a respetiva análise crítica. Tratando-se de uma investigação sobre um fenómeno contemporâneo, este método permite uma abordagem abrangente acerca dos fatores e acontecimentos que nele interferem, propiciam ou condicionam.

Como esforço de pesquisa, o estudo de caso contribui, de forma inigualável, para a compreensão que temos dos fenómenos individuais, organizacionais, sociais e políticos. [...] a clara necessidade pelos estudos de caso surge do

---

<sup>2</sup> APA style (estilo APA) – American Psychological Association style, 6.ª edição. (vd. <http://www.apastyle.org/>)

desejo de se compreender fenómenos sociais complexos. Em resumo, o estudo de caso permite uma investigação para se preservar as características holísticas e significativas dos eventos da vida real - tais como ciclos de vida individuais, processos organizacionais e administrativos, mudanças ocorridas em regiões urbanas, relações internacionais e a maturação de alguns setores. (Yin, 2001, p. 21)

David E. Gray caracteriza o estudo de caso como um método dedutivo: “Therefore case studies benefit from the prior development of a theoretical position to help direct the data collection and analysis process. Note, then, that the case study method tends to be deductive rather than inductive” (Gray, 2004, p. 124). Configurado dessa forma, o estudo de caso permite obter premissas gerais a partir de premissas particulares, sendo que, quanto mais lógica for a relação estabelecida entre as premissas gerais e os factos observados, maior a probabilidade de encontrar, tanto uma hipótese válida, como conclusões corretas.

Segundo Yin (2001), o método de estudo de caso, como estratégia de pesquisa, pode aplicar-se em estudos organizacionais. Esta investigação pode configurar-se como um estudo de caso no sentido em que pretende responder a questões «como» e «porquê» das bibliotecas digitais, tendo como contexto acontecimentos contemporâneos e múltiplos sobre os quais o investigador não exerce controlo: “O estudo de caso é a estratégia escolhida ao se examinarem acontecimentos contemporâneos, mas quando não se podem manipular comportamentos relevantes” (Yin, 2001, p. 27). Gray aponta no mesmo sentido ao afirmar que “the approach is particularly useful when the

researcher is trying to uncover a relationship between a phenomenon and the context in which it is occurring” (Gray, 2004, p. 124).

O estudo de caso tem sido considerado um método suficientemente maleável, no sentido em que pode incluir “observational case studies, interview case studies, organizational case studies, life history case studies, and multi-site and comparative case studies” (Gorman, & Clayton, 2005, pp. 47-48).

O processo envolvido compreende: a recolha sistemática dos dados relativos ao objeto de estudo, sondagem; a análise e seleção do material e da informação recolhidos; e, por fim, a elaboração de um relatório, que consiste na tese a apresentar. Estas etapas processuais convergem para dois registos de investigação complementares, ou seja, um enquadramento teórico e um registo empírico dos factos.

Case study research is a method designed to study the particular within context and has a very specific purpose. That is to say a system that operates within well defined boundaries; the size and nature of that system are not the issue, what dictates the case is the purpose of the investigation. (Pickard, 2007, p. 85)

O método de estudo de caso concilia a descrição e análise de cada situação particular com o objetivo de estudar o fenómeno de uma forma transversal e integrada. As BDLA e os fundos patrimoniais digitalizados (incunábulo e livro antigo) disponibilizados em linha constituem a unidade de análise desta investigação, contemplando situações como a utilização real e as expectativas e necessidades sentidas pelos seus utilizadores e investigadores no domínio da investigação em Humanidades.

Dado que o estudo de caso admite um conjunto de estratégias e instrumentos articulados para explorar, descrever e explicitar um fenómeno (Coutinho, 2011), este método configura-se como adequado ao objetivo de realizar uma investigação holística (sistémica, ampla, integrada), que nos conduza à compreensão das BDLA, no seu todo e na sua unicidade.

As BDLA selecionadas para constituir o caso de estudo são:

- Bibliotecas portuguesas:
  - Biblioteca Nacional Digital (BND);
  - Alma Mater;
- Bibliotecas estrangeiras:
  - Gallica;
  - Internet Archive (IA);
- Bibliotecas de referência para as Humanidades:
  - Perseus Digital Library;
  - Bibliothèques Virtuelles Humanistes (BVH).

Para efetuar a seleção das bibliotecas digitais portuguesas consultou-se o Registo Nacional de Objectos Digitais (RNOD)<sup>3</sup>. Encontram-se registadas 32 instituições, das quais, 20 possuem objetos digitais e 8 possuem livro antigo. Das que possuem livro antigo, 3 (Comissão para a Cidadania e a Igualdade de Género; Exército – Biblioteca; Teatro Nacional D. Maria II – Biblioteca/Arquivo) apresentam as obras em PDF, criado a partir das imagens digitalizadas, sem ter sido aplicado às imagens o reconhecimento

---

<sup>3</sup> Dados verificados a 10/07/2016. Existe alguma discrepância entre os números exibidos pelo RNOD e as obras efetivamente disponibilizadas pelas organizações, dado que a cedência dos dados para o RNOD é realizada de forma voluntária e não automática.

ótico de caracteres e diretamente associadas ao respetivo registo bibliográfico. A biblioteca digital da Universidade Aberta exhibe as obras em JPG num sítio eletrónico rudimentar, sem endereço persistente, cuja última atualização data de 2007. Outras 4 instituições (BNP; Biblioteca Pública Municipal do Porto; Universidade de Coimbra; Universidade de Lisboa) disponibilizam livro antigo, digitalizado e em linha, com as obras devidamente descritas e estruturadas. Na Biblioteca Pública Municipal do Porto, o livro antigo é residual e o acesso é feito apenas pelo catálogo. As restantes (BNP; Universidade de Coimbra; Universidade de Lisboa) disponibilizam coleções digitalizadas de códices, incunábulos ou livro antigo<sup>4</sup>. A Universidade de Lisboa mantém uma coleção digital em linha, cujo acesso é feito através de páginas estáticas, e não tem vindo a incrementar os conteúdos. Para análise, foram selecionadas as bibliotecas da BNP (Biblioteca Nacional de Portugal) com a Biblioteca Nacional Digital (BND) e da Universidade de Coimbra (Alma Mater), que continuam a atualizar-se, tanto a nível de estrutura, como de conteúdo, e onde a presença do livro antigo é relevante. Considerando a missão normalizadora atribuída à BNP, constituindo o paradigma que orienta as bibliotecas digitais portuguesas, o facto de ter desenvolvido um conjunto de programas específicos em código aberto e por ser a mais relevante biblioteca digital em Portugal, com o maior volume de conteúdos, a BND é objeto de uma análise mais exaustiva, incluindo o respetivo historial.

Os critérios definidos para a seleção das bibliotecas internacionais com livro antigo são: o impacto e relevância para a investigação; o volume de dados disponibilizados; e

---

<sup>4</sup> Definem-se como códices os livros manuscritos, anteriores à impressão (Faria & Pericão, 2008); como incunábulos os primeiros livros impressos entre 1455 e 1500 (Faria & Pericão, 2008); e como livro antigo as obras que foram impressas entre 1501 e 1800 (Faria & Pericão, 2008, p. 764).

o conjunto de soluções técnicas que apresentam, quer para edição digital, quer para a disponibilização e recuperação da informação. No intuito de diversificar a amostra, propõe-se o tipo de filiação e a localização como critérios adicionais. As duas bibliotecas digitais analisadas são: a Gallica, biblioteca digital da Bibliothèque nationale de France (BnF), na Europa; e a Internet Archive, biblioteca digital gerida por uma organização privada sem fins lucrativos, sediada em San Francisco, nos Estados Unidos da América.

As bibliotecas específicas para a investigação digital em Humanidades foram escolhidas em função do impacto e relevância para a investigação, da coerência da estrutura e da apresentação dos conteúdos, da respetiva creditação pela comunidade científica: a Perseus Digital Library, das mais antigas, criada em 1987, pelo departamento de estudos clássicos (Classics) da Tufts University, universidade privada localizada em Somerville/Medford, Massachusetts, nos Estados Unidos; a Bibliothèques Virtuelles Humanistes, desenvolvida, desde 2002, na Université François-Rabelais, em Tours, França, pelo Centre d'Études Supérieures de la Renaissance em colaboração com o Institut de Recherche et d'Histoire des Textes, e que, em 2014, recebeu o prémio Succeed, da União Europeia e do Impact Centre of Competence, pelo trabalho de pesquisa em digitalização de documentos históricos.

A avaliação das BDLA inclui um conjunto de procedimentos e instrumentos: observação direta (Jorgensen, 1998) e respetivo registo; organização dos dados; processamento da informação e análise crítica dos factos observados. Embora a observação direta seja uma técnica essencialmente aplicada ao comportamento humano, é aqui usada para a descrição e análise da operabilidade das bibliotecas

digitais, com idênticos objetivos: “Our analytic strategy [participant observation] is to identify and label a phenomenon in terms of its basic components” (Jorgensen, 1998, p. 108).

Neste âmbito, foi efetuada uma análise exploratória que permitiu a definição da amostra e a elaboração da grelha de observação, contribuindo também para a elaboração do guião da entrevista no grupo focal.

A observação como técnica de recolha de dados foi considerada adequada ao âmbito desta investigação, mesmo com o risco de as perceções do investigador serem influenciadas por fatores pessoais, reduzindo o índice de fiabilidade (Denscombe, 2014), obrigando a uma constante vigilância, objetividade e rigor na recolha. As vantagens desta técnica são: a recolha direta dos dados através da observação do meio natural (observação direta); a possibilidade de sistematização (observação sistémica); o posicionamento discreto do investigador, não ativo, nem participativo (observação não participante); produção de dados quantificados para análise; o registo dos dados em instrumentos idênticos (grelha de observação); a apresentação de relatórios (Denscombe, 2014; Gorman & Clayton, 2005). A ação participante levada a cabo limitou-se à pesquisa e navegação nos sítios eletrónicos, sem efetuar qualquer tipo de interação ou contacto com as instituições, à exceção dos pedidos de subscrição de listas, efetuados apenas depois de realizada toda a observação.

A observação consistiu numa consulta exaustiva dos sítios eletrónicos das bibliotecas digitais da amostra, feita na ótica do utilizador, onde se inclui a navegação nos diversos conteúdos, a exploração das funcionalidades, pesquisas simples e avançadas, a leitura das obras em linha, a transferência de ficheiros, a criação e a subscrição de listas.

O registo da observação foi feito através do preenchimento de uma grelha de observação (vd. 1.3.1 e Anexo II), na qual foram anotados os elementos de identificação da biblioteca e os vários componentes e funcionalidades que contribuem para a sua caracterização. Todos os dados recolhidos ao longo da observação direta foram atualizados entre julho e agosto de 2017.

Dado que esta técnica tem subjacente o princípio de “constant comparative method” (Glaser & Strauss, 2006), essencial para a compreensão do fenómeno, os vários itens das grelhas de observação foram analisados no sentido de identificar conexões e tendências, ou categorias, submetidas depois a uma análise comparativa entre as várias unidades analisadas. “This strategy involves taking one piece of data and comparing it with all others that may be similar or different in order to develop conceptualizations of the possible relations between various pieces of data” (Pickard, 2007, p. 241). Por conseguinte, os dados obtidos são processados, de forma a descrever cada uma das BDLA e a estabelecer relações de afinidade entre as unidades observadas, permitindo descrever os princípios genéricos que caracterizam as bibliotecas digitais.

Considerando que o objeto de estudo se centra em bibliotecas digitais que se caracterizam pela relevância dos acervos de livro antigo, optou-se por desenvolver uma triangulação, ou articulação de métodos de investigação (Flick, 2009). Günther (2006) defende que a triangulação consiste na utilização de diferentes abordagens e perspetivas do objeto empírico para prevenir a ocorrência de distorções decorrentes da utilização de um método único ou de uma investigação parcelar, enquanto Denzin e Lincoln (2008) consideram que o uso de múltiplos métodos pode contribuir para a

compreensão de um determinado fenómeno em profundidade. “The combination of multiple methodological practices, empirical materials, perspectives, and observers in a single study is best understood, then, as a strategy that adds rigor, breadth, complexity, richness, and depth to any inquiry” (Denzin & Lincoln, 2008, p. 7). Também neste ponto, se sublinha que os vários componentes de investigação estabelecem uma relação de complementaridade e não de alteridade ou dicotomia.

Para uma melhor compreensão do objeto desta investigação, o estudo de caso complementa-se com a análise do livro antigo, através dum estudo descritivo do processo de edição digital e da caracterização desta coleção na BNP e das edições digitais de Aldo Manuzio, enquanto a análise do livro antigo na BND segue procedimentos de análise quantitativa, com tratamento estatístico dos dados recolhidos (Huot, 2002, p. 60).

Foram recolhidos dados relativos a: dados bibliográficos (autor, título, local de produção, data, língua); características físicas das obras suscetíveis de influenciar a edição digital; características da edição digital. Os dados foram registados numa folha de cálculo no programa Microsoft Excel™, sendo, posteriormente, analisados e estruturados em padrões, através de tabelas e gráficos, proporcionando um relatório com a descrição das principais características observadas, extraíndo informações sobre a tendência central e a dispersão dos dados.

### **1.2.3. Grupo focal**

A avaliação abrange também as necessidades de informação sentidas pelos utilizadores reais ou potenciais destas bibliotecas (Reeves, Apedoe, & Woo, 2005), tendo em vista a otimização dos serviços oferecidos ou a disponibilizar. A investigação

“must take place under strict rules and standards that guarantee user access to high quality, appropriate-for-their-needs information” (Fuhr et al., 2007, p. [6]). Assim, uma das componentes fundamentais desta investigação é o levantamento das modalidades de utilização e das carências e expectativas de informação sentidas, ou não, pelos utilizadores das BDLA.

No âmbito deste estudo, em que se encontra definido o perfil do utilizador das BDLA, consideramos, como público-alvo, os investigadores em atividade na área das Ciências Sociais e Humanas. Para isso, torna-se necessário sondar as suas opiniões, confrontá-los com a realidade existente e apelar à apresentação de novas soluções. Pretende-se recolher informação acerca da metodologia de investigação em Ciências Sociais e Humanas e dos meios tecnológicos existentes, a fim de perceber em que medida isso se reflete nas bibliotecas digitais existentes, e de apresentar propostas para a respetiva otimização.

O tema das bibliotecas digitais, sendo ainda recente, tem sido objeto de estudo recorrente, produzindo um volume crescente e já substancial de informação teórica, metodológica e empírica. Porém, conforme se observa através da revisão da literatura, os estudos incidem sobretudo nas funcionalidades e na forma como disponibilizam os conteúdos, mas é mais incipiente no que respeita ao comportamento dos utilizadores e, em particular, dos investigadores em Humanidades, e à forma como a disponibilização da informação nas bibliotecas digitais se adequa aos parâmetros da sua atividade.

O grupo focal (*focus group*<sup>5</sup>) surge, neste contexto, como a técnica de recolha de dados mais adequada para efetuar este levantamento.

Focus group may be useful at virtually any point in a research program, but they are particularly useful for exploratory research when rather little is known about the phenomenon of interest. [...] Focus groups also have a place as a confirmatory method that may be used for testing hypotheses. (Stewart & Shamdasani, 2015, p. 44)

O grupo focal é uma técnica de investigação qualitativa utilizada em Ciências Sociais e estudos comerciais para recolha das opiniões ou atitudes de um determinado grupo, utilizando a comunicação verbal, e que permite observar e registar as experiências e as reações individuais dos participantes, possibilitando a análise das várias perspetivas sobre o mesmo problema (Krueger & Casey, 2009). Ao optar por esta técnica, pretende-se “[...] to enable a range of perceptions, feelings and attitudes from participants across a range of issues to be explored” (Pickard, 2007, p. 220).

A técnica de grupo focal consiste na reunião de um grupo de indivíduos, criteriosamente selecionados pelas suas características homogéneas ou heterogéneas e em função do tema a discutir, e um moderador que introduz as questões e dirige a discussão, em sessões informais, livres e abertas, as quais não devem exceder 1h:30min a 2h. O grupo deve ter entre seis e doze participantes (Gorman & Clayton, 2005; Krueger & Casey, 2009; Moore, 2006; Pickard, 2007). Os participantes são

---

<sup>5</sup> *Focus group* é a forma abreviada de *focused group* (grupo focalizado) ou *focused interview* (entrevista focalizada). Outras designações são: grupo de foco, grupo de discussão ou sessões de grupo.

selecionados em função de características comuns que os relacionam com o tema em análise.

Esta técnica permite: recolher dados qualitativos; gerar hipóteses que advêm de perspectivas e opiniões do grupo; obter uma maior abrangência em termos de tópicos; obter informação de retorno de estudos prévios (Stewart & Shamdasani, 2015, pp. 44–45). Pressupõe-se que, em complemento à informação obtida através da pesquisa documental, a discussão com os investigadores e com os técnicos abra outras perspectivas, novas e enriquecedoras.

Numa fase preliminar, foi feita uma sondagem (vd. 1.3.3 Guião do grupo focal; Anexo II) de carácter exploratório através da rede social Facebook, no grupo *Bibliotecas digitais*, e do blogue *Bibliotecas digitais para as Humanidades* (2013, 21 jun.), solicitando sugestões de pontos a debater no âmbito das BDLA e respetiva utilização. Esta etapa corresponde ao “brainstorm” definido por Krueger e Casey (2009) para “to throw out ideas that should be asked” (p. 52). Na sondagem, foram sugeridas algumas hipóteses, mas era permitido aos respondentes adicionar novas opções e observações. Este processo de prospeção prévia, junto de públicos segmentados com afinidades ao tema, permitiu aferir os assuntos considerados mais sensíveis ou polémicos neste domínio, mas também decifrar os pontos fortes e fracos dos sistemas atuais.

Os resultados obtidos refletem-se no guião da entrevista a aplicar no grupo de foco (vd. 1.3.3 Guião do grupo focal; Anexo IV), depois de múltiplos refinamentos, no sentido de uma maior coerência, sequência lógica e síntese.

O grupo focal foi constituído por investigadores em Humanidades, utilizadores das BDLA. O recrutamento dos participantes teve em conta a homogeneidade de formação

de base (Ciências Sociais e Humanas) e a heterogeneidade das especializações em História, Linguística, Comunicação, Estudos da Cultura, Antropologia e Património Cultural. Assim, os participantes foram selecionados em função do respetivo domínio de investigação na área das Humanidades: História e História da Arte, Comunicação, Estudos Linguísticos, Estudos da Cultura. Pela afinidade com este domínio, foram também incluídos investigadores na área das Ciências Sociais, no âmbito da Antropologia Cultural e da Etnologia. Os participantes do grupo focal foram recrutados nas seguintes instituições: Universidade de Lisboa; Universidade Nova de Lisboa; Universidade Católica Portuguesa; Universidade Europeia – Laureate International Universities. Os critérios subjacentes a esta seleção foi a representatividade de um largo espectro de domínios de investigação, bem como de instituições universitárias públicas e privadas de reconhecido mérito nestas áreas. Houve, ainda, a preocupação em garantir que os participantes não se conhecessem ou, nos casos em que pertenciam à mesma instituição, o seu contacto não fosse regular ou mantivessem um relacionamento de convívio ou de conflito, a fim de garantir a fiabilidade das respostas.

A reunião do grupo focal foi estruturada em quatro etapas:

- Introdução, aproximadamente 10 minutos: o moderador fez a apresentação do tema a discutir, dos objetivos e dos participantes, indicando as respetivas linhas de investigação, garantiu o anonimato das respostas e pediu autorização para efetuar o registo áudio da sessão;
- Desbloqueamento, aproximadamente 10 minutos: conversa informal, para introduzir a discussão;

- Discussão, aproximadamente 60 minutos: o moderador fez perguntas relacionadas com o objetivo principal do grupo focal, seguindo o guião da entrevista, mas permitindo algumas reflexões marginais, dentro do tema principal das bibliotecas digitais, mas não diretamente relacionadas com a edição digital do livro antigo;
- Conclusão, aproximadamente 10 minutos: o moderador apresentou um resumo da informação obtida, havendo espaço para a confirmação das ideias expressas pelos participantes.

Com a permissão de todos os participantes, a sessão foi gravada, pelo que não se considerou necessário proceder a anotações no decurso da entrevista. Por outro lado, não ocorreram quaisquer interferências externas, nem comportamentos ou reações que perturbassem a entrevista, ou alterassem o sentido da comunicação oral, e não fossem captadas no registo áudio.

O tratamento dos dados não difere do tratamento dado à entrevista, no sentido em que o carácter diferenciador do grupo focal reside na técnica de recolha dos dados e não na sua análise (Hoque, Parker, Covalski, & Haynes, 2017, p. 383). A discussão foi integralmente transcrita (Anexo V), à exceção de algumas exposições marginais (por ex., comentários relativos aos arquivos paroquiais) que, por serem extrínsecas ao tema, ficam apenas assinalados no registo oral. Da mesma forma, por não serem significativas, na transcrição não se inseriram expressões de oralidade (pausas, tautologias e redundâncias).

A codificação dos resultados pretende identificar padrões ou regularidades, os temas e as relações no discurso dos vários participantes, construindo afirmações genéricas

(Denscombe, 2007; Flick, 2009; Saldaña, 2009). Em termos genéricos, “to codify is to arrange things in a systematic order, to make something part of a system or classification, to categorize” (Saldaña, 2009, p. 8), o que se aplica ao tratamento de dados obtidos através da reunião de grupo focal (Saldaña, 2009, p. 100) e, nomeadamente, à sua transcrição (Saldaña, 2009, p. 132). Assim, a partir da transcrição, os dados foram estruturados em dois níveis: primeiro, fazendo a correspondência com as perguntas do guião, eliminando as partes menos significativas; segundo, por categorias, ou temas, dentro de cada uma das perguntas (Anexo VI).

Os resultados são apresentados em relatório (Krueger, 2002; Krueger & Casey, 2009), cuja estrutura segue a ordem das questões, em estilo narrativo e ilustrado por citações consideradas pertinentes, sem identificar o autor por questões éticas de confidencialidade assumidas no início da sessão. A partir deste levantamento, definir-se-á um conjunto de procedimentos a seguir pelas BDLA, tendo em vista a disponibilização da informação útil aos utilizadores e em moldes que estes considerem adequados.

#### **1.2.4. Inquérito por questionário**

A realização do inquérito foi necessária para validar os resultados obtidos no grupo focal.

O inquérito por questionário foi respondido por um painel de especialistas, no caso investigadores em Humanidades, assumindo que as suas opiniões são mais precisas, fundamentadas e credíveis do que as emitidas num grupo de não especialistas, à semelhança do que se aplica no método Delphi. “Its [Delphi study] object is obtain the

most reliable consensus of opinion of a group of experts” (Dalkey & Helmer, 1963, p. 458). Aplicaram-se os questionários em rondas sucessivas. Depois de obtidas as respostas da primeira ronda e analisados os dados obtidos, é feita nova consulta ao mesmo grupo, através de um novo questionário acompanhado dos resultados organizados da consulta anterior. Assim, a segunda ronda implica o tratamento prévio dos resultados da primeira. Embora o grupo consultado tenha a mesma composição, o questionário é alterado em função dos resultados anteriores, retirando as questões que tenham sido consideradas menos relevantes. O procedimento pode repetir-se as vezes consideradas necessárias até obter o nível de consenso pretendido, permitindo a elaboração de conclusões e a análise quantitativa dos dados obtidos.

Previamente, aplicou-se o questionário-piloto a um pequeno grupo de investigadores em Humanidades e utilizadores de BDLA para comprovar a precisão, clareza e imparcialidade das perguntas, a partir do qual se fizeram os ajustes considerados necessários na respetiva formulação. Este questionário serviu, igualmente, para estabelecer o grau de previsibilidade das opções de resposta e eliminar as perguntas que não conduzissem a resultados relevantes, mas de forma a evitar a exclusão de possibilidades relevantes (Ampudia de Haro et al., 2016; Bell, 2004; Hill & Hill, 2002).

Os questionários (vd. 1.3.3 e 1.3.4 Questionário; Anexo VII e VIII) foram elaborados no sistema SurveyMonkey (<https://pt.surveymonkey.com/>). Esta solução, entre as várias que se encontram disponíveis em linha, foi escolhida por ser fiável e de fácil utilização e permitir a realização de pré-avaliação, visualização e teste do questionário. A SurveyMonkey compromete-se a não coletar informações pessoais ou dados de terceiros sem autorização, nem a facultar os resultados a terceiros, garantindo que as

pesquisas são administradas, exclusivamente, pelos respectivos criadores. A criação do questionário nesta plataforma é amigável, beneficiando de ajuda interativa. O aspeto gráfico do questionário é personalizável. Esta plataforma permite, ainda, avaliar o questionário e enviá-lo a respondentes para teste, sem gravar os dados.

O questionário feito através do SurveyMonkey que calcula o tempo estimado para o concluir e permite um número ilimitado de pesquisas, de perguntas de vários tipos e de participantes. O inquérito foi partilhado em linha, através de uma hiperligação para os endereços de correio eletrónico. As respostas ficam acessíveis em tempo real, com uma análise automática dos resultados em gráficos, sugerindo o tipo mais adequado. Adicionalmente, optou-se por transferir as respostas individuais, os dados de resumo e os respetivos gráficos para arquivos Microsoft Excel™.

Assim, a análise dos resultados, em parte, foi fornecida pelo sistema. Os dados das baterias de perguntas que constituíam o corpo do questionário, obtidos na escala Likert, foram processados por meio de estatística descritiva usando tabelas e gráficos (Günther, 2003), com o suporte informático do SurveyMonkey e do Microsoft Excel™.

Os dados foram tratados com a confidencialidade assumida, sendo que o próprio sistema não permite a identificação dos respondentes e os dados crus, não processados, não foram recolhidos o que equivale à respetiva destruição (Moore, 2006).

O inquérito por questionário possui a vantagem de não haver interação entre os intervenientes, de poder ser efetuado à distância, do conhecimento da área pelos especialistas e de ajudar os participantes a estarem focados nas questões quando estão a responder. A participação é anónima, o que permite uma maior abertura por

parte dos intervenientes. Permite, ainda, a avaliação estatística das respostas, para identificar os padrões de concordância.

Pickard (2007) estabelece as seguintes etapas para a correta aplicação do inquérito por questionário: identificar o assunto principal do inquérito em função do projeto, realizar o estado da questão relativamente ao tópico eleito, na literatura, estabelecer hipóteses ou definir objetivos, identificar a população adequada, escolher a técnica que melhor se adapte aos respondentes, desenhar o instrumento de recolha de dados e, em seguida, testar, recolher os dados, analisar, apresentar resultados e extrair as respetivas conclusões (Pickard, 2007, pp. 97-100).

Em conformidade com estes princípios, foram convidados a integrar o painel apenas investigadores em Humanidades (Ciências Sociais e Humanas), com carreira de investigação científica reconhecida e ativa. A participação no painel foi mantida anónima e nenhum dos participantes conhecia a identidade dos restantes membros. Embora os resultados da primeira ronda fossem bastante expressivos e consensuais, foi efetuada uma segunda ronda, eliminando os pontos considerados menos relevantes. Os dados coletados recolhidos na primeira ronda foram registados e organizados de forma a ser entregues com o questionário da segunda ronda.

### **1.3. Instrumentos da investigação**

Para responder às questões de investigação há que recolher dados. Para tal o investigador recorre aos instrumentos de investigação. Foram construídas duas grelhas de observação, uma para as bibliotecas digitais e outra para os incunábulo e livro antigo, um guião para o grupo focal e um questionário para o validar.

### 1.3.1. Grelha de observação das BDLA

As bibliotecas digitais, enquanto sistemas de informação, podem e devem ser avaliadas. Além disso, como são realidades complexas e estão em constante evolução, acompanhando as mudanças das tecnologias de informação e comunicação, devem ser avaliadas sob várias perspetivas (Fuhr et al., 2007; Reeves et al., 2005), mas partindo da premissa de que o seu propósito é servir os utilizadores, facilitando o acesso à informação e a eficácia dos serviços (Xie, 2006).

Segundo o modelo de *DELOS Network of Excellence* (Candela et al., 2007), podem considerar-se os seguintes parâmetros de avaliação das BDLA:

- Facilidade de utilização (*usability*) – qualidade de interação entre o utilizador e o sistema (*user-system*);
- Utilidade – relevância dos conteúdos (*user-content*);
- Performance – relação entre conteúdos e sistema “o desempenho do sistema depende fortemente dos formatos, estruturas e representações do conteúdo” (*content-system*).

No que respeita à tipologia das BDLA, a grelha de observação inclui dados relativos à forma de construção, às instituições contribuidoras, à quantidade e tipologia de conteúdos, às condições de utilização e reutilização das obras e à disseminação do conhecimento.

Assim, a grelha de observação (Anexo II) apresenta três componentes, relativos à identificação, aos conteúdos e às funcionalidades.

Na identificação são registados os seguintes itens: designação; instituição; país; endereço eletrónico; data de início.

Nos conteúdos, inclui-se informação relativa a: número de obras; número de livros; número de espécimes de livro antigo; estatísticas de obras disponibilizadas; obras com verbetes ou sumários.

Nas funcionalidades, inclui-se dados relativos a: navegação (por autor, título, data, assunto, totalidade da coleção, ou tipo); pesquisa (estática e dinâmica); formatos de disponibilização dos ficheiros (JPG; PDF; PDF com texto pesquisável; GIF; PNG; ePub; Dejavu); partilha (por correio eletrónico; através das redes sociais); tutoriais; ajuda interativa; tipologia de menus (vertical; horizontal); página de novidades; gestores de referências; alertas (primeira página; personalizáveis); estatísticas de uso; informação técnica.

### **1.3.2. Grelha de observação dos incunábulos e livro antigo**

Foi elaborada uma grelha de avaliação para os incunábulos e livro antigo (Vd. Anexo I). Primeiro, incluíram-se os elementos da ficha bibliográfica das obras como língua, local e data de edição; em seguida, foram avaliados os aspetos materiais das obras, como o tipo de letra, a foliação, as assinaturas, a existência ou não de ilustrações ou desdobráveis, as capitais iluminadas ou decoradas, os caldeirões ou corações, os encabeçamentos, os reclusos; por fim, foram analisados a edição digital, os sumários e a pesquisa textual nos documentos.

### **1.3.3. Guião do grupo focal**

O guião do grupo focal foi precedido de uma sondagem exploratória (Vd. Anexo II) em forma de tabela com um elenco de 10 opções, ou hipóteses, de escolha e um campo

aberto para comentários. Os respondentes podiam responder a mais do que uma opção e acrescentar opções.

Os campos tabelados eram relativos às funcionalidades das BDLA: pesquisa (através do catálogo em linha; através dos motores de busca); disponibilização das obras (formato texto; formato imagem; formato imagem e texto); hiperligações a bibliografias; apontadores para o conteúdo da obra; navegação por texto e imagem; pesquisa nas imagens; exportação direta para gestores de referências (Mendeley, Zotero, etc.). Um respondente acrescentou uma opção relativa à apresentação das obras em formato acessível (norma W3C).

As questões do guião do grupo focal foram sendo reafinadas, a partir da triangulação entre a revisão da literatura, a observação das BDLA e os resultados obtidos na sondagem, quer através das opções votadas, quer dos comentários.

O guião (Vd. Anexo IV) foi estruturado em duas partes: introdução, com a apresentação do tema da investigação, os procedimentos e objetivos do grupo focal, a duração prevista e considerações éticas (Pickard, 2007, p. 221), relativas à confidencialidade das respostas e utilização dos dados; seis perguntas abertas. “Open-ended questions allow the respondents to determine the direction of the response. The answer is not implied and the type or manner of response is not suggested” (Krueger & Casey, 2009, p. 53).

As perguntas incidiam sobre três componentes: expectativas do utilizador ao utilizar uma BDLA; perceção das BDLA (vantagens e inconvenientes das BDLA face às bibliotecas físicas; vantagens para o ensino e para a investigação); propostas de melhoria (funcionalidades necessárias para a maior eficácia das BDLA).

Ao longo da reunião do grupo focal, o moderador seguiu o guião de forma flexível, dando liberdade de resposta aos respondentes, sem prejuízo do recurso a estímulos, ou comentários circunstanciais, para manter a discussão ou reconduzi-la ao tema.

#### **1.3.4. Questionário**

Os inquéritos por questionário, enviados a um grupo pré-selecionado de especialistas, são o instrumento convencional para a obtenção de dados.

As perguntas da primeira ronda foram formuladas na sequência da reunião do grupo focal.

Após a introdução, o questionário (Anexo VII) foi estruturado em quatro blocos de perguntas relativas ao tema e um bloco final de dados sociodemográficos. A estrutura e a sequência do questionário foram elaboradas segundo uma ordem lógica, do geral para o particular, e agrupando as questões por assuntos e temas (Denscombe, 2007; Günther, 2003). A formalização das questões teve em consideração os seguintes aspetos: simplicidade e brevidade; precisão e clareza; imparcialidade; discrição (Ampudia de Haro et al., 2016).

As informações prestadas na introdução incluíam: uma explicação sintética da pesquisa, referindo a importância da colaboração; a identificação do investigador e das instituições que enquadram a investigação; a garantia de anonimato e confidencialidade das respostas; e, dado que o questionário se destinava a ser preenchido em linha, uma explicação sobre o funcionamento do questionário.

Cada bloco de perguntas correspondia a um enunciado:

1. Ao aceder a uma biblioteca digital de livro antigo o que espera encontrar?

2. Quais as vantagens da biblioteca digital de livro antigo, face às bibliotecas físicas?
3. Quais as vantagens da biblioteca digital de livro antigo, para o ensino e para a investigação?
4. Quais as suas expectativas de mudança/otimização nas bibliotecas digitais de livro antigo?

Optou-se por baterias de perguntas fechadas com recurso a escala de medida nominal do tipo Likert, solicitando ao respondente que avaliasse a opção de resposta proposta numa escala com 5 opções: na pergunta 1, “nunca”, “quase nunca”, “por vezes”, “quase sempre”, “sempre”; e, nas restantes, “discordo totalmente”, “discordo em parte”, “indiferente”, “concordo em parte”, “concordo totalmente”. Em cada pergunta, só era permitida a escolha de uma opção. A ordenação da escala em sentido positivo foi mantida em todas as baterias dos quatro blocos. A cada bloco, correspondia um número variável de perguntas: no primeiro, 8; no segundo, 11; no terceiro, 5; e, no último, 18. Dada a especificidade de cada uma das perguntas, não se considerou relevante o risco de pseudo-opiniões, devido ao carácter repetitivo das baterias de perguntas.

No bloco relativo aos dados sociodemográficos, era solicitada informação relativa à idade, ao género e à formação académica, em perguntas de escolha múltipla, e à área de investigação, em pergunta de escolha múltipla, com possibilidade de preenchimento de um campo em aberto.

Para a segunda ronda, o questionário (Anexo VIII) foi reformulado e sintetizado em função do peso das respostas. Para calcular este valor, foi atribuída, a cada opção da

escala Likert, um peso numérico de 1 a 5, no sentido ascendente do grau frequência (pergunta 1) ou de concordância (perguntas 2 a 4). Depois de hierarquizadas as respostas por ordem de relevância, foram excluídas perguntas dos blocos 1 e 3, que ficaram, respetivamente, com 3 e 4 perguntas. Em cada um dos blocos, a ordem das perguntas foi alterada da mais pontuada para a menos pontuada. Associado ao questionário, foi implementado um glossário (Anexo IX) dos termos que, feita a análise das respostas, poderiam ter comprometido a compreensão das perguntas. O modelo, a estrutura e o modo de implementação do questionário na web não foram alterados. A introdução do questionário foi alterada, para apresentar uma síntese dos resultados obtidos na primeira ronda e uma informação relativa ao glossário, com a respetiva ligação. No texto das perguntas havia, também, hiperligação entre os termos e o correspondente verbete no glossário. Os dados sociodemográficos foram, igualmente alterados em função das respostas obtidas.

#### **1.4. Amostra**

O universo em estudo é constituído por todas as ocorrências; a população está incluída no universo, mas limita-se a um conjunto com determinadas características; e, por fim, a amostra corresponde à parte da população sobre a qual incide o estudo (Ampudia de Haro et al., 2016, p. 133). Ou como afirma Clara Coutinho “É um subconjunto da população que terá de representar ou seja reflectir os seus traços” (2011, p. 85). Foram definidas quatro amostras, relativas às áreas a investigar: os incunábulos e livro antigo; bibliotecas digitais; as obras de Aldo Manuziu e os investigadores em Humanidades.

#### **1.4.1. Bibliotecas digitais**

A seleção das BDLA a analisar será feita em função dos seguintes parâmetros:

- Disponibilizam coleções próprias digitalizadas de livro antigo;
- Possuem representatividade local, nacional ou internacional.

Encontram-se, dentro destes parâmetros, os projetos: a Alma Mater (biblioteca digital da Universidade de Coimbra), a Biblioteca Nacional Digital, a Gallica (Biblioteca Nacional Digital francesa), o Internet Archive, as Bibliothèques Virtuelles Humanistes e o Perseus Digital Library.

#### **1.4.2. Incunábulos e livro antigo**

Para se conhecer as características do livro antigo, efetuou-se o levantamento das obras disponibilizadas em linha pela Biblioteca Nacional Digital<sup>6</sup>, datadas entre 1450 e 1800 (Vd. Anexo I), onde se inclui o acervo disponibilizado em linha pela Faculdade de Letras da Universidade de Lisboa no âmbito do projeto Patrimónia, num total de 7.500 títulos. A fim de limitar o estudo, o arco temporal considerado foi de 1450 a 1550, os primeiros 100 anos da impressão com caracteres móveis. Na classificação “Livro Antigo”, foram obtidos 368 resultados, englobando folhetos e monografias. Excluídos os folhetos, foram selecionadas as 265 obras que constituem esta amostra.

#### **1.4.3. Edição digital das obras de Aldo Manuzio**

As obras de Aldo Manuzio foram selecionadas no grupo das seis bibliotecas digitais analisadas. A Alma Mater não inclui nenhuma obra de Manuzio na coleção. A obra

---

<sup>6</sup> Disponível em <http://purl.pt/>

existente no Perseus não foi considerada nesta amostra, dado que lhe é feita a transcrição do texto com o tratamento corrente, sem acrescentar significativamente ao que fica reportado na análise desta biblioteca.

Foram encontradas 71 obras, distribuídas da seguinte forma: 1 obra, na BND; 16 na Gallica; 3 nas BVH; 51 no IA.

#### **1.4.4. Caracterização dos grupos focal e do inquérito por questionário**

A sondagem prévia (Anexo II) à reunião do grupo focal foi lançada no grupo de Facebook “Bibliotecas Digitais” e no blogue *Bibliotecas digitais para as Humanidades*, cujos membros e leitores constituem a população a partir da qual se recrutou a amostra constituída por 107 respondentes no Facebook e 60 respondentes no blogue. Responderam à sondagem um total de 167 indivíduos.

O grupo focal foi constituído por 6 investigadores em Humanidades (Ciências Sociais e Humanas), 3 do género feminino e 3 do masculino e com idades compreendidas entre os 40 e os 60 anos. Todos os participantes mantêm ativa a investigação ligada a universidades: Universidade de Lisboa (1); Universidade Nova de Lisboa (1); Universidade Católica Portuguesa (2); Universidade Europeia – Laureate International Universities (2).

O inquérito por questionário foi convocado por correio eletrónico (Anexo X). Na primeira ronda, participaram 19 respondentes com os graus de Mestre (5), Doutor (13, 4 dos quais com pós-doutoramento), com idades acima dos 30 anos, dos quais 11 do género feminino e 7 do masculino, tendo havido um participante que não assinalou nenhuma opção nos campos relativos à formação académica e ao género. As áreas de

investigação compreendidas nesta amostra são: História e Arqueologia; Línguas e Literatura; Arte, na especialidade de História da Arte; Ciências da Comunicação; Humanidades Digitais e Humanística Digital; Filosofia, Ética e Religião; e Sociologia (Antropologia e Etnologia); Ciências da Comunicação e da Informação; Estudos de Cultura; e Museologia e Património. Na segunda ronda, participaram 17 respondentes, com os graus de Mestre (3), Doutor (9), ou Pós-Doutorado (5), com idades acima dos 30 anos, dos quais 11 do género feminino e 6 do masculino. As áreas de investigação compreendidas nesta amostra são: História e Arqueologia; Línguas e Literatura; Arte, na especialidade de História da Arte; Ciências da Comunicação; Humanidades Digitais e Humanística Digital; Filosofia, Ética e Religião; e Sociologia (Antropologia e Etnologia); Ciências da Informação; Estudos de Cultura; e Museologia e Património.

## 2. Humanidades Digitais

### HUMANIDADES DIGITAIS

As Humanidades Digitais aliam um novo método de investigação e uma nova forma de divulgação da informação científica, com vista ao reforço, junto da população em geral e dos decisores em particular, da imagem das Humanidades e da sua importância para a sociedade, fomentando o trabalho colaborativo e a reutilização da informação produzida.

A proliferação dos *media* baseados na *web*, a disponibilização de um crescente volume de conteúdos, a emergência de técnicas de análise automatizada de conteúdos e de ambientes de visualização e a computação na nuvem, entre outros, têm vindo a contribuir para uma revolução epistemológica. Ou seja, a era digital veio alterar o modo como se processa a recolha, a produção, a análise e a divulgação da informação. O computador é assumido como uma ferramenta essencial para ordenar, interrogar e analisar os dados disponibilizados a uma escala cada vez mais alargada (Greenhalgh, 2004).

As Humanidades Digitais visam ajudar a interpretar o impacto cultural e social desta nova realidade. Por um lado, respondem às questões históricas e filológicas que daí emergem (“What Is DH?,” s.d.) e, por outro lado, disponibilizam informação certificada de forma apelativa à população em geral, enquanto favorecem a divulgação e o estudo académico das Humanidades.

O projeto *Index Thomisticus*, desenvolvido desde 1949, por Robert Busa, S.J. (28 de novembro de 1913 – 9 de agosto de 2011) (Figura 1), é considerado a primeira experiência no âmbito da aplicação da computação aos estudos linguísticos, dado que

consistia na indexação e lematização (isto é, a identificação da raiz das palavras) das obras de S. Tomás de Aquino. Este projeto impulsionou o desenvolvimento de um novo ramo de investigação: a computação para as Humanidades (*humanities computing*) (Hockey, 2004, p. 4), primeira designação para as Humanidades Digitais. Além de ser pioneiro, a sua relevância foi também imediatamente reconhecida. Em 31 de dezembro de 1956, o trabalho foi referido num artigo publicado na revista *Time* com o título (“Religion: Sacred electronics” 1956).

Figura 1 – Robert Busa na consola de controlo do IBM 705. Nova Iorque, 1958. [IBM Archives].  
Fonte: <http://priestandpunchedcards.tumblr.com/post/118312350925/roberto-busa-at-the-control-console-of-the-ibm>



Em 1980, foram publicados os 56 volumes da versão impressa do *Index Thomisticus* e, em 1992, foi editada a primeira versão em CD-ROM. Em 2005, foi disponibilizada a versão web<sup>7</sup>, que continua em linha. Entre o início do trabalho e a sua disponibilização em linha, passaram mais de cinquenta anos, pelo que o projeto atravessou todas as

---

<sup>7</sup> Disponível em <http://www.corpusthomicum.org/>

etapas da tecnologia computacional, desde as fichas de papel perfurado até á difusão digital dos textos, sendo que as várias fases do trabalho foram sendo documentadas através de registos fotográficos e divulgadas em artigos (Burdick, Drucker, Lunenfeld, Presner, & Schnap, 2012; Busa, 1980, 2004; Dacos & Mounier, 2014; Gonçalves & Banza, 2013; Hockey, 2004; “Religion: Sacred Electronics,” 1956; Sousa, 2012).

Segundo Kirschenbaum (2012, p. 5) e Fitzpatrick (2012, p. 13), a designação *Digital Humanities* (em português, Humanidades Digitais) é utilizada pela primeira vez em 2004, na obra *Companion to Digital Humanities* (Schreibman et al., 2004a), em alternativa a *humanities computing*, associado à noção restrita de digitalização e, assim, evoluindo de um mero serviço para uma nova prática (Hayles, 2012, p. 43). Humanidades Digitais é um conceito que envolve, não só os processos de execução, mas todas as questões relacionadas com a produção e distribuição em linha (Terras, Nyhan, & Vanhoutte, 2013).

Em termos genéricos, as Humanidades Digitais englobam o conjunto de pesquisas e experiências que visam facilitar a utilização dos recursos digitais no âmbito das Ciências Sociais e Humanas, tornando-os mais intuitivos e acessíveis. Os signatários do *Manifesto*<sup>8</sup>, consideram que é esta uma “[...] transdisciplina, portadora dos métodos, dos dispositivos e das perspectivas heurísticas ligadas ao digital no domínio das Ciências humanas e sociais” [...] (*Manifeste des Digital Humanities*, 2010)<sup>9</sup>.

---

<sup>8</sup> O manifesto surge na sequência do ThatCamp 2010, em Paris, uma não conferência sobre as Humanidades Digitais. Atualmente os signatários, investigadores maioritariamente europeus e respetivas instituições, somavam, até final de 2011, 293 individuais e 10 instituições. A 03-03-2014, encontram-se registados 3.155 pessoas.

<sup>9</sup> A versão original é em francês, mas encontra-se traduzida para português, por Hervé Théry, no sítio eletrónico do ThatCamp, disponível em <http://tcp.hypotheses.org/497>

As Humanidades Digitais, mais do que disponibilizar conteúdos em linha, implicam uma mudança na forma de trabalhar. “Computerization provides humanities with new qualitative dimensions” (Busa, 1980, p. 88). O objetivo das Humanidades Digitais é muito maior do que a transferência de suporte, centrando-se no desafio epistemológico (Gonçalves & Banza, 2013, p. 5), incidindo na forma como se alcança o conhecimento.

As Humanidades Digitais envolvem, ainda, uma alargada comunidade de investigadores noutras áreas afins ou complementares das Humanidades: “disciplinary experts, computer scientists, and library and information studies specialists, have been brought together to consider digital humanities as a discipline in its own right, as well as to reflect on how it relates to areas of traditional humanities scholarship” (Schreibman et al., 2004b, p. XXIII).

O conceito de Humanidades Digitais procura conciliar os conhecimentos e os métodos utilizados nas Ciências Sociais e Humanas com o mundo digital. Numa primeira fase, as ações no âmbito das Humanidades Digitais centraram-se na digitalização e na disponibilização de fontes primárias, começando a impor-se o objetivo de construir e facultar ferramentas para a análise e a visualização dessas fontes, para que a aquisição cognitiva seja mais imediata e intuitiva. “The themes that emerged can be grouped under the following rubrics: scale, critical/productive theory, collaboration, databases, multimodal scholarship, code, and future trajectories” (Hayles, 2012, p. 43). Além disso, a aposta nas Humanidades Digitais tem, também, motivações económicas, (Kirschenbaum, 2012).

Desvalorizadas desde há décadas pela cultura do rendimento imediato, das falácias do “empreendedorismo”, “empregabilidade” e da utilidade e consumo imediatos, as tradicionais Humanidades encontram nas Humanidades Digitais um terreno bastante promissor, com impacto directo na preservação e divulgação do património, mas também na economia.

(Gonçalves & Banza, 2013, p. 5)

Por conseguinte, tem vindo a assumir-se o objetivo de rentabilizar os investimentos realizados ao longo dos anos, tanto na digitalização massiva dos acervos, como na própria investigação. Também por esse motivo, procura ampliar-se a divulgação dessa produção científica, assumindo que se trata de fonte credível e certificada para a investigação.

As Humanidades Digitais consolidam-se como um novo campo de saber e, como tal, apresentam os indicadores de um novo campo firmado através de numerosas associações, centros de estudos, departamentos, programas de ensino, revistas especializadas e congressos:

Existen varios indicadores de la consolidación de un nuevo campo de estudio. Entre ellos: la formación de asociaciones, organizaciones y centros especializados, la creación de programas académicos de enseñanza, la celebración de congresos y la publicación de revistas y libros especializados. El campo de las Humanidades Digitales se consolida cada vez más y existe a nivel internacional una comunidad importante que se identifica como «humanista digital». (Galina Russell, 2011a, p. 4)

Esta nova área de investigação e desenvolvimento, transversal a áreas como as Humanidades e as Ciências da Informação e da Computação (Borgman, 2009; Evans & Rees, 2012; Schreibman et al., 2004b), orienta uma parte substancial da pesquisa para a identificação, a definição e a descrição dos métodos, bem como das boas práticas e das normas que os sustentam.

## **2.1. Desenvolvimento das Humanidades Digitais**

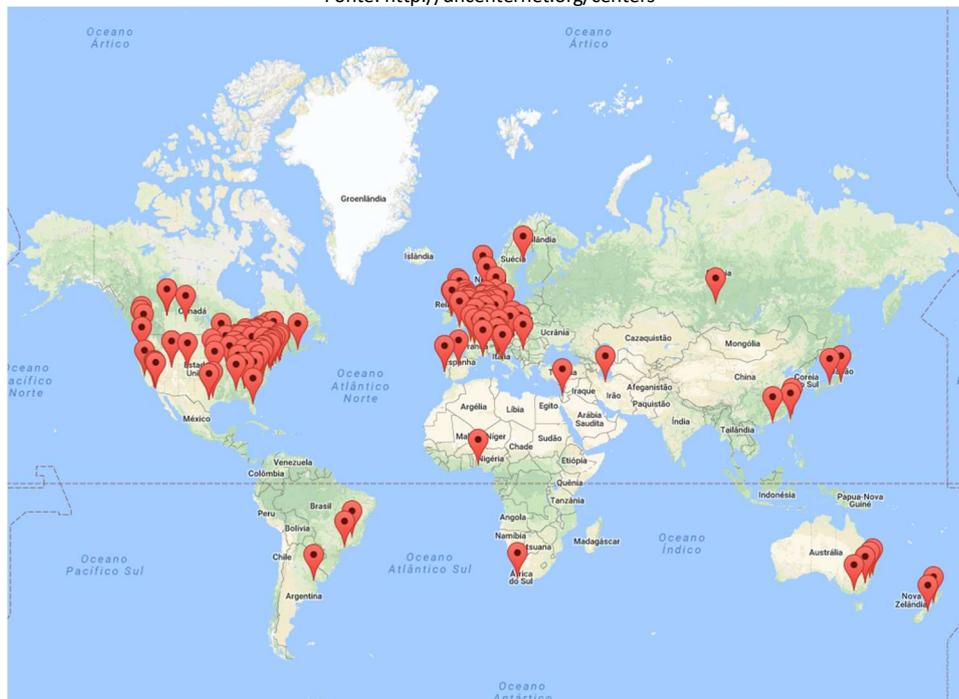
A utilização dos computadores nos trabalhos de investigação em Humanidades acompanhou a evolução tecnológica. À medida que os computadores e os respetivos programas ficavam mais amigáveis, passando a constituir-se como uma ferramenta comum e cada vez mais indispensável no quotidiano, banalizando-se inclusivamente no uso doméstico, a sua utilização pelos investigadores das Ciências Sociais e Humanas também crescia, ao mesmo tempo que aumentavam e se tornavam mais complexas as suas exigências e expectativas em relação ao mundo digital (Hockey, 2004).

A aplicação das Humanidades Digitais tem sido potenciada no âmbito da educação (Greenhalgh, 2004). Em particular, as universidades com ensino no âmbito das Humanidades têm vindo a reconhecer a relevância deste domínio científico através da criação de departamentos de Humanidades Digitais. “Il est un point sur lequel les États-Unis ont pris une avance considérable, toutefois, c’est la construction des équipes de recherche en humanités numériques, grâce à la création de structures très particulières: les centres d’humanités numériques” (Dacos & Mounier, 2014, p. 24). Os centros de estudo constituem o grande motor de desenvolvimento das Humanidades Digitais. Entre outros, podem destacar-se os centros de estudo dedicados às Humanidades Digitais das universidades de Harvard, de Stanford e de Maryland, nos

Estados Unidos da América, e o King's College London, as universidades de Oxford e de Sheffield e o University College London, no Reino Unido. O levantamento dos centros de estudo que privilegiam ou que são dedicados às Humanidades Digitais ainda não está completamente realizado. Em cada trabalho, obtém-se uma imagem parcelar. No entanto, existem vários estudos que concorrem para uma imagem de conjunto. O centerNet<sup>10</sup>, que “établissait une belle carte des centres spécialisés en Digital Humanities, très riche et utile, mais se positionnant au niveau des organisations et non des individus” (Dacos, 2013), em 2014, referenciava 185 centros, concentrados em 19 países (Guerreiro & Borbinha, 2014, p. 66), sendo que a grande maioria se localizava na América do Norte e Europa, enquanto os centros da América Central e do Sul, África e Ásia, praticamente, não eram referenciados (Figura 2).

Figura 2 – Distribuição global dos centros de estudos em Humanidades Digitais. CenterNet, 2017.

Fonte: <http://dhcenternet.org/centers>



<sup>10</sup> Disponível em <http://www.dhcenternet.org/>

Portugal e Espanha apenas possuem uma referência cada um, o que não corresponde à realidade. Este sítio encontra-se em permanente atualização, mas, em 2017<sup>11</sup>, os números permanecem idênticos, referenciando 193 centros. Em Portugal, são referenciados dois, mas, neste caso, há uma duplicação de registos, dados que ambos se referem ao Instituto de História Contemporânea (IHC), em Lisboa.

O trabalho de Melissa Terras (2012), publicado em 20 de janeiro de 2012 e apresentado na forma de infográfico foi, também, alvo de algumas críticas, por parte da comunidade francófona:

[...] la proposition de quantification des Digital Humanities réalisée par Melissa Terras en 2011 ne privilégiait pas les facteurs géographiques et linguistiques. Elle se concentrait, de facto, sur des dispositifs, organisations et indicateurs essentiellement anglophones. Elle décrivait ses Digital Humanities et j'avais l'intuition que ses Digital Humanities n'étaient pas mes Digital Humanities, ne serait-ce que parce qu'elle ignorait les listes de discussion non anglophone ou les financements hors UK et ÉU. (Dacos, 2013)

Embora haja uma discrepância de números, dado que Melissa Terras (2012) regista 114 centros em 24 países, mantém-se a disparidade geográfica. A maioria dos centros encontra-se na América do Norte (11 no Canadá e 44 nos Estados Unidos da América), seguida da Europa, com 41 centros (14 dos quais no Reino Unido), e da Austrália, com 7 centros. Não está referenciado nenhum centro em Portugal. Na América do Sul, África e Ásia são referenciados apenas 10 centros, o que é quase residual em comparação com a área geográfica e a população envolvidas.

---

<sup>11</sup> Dados atualizados em 07/08/2017.

O projeto “Atlas de Ciencias Sociales y Humanidades Digitales”, do GrinUGR<sup>12</sup> (Grupo de Internet de la Universidad de Granada), da Universidade de Granada, em 2014<sup>13</sup>, tinha registados 29 centros de estudo, 22 na Europa, 4 na América do Norte e 3 na América do Sul. Este levantamento tem o inconveniente de a submissão ter sido autoproposta. Em 2017<sup>14</sup>, encontram-se referenciados 31 centros de estudos e 97 projetos. Em Portugal, continua a não haver centros registados, embora estejam 6 projetos referenciados.

Num artigo publicado por Marin Dacos, em 2013<sup>15</sup>, é apresentado um estudo sobre as Humanidades Digitais que, através do levantamento dos investigadores que se assumem como humanistas digitais e das línguas utilizadas, fornece uma cartografia um pouco mais abrangente, emergindo o trabalho que está a ser desenvolvido na América Latina (Brasil, México e Perú). A Europa aparece representada com mais detalhe, o que se justifica pelo contexto de origem deste trabalho.

Um ano depois, num artigo intitulado “Una visión de las humanidades digitales a través de sus centros”, Romero Frias e Del-Barrio-García (2014) tentam efetuar o rastreio dos centros existentes, do trabalho desenvolvido com vista à caracterização das Humanidades Digitais. No entanto, a distribuição geográfica dos centros mantém os mapas anteriores, uma vez que foi realizado com base nos dados do CenterNet.

Analisando os elementos comuns aos vários estudos, todos apontam para a existência de um maior número de centros na América do Norte, Reino Unido, França e

---

<sup>12</sup> Disponível em <http://grinugr.org/>

<sup>13</sup> Dados recolhidos em 11/03/2014.

<sup>14</sup> Dados atualizados em 07/08/2017.

<sup>15</sup> Este trabalho foi a resposta ao repto lançado no DH2012 em Hamburgo. Foi disponibilizado, um inquérito em linha em quatro línguas (Francês, Inglês, Alemão e Espanhol) durante seis meses, de março a outubro de 2012. Obtiveram 851 respostas de 55 países (Dacos, 2013).

Alemanha. No entanto, o rastreio das outras zonas do globo ainda está, maioritariamente, por fazer.

A par dos centros de estudos, as associações criadas em torno das Humanidades Digitais têm vindo a ganhar relevância e a manter ativa a investigação e a reflexão neste domínio. As associações “[...] savantes sont nombreuses dans le domaine des humanités numériques” (Dacos & Mounier, 2014, p. 34). De uma forma genérica, a missão das várias associações é promover e apoiar a investigação e o ensino das Artes e Humanidades no âmbito das Humanidades Digitais, incentivando e promovendo a publicação, a colaboração e a formação. No entanto, mantêm-se as assimetrias no que se refere às respetivas áreas de influência geográfica.

Figura 3 – Símbolo ADHO



A Alliance of Digital Humanities Organizations (ADHO) é o consórcio mundial destas organizações. O logotipo possui fundo branco e negro, sobre o qual sobressaem as iniciais “ADHO” nas cores azul, amarelo, verde e vermelho, evocando o caráter agregador da sua missão e o âmbito alargado dos seus projetos, dado que todas as bandeiras do Mundo possuem pelo menos uma destas cores, pretendendo, com isso, significar que todas as nações estão aqui representadas (Figura 3).

The Alliance of Digital Humanities Organizations (ADHO) is an umbrella organisation whose goals are to promote and support digital research and teaching across arts and humanities disciplines, drawing together humanists engaged in digital and computer-assisted research, teaching, creation,

dissemination, and beyond, in all areas reflected by its diverse membership.

(“ADHO - Alliance of Digital Humanities Organizations,” s.d.)

As organizações que fazem parte da ADHO são também consórcios de organizações de carácter mais localizado, nomeadamente:

- European Association for Digital Humanities (EADH);
- Association for Computers and the Humanities (ACH);
- Canadian Society for Digital Humanities/Société canadienne des humanités numériques (CSDH/SCHN);
- centerNet;
- Australasian Association for Digital Humanities (aaDH);
- Japanese Association for Digital Humanites (JADH).

A ADHO permite o acesso livre à maioria das suas publicações (jornais, atas de conferências, livros, manuais, etc.) e disponibiliza, igualmente, informação referente aos eventos, passados e futuros, e às investigações em desenvolvimento. No entanto, o acesso à revista LLC/DSH é pago, sendo através da respetiva assinatura que as instituições ou elementos individuais se associam à ADHO.

## 2.2. Infraestruturas Europeias

### INFRAESTRUTURAS EUROPEIAS

- **ESF** - European Science Foundation;
- **NeDiMAH** - Network for Digital Methods in the Arts and Humanities;
- **CLARIN-ERIC** - Common Language Resources and Technology Infrastructure;
- **DARIAH** - Digital Research Infrastructure for the Arts and Humanities (ROSSIO - português);
- **CESSDA** - O Council of European Social Science Data Archives.

A nível europeu, com o objetivo de fomentar a investigação e a divulgação dos projetos nesta área, foi criado, em maio de 2011, o Network for Digital Methods in the Arts and Humanities (NeDiMAH)<sup>16</sup>, uma rede de especialistas dos vários países da União Europeia (incluindo Portugal) e financiada pela European Science Foundation (ESF). Integra seis grupos de trabalho, cada um com temática específica: espaço e tempo; informação e visualização; *linked data* e métodos ontológicos; construção e desenvolvimento de coleções digitais para a investigação; utilização de textos em larga escala; edições digitais. Integra, ainda, três grupos transversais, cada um deles, com uma atribuição específica: desenvolvimento de métodos e taxonomias da Tecnologia de Informação e Comunicação (TIC); impacto da tecnologia nos métodos de investigação e nas publicações académicas; outras atividades desenvolvidas pelo NeDiMAH, nomeadamente, os métodos digitais para Artes e Humanidades. A investigação e outras atividades, nomeadamente, de divulgação levadas a cabo pelo NeDiMAH têm o objetivo de contribuir para a classificação das Artes e Humanidades Digitais, através de três vias principais: um mapa de visualização da investigação digital

---

<sup>16</sup> Disponível em <http://www.nedimah.eu/>

na Europa; uma ontologia de métodos de investigação digital; um fórum em linha, de carácter interativo e colaborativo, para a comunidade de profissionais europeus ativos nesta área.

Além desta, destacam-se, ainda, outras importantes infraestruturas e projetos europeus.

Dans le domaine des humanités numériques, l'Union européenne est en effet en train de construire deux infrastructures à l'échelle du continent : Clarin et Dariah. Clarin, pour Common Language Resources and Technology Infrastructure, est spécialisé dans les ressources linguistiques. Dariah, pour Digital Research Infrastructure for the Arts and Humanities, est plus généraliste. D'autres infrastructures existent pour les données quantitatives (CESSDA, ESS par exemple). Clarin et Dariah sont des consortiums (avec statut d'ERIC, qui est une structure juridique créée en 2009 au sein de l'Union européenne destinée à faciliter la coopération à l'intérieur de l'Union). (Dacos & Mounier, 2014, p. 33)

O Council of European Social Science Data Archives (CESSDA)<sup>17</sup>, criado em junho de 2013, é uma estrutura pan-europeia, com o objetivo de fornecer um acesso integrado e continuado a dados em larga escala. Nesse sentido, pretende incentivar a normalização de dados e metadados, a difusão de dados e a mobilidade do conhecimento em toda a Europa. O CESSDA desempenha um papel ativo no desenvolvimento de normas para documentar e divulgar os inventários de dados de

---

<sup>17</sup> Disponível em <https://www.cessda.eu/>

pesquisa disponíveis a partir de recursos transnacionais na Europa. A sede localiza-se em Bergen, na Noruega.

O Common Language Resources and Technology Infrastructure (CLARIN-ERIC)<sup>18</sup>, sediado em Utrecht, na Holanda, é um projeto integrado na União Europeia e tem como objeto a criação de uma infraestrutura pan-europeia que impulse a pesquisa na área multicultural e multilingue que é a Europa. É uma infraestrutura de dados distribuídos, sediada em instituições europeias, e da qual fazem parte universidades, instituições de pesquisa, bibliotecas e arquivos públicos. Os membros do CLARIN são os governos, ou os seus representantes, e organizações intergovernamentais da União Europeia. Portugal está representado através da Universidade de Lisboa. O principal objetivo deste projeto é desenvolver a investigação em Ciências Sociais e Humanas, facultando, aos investigadores, o acesso a uma plataforma integradora de recursos baseados na linguagem e ferramentas avançadas a nível europeu.

A Digital Research Infrastructure for the Arts and Humanities (DARIAH)<sup>19</sup>, sediada em Paris, é uma infraestrutura pan-europeia vocacionada para o ensino das artes e humanidades com recurso a métodos digitais. É constituída por organismos intergovernamentais europeus.

Portugal está representado pelo Ministério da Ciência, Tecnologia e Ensino Superior, sendo o coordenador nacional o ROSSIO, infraestrutura de investigação de referência para as Ciências Sociais, Artes e Humanidades, coordenado pela Faculdade de Ciências Sociais e Humanas da Universidade Nova de Lisboa, e tendo como parceiros, além

---

<sup>18</sup> Disponível em <https://www.clarin.eu/>

<sup>19</sup> Disponível em <http://www.dariah.eu/>

desta, a Câmara Municipal de Lisboa, a Fundação Calouste Gulbenkian, a Fundação Mário Soares, os Sistemas de informação Património Arquitetónico da Direção Geral do Património Cultural, a Cinemateca, o Centro Lusíada de Investigação em Serviço Social e Intervenção Social da Universidade Lusíada de Lisboa e o Teatro Nacional D. Maria II. A rede DARIAH é uma estrutura ERIC (European Research Infrastructure Consortium), o que permite o seu financiamento a longo prazo. Atualmente, a DARIAH apoia vários projetos como, por exemplo, o OpenEdition, uma plataforma para o acesso livre à informação.

### **2.3. Comunidades de interesses**

#### **COMUNIDADES DE INTERESSE**

- **Global Outlook::Digital Humanities (GO::DH);**
- **4humanities ;**
- **Humanities, Arts, Science, and Technology Alliance and Collaboratory (HASTAC);**
- **Red de Humanidades Digitales de México (Red-HD);**
- **Associação das Humanidades Digitais (AHDig);**
- **dh+lib.**

Existem outras organizações, criadas com a missão de promover, divulgar e apoiar a investigação na área de Humanidades Digitais e de fomentar a comunicação entre os vários intervenientes, sejam investigadores das áreas das humanidades, da computação, ou da comunicação. Destacamos, neste âmbito, alguns grupos, associações ou comunidades de investigadores com interesses comuns nesta área, tais

como: Global Outlook::Digital Humanities (GO::DH)<sup>20</sup>; 4humanities<sup>21</sup>; Humanities, Arts, Science, and Technology Alliance and Collaboratory (HASTAC)<sup>22</sup>; Red de Humanidades Digitales de México (Red-HD)<sup>23</sup>; Associação das Humanidades Digitais (AHDig)<sup>24</sup>; dh+lib<sup>25</sup>.

O GO::DH é uma SIG (Special Interest Groups) da ADHO, que assume o propósito de quebrar os entraves à comunicação e colaboração entre investigadores nos domínios das Artes e Humanidades Digitais e do património em países de alto, médio ou baixo rendimento económico, embora não se constitua nos moldes tradicionais dos programas de ajuda. “Global Outlook a pour objectif d’étendre la coopération et d’intensifier les échanges autour des humanités numériques par-delà les frontières qui séparent les spécialisations mais aussi les pays et les zones linguistiques” (Dacos & Mounier, 2014, p. 35). Coloca em comunicação estudantes, investigadores e instituições, dá a conhecer os respetivos projetos e fomenta a partilha de conhecimentos, recursos, experiências e problemas.

Outras organizações, como a 4humanities, a HASTAC, a RedHD, ou a AHDig, defendem genericamente os mesmos propósitos. A 4humanities foi fundada em 2010 por um conjunto de investigadores dos Estados Unidos, Canadá, Reino Unido e Austrália, com o objetivo de incentivar e divulgar os recursos e tecnologias na área das Humanidades Digitais, compensando o crescente desinvestimento económico sentido neste domínio.

---

<sup>20</sup> Disponível em <http://www.globaloutlookdh.org/>

<sup>21</sup> Disponível em <http://4humanities.org/>

<sup>22</sup> Disponível em <https://www.hastac.org/>

<sup>23</sup> Disponível em <http://www.humanidadesdigitales.net/>

<sup>24</sup> Disponível em <https://ahdig.org/>

<sup>25</sup> Disponível em <http://acrl.ala.org/dh/>

A HASTAC<sup>26</sup> é uma associação com mais de 14.000<sup>27</sup> membros individuais (investigadores em Ciências Sociais, humanistas e artistas) e institucionais focados no desenvolvimento de métodos de pesquisa e ensino inovadores no âmbito da tecnologia digital. A Red-HD, especialmente dirigida para falantes de língua espanhola, foi criada no México, em 2011, por um grupo de investigadores, académicos, com o objetivo de apoiar e impulsionar a investigação em Humanidades Digitais.

A AHDig centra a sua atuação no universo da língua portuguesa, a quarta língua mais falada no mundo, com 244 mil milhões de falantes. Foi criada em São Paulo, Brasil, a 23 de outubro de 2013, definindo, como principal objetivo, o seguinte:

Nosso principal objetivo é fortalecer as iniciativas em Humanidades Digitais já ativas no universo dos falantes do português, e promover novas iniciativas nesse campo entre eles. Este será um fórum onde esses pesquisadores poderão debater e compartilhar suas experiências, tecendo redes no interior dessa esfera cultural e linguística, que possam levar ao trabalho colaborativo e ao aumento da visibilidade internacional dos projetos em Humanidades Digitais conduzidos em português. (“AHDig,” s.d.)

É uma rede em crescimento, que funciona com a colaboração dos intervenientes, estando aberta à participação de todos os investigadores, não só falantes de língua portuguesa, mas que também desenvolvam investigação noutros domínios de expressão cultural em português.

---

<sup>26</sup> O acrónimo HASTAC joga com a pronúncia de hashtags (#), hiperligações indexáveis pelos mecanismos de busca em linha.

<sup>27</sup> Segundo dados recolhidos no respetivo sítio eletrónico em 07/08/2017.

O dh+lib é, essencialmente, um sítio eletrónico criado, em 2012, e desenvolvido pela Association of College and Research Libraries (ACRL), uma divisão da American Library Association (ALA). Este sítio eletrónico foi construído com o objetivo de disponibilizar um espaço de discussão onde os bibliotecários, arquivistas, estudantes e especialistas em informação pudessem partilhar ideias e projetos, contribuindo para uma reflexão alargada sobre Humanidades Digitais e a sua influência nas bibliotecas, arquivos e centros de documentação.

#### **2.4. Linhas de investigação**

##### **LINHAS DE INVESTIGAÇÃO**

- Espaço e tempo;
- Informação e visualização;
- *Linked data* e métodos ontológicos;
- Construção e desenvolvimento de coleções digitais para a investigação;
- Utilização de textos em larga escala;
- Edições digitais académicas;
- Preservação da informação;
- Direito de autor.

As várias linhas de investigação em Humanidades Digitais, além da questão de fundo acerca da relação entre Humanidades e tecnologia digital, têm vindo a estudar temas como a otimização dos acervos digitais nas áreas da visualização da informação, a construção de semânticas coerentes, o tratamento e indexação de coleções de textos, a representação do espaço e do tempo e a edição digital para utilização académica e escolar.

Os grupos de trabalho integrados no NeDiMAH<sup>28</sup> indicam uma possível tipologia da investigação em Humanidades Digitais:

- Espaço e tempo (*Space and Time*) – estudos versados sobre a representação visual de dados complexos (*webmapping*), onde se conjuga a análise do espaço e do tempo, a referenciação geográfica (GIS) e os indicadores estatísticos. Inclui, ainda, a análise de redes, a computação móvel, a realidade aumentada e a anotação semântica de lugares, épocas e eventos. Estes temas de investigação cruzam os interesses de disciplinas como a História, a Geografia, a Estatística e a Computação e, com frequência, são desenvolvidos em parceria com organizações do património cultural e turismo. Com um elevado potencial para abrir novos caminhos de pesquisa, esta constitui uma das linhas mais ativas da investigação e é, provavelmente, aquela que provoca maior impacto, não só na comunidade académica, como junto da população em geral.
- Informação e visualização (*Information Visualisation*) – A visualização, neste âmbito, integra diferentes tipos de interação, tecnologias e materiais, como conjuntos de dados em formato texto, imagem e vídeo, dados geográficos, gráficos, tabelas, material de arquivo, imagens de alta resolução e representações 3D (computação gráfica tridimensional). Genericamente, caracteriza-se por ser fidedigna e rica em detalhes, com alta definição, e por permitir interação. A visualização, utilizada com fins descritivos ou analíticos, é de grande aplicabilidade em projetos de investigação em Humanidades e

---

<sup>28</sup> Disponível em <http://www.nedimah.eu/workgroups>

constitui uma linha de investigação relevante, igualmente, com impacto na sociedade.

- *Linked data* e métodos ontológicos (*Linked Data and ontological methods*) – *Linked Data* (ou dados ligados entre si) é um método de publicação de dados estruturados para que possam ser interligados e consultados através de ontologias comuns ou compatíveis. Esta linha de investigação inclui o desenvolvimento de ontologias que forneçam as definições semânticas e as clarificações necessárias para agregar fontes diversas e dispersas num conjunto de dados coerentes. O objetivo é, essencialmente, permitir a agregação de um grande número de fontes de dados em ambiente aberto, colaborativo e acessível, em particular na web, enriquecendo as pesquisas e a recuperação da informação, tornando os padrões e as ligações mais visíveis. A ontologia aplicada aos projetos, em particular, no que se refere à escolha das semânticas a utilizar e à respetiva normalização, tem vindo a criar ferramentas facilitadoras para a investigação em Humanidades.
- Construção e desenvolvimento de coleções digitais para a investigação (*Building and Developing Collections of Digital Data for Research*) – Esta linha de investigação incide especificamente sobre a construção de bibliotecas digitais vocacionadas para a investigação, onde se inclui a utilização e a reutilização da informação para a criação de novo conhecimento. Aborda, em particular, a descrição e a estruturação do objeto digital, incluindo a conservação e a preservação da informação ao longo do tempo. A disponibilização em linha de acervos digitais ou digitalizados, permitindo que a informação possa ser utilizada e reutilizada, envolve a utilização de técnicas TIC

(Tecnologias da Informação e Comunicação) em todas as fases da edição digital. Por outro lado, a criação de uma biblioteca digital implica, além da gestão, da preservação e da disseminação da informação, a descrição e a estruturação dos recursos digitais, bem como o acesso a esses recursos, em termos de autorização e de interoperabilidade.

- Utilização de textos em larga escala (*Using Large-Scale Text Collections for Research*) – As TIC, enquanto métodos de recuperação e extração de dados, aplicadas a grandes e complexos conjuntos de dados (Big Data), têm grande aplicabilidade na investigação em Humanidades, ao permitir tarefas como a análise, a pesquisa, a preservação, a partilha, o armazenamento, a transferência, ou a visualização dos dados, resumindo-os num formato útil. Esta linha de investigação estuda a utilização de novas ferramentas e métodos para a análise e mineração de grandes massas textuais, permitindo a extração de padrões e de dados ocultos. Além disso, estuda projetos desenvolvidos em *corpora* linguísticos, a fim de conseguir uma melhor compreensão acerca da forma como as grandes coleções de textos podem ser usadas na investigação.
- Edições digitais académicas (*Scholarly Digital Editions*) – A edição digital é um procedimento cada vez mais comum, decorrente de programas de digitalização em larga escala ou de iniciativas editoriais de apoio à investigação. Da mesma forma que as TIC's têm proporcionado novas técnicas de análise e processamento dos dados, os investigadores têm vindo a sentir alguma dificuldade em encontrar recursos digitais, como a apresentação dinâmica dos resultados, adequados a uma edição melhorada dos dados obtidos, para lá das soluções usuais ou do mero espelhamento da edição analógica. Esta linha de

investigação promove, por conseguinte, o uso das tecnologias digitais na produção e disseminação de edições acadêmicas em formatos inovadores, envolvendo especialistas de várias áreas, ao mesmo tempo que procura estabelecer o estado da arte e recomendar um conjunto de práticas tendentes a garantir a máxima interoperabilidade e acessibilidade dos dados digitais.

Ainda que os grupos de trabalho do NeDiMAH configurem as principais linhas de investigação em Humanidades Digitais, há outros campos de estudo em desenvolvimento, abordando questões afins, como a preservação da informação, ou outras que, não sendo novas, revelam agora novas perspectivas e implicações, como as relacionadas com os direitos de autor.

Os projetos de digitalização em larga escala, a proliferação de programas e produtos proprietários e as sucessivas versões de um mesmo produto, torna imperiosa a definição de políticas de preservação, sob pena de se perder o acesso à informação. O reverso do desenvolvimento da tecnologia é a inevitabilidade da sua própria obsolescência, pelo que a investigação procura soluções para a preservação da autenticidade e integridade do objeto digital nas suas dimensões física e conceptual (Ferreira, 2011).

As possibilidades de acesso e disseminação da informação, nomeadamente, através das redes sociais, têm vindo a pôr em causa algumas questões relacionadas com os direitos de autor. A utilização e reutilização da informação coloca problemas na atribuição e administração do direito de autor, que necessita de uma nova configuração.

O desenvolvimento de ferramentas informáticas constitui uma linha de investigação crucial no âmbito das Humanidades Digitais. Tendo em vista a rentabilização da utilização dos computadores na área de Humanidades têm vindo a ser desenvolvidas ferramentas para analisar dados, gerir informação bibliográfica, gerenciar tarefas, desenvolver o trabalho colaborativo em rede, organizar os materiais de pesquisa, publicar e partilhar informações, construir coleções, pesquisar visualmente, acompanhar o desenvolvimento de uma pesquisa, transcrever textos manuscritos ou falados, editar imagens, realizar mapas dinâmicos, visualizar dados, etc. As ferramentas desenvolvidas visam servir, não apenas os investigadores de Humanidades, como a generalidade da comunidade científica. Cita-se o exemplo do Mendley<sup>29</sup> ou do Zotero<sup>30</sup>, duas aplicações informáticas que permitem a gestão de bibliografias e a partilha de trabalhos realizados com a sua comunidade de utilizadores. Outros exemplos são os analisadores de conteúdos como o TagCrowd<sup>31</sup> ou Wordle<sup>32</sup>, usados para gerar nuvens de palavras ou etiquetas (*tag-clouds*), ou seja, representações visuais das palavras existentes em determinado texto, em que o tamanho do tipo aumenta em função da respetiva frequência.

Todas estas ferramentas, se utilizadas de forma adequada, constituem um precioso auxílio à investigação. Porém, o entusiasmo que provocam, tendo em conta os benefícios que oferecem, pode conduzir a atitudes de facilitismo, sendo, por isso, fundamental manter a vigilância, individual e coletiva, no sentido de garantir um comportamento ético e rigoroso. Apesar de os produtos digitais se apresentarem de

---

<sup>29</sup> Disponível em <https://www.mendeley.com/>

<sup>30</sup> Disponível em <https://www.zotero.org/>

<sup>31</sup> Disponível em <http://tagcrowd.com/>

<sup>32</sup> Disponível em <http://www.wordle.net/>

forma bastante apelativa, isso não dispensa a realização de uma investigação aprofundada acerca das suas reais funcionalidades. Por outro lado, torna-se essencial que, ao apresentar o produto final, ou o resultado da investigação, se disponibilize uma informação completa acerca dos métodos e das técnicas utilizadas nas várias fases do trabalho.

## **2.5. Projetos**

Ainda que as linhas de investigação em Humanidades Digitais estejam bem definidas e seja possível identificar alguns especialistas para cada uma delas, os projetos que têm vindo a ser desenvolvidos não assentam na especialidade, mas na interdisciplinaridade e no trabalho colaborativo. No entanto, é possível constatar que os projetos, nas várias vertentes das Humanidades Digitais, revelam ainda fragilidades em relação à persistência e preservação dos resultados, além de que, muitos deles, também não disponibilizam as fontes de informação utilizadas, nem referem os tratamentos aplicados aos dados para se obter os resultados apresentados.

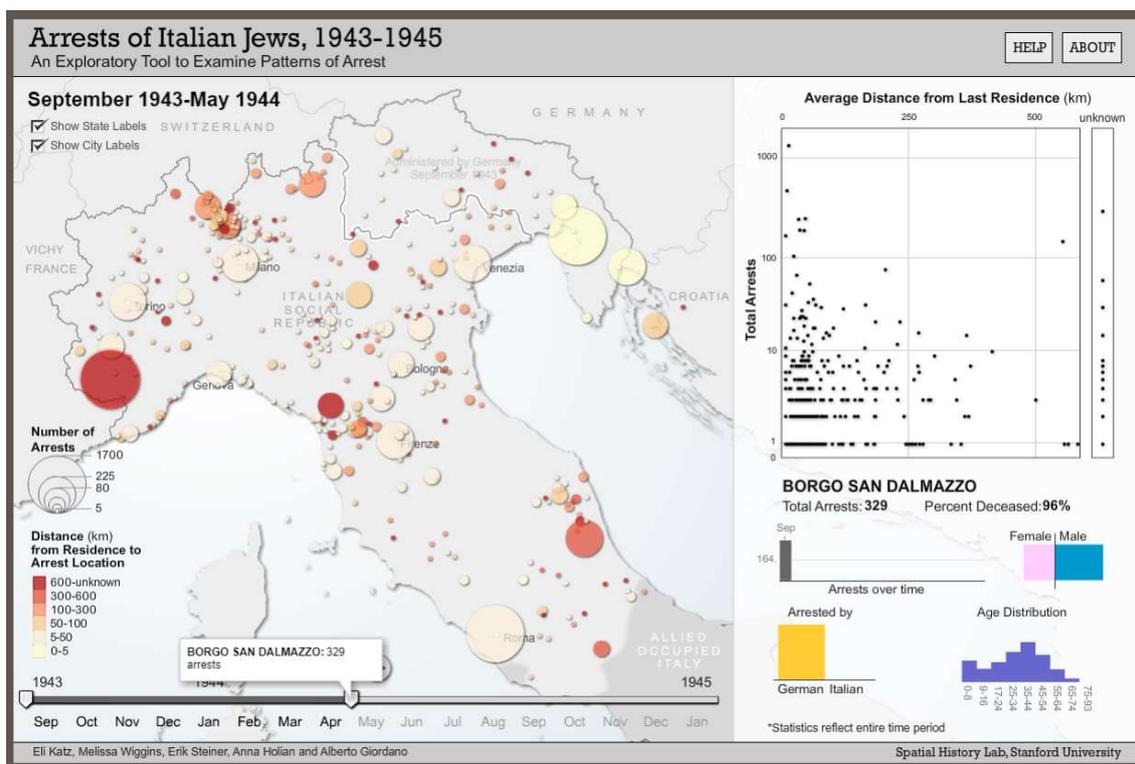
A representação visual de dados complexos (*webmapping*), que conjuga a análise do espaço e do tempo, a referenciação geográfica (GIS) e os indicadores estatísticos, constitui uma das linhas mais ativas de investigação, e provoca grande impacto, sendo particularmente útil para o ensino e para produtos de divulgação de património cultural. Estes desenvolvimentos cruzam os interesses de disciplinas como a História, a Geografia, a Estatística, as Ciências da Informação e a Computação. Um dos projetos

que se enquadra nesta linha é *The Spatial History Project*<sup>33</sup>, desenvolvido na Universidade de Stanford, Estados Unidos da América.

Como exemplo das suas potencialidades, pode referir-se a linha de investigação seguida no *Holocaust Geographies*<sup>34</sup> (Figura 4) que permite visualizar, de forma dinâmica, os dados referentes à prisão de judeus italianos entre 1943 e 1945.

Figura 4 – The Spatial History Project: Holocaust Geographies: Arrests of Italian Jews.

Fonte: [http://web.stanford.edu/group/spatialhistory/cgi-bin/site/viz.php?id=383&project\\_id=0](http://web.stanford.edu/group/spatialhistory/cgi-bin/site/viz.php?id=383&project_id=0)



Os dados apresentados são: número de pessoas detidas, idades, sexo, nacionalidade de quem os prendeu, localidades onde foram detidos, distância do local onde foram detidos à residência e a percentagem de falecidos. Os dados são interatuantes: a informação principal aparece sobre o mapa de Itália de forma gráfica (círculos com área proporcional ao número de prisioneiros) e os valores no canto inferior direito

<sup>33</sup> Disponível em <http://web.stanford.edu/group/spatialhistory/cgi-bin/site/index.php>

<sup>34</sup> Disponível em [http://web.stanford.edu/group/spatialhistory/cgi-bin/site/viz.php?id=383&project\\_id=0](http://web.stanford.edu/group/spatialhistory/cgi-bin/site/viz.php?id=383&project_id=0)

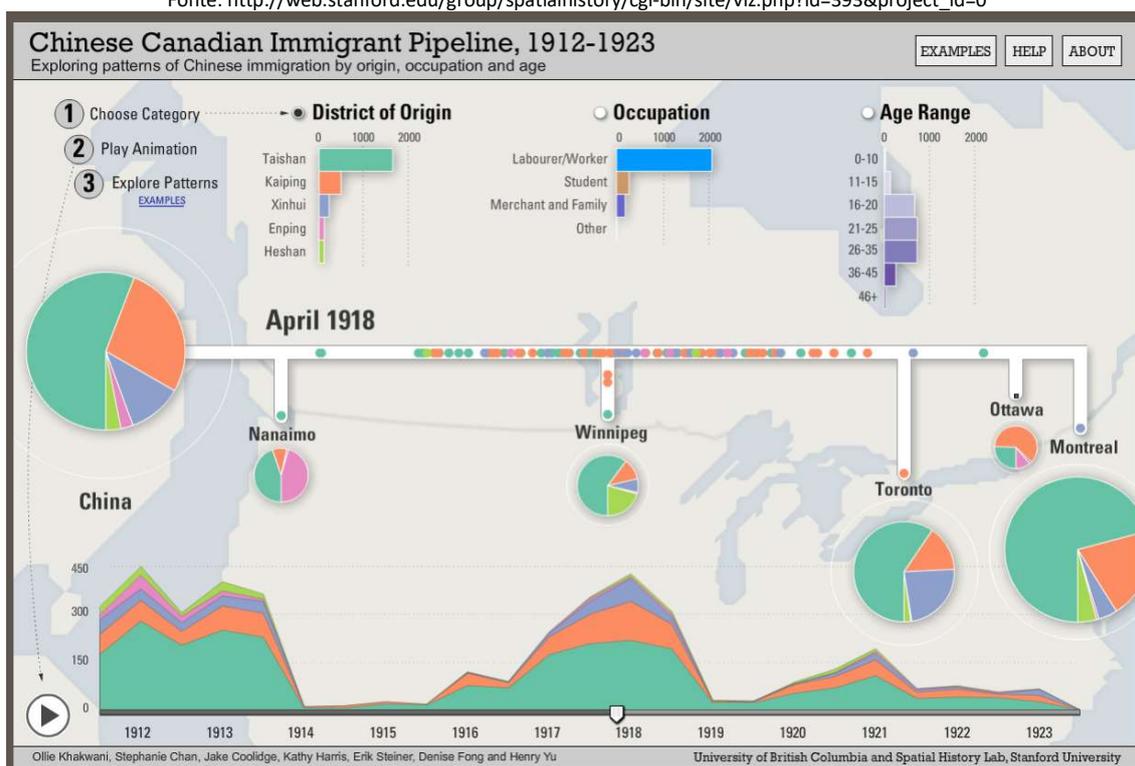
correspondem à localidade sobre a qual se encontra o cursor comandado pelo utilizador.

No gráfico do canto superior direito, é possível visualizar a totalidade das detenções por localidade. Este projeto foi considerado “a humble and modest – and expensive – attempt to do history in a different way” (White, 2010). A sua complexidade deriva do facto de realizar uma representação simultânea do espaço e do tempo, com associação aos eventos relacionados, ao mesmo tempo que integra a informação de arquivo e permite a respetiva visualização de forma muito clara e acessível (Shnayder, 2010), embora apenas seja possível visualizar os dados no ecrã, uma vez que a aplicação não permite a extração dos dados, nem em modo texto, nem em imagem. No entanto, constitui um exemplo de uma fonte credível de informação, referenciada e muito apelativa para o ensino. Foi construído por uma equipa multidisciplinar, constituída por historiadores, linguistas, geógrafos, especialistas na representação espacial de dados e em multimédia, num modelo de projeto colaborativo.

Outro exemplo, é o projeto *Chinese Canadian Immigrant Pipeline, 1912-1923* (Figura 5), que pretendia facultar a visualização do fluxo migratório da China para o Canadá entre 1912 e 1923. Para tal, foram analisados os boletins de imigração (dados de arquivo) de cidadãos chineses nesse período.

Figura 5 – The Spatial History Project: Chinese Canadian Immigrant Pipeline, 1912-1923.

Fonte: [http://web.stanford.edu/group/spatialhistory/cgi-bin/site/viz.php?id=393&project\\_id=0](http://web.stanford.edu/group/spatialhistory/cgi-bin/site/viz.php?id=393&project_id=0)



Este projeto exibe o conteúdo de forma dinâmica. No arco cronológico de 1912 a 1923, é possível ver, por meses, a província de origem, a cidade de destino, a idade e a função dos imigrantes ao entrar no Canadá e colocar em evidência o respectivo distrito de origem, a ocupação ou o nível etário. Apesar de nem os dados originais, nem o tratamento que sofreram estarem disponíveis, há uma descrição mínima do projeto e os tutoriais são explicativos.

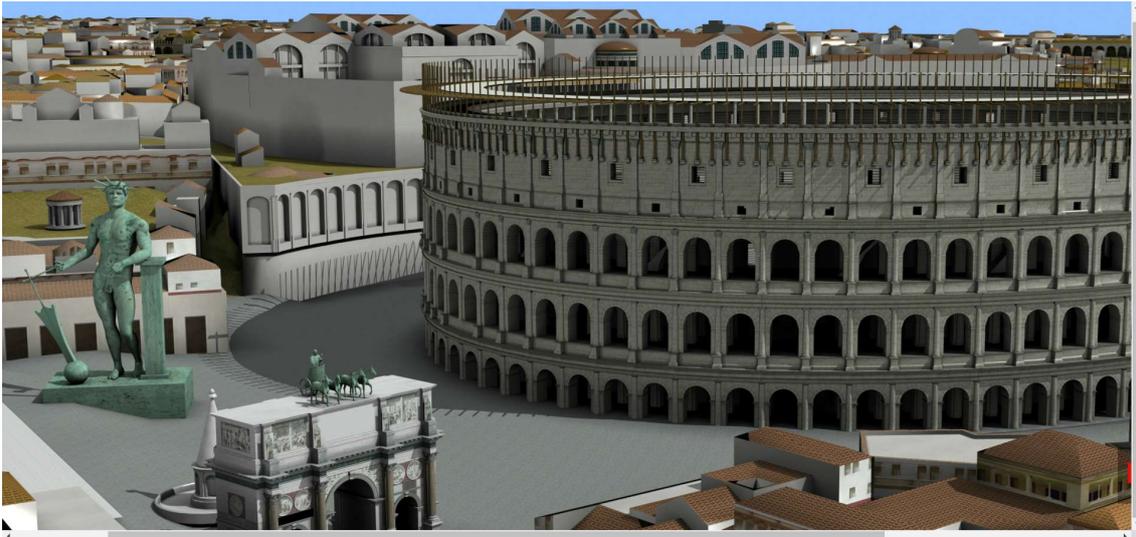
A tecnologia de representação em 3D constitui outra linha de investigação importante. A construção dos modelos em 3D pode integrar imagens, dados geográficos, gráficos, tabelas, referências a material de arquivo ou biblioteca, etc. A respectiva visualização pode ser feita para permitir uma análise descritiva e analítica dos locais representados.

Um dos projetos que utiliza esta tecnologia é o *Rome Reborn: a digital modelo of ancient Rome* (Dylla, Frischer, Mueller, Ulmer, & Haegler, 2009), desenvolvido pelo

Virtual World Heritage Laboratory da Universidade de Virgínia, a partir de uma iniciativa de âmbito internacional, com o objetivo de recriar virtualmente o plano urbanístico e arquitetônico da cidade de Roma entre a Idade de Bronze tardia (c. 1000 a.C.) e a alta Idade Média (c. 550 d.C.) (Figura 6).

Figura 6 – Rome Reborn 1.1 (Estátua colossal do deus Sol e Anfiteatro de Flávio).

Fonte: <http://romereborn.frischerconsulting.com/gallery-archive.php>



Mais uma vez, trata-se de um projeto multidisciplinar, com fins educativos e culturais. O aplicativo podia ser sobreposto ao *GoogleEarth*, a fim de elaborar um estudo comparativo com a atualidade. Apesar de incluir dados referentes à arquitetura, os edifícios não são descritos e analisados no contexto da História da Arte, servindo apenas como complemento ao estudo na perspectiva do urbanismo. Por questões financeiras deixou de estar disponível sobre o *GoogleEarth*, estando apenas disponíveis imagens fixas e vídeos com renderizações do projeto.

As edições digitais escolares também constituem outra linha de investigação, assumindo como objetivos a definição do atual estado da edição digital e a elaboração de um conjunto de boas práticas para a preservação da informação no tempo e a

respetiva interoperabilidade. Assim como a disponibilização em acesso livre da investigação credenciada, sobretudo em Humanidades.

A OpenEdition<sup>35</sup>, desenvolvida pelo Centre pour l'Édition Électronique Ouverte (CLÉO), unidade que congrega o Centre National de la Recherche Scientifique (CNRS), a École des Hautes Études en Sciences Sociales (EHESS), a Universidade Aix-Marseille e a Universidade de Avignon, oferece à comunidade académica uma plataforma internacional de publicação e divulgação da informação em Ciências Sociais e Humanas. É constituída pelas seguintes plataformas: Hypotheses, que alberga blogues denominados cadernos de investigação; Revues, para os periódicos; Calenda, com a agenda e a divulgação de eventos; e os e-book's.

Os blogues que integram a Hypotheses são pré-avaliados por uma comissão académica e podem receber um ISSN (*International Standard Serial Number*), código de registro internacional das publicações periódicas. A atribuição de ISSN a blogues, reconhecendo-os como publicações científicas, é inédita.

A Revues é um portal de revistas científicas em linha no domínio das Humanidades e das Ciências Sociais. Privilegia a disponibilização de artigos em texto integral e em acesso livre. As revistas são selecionadas por um conselho científico. Cada revista tem uma formatação individualizada, mas na página de acolhimento de cada uma, além da pesquisa por palavra, é dado destaque ao último número e apresentam-se os índices por autor (ordem direta) e por palavra-chave, o histórico da revista, as informações, etc. Em cada artigo pode consultar-se: o resumo; o índice (palavras-chave); o mapa (sumário); o texto; a bibliografia; os anexos; as notas; as ilustrações; a citação (como

---

<sup>35</sup> Disponível em <https://www.openedition.org/>

citar o artigo); e os dados biográficos do autor. O texto está disponível em PDF e em HTML. Os parágrafos estão numerados, o que, por se manterem independentemente do tamanho do ecrã em que o artigo é visualizado, facilita a referência.

## 2.6. Publicações

### PUBLICAÇÕES

- *Literary and Linguistic Computing/ Digital Scholarship in the Humanities;*
- *Digital Studies: Le Champ Numérique;*
- *International Journal of Humanities and Arts Computing;*
- *Journal of Text Encoding Initiative;*
- *Journal of Digital Humanities;*
- *Computing in the Humanities Working Papers;*
- *dhq – Digital Humanities Quarterley;*
- *DHCommons jornal;*
- *Digital Humanities Now.*

A informação relativa à área de Humanidades Digitais encontra-se distribuída por monografias, atas de seminários e congressos e compilações de artigos publicados em revistas científicas. Parte substancial da literatura produzida, tanto em periódicos como em monografias, analisa ou descreve projetos no domínio das Humanidades Digitais. Porém, existe uma outra linha mais analítica, conceptual ou teórica, a qual procura equacionar os “como” e “porquê” do desenvolvimento na área de Humanidades Digitais (Berry, 2012; Borgman, 2009; Gold, 2012; Vinopal & McCormick, 2013).

Os primeiros artigos surgiram nos periódicos científicos das áreas de linguística, computação, história e literatura. Na atualidade, existem vários periódicos da

especialidade, em inglês ou francês, a maioria com revisão pelos pares (*peer-review*), em linha e gratuitos, obedecendo aos princípios do acesso aberto (*open access*).

Destacam-se os seguintes:

- *Literary and Linguistic Computing*, fundado em 1986, alterou o título *DSH - Digital Scholarship in the Humanities*, em 2014, em inglês e com revisão por pares. Tem acesso condicionado, estando apenas disponível para os membros institucionais e individuais da ADHO ou das organizações que a ela pertençam.
- *Digital Studies: Le champ numérique = Computing in the Humanities Working Paper*, disponibiliza em linha os números desde 1992, em inglês e francês, de acesso livre e com revisão por pares;
- *International Journal of Humanities and Arts Computing* disponibiliza em linha os números desde 1994, em inglês e francês, de acesso livre e com revisão por pares;
- *Journal of Text Encoding Initiative* disponibiliza em linha os números desde 2011, em inglês e francês, de acesso livre e com revisão por pares;
- *Journal of Digital Humanities*, disponibiliza em linha os números desde 2011, em inglês, de acesso livre e com revisão por pares;
- *Computing in the Humanities Working Papers* disponibiliza em linha os números desde 1992, em inglês e francês, de acesso livre e com revisão por pares;
- *dhq – Digital Humanities Quarterley* em linha desde 2007, em inglês, de acesso livre e com revisão por pares;
- *2015 DHCommons journal*, um dos mais recentes, o primeiro número foi editado em julho de 2015, em inglês, de acesso livre e com revisão por pares.

A publicação em linha *Digital Humanities Now* é um projeto inovador que, com recurso a um programa informático, realiza uma pré-seleção dos conteúdos, em cerca de 500 blogues subscritores. Desenvolvido pelo Center for History and New Media (CHNM), o programa *Press Forward* permite "mettre en place une chaîne complète de sélection de contenus, depuis l'agrégation de billets de blogs mis en ligne sur le web jusqu'à la publication d'une revue à comité de lecture composée d'articles issus de ces billets sélectionnés" (Dacos & Mounier, 2014, p. 39). A informação é recolhida de forma automática, via *fede reader* e marcadores, após o que é avaliada pelos colaboradores e, por fim, em função da avaliação recebida, é, ou não, disponibilizada em linha em acesso livre, mas com a garantia de informação certificada. Alguns artigos podem ser publicados no *Journal of Digital Humanities*, que também aceita artigos que lhe são submetidos diretamente.

A temática das Humanidades Digitais também aparece refletida nos periódicos das editoras comerciais, os quais têm a vantagem de ser indexados, mas onde o acesso ao texto integral dos artigos é, geralmente, pago.

A investigação em Humanidades alterou-se com os desenvolvimentos da tecnologia, surgindo novas oportunidades de aceder, processar e divulgar a informação, no sentido em que "technological advances have altered the way that researchers capture, curate, analyze and visualize data at every scale" (Collins & Jubb, 2012). A área das Humanidades Digitais surge como uma nova forma de solucionar problemas da investigação em Humanidades, mas preservando a tradição, dando atenção à complexidade e mantendo as modalidades de análise em profundidade, crítica e interpretação (Evans & Rees, 2012).

A área de Humanidades Digitais, no entanto, ainda é encarada como uma “caixa negra” (Rieder & Rohle, 2012), da qual tudo pode sair, sem que se perceba exatamente como é que foi produzido. Ou seja: conhecemos os resultados, mas apenas intuímos os métodos. Há, por isso, que criar conteúdos apelativos e certificados para o ensino, para a divulgação junto do grande público, e torna-se imperioso que, em simultâneo, sejam disponibilizadas as fontes ou dados utilizados e se refiram as leituras e as metodologias efetuadas.

As Humanidades Digitais afirmam-se como um campo fértil da investigação, mas o principal desafio é ainda a definição de modelos genéricos para a sistematização, partilha e reuso da informação. As Humanidades Digitais estão firmadas como uma área de saber multidisciplinar que trouxe novos métodos às Humanidades, permitindo um novo olhar sobre os dados e fontes de informação. O trabalho colaborativo, mesmo à distância, está facilitado, permitindo congrega valências em torno de um mesmo projeto. O retorno dos resultados da investigação à sociedade pode ser realizado através de produtos apelativos, tentando relançar a importância e a oportunidade de investigar na área das Humanidades.

O termo Humanidades Digitais está consagrado, mas, como referem Schreibman, Siemens e Unsworth (2016), no prefácio da obra *A New Companion to Digital Humanities*, daqui a uma década ou duas, associar a palavra digital às Humanidades será um pleonasma:

What is important today is not that we are doing work with computers, but rather that we are doing the work of the humanities, in digital form. The field is now much broader than it once was, and includes not only the

computational modeling and analysis of humanities information, but also the cultural study of digital technologies, their creative possibilities, and their social impact. (*A new companion to Digital Humanities*, 2016, p. xvii)

Ou seja, o que é importante é que se deixe de dizer que, no âmbito da investigação em Humanidades se faz um trabalho com computadores, passando a assumir que se faz o trabalho das Humanidades em formato digital e, portanto, estamos a utilizar métodos digitais para estudar as criações humanas. “One of the major challenges of the digital humanities is therefore to retain the open nature of the web and its resources while providing some degree of expert guidance in its use and in the materials accessed” (Gardiner & Musto, 2015, p. 21). Para que as Humanidades Digitais continuem em crescimento há que assegurar, por um lado, a neutralidade da web e, por outro lado, providenciar para que o conhecimento seja amplamente difundido, garantindo a permanência e a disponibilidade da informação.

### 3. Bibliotecas Digitais

**Nado digital** - criado com ferramentas ou processos digitais.

**Metadados** - dados acerca de outros dados. Refere-se aos dados utilizados para ajudar a identificar, descrever e localizar os recursos, nomeadamente da web.

**Obras em domínio público** - não está abrangida pelo direito de autor, nem há restrições à utilização da obra.

#### Termos

**Biblioteca virtual** - não tem existência material (coleção física).

**Biblioteca híbrida** - gere coleções físicas e digitais.

**Repositórios** - disponibilizam a produção científica da respetiva entidade.

**Corpus** - conjunto de textos escritos ou orais, reunidos com critério, de forma a constituir um conjunto coerente para a análise linguística.

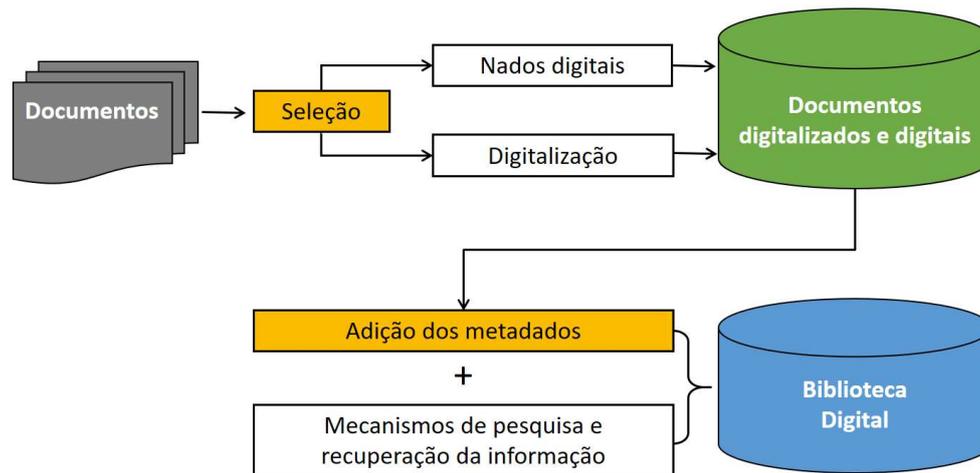
**Agregadores** - bibliotecas digitais que alojam metadados dos objetos digitais.

Quando se aborda o tema das bibliotecas digitais, surgem na literatura vários termos ou expressões: biblioteca virtual, biblioteca híbrida, repositório, *corpora* ou agregador de conteúdos. Não são sinónimos, mas se se considerar que a biblioteca digital se pode definir como “a focused collection of digital objects, including text, video, and audio, along with methods for access and retrieval, and for selection, organization, and maintenance of the collection” (Witten et al., 2009), todos, de alguma forma, se podem considerar e contribuir para uma melhor compreensão do fenómeno. As várias designações têm, em comum, os conceitos subjacentes de seleção, organização, manutenção e disponibilização de coleções, que são tarefas inerentes à biblioteca, seja o acervo digital ou tradicional.

Para se constituir uma biblioteca digital é necessário, em primeiro lugar, um acervo documental; em seguida, são selecionados os documentos e os que não sejam nados

digitais são digitalizados; são adicionados os metadados aos documentos (ficheiros em formato digital); por fim, os ficheiros e a respetiva metacodificação são integrados num sistema de gestão de dados, a que são associados mecanismos de pesquisa e recuperação da informação (Figura 7).

Figura 7 – Esquema da constituição da biblioteca digital. Esquema: DG, 2017.



Porém, nas várias definições, existem algumas diferenças que apontam para diferentes tipologias.

A biblioteca virtual não tem existência material, no sentido em que a coleção analógica, que está na sua origem, pode encontrar-se dispersa por várias instituições e em diferentes localizações, e apenas se encontra organizada e disponível no mundo virtual. São exemplos destas bibliotecas, a Europeana<sup>36</sup>, a World Digital Library (WDL)<sup>37</sup> ou a Digital Public Library of America (DPLA)<sup>38</sup>. A Europeana, biblioteca digital europeia, agrega e disponibiliza em linha as coleções digitais provenientes dos países da União Europeia. A WDL, biblioteca digital mundial, disponibiliza, em linha e através de uma

---

<sup>36</sup> Disponível em <http://www.europeana.eu/>

<sup>37</sup> Disponível em <https://www.wdl.org/pt/>

<sup>38</sup> Disponível em <http://dp.la/>

interface multilíngue, obras provenientes de 193 países. A DPLA, biblioteca pública digital americana, reúne e disponibiliza em linha coleções digitais de vinte e oito instituições, distribuídas pelos Estados Unidos da América.

As bibliotecas híbridas, como a BNP ou a BnF, gerem coleções digitais e analógicas.

Os repositórios têm uma natureza diferente. “Institutional repositories (IRs) are some of the most innovative and creative components of digital libraries. They are intended to be a showcase for the research output of an academic or research institution” (Cassella, 2010, p. 210). Os repositórios de carácter académico, institucional e científico disponibilizam a produção científica da respetiva entidade<sup>39</sup> e, regra geral, têm como objetivo, recolher, agregar e indexar os seus conteúdos científicos. “An institutional repository is a database with a set of services to capture, store, index, preserve and redistribute a university’s scholarly research in digital formats” (Barton & Waters, 2004, p. 10). A produção científica recolhida nos repositórios pode ser de variada tipologia e incluir materiais didáticos, ou de apoio à atividade letiva.

Aside from article preprints (i.e., the unedited, non-peer-reviewed versions of articles) and postprints (final versions of the article created by the publisher or author), authors may also self-archive a wide variety of other types of digital scholarly works, such as books, presentations, teaching materials, technical reports, and theses and dissertations. (Bailey Jr, 2008, p. 5)

---

<sup>39</sup> No caso português os vários repositórios institucionais estão reunidos no RCAAP - Repositório Científico de Acesso Aberto de Portugal. É possível a pesquisa em cada instituição ou em todo o repositório, havendo ainda a hipótese da pesquisa se realizar só em Portugal ou também no Brasil, nas instituições aderentes.

Da mesma forma, há também várias modalidades de introdução de conteúdos, em autossucessão, em que os conteúdos são incorporados pelos próprios autores, ou por equipes especializadas vinculadas ao serviço. Atualmente, tende a privilegiar-se a autossucessão, desde que os sistemas permitam aferir a qualidade dos conteúdos e controlar a uniformidade da informação e dos metadados adicionados. Em regra, os repositórios disponibilizam “conteúdos de natureza científica de forma fácil, rápida, eficaz, em formato digital e em tempo real, ou seja, permitem disponibilizar a informação logo após a sua produção e por intermédio dos seus próprios produtores” (Rodrigues, Amante, Pais, Segurado, & Lopes, 2014, p. 16). Além disso, “agregam dados, metadados e documentos num único local, o que apresenta vantagens, quer para as instituições/organizações, quer para os investigadores a nível individual” (Rodrigues et al., 2014, p. 17). Os repositórios possuem, regra geral, políticas de persistência e preservação das coleções.

Uma das principais características dos repositórios é a disponibilização da informação em acesso aberto. No entanto, existem repositórios cujo acesso gratuito é restrito a membros; caso contrário, o acesso é pago.

Os primeiros *corpora* surgiram, ainda sem suporte informático, na área da linguística (Sardinha, 2000), sendo que o início das bibliotecas digitais se refere, habitualmente, ao *Index Thomisticus*, o *corpus* dos textos de São Tomás de Aquino, trabalhado por Robert Busa (vd. 2.1 Desenvolvimento das Humanidades Digitais).

Un corpus lingüístico es un conjunto de datos lingüísticos (pertenecientes al uso oral o escrito de la lengua, o a ambos), sistematizados según determinados criterios, suficientemente extensos en amplitud y profundidad

de manera que sean representativos del total del uso lingüístico o de alguno de sus ámbitos y dispuestos de tal modo que puedan ser procesados mediante ordenador con el fin de obtener resultados varios y útiles para la descripción y el análisis. (Sánchez, 1995, pp. 8–9)

Um *corpus* é constituído por um conjunto de textos escritos ou orais, reunidos criteriosamente, de forma a constituir um conjunto coerente para a análise linguística. Os documentos são estruturados de forma a ser possível realizar o reconhecimento e a extração automática da informação, relativa tanto ao conteúdo, como às estruturas linguísticas. “A Linguística de Corpus é uma metodologia [...] para os estudos da língua e pode ser aplicada em muitas áreas linguísticas, tais como: sintaxe, semântica, fonética e sociolinguística, dentre outras” (Kader & Richter, 2013, p. 14). Os textos que integram os *corpora* são naturais, ou seja, “autênticos”, isto é, são aqueles que existem na linguagem e que não foram criados com o propósito de figurarem no *corpus*” (Sardinha, 2000, p. 336), devendo constituir uma amostra da realidade e ser representativos.

O *corpus* é uma biblioteca digital construída com objetivos específicos. Inicialmente, os modelos automáticos de análise linguística limitavam-se a realizar indexações e extrações de frequências de ocorrências e concordâncias, mas têm evoluído no sentido de construir sistemas eficazes para o tratamento de textos reais, ou naturais, permitindo o desenvolvimento de ferramentas para vários níveis de análises e descrições linguísticas em *corpora* extensos e enriquecidos com diversos tipos de anotações (Nascimento, 2000). Por conseguinte, a análise de *corpora* “dedica-se à exploração da linguagem através de evidências empíricas, extraídas por meio de

computador” (Sardinha, 2000, p. 325). Os *corpora* são também utilizados por outras disciplinas como, por exemplo, a análise do discurso (Léon, 2006).

Os agregadores são bibliotecas digitais que alojam no seu sistema os metadados dos objetos digitais fornecidos por outras entidades. “Agregador (‘Aggregator’) – entidade que, ao mesmo tempo, é fornecedor de serviços e fornecedor de dados. Como fornecedor de serviços, agrega fontes de dados que posteriormente irão ser fornecidas a outros fornecedores de serviços” (Pedrosa, Edmundo, Borbinha, & Freire, 2012).

Um exemplo de biblioteca agregadora é a Europeana (como, de resto, também o são a WDL e a DPLA). A Europeana apresenta-se como uma plataforma de pesquisa multilingue que agrega os metadados das coleções digitais, que se encontrem em domínio público, provenientes de museus, bibliotecas ou arquivos, e consideradas representativas do património cultural e científico da Europa. A Europeana define-se, por conseguinte, como uma biblioteca em sentido lato, congrega e disponibiliza em linha 53.240.098<sup>40</sup> objetos de arte, artefactos, livros, documentos, imagens fixas e em movimento e arquivos sonoros, sejam originais ou reproduções de objetos com teor patrimonial nos domínios da arte, ciência e cultura.

As diferenças entre as várias tipologias de bibliotecas com acervos digitais podem ser observadas na Tabela 2.

---

<sup>40</sup> Número atualizado em 07/08/2017.

Tabela 2 – Termos associados a bibliotecas com acervos digitais

Designação	Obras em domínio público	Obras acabadas de editar	Metadados das obras	Coleção da instituição	Coleções de várias instituições	Marcação e anotação
BDLA	X		X	X		
B Virtuais	X		X		X	
Repositórios		X	X	X		
Agregadores	X		X		X	
<i>Corpus</i>	X		X			X

Os repositórios disponibilizam a produção científica atual das instituições, as BDLA, as bibliotecas virtuais e os agregadores apenas colocam em linha as obras que se encontram em domínio público<sup>41</sup>, sendo as restantes obras acessíveis nas bibliotecas, por acesso digital ou analógico. Os *corpora* são constituídos para fins específicos na linguística e na análise do discurso.

Em contrapartida, o conceito de biblioteca digital, embora seja equivalente ao de biblioteca virtual, é mais complexo e implica a existência de todos os serviços inerentes a uma biblioteca física. “In general, researchers focus on digital libraries as content collected on behalf of user communities, while librarians focus on digital libraries as institutions or services” (Borgmann, 2000, p. 35). São as tarefas atribuídas aos bibliotecários que definem a missão da biblioteca “as organizations that select, collect, organize, conserve, preserve, and provide access to information on behalf of a community of users” (Borgmann, 2000, p. 38), e que são transferidas, ou espelhadas, na biblioteca digital, e é com base nestes pressupostos que a Organização das Nações

---

<sup>41</sup> Em Portugal de acordo com art.º 31.º do Decreto-Lei n.º 16/2008, de 1 de abril, as obras entram em domínio público 70 anos após a morte do autor, se não penderem sobre a obra direitos conexos.

Unidas para a Educação, a Ciência e a Cultura (UNESCO) e a International Federation of Library Associations and Institutions (IFLA) definem biblioteca digital.

A UNESCO aprovou o *Manifesto da IFLA/UNESCO sobre as bibliotecas digitais* (IFLA & UNESCO, 2011), na 36.ª Conferência Geral (Paris, 25 de outubro a 10 de novembro de 2011). Neste texto, define-se que a biblioteca digital, com a missão de fornecer o acesso direto à informação, tem de ser constituída de acordo com as normas internacionais, ser persistente no tempo e manter os conteúdos devidamente organizados e estruturados, usando as novas tecnologias no acesso e na recuperação da informação.

Una biblioteca digital es una colección en línea de objetos digitales de buena calidad, creados o recopilados y administrados de conformidad con principios aceptados en el plano internacional para la creación de colecciones, y que se ponen a disposición de manera coherente y perdurable y con el respaldo de los servicios necesarios para que los usuarios puedan encontrar y utilizar esos recursos.

La misión de las bibliotecas digitales consiste en proporcionar acceso directo a recursos informativos, digitales y no digitales, de manera estructurada y fiable, para de esa manera vincular la tecnología de la información, la educación y la cultura en las bibliotecas actuales. (IFLA & UNESCO, 2011)

As bibliotecas digitais surgiram na década de 1990, quando a tecnologia o permitiu. Até ao final do século XX, houve um período experimental, durante o qual foram sendo definidas as boas práticas para a digitalização. Como um dos grandes marcos dessa

época, refere-se a Library of Congress<sup>42</sup> que, em 1994, iniciou o programa de digitalização<sup>43</sup>.

O termo biblioteca digital (*digital library*) começou a ser utilizado em 1996, pelos investigadores de Ciências Sociais (Besser, 2002). Na maioria dos casos, as bibliotecas digitais foram criadas por bibliotecários, arquivistas (Lucía Megías, 2010, p. 373) e informáticos, a quem era pedido, sobretudo, a transposição de suportes, mantendo os atributos correntes da biblioteca física convencional. Um reflexo desta influência é ainda hoje perceptível na pesquisa da informação, feita essencialmente através dos elementos da ficha bibliográfica (autor, título, editor, etc.). Esta fase marca o início da disponibilização em linha das obras digitalizadas na íntegra, maioritariamente em formato imagem, e ainda com um carácter embrionário e experimental (Choudhury & Seaman, 2008).

A partir da década de 2000, procedeu-se à consolidação das bibliotecas digitais, com a criação de normas (Anexo XI) e boas práticas. Em 2005, começou a discussão acerca da Biblioteca Digital Europeia que, em 2008, foi colocada em linha com a designação de Europeana (“Europeana Collections”, 2008).

Atualmente, podemos distinguir dois tipos de bibliotecas digitais:

- As de serviços básicos, em que o acesso se realiza através da pesquisa bibliográfica e os acervos se encontram, maioritariamente, em formato imagem (PDF, JPG, PNG, etc.);

---

<sup>42</sup> Disponível em <https://www.loc.gov/>

<sup>43</sup> Disponível em <https://www.loc.gov/collections/>

- As de serviços de valor acrescentado, que disponibilizam as obras estruturadas (como, por exemplo, o texto transcrito e marcado), permitindo uma pesquisa mais elaborada, quer através da ficha bibliográfica ou dos metadados, quer em texto livre, e podem disponibilizar ferramentas que facilitam a análise dos documentos.

As bibliotecas digitais, “cuja acessibilidade tornou mais rápida e menos árdua a actividade do investigador, que deixou de estar condicionado aos horários de arquivos e bibliotecas convencionais” (Gonçalves & Banza, 2013, p. 2), apresentam algumas vantagens face às bibliotecas convencionais. “Digital libraries offer unique ways of recording, preserving, and propagating culture in multimedia form” (Witten et al., 2009, p. 6). As bibliotecas digitais vocacionadas para a investigação devem dar particular atenção à utilização e à reutilização da informação para a criação de novo conhecimento, sem esquecer a descrição e a estruturação do objeto digital e a conservação e a preservação da informação ao longo do tempo (Borges, 2006; Freitas, 2010; Neves & Borges, 2011).

As bibliotecas digitais, por um lado, facultam o acesso imediato à totalidade da obra, a obras raras ou reservadas que por motivos de preservação são de difícil acesso, enquanto que, por outro lado, permitem que vários leitores em sítios distintos consultem a mesma obra em simultâneo, sem os constrangimentos impostos pelo horário e regras da biblioteca física convencional. Além disso, as bibliotecas digitais disponibilizam cada vez mais recursos, serviços e funcionalidades que facilitam a consulta do exemplar em moldes vantajosos face à consulta presencial. Por tudo isso, aumentam a visibilidade das coleções, bem como a sua difusão, acesso e preservação.

As bibliotecas digitais disponibilizam a informação de forma estruturada e organizada. As obras que colocam em linha estão devidamente identificadas, através da catalogação bibliográfica e dos metadados. Tal como numa biblioteca convencional, o utilizador tem a garantia da integridade e fiabilidade das obras que lhe são disponibilizadas. Uma biblioteca digital, tal como as restantes, exige preservação, para que no futuro seja possível continuar a aceder aos recursos, mas, neste caso, o processo é mais complexo, pois implica a preservação dos ficheiros, uma eventual migração para novos formatos atualizados, assim como a atualização dos programas de gestão e do equipamento (Aalberg, Papatheodorou, Dobрева, Tsakonas, & Farrugia, 2013; Bloch & Hesse, 2015; Lucía Megías, 2010; Witten et al., 2009).

A preservação digital assegura que a informação digital se mantenha acessível, interpretável e autêntica. Tal como a congénere física, está em constante crescimento, pressupondo-se que o novo material é adicionado de forma mecânica (Witten et al., 2009). É a organização, preservação, pesquisa e recuperação da informação que distingue as bibliotecas digitais do ambiente quase caótico da World Wide Web (Witten et al., 2009, p. 9), facilitando o acesso a informação organizada (Borgmann, 2000; Lesk, 2005; Witten et al., 2009).

As bibliotecas digitais e, dentro destas, de forma particular, as BDLA, estabelecem critérios para a constituição e gestão da coleção. A definição de critérios, sendo essencial em qualquer biblioteca, é indispensável nas bibliotecas digitais, sob pena de se perder acesso aos seus conteúdos. Estes critérios, ou políticas, aplicam-se à seleção das obras, à escolha do equipamento, às normas de digitalização a seguir, ao esquema de metadados e à preservação a implementar, assim como, à disponibilização em linha

e à divulgação (Fox, 1993; Witten et al., 2009). É através da definição e aplicação de critérios estritos que se constroem bibliotecas digitais coerentes.

Em contrapartida, as bibliotecas digitais devem, também, ser construídas de forma a que a sua utilização seja intuitiva. “Digital libraries should be easy to learn, to use, and to relearn, and should be flexible in adapting to a more diverse user population” (Borgman, 2000, p. 119). Além de serem fáceis de utilizar, devem adaptar-se a mais do que um tipo de utilizador.

De tudo isto, decorre que o tema da biblioteca digital pode ser analisado através de vários prismas, incluindo as questões de planificação, implementação, avaliação e impacto junto dos utilizadores preferenciais que, no âmbito deste estudo, são os investigadores em Humanidades. A biblioteca digital pode definir-se como um somatório de camadas, sendo que nem todas integram a sua totalidade. O ponto de partida é a obra analógica, neste caso, o livro, e os critérios que levaram à sua seleção. Segue-se a digitalização e o conjunto de políticas a seguir, onde se inclui a definição dos formatos de arquivo e de consulta, a resolução da imagem digitalizada e os respetivos metadados. As camadas seguintes correspondem à edição digital e à disponibilização em linha. Nesta, podem ser definidas várias subcamadas: forma de disponibilização, através de páginas HTML estáticas ou dinâmicas, de OPAC (*online public access catalog*, ou catálogo em linha), ou de sistemas de informação que, além de permitirem a pesquisa em catálogo, organizam a informação por diferentes categorias de itens, facilitando a pesquisa e a recuperação da informação.

O levantamento de bibliografia publicada sobre o tema permite observar que os estudos relacionados com as bibliotecas digitais constituem um campo de investigação

em aberto. Há uma assinalável preponderância de estudos e soluções tecnológicas aplicadas a conjuntos muito específicos de dados (por exemplo, dados sísmicos). As normas a utilizar são definidas de forma pontual e dispersa, ainda que se detete um processo cumulativo que aponta para a respetiva evolução no sentido da generalização e uniformização.

A análise dos artigos publicados, desde 2001, em revistas científicas mostrou continuam a surgir novos parâmetros de estudo, em simultâneo com soluções tecnológicas inovadoras. Além disso, as normas para a construção de uma biblioteca digital têm vindo a acompanhar a evolução técnica. No entanto, ao longo da análise dos vários artigos, é perceptível que não se alcançou ainda a formulação de um modelo satisfatório para a pesquisa e recuperação da informação, atendendo, sobretudo, às exigências e expectativas dos investigadores em Humanidades. Em contrapartida, a maioria dos artigos selecionados contribui de forma positiva para a definição de um conjunto de linhas orientadoras para o desenho desse modelo. São, assim, apresentados projetos ou programas que, de alguma forma, podem vir a contribuir para a alteração, atualização ou correção, dos atuais parâmetros das bibliotecas digitais, ao longo das várias fases de construção: digitalização e preservação; metacodificação; conversão de imagens em texto; aplicativos para a construção de bibliotecas digitais; pesquisa e recuperação da informação.

### 3.1. Digitalização e preservação

#### Formatos

**TIFF** - adequado para as imagens matriz, formato de preservação.

**JPG2000** - ainda não é completamente compatível com um formato para preservação.

**PDF-A** - adequado para obras nadas digitais.

**PDF** - em muitas BD não possuem dados estruturais, relativos à estrutura física da obra.

**Perfil de cor** - é a informação, que descreve a forma como o equipamento gere a cor, a gama de cores com que o equipamento funciona e quais as aproximações realizadas.

A digitalização e colocação em linha dos acervos é uma atividade corrente em bibliotecas e instituições afins. A criação de uma biblioteca digital, seja de livro antigo ou de outras espécies bibliográficas, museológicas ou arquivísticas, requer o planeamento de todo o processo (Dawei & Yigang, 2010; Garrison, 2001; Kott et al., 2006), entre a definição, ou redefinição, do fluxo de trabalho e a colocação e manutenção em linha.

O processo de seleção de obras que constituem o *corpus* da biblioteca digital é uma fase crucial do processo. Não obstante, não é especificamente apresentado nos vários artigos analisados, nomeadamente, naqueles que apresentam estudos de caso. Isto poderá justificar-se pelo carácter específico e único deste processo, sem permitir generalizações, nem extrapolações, dependendo do conteúdo dos acervos, do seu estado de conservação, das condições materiais para a execução do projeto e dos objetivos singulares que o norteiam.

Na sequência do processo de seleção das obras, a escolha dos formatos dos ficheiros é crucial, quer para uma adequada política de preservação digital, quer para corresponder às expetativas e contingências dos utilizadores.

O formato TIFF é comumente definido como o mais adequado para as imagens matriciais de preservação. A escolha do formato JPG2000 em substituição do TIFF, para ficheiros de preservação apresenta algumas fragilidades e ambiguidades. Knijff (2011) analisa o comportamento do JPG2000 em procedimentos de preservação, como a emulação ou a migração dos dados para novos formatos. Salienta, nomeadamente, o facto de este formato não guardar o perfil de cor<sup>44</sup> em todas as situações, dado que a gestão da cor só é possível no espaço de cor RGB. Tendo sido realizadas experiências de conversão de TIFF com vários perfis de cor para JPG2000, foi possível concluir que este formato não é adequado à preservação, uma vez que o espaço de cor e a resolução se podem perder durante o processo de migração, comprometendo a correta visualização da imagem (Knijff, 2011, p. 6). No entanto, se o formato for futuramente alterado, corrigindo estas situações, o JPG2000 poderá ser encarado como ficheiro de preservação (Knijff, 2011).

O formato PDF-A é adequado para as obras nascidas digitais, porque tem associados os metadados e a estrutura da obra (Noonan, McCrory, & Black, 2010). Contudo, no caso das obras digitalizadas, é necessário realizar um trabalho póstumo de edição para lhes agregar os metadados e os marcadores da respetiva estrutura (Noonan et al., 2010). Embora os ficheiros PDF sejam largamente utilizados nas bibliotecas digitais, muitos deles não possuem dados estruturais, relativos à estrutura física da obra (linhas, colunas, etc.), nem lógicos (cabeçalho, tabelas, legendas das ilustrações, capítulo, etc.), apesar de o formato o permitir (Gao et al., 2011), por carência de edição digital. Gao et al. (2011), salientando que a falta da estrutura lógica e física compromete a utilização

---

<sup>44</sup> O ICC profile, definido pelo International Color Consortium, define os atributos de cor dos equipamentos.

de algoritmos, propõe o sistema Structure Extraction from Books (SEB) para extração automática da estrutura do PDF e dos respetivos metadados. Apesar de a sua utilização estar vocacionada para a produção massiva de e-books, considera-se que este sistema também poderia ser útil para enriquecer os PDF de obras digitalizadas em bibliotecas digitais.

Neste levantamento, não são frequentes artigos sobre o formato das imagens, o que se pode justificar pelo facto de ter sido normalizado antes do período abordado (Anexo XI). É genericamente aceite pela comunidade das bibliotecas digitais que, para documentos impressos em papel, a digitalização deve de ser realizada em formato TIF, a cores e com a resolução de 300 ou 600 ppp. A partir destes ficheiros que constituem a imagem matriz, criam-se cópias de consulta, em variados tipos de formato e resolução, consoante o fim a que se destinam.

### 3.2. Metadados

**Metacodificação** - ação de adicionar metadados.

**Dublin Core** - esquema de metadados para descrever objetos digitais.

**MODS** - esquema de metadados extraídos a partir do registo bibliográfico Marc.

**TEI** - Text Encoding Initiative, consórcio que desenvolve e mantém uma norma para a representação de textos em formato digital.

**TEI tag set** - esquema de metadados a aplicar aos textos digitais definidas pelo TEI.

**Dados bibliográficos** - permitem uma pesquisa pobre.

**Ontologias** - modelo de dados que representa formalmente um conjunto de conceitos num domínio e os relacionamentos entre estes. Problema decorrente: a sua grande proliferação.

**Endereços persistentes** - endereço de um recurso informático que não muda para o utilizador.

**Ferramentas** (para obras em formato texto) - Meta-Dex - índices remissivos.

**Projetos colaborativos** – projetos que englobam áreas do saber.

Se um dos principais objetivos da biblioteconomia é a organização dos materiais, tendo em vista uma melhor localização da informação, nas bibliotecas digitais a organização e a metacodificação é ainda mais pertinente. A pesquisa e a recuperação da informação passam, em larga medida, pelos metadados, incluindo os dados bibliográficos ou outros que possam contribuir para melhorar o acesso ao conteúdo das obras. No conjunto em análise, registou-se um número substantivo de artigos relativos à metacodificação, apresentando estudos de caso, procedimentos ou programas para agilizar o processo de associação de metadados às obras digitalizadas.

No caso da biblioteca digital do Colorado (Garrison, 2001), um dos objetivos era a criação de um catálogo conjunto, que permitisse realizar uma pesquisa simultânea em obras de biblioteca, museu e arquivo. A solução encontrada foi a criação de um catálogo único, convertendo os catálogos parcelares utilizando o esquema Dublin Core, associando a cada obra o respetivo endereço digital. Para enriquecer a pesquisa, foi, ainda, testada a utilização da Classificação Decimal de Dewey (Dewey Decimal Classification), através de conversão automática, a qual se revelou pouco eficaz e, por isso, não teve seguimento. Em contrapartida, a base de dados dos metadados dos objetos digitais foi considerada uma boa solução, permitindo a replicação dos resultados noutras situações tendo sido, posteriormente, adotada na Europeia, em 2008, e na Digital Public Library of America, em 2010.

O projeto Aquifer (Kott et al., 2006), desenvolvido pela Digital Library Federation (DLF), entre 2005 e 2011, altura em que foi descontinuado, direcionou a pesquisa para a

implementação de um conjunto de boas práticas<sup>45</sup> para a construção de uma biblioteca digital, com recursos distribuídos e endereços persistentes para as obras. O projeto incluía a aplicação do OAIster, um catálogo em linha com registos bibliográficos de obras em acesso livre que, para a recolha e recuperação de dados, utiliza o OAI-PMH, um protocolo desenvolvido pela Open Archive Initiative para a recolha e registo de metadados nos repositórios e bibliotecas digitais. Numa tentativa de aumentar a eficiência da pesquisa, diminuindo os recursos humanos utilizados e rentabilizando a utilização dos computadores para associar os metadados às obras, o projeto desenvolveu o MODS, conjunto de metadados extraídos a partir do registo bibliográfico (Marc) existente nos respetivos OPAC.

O projeto Emblematica Online (Cole, Han, & Vannoy, 2012), desenvolvido pela Herzog August Bibliothek e pela Universidade de Illinois, tem como objetivo a construção de um portal (OpenEmblem Portal) onde se apresentam emblemas e respetivas legendas, enriquecidos, ainda, com metadados descritivos. O portal está construído de forma a suportar a notação sistemática e hierárquica Iconclass, com *thesauri* multilingue, de forma a facilitar a pesquisa, a recuperação e a reutilização da informação nos livros de emblemas. Os registos foram construídos de acordo com o Wolfenbüttel Emblem Schema. Os metadados são aplicados em dois níveis, isto é, à obra (volume) e às ilustrações, e utilizam as etiquetas de marcação do TEI (Text Encoding Initiative), para estruturar a obra, e o MODS, como sistema de metadados extraído do registo bibliográfico em formato Marc. Apesar de o processo de disponibilização das obras em

---

<sup>45</sup> Por boas práticas entende-se, neste caso, um conjunto de procedimentos devidamente testados e analisados, que conduzem ao resultado pretendido, sem, no entanto, estarem reconhecidos universalmente e constituídos em norma.

linha ser moroso, uma vez que é realizado manualmente, é viável enquanto procedimento experimental de pesquisa, dado que o universo dos livros de emblemas é diminuto, estimando-se em cerca de 6.500 unidades (Cole et al., 2012). Este projeto assentou num trabalho de investigação colaborativo, envolvendo académicos, nomeadamente, especialistas em emblemática e iconografia, bibliotecários e informáticos. Os resultados revelam-se promissores, sendo viável a aplicação desta metodologia no tratamento de obras com elevado número de ilustrações.

A ferramenta Meta-Dex (Huggett & Rasmussen, 2012) permite a criação de índices remissivos a partir de obras digitalizadas, facilitando o acesso ao respetivo conteúdo. Os índices podem ser criados para cada obra individualmente, ou ser aplicados a um conjunto de obras digitalizadas, constituindo um meta-índice da coleção. As várias tarefas de criação deste meta-índice incluem a extração, a limpeza, a expansão e a agregação dos dados, através da própria ferramenta. Huggett e Rasmussen (2012) apresentam os resultados da aplicação desta ferramenta a uma coleção com obras em texto integral, onde cada obra é apresentada com a respetiva estrutura (página, capítulos, subcapítulos, etc.) e com índices remissivos que são transversais a toda a coleção, otimizando o acesso à informação. A experiência está a ser alargada a um maior número de obras e estão a ser testadas novas interfaces, de forma a tornar o produto final mais amigável ao utilizador. Em certa medida, os princípios norteadores desta experiência estão subjacentes à prática da maioria das bibliotecas digitais das editoras.

A partir da biblioteca digital HathiTrust, constituída com obras provenientes de vários acervos, Fenlon, Fallaw, Cole e Han (2014) estudaram a operabilidade da associação de

metadados às obras digitalizadas a partir do respetivo registo bibliográfico e avaliaram a utilidade dos metadados com base no formato dos registos bibliográficos Marc. Reconhecem que os registos bibliográficos em formato Marc, convertidos em metadados e associados às respetivas obras digitais, acarretam alguns problemas para a recuperação da informação, pelo que não os consideram absolutamente fiáveis. Por exemplo, o nome do autor surge por ordem inversa, apresentando o apelido em primeiro lugar, e os títulos podem apresentar-se em sete campos diferentes<sup>46</sup> e com diferente informação em cada um, o que dificulta a pesquisa. De facto, o formato Marc foi pensado para uma pesquisa rígida por campos, sendo, por isso, pouco maleável em articulação com sistemas mais dinâmicos. Numa avaliação empírica preliminar, os autores sugerem que há espaço para melhorar a precisão de, pelo menos, alguns dos metadados herdados do formato Marc, mas afirmando que são necessários outros metadados para uma pesquisa e recuperação da informação mais assertiva.

As ontologias, enquanto modelos de dados que representam um conjunto de conceitos dentro de um domínio, bem como os relacionamentos possíveis entre estes, têm vindo a ser criadas de forma independente pelos vários investigadores ou centros de estudos. Embora sejam cada vez mais utilizadas para melhorar a seleção da informação, a proliferação dos esquemas provoca alguns estrangulamentos. Jett, Cole, Page e Downie (2016) apresentam um estudo comparativo de várias ontologias e metadados bibliográficos, com o objetivo de aumentar a utilização das bibliotecas digitais e, em particular, da biblioteca digital HathiTrust, pela comunidade académica,

---

<sup>46</sup> Campos de título no Marc 21: 210 Título abreviado; 222 Título-chave; 240 Título uniforme; 242 Título traduzido pelo catalogador/agência bibliográfica; 245 Indicação de título; 246 Formas variantes de título; 247 Título anterior (vd. <http://www.loc.gov/marc/>).

cada vez mais exigente na aplicação dos metadados na constituição das coleções digitais. Dada a proveniência diversificada dos acervos que constituem esta biblioteca, os metadados adicionados seguiram diferentes critérios e o reconhecimento de caracteres foi realizado com diferentes fluxos de trabalho, o que conduziu a resultados heterogêneos. Além disso, apenas uma percentagem reduzida (34%) das obras está em acesso livre. Os autores analisam várias ontologias: MADSRDF/MODSRDF; Bibframe; schema.org; BIBO; e FaBiO. O estudo comparativo é viável, dado que apresentam muitas semelhanças entre si. Alguns dos pontos comuns são relativos a lacunas detetadas no desenvolvimento destas ontologias, nomeadamente, a inexistência de uma adequada revisão e reutilização dos vocabulários e estruturas, e a falta de documentação acerca das estruturas existentes. Sugere-se, por isso, que mais importante do que criar novas estruturas, se devia investir na utilização e melhoria das estruturas existentes. A próxima etapa do desenvolvimento de ontologias deveria passar por uma investigação concertada entre os vários intervenientes, avaliando as sobreposições de ontologias de alto nível, a fim de unificar as classes relevantes e permitir que as ontologias existentes sejam aplicáveis a um determinado conjunto de dados. Ao invés de encarar cada projeto como um nicho, requerendo uma estrutura sob medida, específica de determinado domínio, a proposta de Jett, Cole, Page e Downie (2016) aponta para a reutilização das estruturas existentes, promovendo a partilha de dados interdisciplinares através da adoção de vocabulários e esquemas comuns.

As bibliotecas digitais continuam a utilizar os metadados e as ontologias, mas estas últimas ontologias complementam aqueles, tornando mais eficaz a pesquisa e a recuperação da informação. Embora não exista, ainda, uma normalização para as

ontologias, a capacidade destes modelos de dados para ligar conceitos, permitindo realizar pesquisas mais finas e elaboradas por assunto, indicam-nas como um campo em desenvolvimento.

### 3.3. Conversão de imagens em texto

**OCR** – A transcrição dos textos contidos nas imagens digitais pode ser realizada recorrendo a programas de reconhecimento ótico de caracteres. O texto obtido permite o enriquecimento dos metadados e a pesquisa por palavra nos documentos.

**Programas reconhecimento ótico de caracteres para documentos patrimoniais**

- **Gamera** - utilizado no Perseus Digital Library;
- **AGORA** - utilizado na Bibliothèques Virtuelles Humanistes;
- **Glyph Miner**;
- **Avaliação**;
- **Hidden Markov Model**;

**Ruído no reconhecimento ótico de caracteres:**

- Encadernações apertadas, a tinta irregular e repassada, as anotações dos leitores, os fundos com diferente coloração, imagens desfocadas, inclinadas ou cortadas, repetidas ou em falta.

A aplicação de ferramentas de reconhecimento ótico de caracteres (OCR) às imagens das obras digitalizadas e a respetiva transcrição é um cada vez mais um procedimento recorrente para a sua disponibilização em linha, por permitir efetuar pesquisas por palavra no conteúdo textual da obra, além de contribuir para o enriquecimento dos metadados (Choudhury, DiLauro, Ferguson, Droettboom, & Fujinaga, 2006; Tanner, Muñoz, & Ros, 2009). Os programas comerciais de OCR funcionam bem com a documentação produzida na atualidade, devido à precisão dos caracteres e também pelo facto de utilizar uma linguagem reconhecida pelos dicionários ortográficos dos programas. No entanto, em regra, não efetuam a conversão para formato texto no livro antigo de forma satisfatória. Mesmo nos programas com ferramentas de “aprendizagem” que personalizam um determinado carácter ou conjunto de

caracteres associando cada um deles a uma ou mais imagens, este processo, embora tenha melhorado nos últimos tempos, é ainda muito lento. Apesar destas limitações, as tecnologias de reconhecimento automático de caracteres contribuem para agilizar a pesquisa em grandes massas documentais. Tanner, Muñoz e Ros (2009) descrevem o processo de aplicação de OCR e respetiva transcrição dos documentos num estudo sobre os periódicos do século XIX existentes na British Library. Para lá da importância de estudos estatísticos atualizados acerca da precisão com que a conversão está a ser aplicado, a metodologia seguida neste projeto pode ser replicada ao livro antigo, embora com algumas adaptações, nomeadamente, no que respeita à escolha das zonas a analisar.

O Gamera, programa de aplicação de reconhecimento ótico de caracteres a imagens de obras digitalizadas, é apresentado em dois artigos (Choudhury et al., 2006; Droettboom et al., 2002). Droettboom (2002) apresenta o programa como uma ferramenta em acesso livre, que combina a edição de imagem com o reconhecimento de caracteres e a respetiva transcrição do texto, produzindo documentos estruturados. É de fácil utilização, mesmo por utilizadores que não possuam experiência em programação, mas exige conhecimentos especializados na língua dos documentos em processamento. Foi especificamente desenvolvido para tratamento de diversos tipos de documentos patrimoniais, como os manuscritos medievais, a notação musical antiga, textos gregos, etc., cuja impressão dos caracteres é variada e diferente das atuais. Este programa tem vindo a ser utilizado em vários projetos e, nomeadamente, na biblioteca digital Perseus (Crane, Smith, & Wulfman, 2001). Em contrapartida, na análise que faz ao Gamera, Choudhury et al. (2006) salienta a capacidade de aprendizagem como a característica mais relevante do programa, mas refere a

necessidade de manter a semântica e a estrutura das obras digitalizadas, ao aplicar-lhes o reconhecimento ótico de caracteres e a respetiva transcrição.

A avaliação da transcrição efetuada pelos programas de OCR, seja a uma obra, ou a um conjunto de obras, é um procedimento necessário, ainda que nem sempre seja exequível. Feng e Manmatha (2006) apresentam o modelo Hidden Markov Model (HMM) para avaliação e aferição dos resultados da transcrição aplicados a obras digitalizadas. A deteção da sincronia da conversão da imagem em texto produzida ao longo da obra não é uma tarefa realizável manualmente, de forma rápida e expedita. O HMM verifica se a transcrição foi aplicada de forma homogénea nos vários documentos analisados. No livro antigo, as encadernações apertadas, a tinta irregular e repassada, as anotações dos leitores e os fundos com diferente coloração, dado que o papel amarelece de forma desigual ao longo do tempo, são fatores que adicionam ruído, isto é, um conjunto de elementos indesejáveis, que comprometem a aplicação do reconhecimento ótico de caracteres, dificultando a obtenção de um texto de qualidade, editável, coerente e, sobretudo, pesquisável. Além disso, a própria digitalização também pode gerar imagens com diversos tipos de problemas: desfocadas, inclinadas ou cortadas; repetidas ou em falta. Também nestes casos, a aplicação dos programas OCR é difícil e pode revelar-se ineficaz.

A apreciação da eficácia dos programas de OCR tem de ser avaliada através do estudo estatístico dos resultados obtidos para um conjunto vasto de imagens. O modelo HMM sequencia os dados fornecidos, permitindo obter o grau de erro com que texto foi convertido. A aplicação deste modelo a 147 obras do Projeto Gutenberg permitiu verificar que o grau de exatidão de conversão das imagens em texto é muito elevado,

na ordem dos 95%, o que significa uma perda de 5 palavras em 100 consideradas corretas (Feng, & Manmatha, 2006, p. 116).

O AGORA é um programa específico para o reconhecimento ótico de caracteres (Ramel, Sidère, & Rayar, 2013) que, tendo sido utilizado no processamento das obras do Renascimento que integram o projeto Les Bibliothèques Virtuelles Humanistes (vd. 4.2.4), se encontra disponível para outras utilizações. Este projeto, desenvolvido em colaboração com o Centre d'Études Supérieures de la Renaissance (CESR), consiste na análise da configuração das páginas, na separação entre texto e elementos não textuais, no reconhecimento ótico de caracteres e respetiva conversão e em processos de transcrição de texto de livro antigo. O AGORA foi desenvolvido utilizando a Pattern Redundancy Analysis, um modelo de análise que extrai a informação específica e procede à sua indexação. Além disso, o programa realiza a análise da configuração das páginas, segmentando cada uma das imagens em dois planos: no primeiro, apresenta os conteúdos; e, no segundo, o fundo. Para aplicar o reconhecimento ótico de caracteres, o AGORA recorre ao RETRO, um programa secundário que integra uma biblioteca de caracteres em evolução, permitindo obter uma transcrição de qualidade, tendencialmente automática e com alto nível de aproximação ao conteúdo textual de cada página da obra. A biblioteca digital obtida através da aplicação destes programas permite que sejam aplicadas ferramentas digitais em processos posteriores de análise dos textos.

Budig, Dijk e Kirchner (2016) propõem o sistema semi-automatizado Glyph Miner para efetuar a conversão de imagens de documentos impressos antigos em texto. O sistema identifica um determinado glifo e mostra ao utilizador as várias formas em que ocorre,

por ordem decrescente de similitude. As correspondências corretas são validadas pelo utilizador e aplicadas a todo o *corpus*. Além de ser mais fácil validar os glifos do que as palavras, quanto maior for o catálogo das ocorrências mais fiável é o texto obtido.

O processamento das imagens, tendo em vista a obtenção de um texto editável, continua a exigir uma intervenção humana. No entanto, os programas, com mecanismos de aprendizagem, tornam a tarefa menos fastidiosa e os resultados têm vindo a revelar-se cada vez mais promissores, sendo que já é possível obter textos bastante fiáveis. Por outro lado, ainda que a oferta de programas personalizáveis com um razoável índice de fiabilidade seja vasta, a seleção de um programa de reconhecimento ótico de caracteres depende dos originais, quer no que respeita à língua do texto, quer em função dos caracteres utilizados.

As imagens digitais de livros encadernados, mesmo utilizando equipamento adequado, podem apresentar um excesso de curvatura nas páginas. Esta anomalia aumenta o índice de erros, nos textos transcritos, decorrentes da aplicação do reconhecimento ótico de caracteres, pelo que começam a surgir programas para a correção de curvatura. Rahnemoonfar e Plale (2013) propõem um método de avaliação da qualidade das imagens, recorrendo a algoritmos matemáticos, diretamente a partir das imagens digitalizadas e não a partir da avaliação dos textos produzidos depois da aplicação dos programas de reconhecimento ótico de caracteres, medindo o desvio de cada linha de texto a partir da linha base padrão. Este método pode ser particularmente vantajoso na avaliação de um grande número de imagens digitalizadas.

### 3.4. Aplicativos para a construção de bibliotecas digitais

#### **Programas para a construção de bibliotecas digitais:**

- GreenStone;
- OMEKA.

**METS** - Metadata Encoding and Transmission Standard - criado e desenvolvido para realizar a recuperação e preservação digital dos dados dos objetos digitalizados.

Atualmente, encontram-se disponíveis diversos programas para a construção e manutenção de bibliotecas digitais, entre os quais se refere o GreenStone (Bainbridge, Thompson, & Witten, 2003; Buchanan, Bainbridge, Don, & Witten, 2005), produzido pelo New Zealand Digital Library Project na Universidade de Waikato, Nova Zelândia, que tem vindo a ser desenvolvido e distribuído em colaboração com a UNESCO e a Human Info NGO, da Bélgica.

As bibliotecas digitais devem facilitar a construção de novas coleções e a respetiva disseminação na web (Bainbridge et al., 2003). Nesse sentido, o GreenStone, através de uma interface intuitiva, permite adicionar novos conteúdos e os respetivos metadados, embora exija a definição prévia das políticas a seguir, nomeadamente, os critérios de seleção, a estrutura da coleção, os índices de pesquisa, os metadados a incluir e a forma de navegação. Por omissão, o programa assume os metadados Dublin Core, mas autoriza a utilização de qualquer outro esquema personalizado. As principais vantagens associadas ao Greenstone são: a possibilidade de pesquisa de palavras ou frases no conteúdo textual de todos os documentos em simultâneo; a criação automática de índices personalizáveis, a partir dos metadados; a possibilidade de construção de uma biblioteca digital a partir de conteúdos disponíveis na web ou

criados localmente; a possibilidade de personalização do aspeto final da biblioteca digital.

Buchanan, Bainbridge, Don e Witten (Buchanan et al., 2005) apresentam o GreenStone3, cuja principal alteração é a integração do esquema de metadados METS no processo da construção da biblioteca digital, anexando os metadados aos ficheiros dos dados. O METS de cada documento é guardado numa base de dados MySQL, devido à sua conformidade com a GNU Public License utilizada no GreenStone3. A principal vantagem da utilização do METS é, por conseguinte, a interoperabilidade entre sistemas. O Greenstone3 usa uma arquitetura de construção em sete fases: as três primeiras fases identificam e armazenam cada documento; as três fases seguintes enriquecem e indexam esses documentos; a fase final efetua o controle de qualidade.

A plataforma OMEKA (Kucsma, Reiss, & Sidman, 2010) foi desenvolvida pelo Roy Rosenzweig Center for History e pelo New Media da George Mason University, nos Estados Unidos. Tal como o GreenStone, configura-se como um gestor de conteúdos gratuito e em acesso livre e é também utilizado na construção e manutenção de bibliotecas digitais, nomeadamente nos conteúdos digitais da Metropolitan New York Library Council's (METRO). O OMEKA foi escolhido para este projeto devido à facilidade de instalação e de personalização, aceitando o Dublin Core ou qualquer outro esquema de metadados. Além disso, permite a criação e a visualização de diferentes tipos de ficheiros digitais (texto, imagem, som, vídeo) integrados numa mesma biblioteca, embora a pesquisa e a recuperação da informação possam ser melhoradas (Kucsma et al., 2010). A facilidade de instalação e a versatilidade tornam o OMEKA adequado a pequenas e médias coleções digitais.

Ambos, GreenStone e OMEKA, permitem a criação e a visualização dos conteúdos digitais, são gratuitos e estão em acesso aberto. No entanto, exigem competências informáticas, nomeadamente, conhecimentos ao nível da programação, para otimizar os resultados.

### 3.5. Pesquisa e recuperação de informação

#### Projetos para melhoria da pesquisa

- Event-centric approach;
- German vocabulary database;
- Popular passages;
- Rede concept-in-context (CIC);
- Scatter/gather;
- Search and mining tools with linguistic analysis (SAMTLA);
- Topic models.

Com o aumento dos conteúdos disponíveis em linha, a pesquisa e recuperação de informação é uma questão cada vez mais pertinente no domínio das bibliotecas digitais. Um dos problemas inerentes prende-se com os erros que os metadados podem ter, seja de digitação ou outros, que podem impedir o utilizador de localizar a informação de que necessita. No artigo “That's ‘é’ not ‘p’ ‘?’ or ‘□’: A User-driven Context-aware Approach to Erroneous Metadata in Digital Libraries” (Bainbridge, Twidale, & Nichols, 2011) é apresentado o sistema Computer Says Yes (CSN) para a integração de variantes dos nomes dos autores (autoridades) em sistemas de informação baseados na web. As variantes detetadas por cada utilizador são refletidas no respetivo espaço pessoal, mas um resumo das várias alterações, sugeridas pelos utilizadores, é disponibilizado aos administradores de sistema para validação. O

sistema foi implementado em JavaScript, mas pode correr em qualquer navegador que permita o GreaseMonkey.

A obtenção de resultados positivos é o objetivo dos utilizadores quando utilizam uma biblioteca e, em particular, uma biblioteca digital. A pesquisa por assuntos pode facilitar o refinamento das pesquisas. Hagedorn, Kargela, Noh e Newman (2011) referem as pesquisas facetadas como uma boa forma de ajudar a localizar a informação. Sugerem, além disso, a utilização do sistema de classificação da Biblioteca do Congresso como complemento para o enriquecimento das pesquisas.

A recuperação de informação em textos antigos de alemão, sem ortografia normalizada, foi abordada por Ernst-Gerlach e Fuhr (2007), trabalhando sobre o *German Vocabulary Database*, um dicionário de alemão contemporâneo. Aplicaram à ortografia atual as regras ortográficas em sentido inverso, de forma a obter ortografias obsoletas e, com isso, elaborar um procedimento baseado em algoritmos para melhorar a recuperação da informação em textos antigos. Utilizando apenas a derivação padrão (*standard stemming*), os resultados ficam aquém do pretendido. Assim, foi necessário, em primeiro lugar, mapear os termos de pesquisa e, em seguida, gerar as variantes históricas correspondentes, para que, ao usar um mecanismo de pesquisa padrão, se obtivessem bons resultados. Não foi possível gerar todas as palavras, mas todas as palavras geradas estavam corretas. Em contrapartida, os pontos fracos deste método são a seleção manual dos termos e os erros associados que daí advêm.

Uma forma de navegar num conjunto de obras digitalizadas é através do aplicativo *Popular Passages*, onde a navegação é realizada através das citações e das repetições

dos autores. Com base nas citações, são geradas palavras-chave que permitem a navegação dentro das obras ou entre várias obras. Através da navegação pelas citações, também se pode observar o contexto em que foram utilizadas. Este sistema de navegação é testado na biblioteca digital de livros do Google, pelo que pode ser replicado em qualquer biblioteca digital de obras digitalizadas (Schilit & Kolak, 2008).

Clarkson, Navathe e Foley (2009) analisam modelos de interfaces e os respetivos algoritmos, procurando simplificar a pesquisa e maximizar a recuperação da informação recorrendo à organização por facetadas. Foram analisados oito sistemas, tendo em comum a utilização de sistemas de navegação facetados: Relation Browser, mSpace, Flamenco, Elastic Lists, Humboldt, Freebase Parallax, Bungee View e Nested Faceted Browser. Quando se inicia a consulta de uma determinada base de dados sem conhecer o respetivo conteúdo, a pesquisa é mais simplificada e rentável se existir um conjunto de termos pré-definidos para efetuar a pesquisa, mas pode ser otimizada ao adicionar o método de pesquisa por palavra-chave ou outro. As facetadas com barra de deslocamento (*scrolling*) vertical conseguem disponibilizar mais termos do que aquelas que colocam a barra na horizontal. Outro elemento que facilita a pesquisa é a indicação do número de itens para cada termo.

Apesar de a utilização das facetadas, como auxiliar na pesquisa, ser genericamente positiva, a sua eficácia deve ser aferida. Kules, Capra, Banta e Sierra (2009) avaliam o modo como é feita a pesquisa em catálogos com facetadas. A pesquisa é realizada através de metadados estruturados, os quais fornecem uma visão geral dos resultados, podendo incluir categorias, e permitem refinar os resultados sem ser necessário lançar nova pesquisa. O seu domínio de aplicação é muito vasto, desde os catálogos das

bibliotecas, sobretudo nos OPAC, a catálogos de compras e outros. A utilização das facetas na pesquisa bibliográfica aproxima a pesquisa em biblioteca da pesquisa na web, diminuindo a distância entre a navegação e a pesquisa, e permite uma visão geral dos recursos disponibilizados. Porém, como os termos de pesquisa são pré-definidos, a pesquisa por facetas mostra-se insuficiente para utilizadores ou investigadores com níveis de exigência mais específicos.

Outra forma de pesquisar nas bibliotecas digitais é através de *topic models* (Newman, Noh, Talley, Karimi, & Baldwin, 2010), genericamente definidos como modelos probabilístico que permitem encontrar conceitos num documento ou numa coleção de documentos. Estes modelos podem vir a ter grande impacto na forma como os utilizadores localizam a informação em bibliotecas digitais, através da capacidade de aprender e de aplicar automaticamente etiquetas de assunto a cada item de uma coleção, e à capacidade de criar de forma dinâmica coleções virtuais. Newman, Noh, Talley, Karimi e Baldwin (2010) apresentam um estudo empírico sobre estes modelos, onde avaliam a coerência e a interoperabilidade dos tópicos criados pelo *topic models*.

A recuperação de informação em documentos também é possível através de *event-centric approach* (Strötgen & Gertz, 2012), isto é, através dos elementos geográficos e temporais existentes nas obras. Para isso, é necessário, em primeiro lugar, realizar uma extração normalizada de toda a informação geográfica e temporal disponível na obra e organizar hierarquicamente essa informação, de forma a construir uma lista de eventos associados às respetivas obras. Como resultado, durante a pesquisa, é possível localizar os documentos cujos eventos se realizem numa determinada janela temporal ou geográfica, numa data ou numa localização específica. A visualização dos resultados

pode ser feita através de uma lista de eventos com os documentos que correspondem à pesquisa nos requisitos geográficos e temporais, ou através de um mapa onde os dois tipos de informação são apresentados.

O sistema *Scatter/Gather* (Gong, Ke, Zhang, & Broussard, 2013) pode ser implementado em bibliotecas digitais para melhorar a pesquisa. Em cada interação, o sistema apresenta, ao utilizador, um conjunto de tópicos principais criados a partir das pesquisas prévias que tenha efetuado. O utilizador pode construir conjuntos e subconjuntos que melhor se possam adequar aos seus objetivos de pesquisa. Este sistema é particularmente útil para ajudar a concluir tarefas com as quais os utilizadores estejam menos familiarizados, ou quando as especificações da consulta são difíceis. No entanto, a sua eficácia depende das pesquisas anteriores feitas pelos investigadores.

O sistema Search and Mining Tools with Linguistic Analysis (Samtla) é uma outra ferramenta de pesquisa apresentada por Harris, Levene, Zhang e Levene (2014). Definido como um ambiente de pesquisa integrado, funciona em linha e foi projetado em colaboração com historiadores e linguistas com o objetivo de facilitar o estudo de textos digitalizados escritos em qualquer idioma. O sistema articula-se em três componentes: *search* (pesquisa) dá a resposta às consultas dos utilizadores; *mining* (mineração) contém as ferramentas para mostrar os documentos relacionados e os padrões existentes; e *community* (comunidade) permite ao investigador ver as consultas realizadas por outros investigadores. Ao fazer a pesquisa num fragmento de texto, o sistema devolve, igualmente, os textos onde se encontrem as palavras pretendidas, numa ordem igual ou idêntica à da pesquisa efetuada. A pesquisa é feita recorrendo ao

*statistical language model* (SLM), um modelo matemático que representa a distribuição probabilística de sequências de palavras ou caracteres encontrados em *corpus* de texto em linguagem natural. O Samtla pretende ajudar os investigadores de Humanidades Digitais a encontrarem os textos relacionados com os seus temas de pesquisa, incluindo a descrição precisa, transcrição e análise de *corpora* cultural, histórico, político ou linguístico.

A pesquisa semântica pode ser mais eficiente do que as técnicas tradicionais de pesquisa lexical, dado que estas não correspondem à complexidade da linguagem natural. Hinze, Taube-Schock, Bainbridge, Matamua e Downie (2015) desenvolveram a rede Concept-in-Context (CiC) que tem a vantagem de usar ontologias pré-existentes, evitando, assim, a necessidade de uma marcação semântica integral dos documentos. A pesquisa semântica avançada é semelhante à pesquisa lexicográfica e pode ser usada nas mesmas configurações. Além disso, os textos são analisados automaticamente não apenas por termo, mas pela semântica dos respetivos conteúdos e metadados. Internamente, uma ontologia ou rede concetual é usada para traduzir palavras-chave e conceitos (desambiguação). Na primeira fase (indexação), os documentos são analisados por palavras-chave que indicam conceitos, por comparação com a ontologia existente, a partir dos quais elabora um índice de conceitos. Durante a pesquisa, as palavras-chave do utilizador são traduzidas em conceitos, os quais são igualmente usados para fazer a pesquisa no índice de documentos correspondentes. A vantagem deste processo é a desambiguação, através do sistema CAPISCO, que permite o refinamento da pesquisa.

O constante aumento dos recursos disponibilizados em linha e a crescente dificuldade em efetuar a pesquisa e a respetiva recuperação da informação, tem dado origem a vários projetos, com diferentes metodologias, mas com o objetivo comum de tornar os conteúdos das bibliotecas digitais mais acessíveis. Muitas das soluções apresentadas limitam-se a pequenas coleções, sendo expectável que a aplicação a grandes massas documentais acarrete outros problemas, ainda não detetados na fase experimental. Nos casos particulares em que foram aplicados, estas ferramentas atingem o objetivo de tornar a pesquisa e a recuperação da informação mais eficiente. Contudo, nos artigos consultados, não há referência a nenhum estudo em grande escala. As Humanidades Digitais, ou os novos métodos de investigação em Humanidades, têm vindo a impor a necessidade de criar ferramentas capazes de responder às crescentes e cada vez mais complexas exigências dos investigadores. No caso da pesquisa e recuperação da informação, é requerido um acesso facilitado e assertivo aos textos integrais, pelo que têm vindo a desenvolver-se modelos estatísticos, específicos para linguagem natural. Por outro lado, num ciclo contínuo, as Humanidades Digitais conferem novas competências aos investigadores, desencadeando novas exigências em relação às bibliotecas digitais.

### 3.6. Adequação das bibliotecas digitais às Humanidades

#### **Implementar**

- Sistema de alertas;
- Páginas de novidades;
- Marcação de textos - aplicar etiquetas normalizadas (esquema TEI) ao conteúdo dos documentos, indicando parágrafos, frases, sinais gráficos, etc.

#### **Projetos**

- Anotação das imagens e ilustrações;
- Anotação de fac-símiles - Annotated Facsimile Edition;
- Text-Image-Link-Editor - associar, às imagens das obras digitalizadas, a transcrição do respetivo texto e anotações;
- e-codice - Linked Open Data modelo para publicar dados estruturados.

#### **Biblioteca produtora da informação**

#### **Bibliotecário parceiro nos projetos de Humanidades Digitais**

A investigação em Humanidades exige uma biblioteca digital adequada e atualizada em função do crescente grau de exigência, nomeadamente, no que respeita ao acesso, à informação na íntegra e à adequação dos formatos disponibilizados (Padilla, 2016). Apesar da constante evolução das bibliotecas digitais, ainda não se encontram devidamente estruturadas para a investigação em Humanidades.

Os utilizadores das bibliotecas e em particular das digitais, têm dificuldade em manter-se atualizados face aos novos conteúdos, adicionados ou atualizados, pelo que a criação de sistemas de alertas é pertinente e necessária, não só para os leitores, ou utilizadores, como também para os bibliotecários e administradores de sistema. As bibliotecas digitais, com significativas diferenças entre si, variam, nomeadamente, no que respeita à estrutura interna, aos ficheiros e aos metadados, além de que o sistema de alerta ou de notificação também varia de modelo para modelo. Os alertas através de mensagem são de dois tipos: um, através de mensagem pessoal para o utilizador; e, o outro, através de anúncio na biblioteca digital. Os anúncios abrangem eventos como

a adição de novos conteúdos, a remodelação dos existentes, alterações nos metadados, introdução de novas coleções, entre outros. As bibliotecas digitais que não possuam identificadores persistentes para os respetivos recursos não podem gerar sistema de alerta, dado que, durante pesquisas sucessivas, ao mesmo recurso é associado um URL diferente em cada pesquisa (Buchanan & Hinze, 2005, p. 133). Através da análise a vários serviços de alerta implementados, Buchanan e Hinze (2005) concluem que os alertas publicados nos sítios eletrónicos das bibliotecas digitais são cada vez mais utilizados e que as páginas das novidades estão amplamente difundidas, enquanto as mensagens individualizadas, elaboradas em função do perfil do utilizador, são sobretudo utilizadas nas bibliotecas digitais comerciais. Além disso, enunciam os requisitos conceptuais a considerar neste tipo de serviço: alargar o serviço de alertas a todo o conteúdo, sem se cingir apenas aos novos documentos adicionados; ter em consideração os conceitos tradicionais da biblioteconomia, para definir um serviço de qualidade; desenhar o serviço de alerta em função das especificidades de cada biblioteca digital. Por outro lado, reconhecem que é difícil implementar este serviço em bibliotecas sujeitas a alterações, pela modificação ou exclusão de documentos, apontando para a necessidade de prosseguir a investigação no sentido de encontrar uma solução adequada a este problema.

A anotação das imagens e ilustrações dos textos facilita o seu estudo. Agosti, Ferro e Orio (2005) desenvolveram o projeto colaborativo *Imaginum Patavinae Scientiae Archivum* - Archive of the Patavinian Scientific Images (IPSA) para elaborar uma anotação das iluminuras em torno das iluminuras dos manuscritos, o qual pode ser aplicado aos incunábulos e ao livro antigo com ilustrações, capitulares ornamentadas, tarjas e vinhetas, iconografia, mapas, entre outras. Para cada imagem é possível

estabelecer quatro tipos de ligação, sempre devidamente fundamentadas e com a respetiva autoria: o *link authorship* faz a ligação entre duas ou mais imagens do mesmo autor; o *link typology* faz a associação entre duas ou mais imagens que se enquadrem na mesma tipologia, ou que de alguma forma se relacionam entre si; *link symmetry*, automaticamente adicionado pelo sistema, caso detete alguma relação entre a imagem A e B, estabelece uma relação simétrica entre B e A; *paths*, através das associações anteriores cria relações entre imagens que, aparentemente, não estariam relacionadas. No universo deste projeto, as ligações estabelecidas podem ser de carácter histórico, estético e técnico. Para as clarificar, foi criada uma taxonomia dividida em duas classes, uma, hierárquica e, outra, relacional. As ligações são complementadas com mais duas entidades semiológicas: o significado (*meaning*), relativo à descrição; e o signo (*sign*), relativo ao elemento gráfico ou textual representado. Toda esta cadeia de informação está ligada ao objeto digital representado.

Audenaert e Furuta (2009) desenvolveram o produto Annotated Facsimile Edition (AFED) de anotação de fac-símiles, isto é, a reprodução digitalizada de obras disponibilizadas em linha e em formato de imagem. O AFED fornece um modelo flexível que pode servir como base para o desenvolvimento de ferramentas destinadas a apoiar a edição de documentos visuais para a pesquisa académica. Como tal, este produto pode ter aplicabilidade mesmo em bibliotecas digitais que, embora tenham as obras em formato texto, disponibilizem também o fac-símile.

Dado que permite reordenar as imagens das páginas das obras, o AFED pode ser particularmente útil para o tratamento do livro antigo, onde são frequentes os erros

de encadernação, com os cadernos agrupados numa sequência arbitrária, ou de paginação, com falhas ou incorreções nos números de páginas impressos. Além disso, o AFED permite adicionar outros dados, relativos, por exemplo, à fonte utilizada, ao espaçamento do texto, ou à cor do papel, os quais podem ser elementos relevantes para o estudo da história do livro enquanto objeto físico. Em articulação com a marcação com o TEI, permite fazer a transcrição anotada dos documentos, com a vantagem de poder associar múltiplas transcrições a um mesmo documento. Permite, ainda, editar o texto, fornecendo suporte computacional para a edição de documentos antigos. Ainda que o projeto e implementação do AFED se tenha centrado no objetivo de melhorar a compreensão e modelagem da estrutura de fac-símiles antigos, os resultados podem ser replicados em projetos de outra natureza, constituindo um modelo flexível passível de ser utilizado como base em ferramentas que suportam processos editoriais exploratórios.

O Text-Image-Link-Editor (TBLE) é uma aplicação que permite associar, às imagens das obras digitalizadas, a transcrição do respetivo texto e anotações, ligando os segmentos de texto aos segmentos de imagem correspondentes. Para a transcrição, utiliza o TEI:P5. O TBLE foi desenhado para manuscritos, mas é possível utilizá-lo em impressos (Al-Hajj & Küster, 2013). Audenaert e Furuta (2010) se, por um lado, reconhecem a extrema importância de aceder em linha aos documentos digitalizados, por outro lado, alegam que, em vez de disponibilizar textos fac-similados, seria mais útil a disponibilização de textos transcritos e devidamente marcados. As bibliotecas digitais constituem uma fonte fiável de documentos, alguns dos quais são, já na versão digital, editados pela comunidade académica. No entanto, existe uma grande diferença de escala, dado que, enquanto que uma biblioteca digital pode ter alguns milhares de

itens, os tratados pela comunidade estão na ordem das centenas (Audenaert & Furuta, 2010). Por esse motivo, estes autores consideram que as bibliotecas digitais, com diferentes níveis de sofisticação, precisam de alterar a sua arquitetura a dois níveis, isto é, na representação dos dados e na interface do utilizador (Audenaert & Furuta, 2010).

Num estudo acerca do impacto da digitalização em larga escala, Gooding, Terras e Warwick (2013) começam por diferenciar “digitalização em larga escala” (*large-scale digitization*) e “digitalização em massa” (*mass digitization*). A “digitalização em larga escala” está a ser executado por instituições detentoras do património, como bibliotecas, universidades, centros de estudo, e outras, que digitalizam os próprios acervos, enquanto a “digitalização em massa” é efetuada por grandes companhias, como a Google, com conhecimentos na área da digitalização, as quais digitalizam massivamente o espólio de outras instituições, distinguindo-se quer pela escala da digitalização, quer pela propriedade heterogénea dos acervos.

A digitalização traz benefícios à atividade académica, no sentido que torna mais fácil e alargado o acesso à informação, fazendo mesmo com que surjam novos campos de investigação. Porém, há na comunidade académica quem lhe aponte alguns aspetos menos positivos, tais como: a pouca qualidade dos textos transcritos através dos programas de reconhecimento ótico de caracteres; a pobreza dos metadados; a escolha das normas para os metadados; a pouca qualidade das imagens digitalizadas; e o predomínio da cultura anglo-americana (Gooding et al., 2013). Por outro lado, a resistência à introdução dos métodos quantitativos em Humanidades ainda se faz sentir, enquanto os mais elitistas têm reservas face à democratização dos conteúdos,

acessíveis fora dos círculos habituais da investigação. Gooding, Terras e Warwick, (2013) consideram, por isso, que é necessário incrementar uma política de divulgação acerca das vantagens da digitalização junto dos investigadores em Humanidades.

Outras vantagens reconhecidas nas bibliotecas digitais são a reutilização da informação e a interoperabilidade dos recursos. Apela-se, agora, à utilização integrada de funcionalidades e a um contínuo processo de avaliação para adequar os serviços disponibilizados aos seus utilizadores nas suas múltiplas especificidades (Blandford & Gow, 2006). Como Padilla (2016) refere, apesar de as bibliotecas digitais terem evoluído muito desde a sua criação, ainda não estão devidamente estruturadas para as Humanidades.

Library Humanities data collection models vary widely. In order to support more concerted effort in this space, it is necessary to examine models that were explicitly designed to support computational engagement with collections *as well as* collections and access mechanisms that were not clearly designed for this purpose but nonetheless hold the potential to inform model development. (Padilla, 2016, p. 6)

O aparecimento das bibliotecas digitais alterou as funções tradicionalmente adstritas à biblioteca. Se a biblioteca física é essencialmente mediadora no acesso à informação, a biblioteca digital torna-se, ela própria, produtora de informação, sobretudo junto da comunidade académica. Partindo do pressuposto de que a disponibilização de obras digitalizadas em linha implica um processo de edição digital (Sukovic, 2002), o bibliotecário está particularmente credenciado para o efeito, dado que, no exercício da sua profissão, lhe são exigidas as competências implicadas na codificação dos textos e

na metacodificação. Por conseguinte, cabe ao bibliotecário um papel importante, como parceiro, em projetos de investigação em Humanidades (Padilla, 2016; Sukovic, 2002).

Num estudo comparativo sobre projetos em Humanidades Digitais, realizados nos Estados Unidos da América, Canadá e Reino Unido (Siemens, Cunningham, Duff, & Warwick, 2011), conclui-se que as bibliotecas digitais têm vindo a ser desenvolvidas por equipas pluridisciplinares, mas, em regra, sem a intervenção dos especialistas em Humanidades que poderiam adequá-las aos parâmetros da sua investigação. Os projetos em Humanidades Digitais envolvem competências diversas, nomeadamente, na área da biblioteconomia. Enquanto as competências dos bibliotecários têm vindo a ser aproveitadas para a realização de projetos sustentados e preservados no tempo, assinala-se a falta de conhecimentos especializados nas diferentes áreas da investigação em Humanidades. A colaboração interdisciplinar é dificultada pela existência de diferentes métodos e treino de trabalho colaborativo, nas várias disciplinas. Nas bibliotecas o trabalho é muito mais colaborativo do que em Humanidades. Siemens, Cunningham, Duff e Warwick (2011) sustentam que a construção de bibliotecas ou de outros projetos digitais deveriam basear-se em trabalho colaborativo, envolvendo as áreas da biblioteconomia e da investigação em Humanidades, sugerindo que, para isso, houvesse uma formação específica, levada a cabo nas instituições universitárias, tendo como objetivo o desenvolvimento de competências funcionais para o trabalho em equipa.

Regista-se, por outro lado, o desenvolvimento de novas funcionalidades requeridas às bibliotecas digitais em função do alargamento do universo de utilizadores e ajustando-

se aos vários campos da atividade acadêmica, entre a investigação e o ensino (Kalinichenko, Korenkov, Shirikov, Sissakian, & Sunturaenko, 2003). A utilização de bibliotecas digitais pode, em particular, potenciar o ensino à distância.

Entretanto, o Linked Open Data (LOD), um modelo para publicar dados estruturados em linha e que já é utilizado em bibliotecas para otimizar os respetivos catálogos, abre a perspectiva de novas funcionalidades das bibliotecas digitais e, em particular, das que se orientam para a investigação em Humanidades. Este modelo de dados tem vindo a ser usado para descrever as relações entre imagens, textos e outros recursos, com o objetivo de melhorar a interoperabilidade de repositórios. Foi implementado na biblioteca digital de códices medievais no âmbito do projeto colaborativo *e-codices*, criado, em 2008, na Universidade de Friburgo (Sanderson, Albritton, Schwemmer, & Sompel, 2011). O modelo permitiu, num determinado conjunto de páginas de manuscritos medievais, estabelecer relações entre fragmentos, imagens, páginas desaparecidas, fora de ordem, e descrevê-las, recorrendo a duas funções ORE Aggregations e OAC Annotations. Os resultados obtidos neste projeto apontam para a sua aplicabilidade noutros conjuntos de documentos, manuscritos ou impressos, e na construção de bibliotecas digitais de livro antigo.

### 3.7. Bibliotecas digitais para as Humanidades

#### Humanidades Digitais

- Aumento da informação disponibilizada em linha;
- Divulgação da produção científica de forma apelativa;
- Produção de programas específicos para Humanidades;
- Criação de novas oportunidades de financiamento;
- Estímulo à captação de novos investigadores.

#### Biblioteca Digital

- Criação de um produto acessível;
- Fácil consulta;
- Resultados reutilizáveis;
- Utilizável pelos investigadores em Humanidades, como pelos leitores comuns.

As bibliotecas, na sua forma analógica, assumindo-se como instituições guardiãs da memória e do conhecimento que ia sendo produzido, foram essenciais para o desenvolvimento das Humanidades, ao longo dos séculos. O conceito epistemológico de Humanidades abrange um extenso campo disciplinar cujo objeto é o ser humano e as manifestações humanas relacionadas com a língua, o pensamento, a arte e a cultura, cujo estudo passa pela análise dos documentos textuais e gráficos produzidos ao longo da história. Por conseguinte, a investigação em Humanidades passa, necessariamente, pelos acervos guardados em bibliotecas e que se encontram, agora cada vez mais, amplamente disponibilizados através das bibliotecas digitais.

As we enter the twenty-first century, digital libraries appear to be as critical to humanities scholarship as brick-and-mortar libraries were to scholarship in previous centuries. Not only do digital libraries provide access to original

source material, contextualization, and commentaries, but they also provide a set of additional resources and service. (Besser, 2004, p. 557)

Se os investigadores em Humanidades continuam a necessitar de aceder às fontes, para o desenvolvimento dos seus trabalhos, as bibliotecas digitais oferecem-lhes, agora, a mais valia de novas modalidades de pesquisa e acesso à informação.

[...] given the size and scale of many projects that focus on source material – projects that often measure their duration in decades rather than years – we need to consider the role of digital libraries as evolving resources designed to support the entire life cycle of a research project. This life cycle should include the initial digitization of material, the ongoing analysis of that material and the ultimate publication of long-lived scholarly resources. (Audenaert & Furuta, 2010, p. 291)

O alcance das Humanidades Digitais ultrapassa largamente a mera transferência do analógico para o meio digital, centrando-se no desafio epistemológico e na articulação dos conhecimentos e métodos utilizados nas Ciências Sociais e Humanas com o mundo digital. “Professional research in the humanities required and still requires the study of original texts and objects – ideally in their original contexts – in order to understand who created them, why and in what settings” (Gardiner & Musto, 2015, p. 16). Numa primeira fase, as ações no âmbito das Humanidades Digitais centraram-se na digitalização e na disponibilização de fontes primárias, mas, atualmente, começa a impor-se o objetivo de construir e facultar ferramentas para a análise dessas fontes e para as disponibilizar de forma apelativa, tornando a aquisição cognitiva mais imediata e intuitiva. “The themes that emerged can be grouped under the following rubrics: scale, critical/productive theory, collaboration, databases, multimodal scholarship,

code, and future trajectories” (Hayles, 2012, p. 43). De facto, as Humanidades Digitais têm vindo a incentivar uma maior utilização dos recursos digitais por parte dos investigadores, enquanto que, em simultâneo, criam exigências em relação aos conteúdos das bibliotecas digitais e aos modelos de disponibilização da informação.

As bibliotecas digitais que incluem o livro antigo são relevantes para a investigação em Humanidades, nomeadamente, nas áreas de História, Filosofia, Literatura e Linguística.

“Not only will they provide access to a host of source materials that humanists need in order to do their work, but these libraries will also enable new forms of research that were difficult or impossible to undertake before” (Besser, 2004, p. 557). Não obstante, constata-se que a maioria das bibliotecas digitais não possui os requisitos para a investigação, nem responde às crescentes exigências dos investigadores. “Digital collections are proliferating, but most remain difficult to use, and digital scholarship remains a backwater in most humanities departments with respect to hiring, promotion, and teaching practices” (Borgman, 2009, p. 2). Christinne Borgman, investigadora na área das Ciências da Informação e Documentação, muito crítica em relação à forma como as instituições disponibilizam os respetivos acervos, fundamenta esta observação através de uma constatação direta: “My student’s complaint, ‘So what use are the digital libraries, if all they do is put digitally unusable information on the web?’ nicely captures the challenges facing the humanities today. Digital content, tools, and services all exist, but they are not necessarily useful or usable” (Borgman, 2009, p. 19).

Não obstante, Rydberg-Cox (2006) defende uma contenção na disponibilização de ferramentas, considerando que estas se podem tornar desajustadas ou supérfluas,

acabando por não ser usadas e, mesmo, ao gerar alguma confusão, tornar-se prejudiciais à consulta da obra. Por outro lado, recomenda a criação e o desenvolvimento de boas páginas de ajuda, a realização de ações de esclarecimento e formação e a produção de campanhas de divulgação e publicidade, com o objetivo de potenciar a utilização das bibliotecas de forma rentável e adequada.

Algumas bibliotecas digitais, como o Internet Archive (vd. 4.2.3), as Bibliothèques Virtuelles Humanistes (vd. 4.2.4) e a Perseus Digital Library (vd. 4.2.5) assumiram como princípios orientadores a criação de um produto acessível, de fácil consulta e cujos resultados fossem reutilizados tanto por investigadores em Humanidades, como pelo leitor comum.

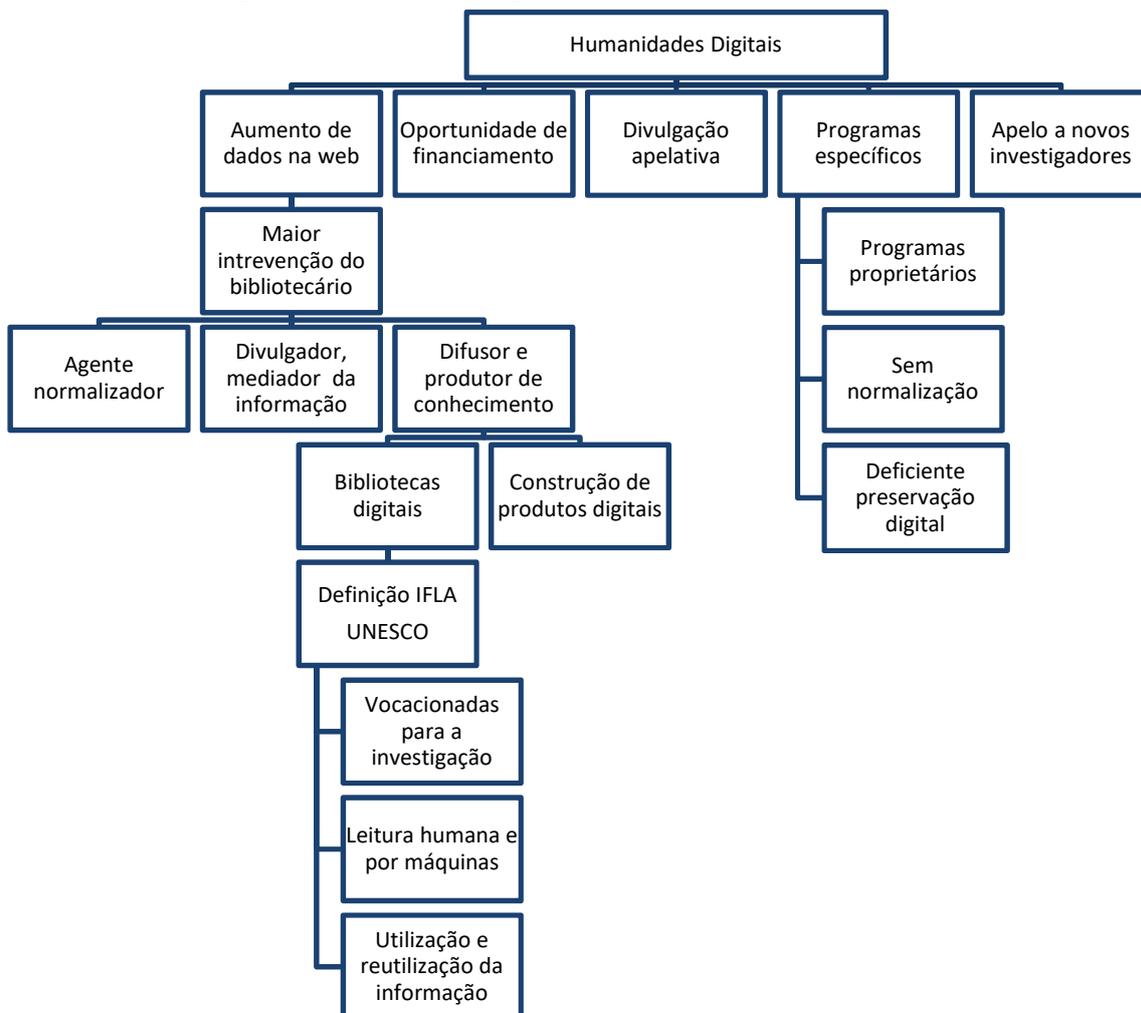
As Humanidades Digitais, de certa forma, condicionam e aumentam as exigências dos professores e investigadores em relação às bibliotecas e em especial às bibliotecas digitais. A dinamização das Humanidades Digitais teve como consequências:

- O aumento da informação disponibilizada em linha;
- A divulgação da produção científica de forma apelativa;
- A produção de programas específicos para Humanidades;
- A criação de novas oportunidades de financiamento;
- Ajuda à captação de novos investigadores.

O aumento da informação disponibilizada em linha leva a uma maior intervenção dos bibliotecários (Figura 8), sobretudo dos que trabalham em bibliotecas universitárias, como agentes normalizadores dos produtos desenvolvidos pelos docentes, durante a fase de produção e disponibilização. A biblioteca universitária, em particular, tende a evoluir para uma configuração próxima do centro de competências. “La biblioteca es

una parte fundamental de la estructura universitaria y puede proporcionar la estabilidad e institucionalidad que se requiere para la manutención de estos recursos académicos” (Galina Russell, 2011b, p. 7). Além da mediação, da organização e de todas as funções inerentes à gestão da biblioteca tradicional e digital, o bibliotecário desenvolve outras atividades e competências, por exemplo, na área da publicação digital, dando apoio na criação de wikis, blogues, etc., ou nas questões inerentes ao direito de autor e à criação de normas e/ou boas práticas para os projetos de Humanidades Digitais.

Figura 8 – Humanidades Digitais e bibliotecas. Esquema: DG, 2017.



Contribuindo com os seus conhecimentos e competências para a divulgação da produção científica, o bibliotecário pode ter a função de mediador cultural da produção da instituição a que pertence. Atualmente, a divulgação da produção científica em Humanidades, de forma apelativa e vocacionada para o ensino, é ainda pouco transparente (Figura 8): as fontes não são enunciadas, ou são-no de forma muito vaga; as opções tomadas na seleção dos dados não são justificadas; frequentemente, são utilizados programas proprietários construídos para determinado projeto, o que dificulta as ações de preservação digital.

No entanto, as Humanidades Digitais captaram novos investigadores, quer os que se interessam por Humanidades, quer os que gostam das novas tecnologias, quer os que apreciam o facto de poder conciliar as duas áreas do conhecimento. A biblioteca digital, tal como é definida pela IFLA/UNESCO (2011), é vocacionada para a investigação. Além dos dados de referência às fontes consultadas, existem ferramentas que permitem extrair dados dos textos dando-lhe novas leituras. Para tal, é necessário que as bibliotecas digitais disponibilizem os dados de forma a serem lidos pelo ser humano e através de algoritmos informáticos.

Surgiram, entretanto, bibliotecas digitais criadas especificamente para responder às novas necessidades dos investigadores em Humanidades, como a Bibliothèques Virtuelles Humanistes, a Perseus Digital Library e a Miguel de Cervantes.

A Perseus Digital Library é uma biblioteca digital hipertextual, em XML e com associação de etiquetas para entidades (pessoas, figuras míticas, personagens, etc.) e lugares. Crane e outros (2001) descrevem a criação da London Collection na Perseus, uma nova coleção formada por livros, imagens e mapas de Londres anteriores ao

século XX. O objetivo era testar os procedimentos anteriores, relativos à constituição da biblioteca dos clássicos, e construir uma biblioteca digital com base em documentos impressos, mas que não se limitasse a estes, organizada espacial e temporalmente. Os autores defendem que a biblioteca digital deve ter alguma dimensão a nível de conteúdos e que as ligações devem ocorrer dentro dos documentos da biblioteca digital, mas também em fontes externas. Assim, foram criadas etiquetas para as referências de espaço e tempo existentes nos textos, estabelecendo ligações internas, dentro da biblioteca digital, e com o Sistema de Informação Geográfica (SIG). No entanto, reconhecem que subsistem alguns problemas, nomeadamente, a falta de uniformização das listas de autoridades e a dificuldade em estabelecer a ligação a bases de referência externas. Apesar disso, a construção desta coleção permitiu aferir que os procedimentos referentes à coleção clássica eram válidos e escaláveis. Num estudo posterior, acerca do público, serviços e coleções da Perseus, Crane e Wulfman (2003, p. 77) registam que a maioria do público (84%) que consulta a biblioteca digital se concentra na coleção clássica. No que se refere ao sistema de gestão e recuperação da informação implementado, é considerado complexo, dado que utiliza múltiplas plataformas: o Fedora e o DSpace, para a gestão de repositórios; o Open Archive Initiative (OAI), para a interoperabilidade e partilha de informação; e o HyRex 1, para a pesquisa e recuperação da informação. A maioria das obras são etiquetadas com o TEI. “[...] cultural heritage digital libraries play a particularly important role for students of the humanities, because the digital library is the humanist’s laboratory, its resources the scholar’s primary data” (Crane & Wulfman, 2003, p. 84). Tendo contribuído para a valorização das coleções patrimoniais, são, agora, as novas combinações de públicos, serviços e coleções que colocam outros desafios conducentes a práticas mais eficazes.

A Biblioteca Virtual Miguel de Cervantes (Bia & Pedreño, 2001), criada em parceria entre a Universidade de Alicante e o Banco Santander Central Hispano, é constituída por obras relevantes para a cultura espanhola, agrupadas da seguinte forma: registos sonoros; monografias e periódicos digitalizados; teses; bibliotecas de autor; ligações a recursos digitais; serviços de comunicação e portais especializados. Assume, como principal objetivo, a promoção da literatura e da cultura espanhola, assumindo-se como uma contrapartida ao predomínio das bibliotecas digitais anglo-americanas. “The Miguel de Cervantes Digital Library hopes to act as inspiration to other non-English speaking cultures to create their own novel digital tools, which can be used by a world-wide multiracial and multilingual student and academic community of Internet users” (Bia & Pedreño, 2001, p. 163). Para além das inerentes funções de classificação e recuperação de documentos na íntegra, aponta para a recuperação do conteúdo de fragmentos de documentos, ou objetos textuais, como poemas, parágrafos, cabeçalhos, frases, ou palavras, e para o respetivo processamento, através de uma análise estatística que permite a geração de novas concordâncias literárias.

O projeto *Early English Books Online* (EEBO), desenvolvido pela ProQuest, não sendo uma biblioteca digital, é referido, aqui, por introduzir uma nova orientação no que respeita à digitalização do livro antigo e das bibliotecas digitais para os investigadores em Humanidades, no sentido em que altera significativamente a relação entre o leitor e a obra (Kichuk, 2007). Reconhecendo que, na maioria dos projetos, a pesquisa no documento e a sua disseminação tem vindo a ser melhorada, tal como o acesso ao seu conteúdo, Kichuk (2007) considera que os dados relacionados com a materialidade da obra, tais como a dimensão do fólio ou as margens, são negligenciados ou mesmo perdidos. Assim, define o conceito de “remediation”, como a aproximação possível

entre a obra física e a correspondente digital, implementada no EEBO. “Remediation acts like a distorting lens or opaque veil through which the scholar ‘sees’ the mediated Early English book” (Kichuk, 2007, p. 296). As obras foram digitalizadas a partir dos microfilmes. No processo de digitalização, são aplicados filtros para detetar eventuais imagens desfocadas, com margens desajustadas ou mesmo inclinadas.

A construção da biblioteca digital é um processo em evolução. Por conseguinte, qualquer que seja o sistema de gestão de dados adotado, este deve ser suficientemente flexível de forma a permitir a incorporação das sucessivas mudanças tecnológicas, quer sejam propostas pelos bibliotecários, ou pelos utilizadores, nomeadamente, os investigadores em Humanidades.

Schöneberg, Schmidt e Höhn (2013) descrevem o fluxo de trabalho de digitalização, metacodificação e disponibilização desenvolvido na biblioteca da Universidade de Würzburg, o qual pode ser considerado como paradigmático. Neste processo, estão incluídos módulos de análise automática das imagens, deteção de erros, prospeção de dados, análise semântica e qualidade, assim como o armazenamento, a publicação nas diversas plataformas e a preservação dos documentos a de longo prazo. O texto, navegável e pesquisável, é transcrito a partir das imagens. Em seguida, é realizada a extração de informação dos textos, como por exemplo as datas, locais e personagens envolvidas, e é criada uma tabela de conteúdos. Por fim, é feita a exportação, tanto para a preservação digital a longo prazo, como para a edição e publicação, usando os modelos TEI:P5 e METS/MODS, nas várias redes. Uma das principais dificuldades que as bibliotecas digitais enfrentam é a grande quantidade de informação, mas este problema é ultrapassado com o sistema distribuído de processadores. “A digitization

process does therefore not end with the production of a source's digital representation, it should cover the retrieval of semantics and metadata as well, which calls for a specialized workflow-system" (Schöneberg, Schmidt, & Höhn, 2013, p. 359).

Os projetos de bibliotecas tendem a dirigir a investigação no sentido de facilitar a disseminação do conhecimento e de o tornar globalmente acessível, cumprindo um conjunto de procedimentos que tem vindo a fixar-se: digitalização da obra ou da coleção; associação do registo bibliográfico às imagens; a correta e detalhada descrição dos objetos segundo os padrões existentes; a adequada metacodificação, para garantir uma efetiva pesquisa e recuperação da informação; o controlo de qualidade das imagens; a aplicação do reconhecimento ótico de caracteres às imagens dos documentos e obtenção de transcrições de qualidade, para que possam contribuir para o refinamento das pesquisas, tornando-as mais assertivas e permitindo a aplicação de algoritmos para a análise quantitativa das obras; etiquetagem dos conteúdos com o TEI e respetiva conversão em TEI-XML; criação de ficheiros HTML e de outros formatos de consulta; disponibilização no servidor para acesso em linha.

Algumas questões continuam em aberto: mesmo com imagens digitais de qualidade e apesar da evolução tecnológica, a aplicação dos programas de OCR continua a exigir a intervenção humana para validar as opções, mesmo que esta intervenção seja cada vez menor e os resultados cada vez melhores; a marcação dos textos facilita a leitura pelo computador, mas deixa de fora toda a materialidade da obra e é também um processo moroso; ainda que a codificação em TEI pareça a melhor solução para etiquetar os conteúdos, não responde às necessidades dos investigadores que analisam a mancha impressa, espaçamentos, tipos de letra, ou mesmo as ilustrações e

mapas, sendo necessário realizar múltiplas aproximações e interações com as coleções.

A nível europeu a ENUMERATE fornece as estatísticas sobre a digitalização, as políticas de preservação digital e as coleções digitais disponibilizadas na Europa. “The primary objective of ENUMERATE was to create a reliable baseline of statistical data about digitization, digital preservation and online access to cultural heritage in Europe” (McKenna, s.d.).

Todos os dados recolhidos encontram-se disponíveis<sup>47</sup> (Digibís, s.d.). Foram publicados três relatórios:

- ENUMERATE Core Survey 1, publicado em 2012 e cuja pesquisa foi realizada de dezembro de 2011 a abril de 2012: obtiveram-se 1951 respostas de instituições de património cultural de 29 países europeus nas áreas de arquivo, museu, biblioteca, incluindo também a digitalização de áudio e de vídeo (Stroeker, Vogels, & Panteia, 2012).
- O ENUMERATE Core Survey 2, publicado em 2014 e cuja pesquisa foi realizada de setembro a dezembro de 2013: obtiveram-se 1400 respostas de instituições de património cultural de 33 países europeus nas áreas de arquivo, museu, biblioteca, incluindo também a digitalização de áudio e de vídeo (Stroeker, Vogels, & Panteia, 2014).
- O NUMERATE Core Survey 3, publicado em 2015 e cuja pesquisa foi realizada de fevereiro a maio 2015: obtiveram-se 1000 respostas de instituições de

---

<sup>47</sup> De acordo com a Diretiva 2003/98/CE, entrou em vigor em 31 de dezembro de 2003. Foi revista pela Diretiva 2013/37/UE, que entrou em vigor em 17 de julho de 2013, relativa à reutilização de informações do sector público.

património cultural de 31 países europeus nas áreas de arquivo, museu, biblioteca, incluindo também a digitalização de áudio e de vídeo (Nauta, Heuvel, & DEN Foundation, 2015).

As instituições que respondem ao inquérito têm vindo a diminuir (Nauta et al., 2015), passando de 1950 para 1030, o que se deverá, por um lado, à extensão do inquérito (32 questões com um tempo estimado de resposta de 30min) e, por outro, à periodicidade com que está a ser implementado. No último relatório, a maioria das 1030 instituições que responderam foram museus (355), seguidas por bibliotecas (346), arquivos (218) e outras instituições (111) (Nauta et al., 2015, p. 9).

De acordo com os dados disponibilizados no relatório 2015 da ENUMERATE (Nauta et al., 2015), no universo das bibliotecas estão digitalizados 19% das coleções, faltando digitalizar 47%, sendo que cerca de 37% não é necessário digitalizar. Por tipologia de objetos, o texto (83%) e as imagens visuais (81%) são as tipologias mais representadas analogicamente. A razão mais apontada para a constituição de coleções digitais é o apoio à investigação e o ensino, seguida da preservação, mas só 68% das bibliotecas têm os metadados disponíveis em linha. A preservação digital é uma preocupação recorrente em todas as bibliotecas, e quase todas (90%) das instituições têm uma solução de preservação no tempo.

As BDLA devem promover a facilidade de utilização, que começa nos conteúdos, não esquecendo a comunidade que vai servir, os seus hábitos e práticas “whether digital library design begins with the content, collections, or the user community to be served, understanding the behavior, context, practices, expertise, and requirements of the prospective users is essential for improving usability” (Borgman, 2000, p. 129).

A comunidade de leitores é amplamente aumentada com a colocação em linha das coleções e está em constante mudança, tornando difícil a sua completa caracterização.

“The next generation of digital libraries must serve a large and diverse community and provide a large and diverse collection of information resources. Though we do not yet know how to build such a system, a starting point is to employ what is known about information-related behaviour in the systems of today” (Borgman, 2000, p. 149).

Espera-se, por conseguinte, que as bibliotecas digitais evoluam de forma a responder às necessidades de quem as procura, ou seja:

A digital library can be searched for any phrase; it can be accessed all over the world; and it can be copied without error. This is why digital libraries are coming. They address traditional problems of finding information, of delivering it to users, and of preserving it for the future. Digital information takes less space than paper information and thus may help libraries reduce costs. But, more important, they can provide a level of service never before attainable — delivery of information to the user’s desk, search capability by individual words and sentences, and information that does not decay with time, whether words, sounds, or images. (Lesk, 2005)

Num futuro, cada vez mais próximo, as bibliotecas digitais patrimoniais irão responder tanto aos seus utilizadores principais, os investigadores, como a qualquer leitor que por curiosidade ou lazer as procure.

## 4. Bibliotecas digitais: estudo crítico comparativo

### **Bibliotecas Digitais/obras em domínio público**

- Organizadas por bibliotecas nacionais ou patrimoniais, com obras fac-similadas e/ou com a respetiva versão em texto;
- Pensadas por humanistas, que dão primazia ao conteúdo e são específicas para os fins que foram criadas;
- Outras instituições, que pretendem divulgar a cultura e torná-la acessível a todos, disponibilizam as obras em vários formatos.

As bibliotecas digitais possuem diversas configurações, no sentido em que, quer a pesquisa, quer a lista dos resultados, divergem de biblioteca para biblioteca, mas, em contrapartida, apresentam também um conjunto de atributos comuns que as caracterizam como um todo.

As Bibliotecas Digitais de Livro Antigo (BDLA) analisadas neste estudo podem dividir-se em três categorias, em função dos instituidores, objetivos e públicos-alvo:

- Bibliotecas organizadas por bibliotecas nacionais ou patrimoniais, com grande tradição biblioteconómica. Disponibilizam obras fac-similadas e/ou com a respetiva versão em texto, em acesso livre e, por norma, dão primazia à quantidade de obras/páginas disponibilizadas em linha, cujas pesquisa e recuperação da informação se processa através dos elementos de catalogação e de gestores de conteúdos.
- Bibliotecas pensadas e implementadas por grupos de investigadores nas áreas das Humanidades. Dão primazia ao conteúdo e são específicas para os fins que foram criadas, mas, nalguns casos, perde-se a relação com o exemplar físico original, nomeadamente, a numeração das páginas. A pesquisa realiza-se

através dos elementos bibliográficos das obras e dos seus conteúdos e permitem a reutilização da informação através de aplicações informáticas. A sua manutenção, onde se inclui, não só a manutenção informática, mas também a gestão da coleção e a adição de novos itens, requer conhecimentos específicos e muito trabalho, pelo que se torna muito onerosa.

- Bibliotecas digitais realizadas por bibliotecas públicas ou outras instituições, públicas ou privadas, que pretendem divulgar a cultura e torná-la acessível a todos. Disponibilizam as obras em vários formatos, a fim de tornar a leitura possível nos vários dispositivos móveis existentes.

As bibliotecas digitais, cujos objetivos têm vindo a ampliar-se para lá da disponibilização de imagens em linha, fulcral nos primeiros projetos, tendem também a tornar-se mais flexíveis. Atualmente, exige-se que as bibliotecas digitais forneçam serviços personalizados e que integrem ferramentas úteis e adequadas para grupos específicos de utilizadores, como os investigadores e os professores.

Em Portugal, são várias as instituições que constituíram bibliotecas digitais a partir da digitalização de obras das suas coleções. Por norma, todas seguem o princípio legal de apenas disponibilizar em linha obras que se encontrem em domínio público, conforme o estabelecido pela Lei n.º 16/2008, de 1 de abril, cujo artigo 31.º determina que “o direito de autor caduca, na falta de disposição especial, 70 anos após a morte do criador intelectual, mesmo que a obra só tenha sido publicada ou divulgada postumamente”. Diferem, em contrapartida, no que respeita aos aspetos formais e processuais: algumas são bibliotecas organizadas, com os objetos digitais estruturados, o que implica o pós-tratamento da informação associada à imagem, nomeadamente, a

metacodificação, enquanto outras se formalizam como meras coleções de imagens digitais; numas, o acesso é realizado através de sistemas de gestão de informação, enquanto noutras, o acesso à imagem é feito através de páginas estáticas. No âmbito deste estudo, excluem-se da análise estas coleções de ficheiros digitais, habitualmente em formato PDF, com acesso através do respetivo registo bibliográfico, e cujo objetivo se centra na disponibilização da informação em linha, sem prestar atenção aos aspetos relacionados com a metacodificação, reutilização ou preservação digital.

#### 4.1. Biblioteca Nacional Digital

##### **BND 1999**

- Obras em domínio público;
- Página de novidades;
- Disponibiliza as obras em formato JPG, PDF e *flash*;
- Acesso: pesquisa e navegação;
- Pesquisa no conteúdo dos documentos, obra a obra.

Facilidade de utilização – média.

Utilidade – alta.

Performance – baixa.

“A Biblioteca Nacional Digital (BND) foi lançada, oficialmente, em fevereiro de 2002 como um projecto de dimensão aberta” (Campos, 2005, p. 23), com o objetivo de disponibilizar em linha os conteúdos mais procurados e de maior relevância na Biblioteca Nacional (BN), mas “integrava já alguns conteúdos anteriormente desenvolvidos (desde 1998)” (Campos, 2005). Efetivamente, acompanhando a evolução internacional, a BN havia criado, em 1998, o GEDE - Gabinete de Edições Eletrónicas, onde implementou a criação de sítios eletrónicos temáticos que constituíram os primeiros esboços da futura Biblioteca Nacional Digital (BND), a par de

atividades propedêuticas nos domínios da digitalização e da edição digital<sup>48</sup>. A designação de “biblioteca digital” (*Biblioteca digital de Almeida Garrett*) surgiu, pela primeira vez, em 1999, no sítio eletrónico institucional, para apresentar uma coleção de obras de Almeida Garrett, por ocasião das comemorações do bicentenário do nascimento do autor, com a reprodução, na íntegra, das suas primeiras edições, manuscritos e periódicos, com leitura assistida por um motor de busca e a possibilidade de impressão.

Ainda nesse ano, surgiu o separador da biblioteca digital (Figura 9), que integrava, além da biblioteca de Garrett, o sítio eletrónico *25 de Abril: da efemeridade à história*.



No ano seguinte, foram adicionados os sítios eletrónicos *António Feliciano de Castilho* e *Eça de Queirós*, com obras e espólios digitalizados. As obras eram disponibilizadas na

<sup>48</sup> A recuperação da informação aqui apresentada é feita através dos endereços eletrónicos <http://lisboa.biblioteca-nacional.pt> (1998), <http://www.biblioteca-nacional.pt> (1999-2002), <http://www.bnd.bn.pt> (2002-2006) e <http://www.purl.pt> (desde 2006) consultados nos bancos de dados *Wayback* do *arquivo.pt* e do *Internet Archive*.

íntegra, encadernação digital em formato HTML, onde a navegação se fazia através do sumário, por capítulos ou partes das obras. A coleção era organizada por autor.

Figura 10 – BND: página de acolhimento, 2002. Fonte: arquivo.pt 2002/03/24



Em 2002, a BND (Figura 10), com endereço eletrónico próprio (<http://bnd.bn.pt>), alcançou um patamar de maturidade e consolidação das experiências anteriores, definindo um sistema integrado de informação, ao mesmo tempo que eram estabelecidas normas de procedimentos para a nomeação de ficheiros e dos requisitos técnicos das imagens e encadernação eletrónica. O núcleo da BND era, então, o sítio eletrónico *Memória* (da Ciência, da Educação, da Língua, da Literatura, do Livro, da Imprensa, da Música, da História, da Arte, do Espaço, do Pensamento), cujas coordenações setoriais foram confiadas a investigadores especializados, com a incumbência de selecionar um conjunto de obras relevantes nos vários domínios do conhecimento e do pensamento, produzidas por autores portugueses. Segundo Fernanda Campos, na altura, subdiretora da BN, a seleção das obras para o núcleo inicial da BND, obedecia ainda a critérios de direitos de autor e estado de conservação dos originais:

Esse critério de relevância é sempre cruzado com os direitos de autor e o estado de conservação da espécie original, por forma a que sejam apenas divulgadas na Web obras que estão em domínio público (ou para as quais exista autorização expressa) e espécies que não possam ser danificadas no processo de digitalização. (Campos, 2005, pp. 24–25)

Em 2004, o sítio da BND foi remodelado, mantendo a denominação *Memória* (atualmente disponível em <http://purl.pt/401/1/>). Ao mesmo tempo que se procedia à automatização de algumas rotinas e ao desenvolvimento e teste dos programas informáticos, em código aberto, de suporte à edição digital (ContentE<sup>49</sup>, SECO<sup>50</sup>, KIWI<sup>51</sup> e PAPAIA<sup>52</sup>) e ao armazenamento das obras disponibilizadas em linha (ARCO<sup>53</sup> e LUSTRE<sup>54</sup>) (Vd. <http://bnd.bn.pt/tec.html>).

Esta reformulação tecnológica conduziu à criação de um novo endereço (<http://purl.pt>), identificador permanente das obras colocadas em linha, inicialmente, de uso interno e consistindo apenas numa lista sequencial de obras digitalizadas.

Em 2006, a infraestrutura e os serviços da BND foram integrados na orgânica da BN. O esquema METS foi adotado em toda a BND, implicando o reprocessamento das obras já disponibilizadas em linha, através do processador automático SECO.

O desenho da BND foi, novamente, reestruturado em 2013, incluindo algumas alterações substanciais: aspeto de visualização, incidindo nos fundos e no esquema de

---

<sup>49</sup> ContentE, editor de conteúdos, efetua a encadernação digital das obras e associa as imagens a metadados (Pedrosa & Borbinha, 2010).

<sup>50</sup> SECO (SERial CONverter), para o processamento automático das imagens matrizes (TIF).

<sup>51</sup> KIWI, indexador de palavras relevantes nas imagens de forma manual.

<sup>52</sup> PAPAIA, para o processamento de páginas digitalizadas.

<sup>53</sup> ARCO, armazenamento e preservação digital.

<sup>54</sup> LUSTRE, armazenamento distribuído, escalável e de alto desempenho.

cores; eliminação das cópias digitais em JPG e dos ficheiros múltiplos em PDF; eliminação dos sumários de acesso à maioria das obras; introdução de cópias em Flash Player. A última atualização, realizada entre 9 e 13 de maio de 2017, repôs a versão em JPG para a generalidade das monografias.

Atualmente, o modelo de dados da BND está a ser consolidado com a adoção dos modelos Premis, para a preservação digital, e ALTO (Analyzed Layout and Text Object)<sup>55</sup>, para as monografias, e com a implementação das licenças CC0<sup>56</sup>, para os metadados descritivos, e *public domain*, para as obras digitalizadas (Patrício, 2015).

Figura 11 – BND: página de acolhimento. Fonte: purl.pt (31 jul. 2017)

The screenshot shows the BND homepage with the following elements:

- Header:** BNP BIBLIOTECA NACIONAL DE PORTUGAL logo on the left; 'biblioteca nacional digital' logo on the right.
- Navigation Bar:** Início | Sítios temáticos | Portuguese Culture | Dic. Historiadores | A Grande Guerra | Cooperação | Sobre a BND
- Left Sidebar:**
  - Coleções BND**
  - Navegar por:
    - Autor
    - Título
    - Data de publicação
    - Todas as obras
  - Pesquisar nas coleções BND
  - RNOD logo
  - Logotipo da rede europeana
- Main Content Area:**
  - Últimas obras disponibilizadas:** (with social media icons)
  - Four document entries, each with a thumbnail and a 'Consultar' link:
    - IRMANDADE DO SANTISSIMO SACRAMENTO (Torrão) Compromisso da Hirmandade do Sanctissimo Sacramento sita na Igreja de Nossa Senhora da Asumpção desta villa do Torrão reformado no anno de mil e seiscentos e tres anos 1603. - [2], 11 f. : pergaminho, il. color. ; 295x200 mm
    - Brazão da Fidalguia e Geracam de Domingos Dantas da Cunha MDCLXXX 7 de Setembro de 1680. - [16] f. (25 l.) : pergaminho, il. color. ; 271x200 mm
    - [Coleção hagiológica] [Entre 1426 e 1475]. - [171] f. (26 l.) : perg., il. color. ; 27 cm
    - [Bíblia]. - [Mosteiro de Santa Maria de Alcobaça (?)]. [1250-1260]. - [4], [1211], [2] p. (2-4 colunas, 46-47 linhas) : velino, il. color. ; 135x110 mm
- Footer:** REPÚBLICA PORTUGUESA logo; navigation links: Sítios temáticos | Portuguese Culture | Dic. Historiadores | A Grande Guerra | Sobre a BND | Contactos e ficha técnica | Mapa; date and time: 2017-07-28T17:25:34

<sup>55</sup> ALTO, combina a descrição em METS com o OCR, permitindo obter obras em vários formatos.

<sup>56</sup> CCO – Licença *creative commons*, para obras sem direito de autor, nem direitos conexos (vd. [https://creativecommons.org/publicdomain/zero/1.0/deed.pt\\_PT](https://creativecommons.org/publicdomain/zero/1.0/deed.pt_PT)).

Na página de acolhimento da BND (Figura 11), surge, em lugar de relevo, o elenco das últimas obras colocadas em linha, apresentando quatro ou cinco referências, dependendo do formato da obra, em retrato ou paisagem.

Na parte superior da página, insere-se o cabeçalho com a barra de acessos, que se mantém constante durante a navegação no sítio eletrónico:

- Sítio institucional da Biblioteca Nacional de Portugal (BNP)<sup>57</sup>;
- “Início”;
- “Sítios temáticos”: obras digitais elaboradas com base nas obras da BNP.

Desdobra-se em:

- Autores e personalidades,
- Exposições e comemorações,
- Outros temas e projetos;
- “Portuguese Culture”: projeto especial, com o patrocínio da Fundação Luso-Americana, que integra obras do espólio da BNP em língua inglesa, disponibilizadas em linha na íntegra;
- “Dici[onário] de historiadores”: ligação ao projeto colaborativo da BNP e do Centro de História da Universidade de Lisboa, “Dicionário de historiadores portugueses: Da Academia Real das Ciências ao final do Estado Novo”, com o elenco de historiadores portugueses ou estrangeiros que escreveram sobre Portugal, com as respetivas biografias;

---

<sup>57</sup> Designação da Biblioteca Nacional portuguesa a partir de 2006 (vd. Diário da República n.º 208/2006, Série I de 2006-10-27 Lei Orgânica do Ministério da Cultura).

- “A Grande Guerra”: sítio temático que assinala o centenário da 1.<sup>a</sup> Grande Guerra, disponibilizando toda a documentação da época relacionada com a participação portuguesa no conflito;
- “Cooperação”: projetos colaborativos e coleções onde a Biblioteca Nacional Digital está representada;
- “Sobre a BND”: informação técnica, cuja última atualização data de 2012 (cfr. consulta efetuada a 31 jul. 2017)

A BND tem os seus conteúdos acessíveis através de outros portais: Europeana; The European Library (TEL); World Digital Library (WDL); Biblioteca Digital do Património Ibero-Americano (BDPI); Repositório Nacional de Objectos Digitais (RNOD); Biblioteca Digital Luso Brasileira (BDLB); Manuscriptorium – European Digital Library of Written Cultural Heritage.

Na BND, o acesso às obras pode fazer-se através da pesquisa no catálogo em linha da BNP (OPAC), pelos elementos constantes da ficha bibliográfica, sendo possível delimitar a pesquisa às obras digitalizadas pela seleção do separador “BND”; na pesquisa avançada, a seleção pode ser feita através dos elementos bibliográficos, restringindo a pesquisa por coleção através das expressões “Documentos Electr[ónicos] Monogr[afia]” e “Documentos Electr[ónicos] Série” ou limitando a pesquisa à BND. No sítio eletrónico da BND, a navegação pode ser efetuada através dos índices de título, autor ou data de publicação, com a reserva de que, na coleção dos periódicos, só é possível navegar pelo título da obra (Tabela 3). Também é possível localizar as obras, através de pesquisa nos motores de busca, como por exemplo no Google, combinando o termo de pesquisa com a expressão “purl.pt”.

Tabela 3 – BND: acesso às coleções

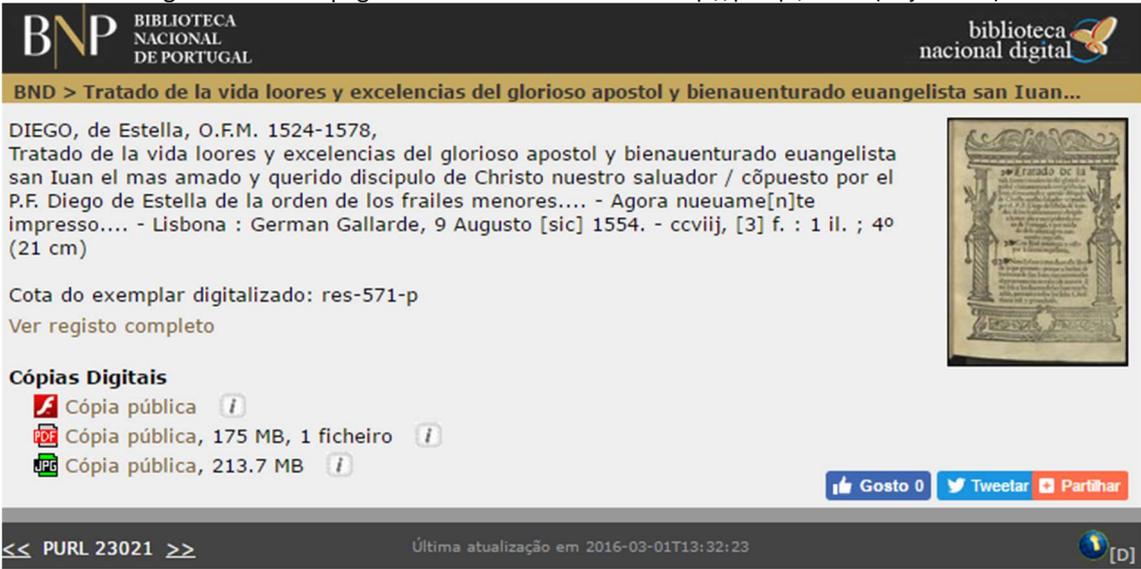
Página BND		OPAC / BND
<b>Coleções</b>	Navegação por:	Pesquisa em:
Livros	Autor Título Data de publicação Todos os livros	Fundo Geral Reservados
Periódicos	Título  Projetos especiais: – Seminário Livre de História das Ideias – <i>Portugaliae Mathematica</i>	—
Iconografia	Autor Título Data de publicação Todas as obras iconográficas	Iconografia
Cartografia	Autor Título Data de publicação Todas as obras cartográficas	Cartografia
Partituras	Autor Título Data de publicação Todas as partituras	Música

No catálogo da BNP, a recuperação da pesquisa nem sempre devolve documentos digitalizados e, em contrapartida, apresenta, em acréscimo, os resultados obtidos na Europeiaana, na Biblioteca Digital do Património Iberoamericano (BDPI), no Google Books, no Google Scholar, no Internet Archive, entre outros.

A denominação das coleções é, em regra, coincidente nos dois acessos. No entanto, a coleção que, no OPAC, é designada por “Música”, surge no sítio da BND na página de acolhimento como “Partituras”, o que, aliás, é redutor face ao conteúdo, dado que,

além de partituras, são disponibilizadas obras sobre a criação e a educação musical em Portugal (vd., por exemplo: *Principios de musica ou exposição methodica das doutrinas da sua composição e execução*, de Rodrigo Ferreira da Costa<sup>58</sup>; o libreto da ópera *Laurianna: opera em 4 actos e 6 quadros*, de Augusto Machado<sup>59</sup>; ou *Arte mínima ...*, de Manuel Nunes da Silva<sup>60</sup>).

Figura 12 – BND: página de acesso à obra. Fonte: <http://purl.pt/23021> (31 jul. 2017)



BNP BIBLIOTECA NACIONAL DE PORTUGAL biblioteca nacional digital

BND > Tratado de la vida loores y excelencias del glorioso apostol y bienaventurado euangelista san Iuan...

DIEGO, de Estella, O.F.M. 1524-1578,  
Tratado de la vida loores y excelencias del glorioso apostol y bienaventurado euangelista san Iuan el mas amado y querido discipulo de Christo nuestro saluador / cõpuesto por el P.F. Diego de Estella de la orden de los frailes menores.... - Agora nueuame[n]te impresso.... - Lisboa : German Gallarde, 9 Augusto [sic] 1554. - ccviiij, [3] f. : 1 il. ; 4º (21 cm)

Cota do exemplar digitalizado: res-571-p  
Ver registo completo

**Cópias Digitais**

- Cópia pública
- Cópia pública, 175 MB, 1 ficheiro
- Cópia pública, 213.7 MB

Gosto 0 Tweetar Partilhar

<< PURL 23021 >> Última atualização em 2016-03-01T13:32:23 [D]

As monografias (Figura 12) são disponibilizadas em formato PDF, JPG e numa aplicação em *flash*, em cuja animação se simula o folhear do livro, mas que apenas permite a leitura em linha. Na maioria dos casos, o PDF disponibilizado não inclui nem a transcrição do texto, nem marcadores. Nas versões em *flash* e em JPG, algumas obras apresentam o sumário com a estrutura do respetivo conteúdo, fornecendo um ponto de acesso direto a uma parte, capítulo ou subcapítulo da obra (vd., por exemplo, *Orthographia da lingoa portuguesa*, de Duarte Nunes de Leão, Figura 13). Contudo, a janela reservada ao sumário da obra é pequena, o que dificulta a navegação. Por se

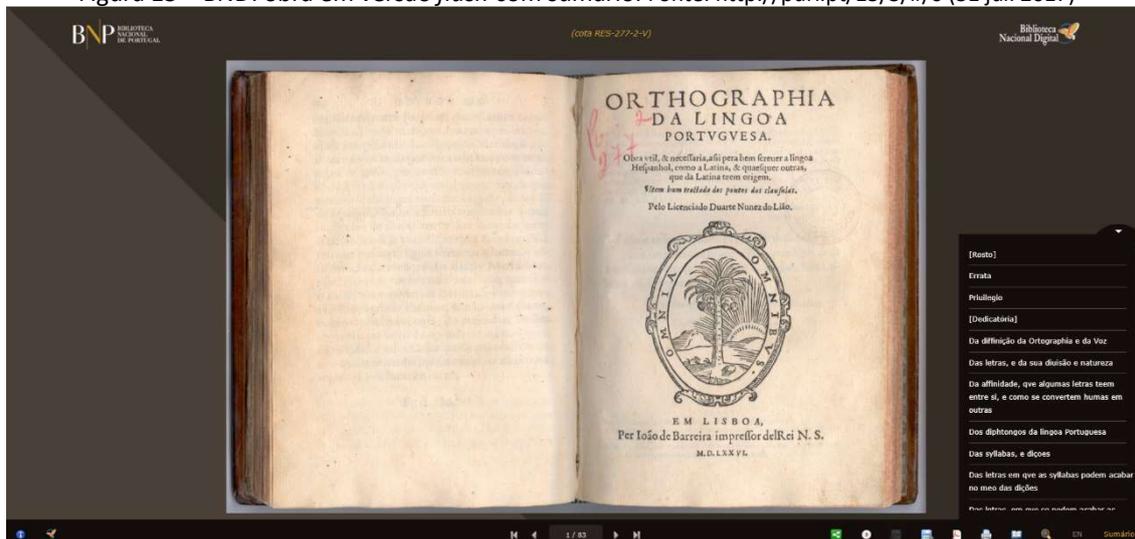
<sup>58</sup> Disponível em <http://purl.pt/169>

<sup>59</sup> Disponível em <http://purl.pt/92>

<sup>60</sup> Disponível em <http://purl.pt/170>

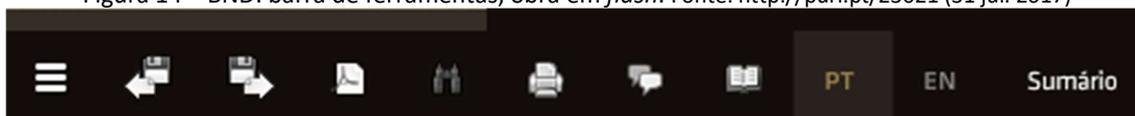
encontrarem em domínio público, não há qualquer constrangimento de acesso às obras de livro antigo disponíveis em linha.

Figura 13 – BND: obra em versão *flash* com sumário. Fonte: <http://purl.pt/15/3/#/0> (31 jul. 2017)



No sítio eletrónico da BND e no OPAC, a informação apenas se encontra disponível em língua portuguesa. Na versão *flash*, ao aceder ao conteúdo da obra, é possível selecionar a língua (português ou inglês) do texto das legendas nas janelas *pop-up* ao passar o cursor sobre os símbolos dos menus ou hiperligações. A sinalética escolhida, apesar do esclarecimento textual destas janelas, não é de apreensão imediata, quer pela pequena dimensão dos sinais, quer por não utilizar os mais comuns e convencionais, como, por exemplo, a utilização de binóculos para a pesquisa no conteúdo das obras.

Figura 14 – BND: barra de ferramentas, obra em *flash*. Fonte: <http://purl.pt/23021> (31 jul. 2017)



Na versão em *flash* (Figura 14), a barra de ferramentas permite aceder às seguintes funções: partilhar, *autoplay*, *download* página esquerda e direita, *download* PDF, pesquisar, imprimir, enviar por mail, visualizar Miniaturas (*thumbnails*), *zoom*,

selecionar versão em português (PT) ou inglês (EN) e sumário. Ao expandir os botões, para aceder às funções de *zoom* e rotação da imagem, partilha nas redes sociais e *autoplay*, estes ficam por cima das setas de navegação.

Figura 15 – BND: barra de ferramentas, obra em JPG. Fonte: <http://purl.pt/23021> (31 jul. 2017)



A versão em JPG (Figura 15) apresenta o mesmo conjunto de funcionalidades, à exceção da pesquisa e do envio por mail. A sinalética nas barras de ferramentas difere nas duas versões. Ao efetuar a consulta em equipamentos com o iOS, o utilizador é reencaminhado, de forma automática e sem aviso, para a versão em JPG.

Ambas as versões permitem visualizar a totalidade da obra em miniaturas (*thumbnail*) e a função *zoom*, apesar de esta se localizar em pontos diferentes na barra de ferramentas. A função de “*autoplay*”, uma passagem sequencial das páginas com o ritmo pré-definido (muda de página a cada 3s), não tem utilidade prática evidente. A versão em *flash* permite a visualização em simultâneo de duas páginas da obra (verso e rosto), reproduzindo o livro aberto, enquanto a passagem das páginas simula o folhear manual.

A navegação nas obras que apresentam sumário é facilitada em relação às restantes, a maioria das quais não foram convertidas em ficheiros pesquisáveis nem fornecem quaisquer outros pontos de acesso ou chaves de leitura.

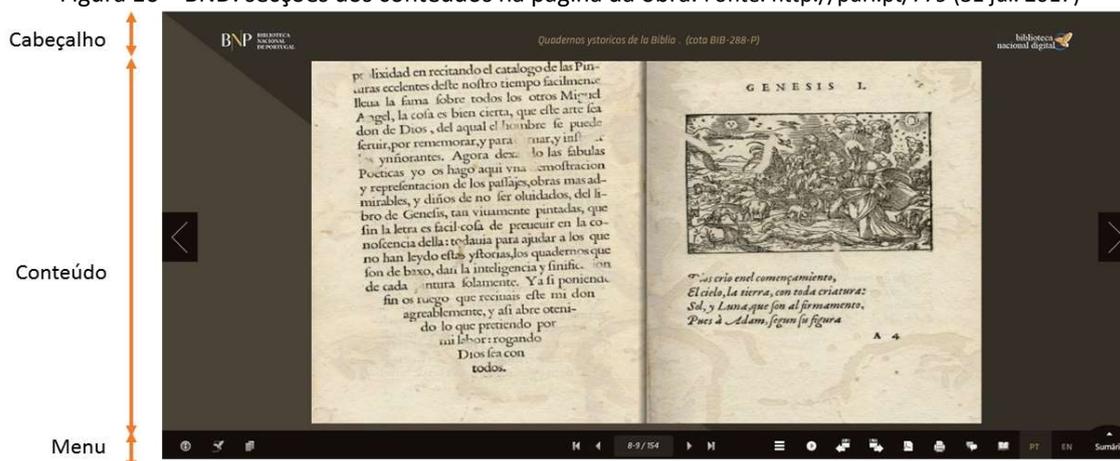
A disponibilização nas redes sociais é possível nas duas versões, na JPG está visível e na versão em *flash* só está disponível quando se expandem os botões. No sítio da BND é possível a partilha direta nas redes sociais, nomeadamente, no Facebook, Twitter e Google+, ou através do correio eletrónico. No entanto, só permite a partilha da

página de acolhimento da obra e, nas versões em JPG ou *flash*, é possível a partilha de uma página específica. Na versão em JPG, não é possível pesquisar no interior do documento, nem enviar por mail.

Só é possível guardar a informação bibliográfica da obra em formato adequado aos gestores de referências e apenas no formato RIS (formato etiquetado e normalizado desenvolvido pela Research Information Systems para citações bibliográficas), ao aceder ao registo completo da obra no catálogo bibliográfico da BNP.

Nas versões em JPG e *flash*, as obras são disponibilizadas numa página estruturada em três secções (Figura 16): cabeçalho, com os ícones de ligação à BNP e à BND, respetivamente, nos cantos superiores esquerdo e direito, e o título da obra em consulta e a respetiva cota, ao centro; conteúdo, onde são exibidas as imagens das obras; e o menu, na barra inferior, onde se localizam as várias opções de navegação e ferramentas.

Figura 16 – BND: secções dos conteúdos na página da obra. Fonte: <http://purl.pt/779> (31 jul. 2017)



O texto, em regra, é legível, embora possa exigir o recurso à ferramenta *zoom*. O campo central, onde é apresentada a obra, é geralmente suficiente para a apresentação de obras com formatos inferiores a *in-octavo* (in-8<sup>o</sup>), mas as páginas das

obras de maior dimensão que, por defeito, também são exibidas na íntegra, não permitem leitura, dado que os caracteres demasiado pequenos se tornam ilegíveis. Neste caso, a leitura implica efetuar *zoom* que, na versão *flash* se mantém constante ao longo da pesquisa numa obra, mas que, na versão JPG é necessário ser feito página a página.

A apresentação do menu, na barra horizontal colocada no registo inferior, também não favorece uma utilização intuitiva, dado que não segue o padrão de leitura ocidental, da esquerda para a direita e de cima para baixo. A localização do sumário abre numa página *pop-up*, no canto inferior direito, e pode sobrepor-se à imagem da página, tapando o texto e inibindo a respetiva leitura. Ao selecionar determinado item no sumário, as respetivas páginas são exibidas, mas não há informação acerca do sítio da obra onde se encontra, como o respetivo item do sumário, ou número de página.

As imagens descarregadas têm denominações diferentes consoante o *download* se faz a partir da versão em JPG, que mantém a cota da obra e a indicação relativa ao número sequencial da digitalização, número de página, formato, resolução e profundidade de cor da imagem, ou da versão *flash*, onde esta informação é substituída pelo respetivo número de PURL e o número sequencial da digitalização.

Assinalam-se outros aspetos formais que comprometem a autenticidade da versão digital da obra. Nalgumas obras, o cromatismo da imagem não é fiável, registando-se desvios mais ou menos acentuados face aos originais, nomeadamente, nos exemplares em que foi inserido um fundo artificial que limita, sobretudo, o texto com rubricas ou

com impressão a vermelho (vd. *Libre apellat Consolat de mar*<sup>61</sup>). Por outro lado, as irregularidades decorrentes do processo de digitalização, quer devido à configuração do equipamento, quer às condições de iluminação, também alteram a qualidade da imagem e a sua correspondência com o original, determinando tonalidades azuladas (vd. *De patronatibus ecclesiarum regiae coronae Regni Lusitaniae*, de Jorge de Cabedo<sup>62</sup>), amareladas, ou demasiado vibrantes, sem correspondência com o original. As obras aparadas digitalmente podem truncar o texto e dificultar a leitura e não permitem validar a integridade face ao original (vd. *Breuiarium bracare[n]se*<sup>63</sup>).

Do ponto de vista do utilizador-investigador, assinala-se a inexistência de mecanismos que permitam a criação de uma área pessoal ou a construção de uma biblioteca digital personalizada. Além disso, o facto de a lista “Últimas obras adicionadas”<sup>64</sup> não indicar a data de atualização, além de que as últimas obras adicionadas não coincidem com o último PURL disponível, dificulta a perceção do fluxo de atualizações por parte do utilizador. Embora a barra de ferramentas da página de acolhimento do sítio eletrónico da BND integre o sinal RSS (Really Simple Syndication), com o texto “Follow on RSS” (em inglês), esta tecnologia não está ativa e, por conseguinte, não emite os alertas para a adição de novos conteúdos, permitindo que os utilizadores fossem informados acerca das atualizações efetuadas.

A BND não disponibiliza qualquer tutorial ou página de ajuda, que oriente os leitores na pesquisa, elucide acerca do significado da sinalética utilizada e respetiva funcionalidade e indique como aceder à informação menos visível ou como efetuar

---

<sup>61</sup> Disponível em <http://purl.pt/22620>

<sup>62</sup> Disponível em <http://purl.pt/14133>

<sup>63</sup> Disponível em <http://purl.pt/14279>

<sup>64</sup> Disponível em <http://purl.pt/index/geral/PT/index.html>

alguns procedimentos mais complexos. Depois de localizada a obra pretendida, a sua leitura ou a cópia para o computador é uma tarefa simples e intuitiva. Contudo, alguns procedimentos afiguram-se particularmente intrincados como, por exemplo, a exportação do registo para os gestores de referência, feita a partir do registo bibliográfico no OPAC e implicando três operações: abrir a informação bibliográfica, seleccionar o registo completo e guardar o ficheiro em formato RIS (Research Information Systems).

A BND não disponibiliza dados estatísticos globais. Para calcular o número de títulos disponibilizados, optou-se por realizar pesquisas por títulos em cada uma das coleções (livros, periódicos, iconografia, cartografia, música) e somar os subtotais obtidos por séculos, o que fornece um total de 27.219 títulos (Tabela 4).

Tabela 4 – BND: número de títulos digitalizados por coleção

<b>Coleções</b>	<b>Número de títulos disponíveis</b>
Livros	7.666
Periódicos	680
Iconografia	16.221
Cartografia	2.082
Música	570
<b>Total</b>	<b>27.219</b>

Uma das fragilidades desta biblioteca digital reside na pesquisa da informação. Ao fazer a pesquisa no OPAC e no separador da BND, os resultados devolvidos, por norma, incluem obras que não correspondem à pesquisa. Ainda que a navegação seja amigável e permita uma visão de conjunto de toda a biblioteca digital, num universo

de mais de trinta mil PURLs<sup>65</sup> torna-se pouco prática. Apesar de haver obras disponibilizadas em texto integral, só é possível pesquisar o conteúdo das obras uma a uma, dado que não existe um mecanismo que percorra a BND com uma pesquisa por palavra nos respetivos conteúdos.

Outros aspetos a melhorar seriam a informação relativa à periodicidade de atualização da página de entrada e a introdução de alertas personalizáveis, permitindo que o utilizador receba informação relativa à disponibilização de novos conteúdos ou à alteração dos existentes e que defina o tipo de obras acerca do qual pretende ser avisado. Além disso, seria pertinente uma maior atualização dos dados estatísticos relativos ao desempenho da BND (número de obras disponibilizadas em linha, número de leitores e respetiva caracterização genérica, etc.), dado que os últimos dados disponíveis são referentes a 2012.

---

<sup>65</sup> Informação atualizada a 31 de julho de 2017: 31.021 PURLs. A discrepância entre este número e o número total de títulos disponibilizados, justifica-se pelo facto de, nestes últimos, não se incluir as obras disponibilizadas nos espólios e nos sítios eletrónicos temáticos.

## 4.2. Breve descrição de outras bibliotecas digitais de referência

### 4.2.1. Alma Mater

#### Alma Mater c. 2006

- Obras em domínio público;
- Disponibiliza as obras em formato em PDF e no *e-reader*;
- Acesso: pesquisa facetada e navegação.

Facilidade de utilização – média.

Utilidade – alta.

Performance – baixa.

A Alma Mater<sup>66</sup> é a biblioteca digital do fundo antigo da Universidade de Coimbra, constituído, maioritariamente por obras anteriores a 1940, os trabalhos preparatórios para a sua construção foram iniciados em 2006. Em 2014, possuía “[...] 5595 obras digitalizadas, que correspondem a 1.250.577 imagens” (Miguéis & Fiolhais, 2014, p. 241). Nem o projeto, nem a Universidade disponibilizam dados estatísticos atualizados.

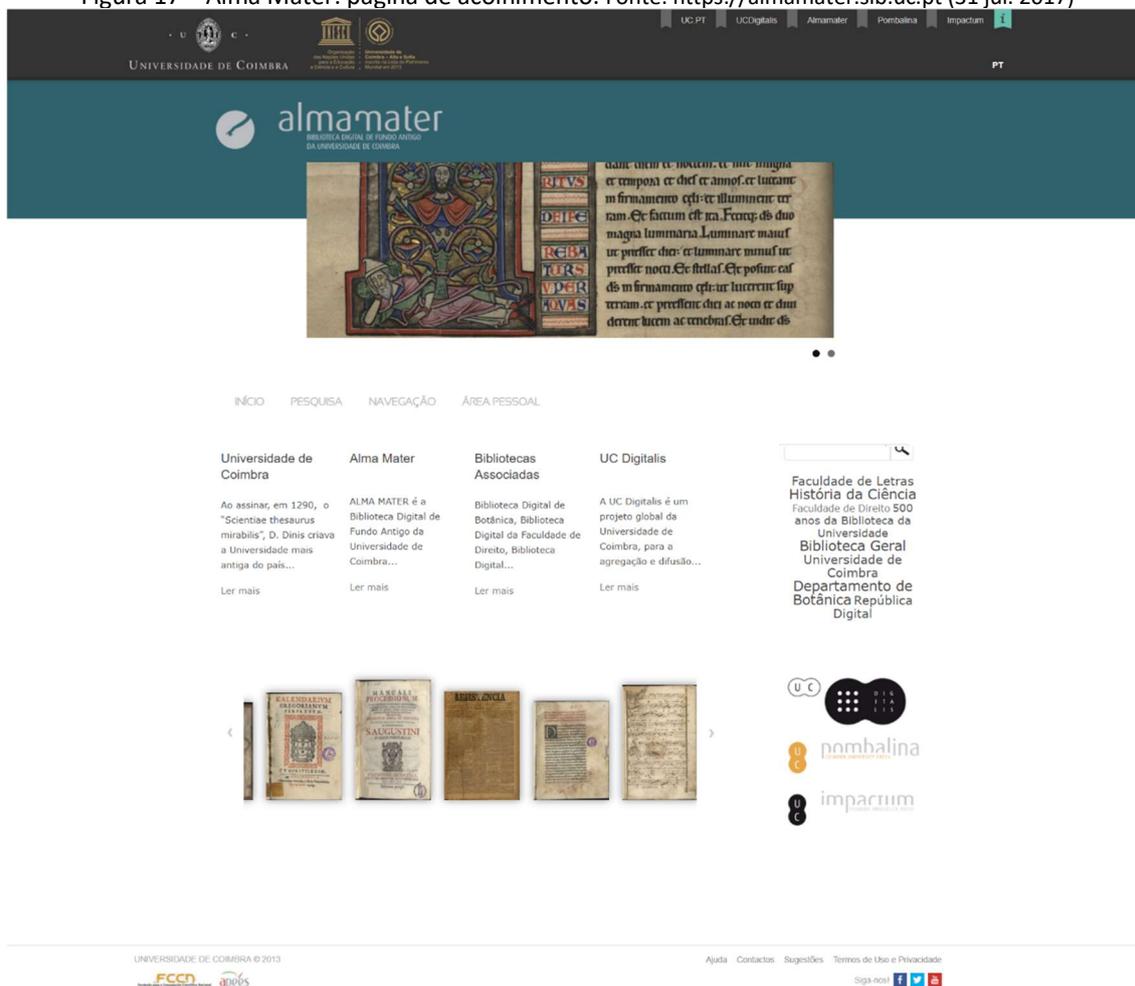
Na página de acolhimento (Figura 17), o cabeçalho, que se mantém em todas as páginas do sítio eletrónico, integra duas barras de navegação: no topo, a ligação ao sítio eletrónico da Universidade de Coimbra, à página da UNESCO relativa à candidatura da universidade a património mundial, e aos projetos UCDigitalis, Almamater, Pombalina e Impactum; a barra relativa ao projeto, com o respetivo logótipo e o acesso à página de acolhimento (Início), às pesquisas simples e avançada (Pesquisa), navegação por autor, periódicos do fundo antigo, título, tipo ou data (Navegação) e ao espaço personalizável pelo utilizador (Área pessoal). Há ainda a

---

<sup>66</sup> Disponível em <https://almamater.sib.uc.pt/>

possibilidade de escolha da língua do texto do sítio: português (PT), inglês (EN), espanhol (ES) e chinês (ZH).

Figura 17 – Alma Mater: página de acolhimento. Fonte: <https://almamater.sib.uc.pt> (31 jul. 2017)



O campo central da página integra os títulos e o início dos textos de apresentação da “Universidade de Coimbra”, “Alma Mater”, “Bibliotecas associadas” e “UC Digitalis”, com ligação aos textos integrais, sob o qual corre um registo dinâmico de destaques com a apresentação de imagens de obras, geralmente da página de título, com acesso ao registo bibliográfico, a partir do qual se faz a ligação ao respetivo conteúdo digitalizado. Ainda nesta página, apresenta-se um carrocel com as imagens de capa, em miniatura (*thumbnail*), de nove obras, sem que seja perceptível o critério seguido para a construção deste destaque.

O acesso às obras realiza-se por pesquisa, ou por navegação no sítio eletrónico, ou por pesquisa no catálogo SIBUC (Serviço Integrado das Bibliotecas da Universidade de Coimbra), ou por pesquisa através dos motores de busca, combinando os termos a pesquisar com a expressão “almamater.sib.uc.pt”.

A pesquisa simples permite efetuar a busca nos elementos bibliográficos das obras, enquanto que, na pesquisa avançada, existe a possibilidade de pesquisa no conteúdo dos documentos, embora esta funcionalidade não seja evidente para o utilizador. A pesquisa simples inclui a pesquisa facetada<sup>67</sup>, ou analítico-sintética, que, além dos assuntos referentes ao conteúdo intelectual da obra inclui outras características ou denominadores comuns no conjunto das obras. Esta pesquisa permite a utilização de termos como, por exemplo, “livro antigo”, conduzindo mais assertivamente à obra ou conjunto de obras pretendidas.

A lista dos resultados da pesquisa pode ser ordenada por relevância, título, tipo, autor, data. Pode, também, ser refinada através da seleção do tipo de documento, origem, coleção digital, autor, área do conhecimento e idioma.

Na primeira fase desta biblioteca digital, as obras eram disponibilizadas nos formatos PDF e JPEG, incluindo, nesta versão, um sumário estruturado do conteúdo da obra. Atualmente, porém, as obras são disponibilizadas apenas em PDF em formato imagem, sem a aplicação de programas de OCR às imagens, ou seja, sem a transcrição do texto dos documentos, e não incluem sumário ou qualquer outro tipo de marcador. Para a leitura em linha, existe um aplicativo *ebook reader*, que se mostra pouco eficaz na

---

<sup>67</sup> A Teoria da Classificação Facetada foi desenvolvida por Shiyali Ramamrita Ranganathan na década de 1930.

consulta, quer pela lentidão no processamento, sobretudo, na visualização de obras de maiores dimensões, quer por não acrescentar outras funcionalidades, como a pesquisa no conteúdo do documento.

Existe a hipótese de registar o utilizador, pressupondo a personalização do perfil e a construção de listas de interesse, mas funciona de forma irregular (nos testes realizados, a percentagem de sucesso na ativação do registo é de cerca de 10%).

É possível colocar, de forma direta, as referências das obras quer em gestores de referências, quer nas redes sociais, mas sem incluir ficheiros de imagem da página de rosto.

Na Alma Mater, a navegação é, em regra, amigável e intuitiva e a pesquisa revela-se eficaz. Em contrapartida, a página de ajuda revela-se muito incipiente e não é disponibilizada a previsualização da lista de termos para a pesquisa facetada. Não existe informação sobre as estatísticas de utilização, ou sobre os programas e procedimentos utilizados na construção da biblioteca digital.

#### **4.2.2. Gallica**

##### **Gallica 1997**

- Obras em domínio público;
- Página de novidades;
- Alertas;
- Disponibiliza as obras em formato TIF, JPG, PDF e ePub;
- Acesso: pesquisa e navegação;
- Pesquisa no conteúdo dos documentos.

Facilidade de utilização – alta.

Utilidade – alta.

Performance – alta.

A Gallica<sup>68</sup>, biblioteca digital francesa, foi criada em 1997, constituída inicialmente pelas obras digitais e digitalizadas a partir das coleções da Biblioteca nacional de França (BnF). “Gallica se define como uma “bibliothèque encyclopédique et raisonnée”, que nació con la finalidad de convertirse en el portal de acceso a las colecciones digitales francesas” (Lucía Megías, 2010, p. 380). Em 2008, a BnF lançou o desafio da digitalização a outras bibliotecas, cujos acervos específicos permitissem completar a oferta da Gallica, aumentando consideravelmente o volume de obras disponibilizadas em linha. Passou, por conseguinte, a definir-se como uma biblioteca digital coletiva e federada que, atualmente, fornece o acesso a recursos digitais de 346 parceiros. Neste universo, são considerados quatro tipos de parceiros: 213 parceiros territoriais (bibliotecas e arquivos de autoridades locais, estruturas de cooperação regional, sociedades científicas); 55 parceiros de instituições do ensino superior e de investigação; outros (72), nos quais se incluem, principalmente, bibliotecas especializadas, mas com diferentes políticas de acesso e divulgação; e 6 bibliotecas estrangeiras.

A página de acolhimento (Figura 18) é uma montra da Gallica, com sítios temáticos, vídeos e notícias, e inclui a barra de acesso ao blogue *Le Blog Gallica*<sup>69</sup>, às pesquisas e a quatro grandes secções: lista de obras recentemente adicionadas (*Les actualités*); seleções de obras por tipo de documentos, áreas temáticas e áreas geográficas (*La sélection*); jogos e manualidades para crianças (*La fabrique*); apresentação dos vários parceiros (*Le regard des partenaires*); lista de anteriores destaques e sítios temáticos (*Déjà vu?*). O cabeçalho, na barra horizontal superior, é persistente em todas as

---

<sup>68</sup> Disponível em <http://gallica.bnf.fr/>

<sup>69</sup> Disponível em <http://gallica.bnf.fr/blog>

páginas do sítio e inclui ligação à BnF e à página de entrada da Gallica, às pesquisas simples e avançada em toda a biblioteca digital, às seleções predefinidas, ao blogue, bem como um botão para a seleção do idioma, cuja funcionalidade ainda não se encontra disponível.

Figura 18 – Gallica: página de acolhimento. Fonte: <http://gallica.bnf.fr/accueil> (31 jul. 2017)

Destaca-se, nesta página, o número de documentos em linha, tendo já ultrapassado os 4 milhões<sup>70</sup>. Este valor está em constante atualização e dá acesso à estatística de obras digitalizadas por tipo de objeto digital, mas sem indicar a data da última atualização.

As seleções predefinidas apresentam uma estrutura hierárquicas (Figura 19).

<sup>70</sup> Valor recolhido no dia 27/06/2017: 4.203.731; valor recolhido no dia 27/07/2017: 4.252.343 (valor confirmado a 31 de julho). Assim, no espaço de um mês, regista-se um incremento de 48.612 documentos em linha.

Figura 19 – Gallica: opções de pesquisa. Fonte: <http://gallica.bnf.fr/accueil> (31 jul. 2017)



As categorias definidas por tipo de documentos são: livros (*Livres*); manuscritos (*Manuscrits*); cartografia (*Cartes*); iconografia (*Images*); periódicos (*Presse et revues*); registros sonoros (*Enregistrements sonores*); partituras (*Partitions*); objetos tridimensionais, em que se inclui a coleção numismática (*Objets*); e vídeos, em construção. As áreas temáticas incluem: artes, lazer e desporto; direito e economia; literaturas; história; ciência. As áreas geográficas das fontes georreferenciadas estruturam-se em: França, África, América e Europa.

A navegação pelas listas dinâmicas das seleções predefinidas, bem como as listas de resultados a partir das pesquisas simples ou avançada, podem ser refinadas por: origem dos dados (*Site de consultation*); tipo de documento (*Type de document*); data de edição digital (*Date d'édition*); tema (*Thème*); versão (*Version*); formato de texto (*Mode texte*); tipo de acesso (*Type d'accès*).

As pesquisas simples e avançada incidem sobre todas as obras digitalizadas e disponíveis em linha. A maioria das obras encontra-se disponível na íntegra e em acesso aberto, sob a reserva de que as obras digitalizadas que se encontram

protegidas pelo direito de autor, apenas poderem ser consultadas no espaço físico da BnF.

No acesso rápido (*Accès rapide*), localizado no registo inferior da página, além da ligação às seleções predefinidas por tipo de documentos, permite o acesso: às primeiras páginas dos jornais do dia da consulta, cem anos antes (*Les unes il y a 100 ans*); aos dados estatísticos (*Gallica en chiffres*); à ordenação cronológica, em ordem descendente, das obras disponibilizadas (*Nouveaux documents*); às obras disponibilizadas em formato ePUB ("*electronic-Publication*", ou publicação eletrónica) arquivo digital padrão específico para *e-books* (*Les EPUB*); e à apresentação dos parceiros (*Tous les partenaires*).

Também na barra do registo inferior, as informações adicionais (*En savoir plus*) incluem: o historial e a descrição da biblioteca digital (*A propos*); as condições de utilização (*Conditions d'utilisation*); perguntas frequentes (F.A.Q. [*Frequently Asked Questions*]); acessibilidade, em construção; e serviço de mensagens eletrónicas (*Nous contacter*). À exceção desta, em cada uma das páginas acedidas a partir desta secção, está acessível um menu de ajuda para a pesquisa na Gallica, para consultar, descarregar, ou partilhar documentos, comprar reproduções e manter-se informado, adotar um livro ou voluntariar-se para ler uma obra para disponibilizar o respetivo registo sonoro (*Adopter un livre / Offrir une voix*), ou para subscrever informação relativa às atualizações (*Flux RSS [Really Simple Syndication]*). Está também acessível uma página com comentários dos utilizadores (*Du côté des Gallicanautes*).

Na secção relativa a outros endereços referentes à Gallica (*Retrouver Gallica sur*), faz-se a ligação às redes sociais Facebook, Twitter ou Pinterest, ao blogue e ao boletim

informativo *Lettre de Gallica*, sendo possível subscrever ou consultar os números anteriores em linha.

Ainda na mesma barra, está disponível a ferramenta para instalar a app Gallica (*Services mobiles: Télécharger l'application Gallica*), para dispositivos Androide e Apple, permitindo a sua consulta em dispositivos móveis. Há, ainda, a ligação a sítios eletrónicos externos (*Autres sites*): à página “*Adopter un livre*”, ao sítio eletrónico dos Amigos da BnF (*les Amis de la Bibliothèque nationale de France*), que convida a patrocinar a Gallica através do pagamento da digitalização de uma obra do catálogo da BnF; ao portal da Europeana.

As ferramentas de pesquisa simples e pesquisa avançada mantêm-se ativas na barra superior de todas as páginas do sítio eletrónico. Na pesquisa simples, é possível lançar um termo de pesquisa conjuntamente com um tipo de obra e aplicar refinadores: sítio eletrónico de consulta da obra (Gallica); tipo de documento; autor; data de edição; assunto; versão ePUB; língua do documento (francês, inglês, alemão, espanhol, italiano, chinês, grego e latim); modo de texto; modo de acesso. Na opção da pesquisa avançada, além da busca nos elementos bibliográficos da obra ou no respetivo conteúdo, caso estejam disponíveis em formato texto, podem ser aplicados os refinadores da pesquisa simples, a que se acrescem os seguintes: ano de edição, através de balizas temporais; versão 3D (para objetos); coleção digital; biblioteca detentora do original; data de colocação em linha. Permite, ainda, fazer pesquisa exata ou aproximada.

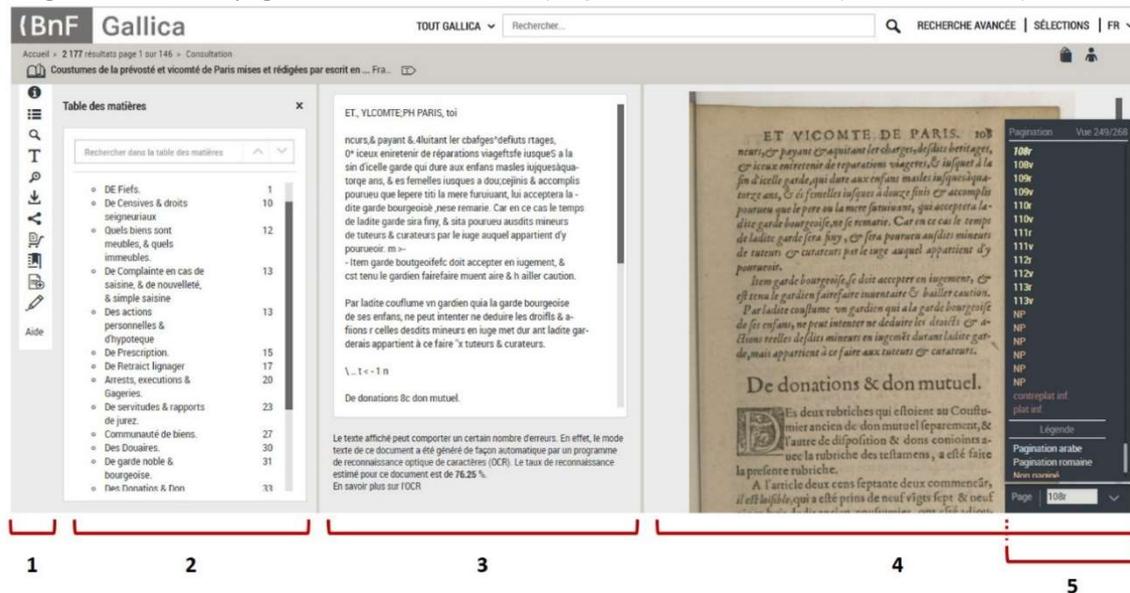
Os resultados podem ser ordenados por: relevância; autor; título; data de edição crescente ou decrescente; qualidade do texto (percentagem de acuidade da transcrição); e data de colocação em linha.

A informação acerca das últimas obras adicionadas é disponibilizada através de vários canais. Essa informação pode ser obtida, como referido, na consulta avançada, indicando a data específica ou um período temporal. Mediante subscrição de RSS, as atualizações das várias coleções são enviadas para a caixa de email indicada. A informação disponibilizada por este mail é mais assertiva do que a disponibilizada através da página de novidades, além de permitir personalização, inserindo centros de interesse específicos, por exemplo por autor, tema, ou tipo de obra. Outras estratégias adicionais são a subscrição do boletim informativo *Lettre de Gallica*, ou acompanhando a Gallica no blogue, ou nas redes sociais Facebook, Twitter e Pinterest.

As obras digitalizadas a partir de originais pertencentes às coleções da BnF são disponibilizadas na Gallica com uma apresentação uniforme, no que se refere à distribuição dos vários elementos gráficos e textuais na composição da página. Em contrapartida, as obras disponibilizadas por parceiros são visualizadas nos respetivos sistemas, pelo que a sua apresentação é muito diversificada.

As obras da BnF são apresentadas numa página dividida em três faixas horizontais, as primeiras, mais estreitas e com informação relativa à obra, e a última, a ocupar o campo principal em quase toda a extensão do ecrã, com a versão digital da obra (Figura 20).

Figura 20 – Gallica: página de consulta. Fonte: <http://gallica.bnf.fr/ark:/12148/bpt6k10404153> (31 jul. 2017)



A faixa superior corresponde ao cabeçalho constante em toda a Gallica. Na seguinte, os elementos básicos da obra a consultar, o acesso ao cesto de compras (para adquirir, por exemplo, reproduções de alta qualidade) e ao espaço pessoal.

No terceiro registo, é apresentada a obra numa estrutura compositiva que pode ter até cinco divisões verticais. À esquerda o menu, fixo e sempre visível, numa faixa estreita e com ícones de funcionalidades, com ajuda contextual em janela *pop-up* (Vd. Figura 20: 1): informações; sumário; pesquisa no documento; modo de apresentação; *zoom*; transferência ou impressão; partilha ou envio por correio eletrónico; compra de impressões; marcação de página; inserção da obra no espaço pessoal; e ajuda. Segue-se a faixa com o sumário da obra, a pesquisa no texto da obra ou a ficha bibliográfica (Vd. Figura 20: 2). A obra é reproduzida nas seguintes versões: a versão texto, caso exista (Vd. Figura 20: 3); a versão imagem que ocupa a maior parte do espaço, com os ícones de recuo e avanço (Vd. Figura 20: 4). Na última faixa, em *pop-up*, é indicada a paginação de acordo com o original (Vd. Figura 20: 5).

A consulta e navegação na Gallica é suportada por uma página de ajuda (F.A.Q., na secção “*En savoir plus*”), acedida igualmente a partir da página de visualização de cada obra. Além disso, todo o sítio eletrónico possui ajuda contextual com janela *pop-up*.

No espaço pessoal, é possível: adicionar obras ou páginas, construindo uma biblioteca digital personalizada, e pesquisas realizadas; definir as preferências, como, por exemplo, o formato do ficheiro a transferir; e efetuar aquisições e reservar lugar para leitura presencial. Como fator negativo, aponta-se a impossibilidade de exportar os registos bibliográficos para os gestores de referências.

A documentação técnica referente à Gallica encontra-se disponível no sítio eletrónico da BnF na secção “*Innovation numérique à la BnF*” (BnF, 2016). As estatísticas são disponibilizadas, também no sítio da BnF, na secção “*Gallica et la politique de diffusion numérique des collections*”, onde se destaca a apresentação dos relatórios anuais sobre toda a atividade da BnF e da Gallica (BnF, 2017).

#### **4.2.3. Internet Archive**

##### **Internet Archive 1996**

- Obras em domínio público;
- Alertas;
- Disponibiliza as obras vários formatos;
- Acesso: pesquisa e navegação;
- Pesquisa no conteúdo dos documentos.

Facilidade de utilização – média.

Utilidade – alta.

Performance – média.

O Internet Archive (IA)<sup>71</sup> é uma biblioteca digital sem fins lucrativos, fundada por Brewster Kahle em 1996, em São Francisco, Califórnia, nos Estados Unidos da América, com a missão de preservar e fornecer o acesso gratuito e universal ao conhecimento. Inclui ficheiros de texto (monografias, periódicos, documentação de arquivo, etc.) registos audiovisuais, material informático (programas, jogos, etc.) e o arquivo de páginas web desde 1996. É financiado por várias instituições, como Alexa Internet, HP Computer, Library of Congress, LizardTech, National Science Foundation, Prelinger Archives, Sloan Foundation e Kahle/Austin Foundation, e mecenas particulares.

O IA disponibiliza mais de 14 milhões de livros e textos em linha, além de registos de áudio (4 milhões), vídeos (3,5 milhões), imagens (1,5 milhões) e programas de computador (183.000)<sup>72</sup>. Um dos serviços mais relevantes é o Wayback Machine, que, atualmente, recupera o acesso a mais de 302 biliões<sup>73</sup> de páginas web arquivadas, num processo em conformidade com o Protocolo de Exclusão de Robôs (robotx.txt)<sup>74</sup>. Além destes, fornece serviços de arquivo e disponibilização de acervos, através do projeto Archive-it.

Os conteúdos são disponibilizados de forma gratuita e universal, provenientes, quer da própria atividade de digitalização, quer do contributo dos parceiros de projeto: Library of Congress, Harvard University, the New York Public Library, Smithsonian Institution, The Getty Research Institute, University of California, University of Toronto, Biodiversity Heritage Library, Boston Library Consortium, C.A.R.L.I., Johns Hopkins

---

<sup>71</sup> Disponível em <http://archive.org/>

<sup>72</sup> Ou seja 14.117.276, segundo informação recolhida a 31/07/2017, na página de acolhimento do sítio do IA.

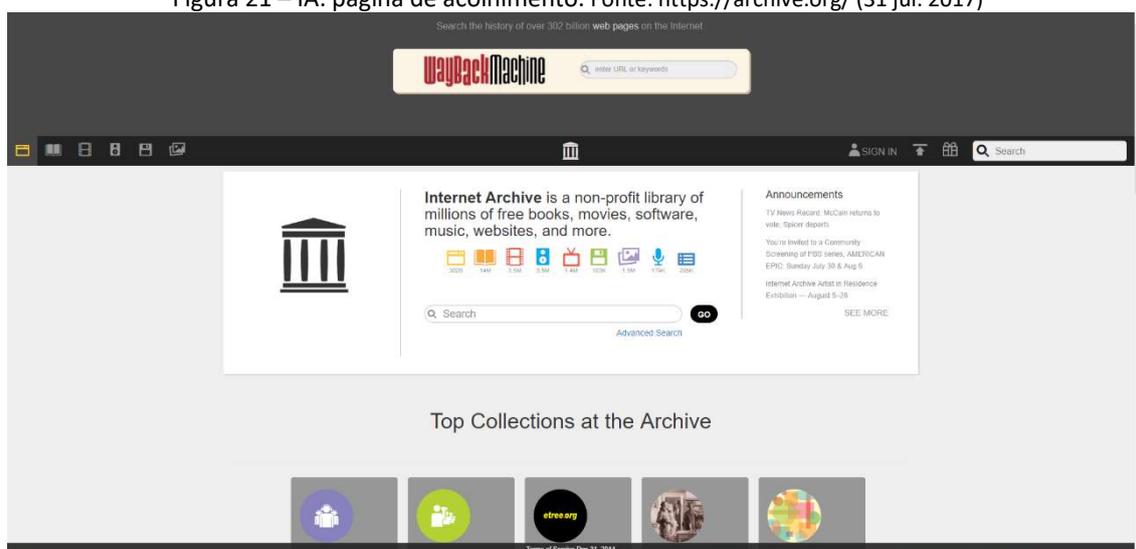
<sup>73</sup> Os dados numéricos seguem a notação americana, segundo informação recolhida a 31/07/2017, na página de acolhimento do sítio do IA.

<sup>74</sup> Disponível em <http://www.robotstxt.org/>

University, Allen County Public Library, Lyris, Massachusetts Institute of technology, State Library of Massachusetts, entre outros, instituições que aderiram ao Open Content Alliance OCA<sup>75</sup>.

A página de acolhimento (Figura 21) apresenta ferramentas de pesquisa e dados relativos à missão, ao desempenho e aos conteúdos da biblioteca.

Figura 21 – IA: página de acolhimento. Fonte: <https://archive.org/> (31 jul. 2017)



No cabeçalho, à esquerda, uma barra de ícones, com ajuda contextual em janela *popup*, dá acesso às várias tipologias de documentos, (arquivo web, textos, vídeos, registos áudio, programas informáticos e iconografia). Ao selecionar cada uma das tipologias, a barra superior, em destaque, muda em conformidade, permitindo o acesso direto ao arquivo da web (Waybackmachine), inserindo o endereço do sítio eletrónico pretendido, ou através da consulta por palavras-chave; nos restantes acessos, destacam-se duas coleções (nos textos, a coleção “eBooks & Texts” e a ligação à Open Library), e listas das principais coleções ou parcerias (Featured; Top). À direita,

<sup>75</sup> O Open Content Alliance - OCA (<http://www.opencontentalliance.org>) é um consórcio colaborativo, sem fins lucrativos, de organizações governamentais e grupos culturais e tecnológicos, que, em todo o mundo, contribuem para a disponibilização de textos e material multimédia de forma permanente e em acesso livre.

encontram-se os ícones referentes à intervenção do utilizador: registo para a construção de uma área pessoal; transferência de conteúdos para a biblioteca (*upload*); patrocínio financeiro; e pesquisa simples. Ao centro, o ícone do IA estabelece a ligação à página de acolhimento. Segue-se uma barra, pouco perceptível, em cinzento sobre cinzento-escuro, com ligação a informações relativas ao projeto (*About*), aos contactos (*Contact*), ao blogue (*Blog*), aos vários projetos (*Projects*), à página de ajuda (*Help*), ao apelo a doações (*Donate*), às políticas seguidas (*Terms*), às ofertas de trabalho efetivo (*Jobs*) ou voluntário (*Volunteer*) e à apresentação da equipa de trabalho (*People*). Este cabeçalho mantém-se ativo em todas as páginas do sítio eletrónico.

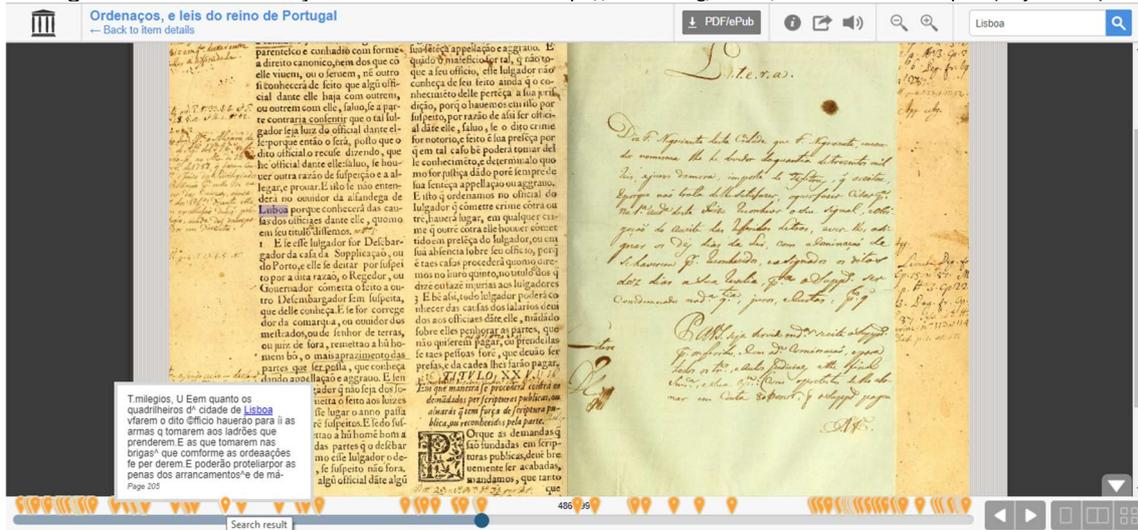
O campo principal apresenta: o símbolo do Internet Archive; a definição do projeto, com informação estatística relativa a cada uma das tipologias de documentos já enunciadas no cabeçalho e a que acrescem programas de televisão, concertos e coleções; um campo de pesquisa simples e ligação à pesquisa avançada; uma barra dinâmica de notícias (*Announcements*). A seguir a este campo, apresenta-se, numa disposição em mosaico, o acesso direto às várias coleções por ordem de relevância (*Top Collections at the Archive*).

O acesso às obras é feito através da pesquisa no próprio portal ou pela navegação nas diversas coleções. As espécies bibliográficas são catalogadas no formato Marc (Machine Readable Cataloging), a que se adiciona o esquema de metadados Dublin Core. Por conseguinte, no portal, a pesquisa de livros é efetuada, quer através dos elementos bibliográficos da obra, quer através do conteúdo textual dos documentos ou dos metadados adicionados.

As obras são visualizadas em *e-Reader*, permitindo navegar pelo número de página, passando o rato pelo corte do livro, o que não funciona bem em obras volumosas. Sob a imagem da obra, encontra-se, à esquerda, o respetivo registo bibliográfico com notas extensas e a possibilidade de aceder ao registo completo em formato Marc, bem como a ficha técnica da edição digital. Os dados do registo bibliográfico estão formatados em hipertexto, permitindo lançar pesquisas a partir dos respetivos termos: autor, ano de publicação, assunto, impressor, coleção, entidades patrocinadora e fornecedora da digitalização, língua do documento. À direita, encontram-se outros dados relativos à obra digital: os ícones que permitem marcar a obra como favorita, partilhar o ficheiro, por correio eletrónico ou nas redes sociais, e assinalar se a obra tem conteúdos visuais de violência ou de carácter sexual, se é publicidade (*spam* e *scam*) ou fraude, ou se os conteúdos estão danificados; uma lista de opções de formato para a transferência de ficheiros, variando consoante o tratamento aplicado à obra; a coleção em que a obra se insere; a identificação do agente que disponibilizou a obra; o número de visualizações e a respetiva marcação como favorito. Permite, ainda, inserir uma revisão da obra ou visualizar revisões inseridas. No registo inferior, apresenta as miniaturas das imagens (*thumbnails*) e informação básica acerca de itens similares, selecionados em função dos respetivos metadados, com ligação direta às obras digitalizadas.

Na visualização em ecrã inteiro (Figura 22), a obra aparece entre duas barras horizontais. Na barra superior, à direita, a navegação no sítio eletrónico permite o acesso: à página de acolhimento, a partir do símbolo do IA; aos detalhes da obra (página de entrada na obra), a partir do respetivo título.

Figura 22 – IA: visualização ecrã inteiro. Fonte: <https://archive.org/details/ordenaoseleisdor00port> (31 jul. 2017)



À esquerda, ferramentas para utilização e navegação na obra permitem o acesso a: transferência em PDF ou formatos *e-Reader*; informação bibliográfica; partilhas da obra por correio eletrónico ou nas redes sociais; leitura áudio do texto; ferramentas para aumentar ou diminuir o tamanho da imagem visualizada no ecrã; e, por fim, pesquisa textual no documento, com ajuda interativa. Sob a imagem, apresenta-se uma barra horizontal com a indicação do número das páginas em visualização e com um marcador da respetiva posição relativa no conjunto da obra. É nesta barra que surgem os marcadores referentes aos resultados da pesquisa no conteúdo da obra; em cada marcador, além de abrir a página, é possível abrir uma janela *pop-up* com a transcrição do texto em que o termo pesquisado se insere (Figura 22). Na imagem da obra, os termos pesquisados são realçados. Ao lado desta barra, surge um conjunto de botões de auxílio à navegação e visualização da obra: setas de recuo e avanço nas páginas; modo de visualização em uma ou duas páginas (verso-rostro) ou em mosaico de miniaturas (*thumbnail*); seta para esconder as barras de navegação.

As opções de leitura em linha e em PDF apresentam um nível de qualidade muito variável. As versões baseadas na transcrição de texto recorrendo ao OCR, sobretudo

em obras mais antigas, têm um desempenho bastante inferior com um elevado índice de erro na transcrição. Porém, dado que apresenta sempre uma versão fac-similada das obras, é possível comparar a versão transcrita com o original e corrigir os erros existentes.

A exportação dos dados bibliográficos das obras para os gestores de referências não é direta, sendo necessário abrir a referência no WorldCat e, a partir daí, transferir o ficheiro para o gestor de referências.

Através da pesquisa avançada, o IA permite a subscrição de informação relativa às atualizações, selecionando a opção RSS (*RSS format*). O *Internet Archive Blogs*<sup>76</sup>, os fóruns e as redes sociais do IA facultam informação atualizada acerca dos conteúdos, mas o acesso a estas páginas não está visível na página de acolhimento.

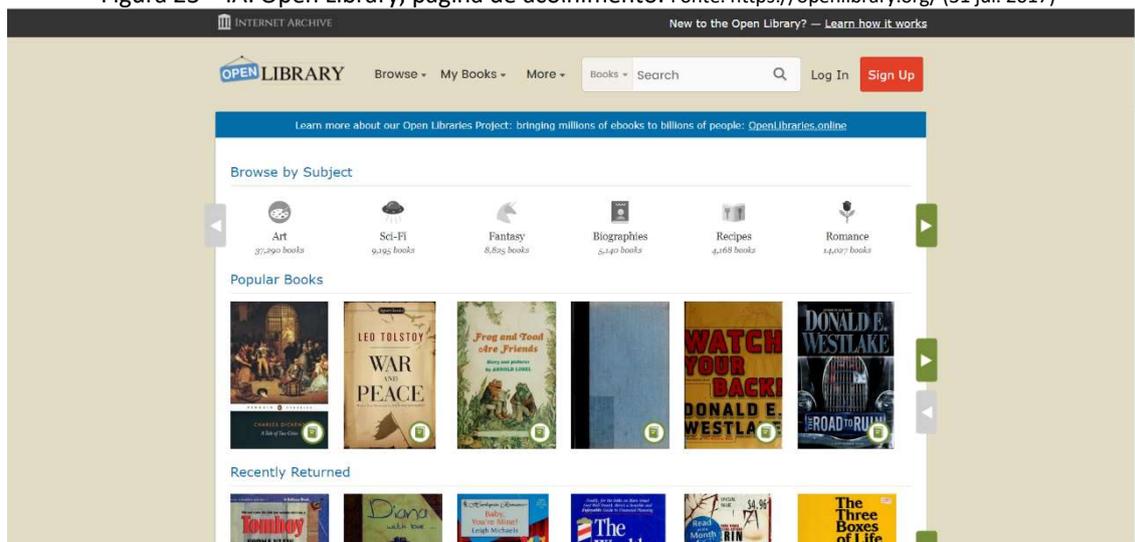
Outra modalidade de pesquisa de conteúdos pode ser efetuada através da Open Library<sup>77</sup>, um projeto do IA, que se define como catálogo que agrega os registos bibliográficos desta e de outras bibliotecas parceiras em todo o mundo, permitindo, sempre que disponíveis em livre acesso, a ligação aos conteúdos digitalizados das obras; caso estes não se encontrem em livre acesso, a Open Library faculta gratuitamente o empréstimo da respetiva versão em formato digital, mediante o registo do utilizador como leitor, replicando a prática comum numa biblioteca física. A Open Library, tendo como objetivo “[...] to provide a page on the web for every book ever published” (Gee & West, 2008), encontra-se disponível numa plataforma aberta ao contributo de todos os utilizadores.

---

<sup>76</sup> Disponível em <http://blog.archive.org/>

<sup>77</sup> Disponível em <https://openlibrary.org/>

Figura 23 – IA: Open Library, página de acolhimento. Fonte: <https://openlibrary.org/> (31 jul. 2017)

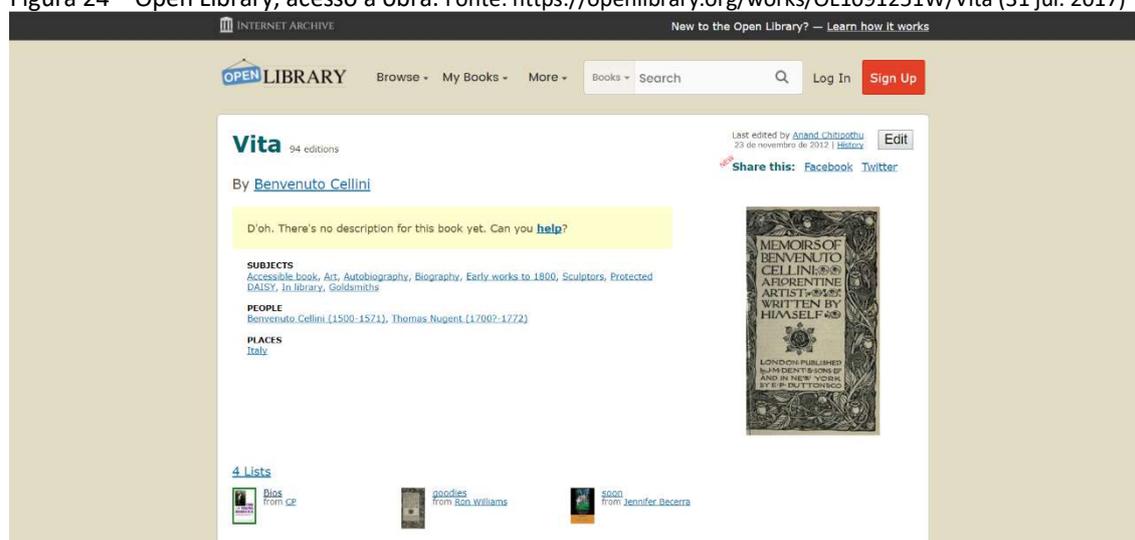


Na página de acolhimento (Figura 23), a pesquisa (*Browse*) é orientada por assuntos: arte (*Art*), ficção científica (*Sci-Fi*), fantasia (*Fantasy*), biografias (*Biographies*), culinária (*Recipes*), romance (*Romance*), textos (*Textbooks*), infantis (*Children*), história (*History*), medicina (*Medicine*), religião (*Religion*), mistério (*Mysteries*), teatro (*Plays*), música (*Music*) e ciência (*Science*). Inclui, também, listas de obras com a imagem da capa em miniatura (*thumbnail*), com os mais populares (*Popular Books*), adições recentes (*Recently Returned*) e outros temas, como literatura clássica (*Classic Literature*), ou obras acessíveis a cegos (*Print Disabled (Blind Accessible)*), entre outros. Sobre cada capa, um ícone indica as obras em acesso livre para leitura em linha, as que apenas se encontram disponíveis para empréstimo, ou as obras (livros falados ou em *braille*) integradas no programa NLS (National Library Service for the Blind and Physically Handicapped), da Library of Congress, e que requerem um registo de deficiência física, de baixa visão ou cegueira.

Na Open Library (Figura 24), a pesquisas simples e avançada é feita nos elementos bibliográficos e no conteúdo dos documentos, sendo possível restringir os resultados a

livros eletrônicos (*e-books*). As obras disponíveis para leitura permitem a transferência para equipamentos pessoais.

Figura 24 – Open Library, acesso à obra. Fonte: <https://openlibrary.org/works/OL1091251W/Vita> (31 jul. 2017)



As obras são disponibilizadas em múltiplos formatos: leitura em linha HTML, PDF, PDF com transcrição do texto associada, ePUB, Kindle (Mobi), Daisy, texto e DjVu. O formato ePUB é um formato textual aberto e normalizado que tem a vantagem de permitir a leitura dos documentos em vários tipos de ecrã. É possível, também, criar listas personalizadas de obras, as quais podem ser disponibilizadas e associadas à obra, tornando-se um auxiliar de pesquisa para a investigação, em complemento aos respetivos descritores por assunto (*subjects*), pessoas (*people*) e lugares (*places*).

O elemento diferenciador deste projeto é a atualização das funcionalidades da biblioteca analógica às virtualidades das novas tecnologias e, em particular, a política de empréstimo de obras em formato digital.

#### 4.2.4. Bibliothèques Virtuelles Humanistes

##### **Bibliothèques Virtuelles Humanistes 2002**

- Obras em domínio público;
- Disponibiliza as obras em formato PDF e etiquetadas com TEI;
- Acesso: pesquisa e navegação;
- Pesquisa no conteúdo dos documentos.

Facilidade de utilização – média.

Utilidade – alta.

Performance – média.

O projeto universitário Bibliothèques Virtuelles Humanistes (BVH)<sup>78</sup> é desenvolvido, desde 2002, pelo Centre d'Études Supérieures de la Renaissance (CERS), da universidade François-Rabelais de Tours, com a colaboração da secção de humanismo do Institut de Recherche et d'Histoire des Textes (IRHT) do Centre National de la Recherche Scientifique (CNRS), em França.

Concebido como uma biblioteca digital de documentos patrimoniais renascentistas, a BVH disponibiliza uma seleção de obras fac-similadas, uma base de dados textuais com transcrições em TEI-compliant XML<sup>79</sup> (Text Encoding Initiative – compliant Extensible Markup Language) e transcrições ou análises de minutas notariais e manuscritos renascentistas.

No cabeçalho da página de acolhimento (Figura 25), onde é apresentado o logótipo, a informação é estruturada em quatro secções: obras fac-similadas (*Fac-similés*);

---

<sup>78</sup> Disponível em <http://www.bvh.univ-tours.fr/>

<sup>79</sup> TEI - Text Encoding Initiative. O TEI foi criado em 1987 por um consórcio das associações académicas, Association for Computers and the Humanities (ACH), Association for Computational Linguistics (ACL) e Association for Literary and Linguistic Computing (ALLC), que o propuseram como o resultado da pesquisa efetuada para desenvolver um esquema de metadados que respondesse cabalmente à colocação em linha de textos digitais, nos seus múltiplos formatos.

documentos transcritos em TEI (*Epistemon (textes)*); manuscritos (*Manuscripts*); iconografia (*Iconographie*); e a apresentação do projeto (*Project*). Este cabeçalho é persistente em todas as páginas do sítio eletrónico até à abertura da obra, que se efetua num novo separador.

Figura 25 – BVH: página de acolhimento. Fonte: <http://www.bvh.univ-tours.fr/> (31 jul. 2017)

The screenshot shows the homepage of the BVH website. At the top, there is a header with the BVH logo on the left, a central illustration of a scholar, and the text 'Les Bibliothèques Virtuelles Humanistes' on the right. Below the header is a navigation menu with five items: 'Fac-similés', 'Epistemon', 'Manuscrits', 'Iconographie', and 'Projet'. The main content area is divided into several sections. On the left, there is a search bar with the text 'Recherche simple' and a search button labeled 'ok'. Below the search bar, there are options for 'Choix du corpus' and 'Options'. The 'Présentation' section contains a text introduction and a list of bullet points. The 'Actualités' section contains a list of news items. The 'Dossiers | Projets' section contains a list of projects.

**Recherche simple**   > RECHERCHE AVANCÉE

Choix du corpus :  Fac-similés  Epistemon (textes)  Manuscrits

Options :  Notices  Plein-texte

**Présentation**

Le programme « Bibliothèques Virtuelles Humanistes » diffuse des documents patrimoniaux et poursuit des recherches associant des compétences en sciences humaines et en informatique. Il agrège plusieurs types de documents numériques :

- Une sélection de [fac-similés](#) d'ouvrages de la Renaissance numérisés en Région Centre et dans les établissements partenaires
- La base textuelle [Epistemon](#), qui offre des éditions numériques en XML-TEI
- Des transcriptions ou analyses de [minutes notariales](#) et des [manuscrits](#)

Responsable : Chiara Lastraioli [Lire la suite](#)

**Dossiers | Projets**

- [FOnds Rabelais et ses Sources En ligne \(FORSE\)](#)
- [Projet ReNOM Renaissance / Site ReNom](#)

**Actualités**

- Offre d'emploi : [Assistant ingénieur pour l'informatisation des Catalogues Régionaux des Incunables des Bibliothèques publiques de France \[2 postes à pourvoir\]](#), projet CRII | CESR, Tours [Candidatures jusqu'au 10 juil. 2017]
- École d'été [Bibliotheca Digitalis](#) : [Conférences publiques](#) | Médiathèque Louis-Aragon, Le Mans [4-6-7 juil. 2017]
- Formation : [Initiation à l'encodage XML-TEI des textes patrimoniaux](#), CESR, Tours | Inscriptions ouvertes [23, 24 & 25 oct. 2017]
- Epistemon / Montaigne : [Édition numérique XML-TEI du Manuscrit du Discours de la Servitude Volontaire d'Étienne de La Boétie](#) [22 juin 2017]

A página de acolhimento apresenta um quadro de pesquisa simples e ligação à pesquisa avançada, o texto de apresentação do projeto (*Présentation*), uma resenha de atualidades (*Actualités*), pastas de projetos associados (*Dossiers*) e ligações ao boletim informativo *La Lettre des BVH*, permitindo a respetiva subscrição, ao blogue do

projeto *Les Bibliothèques Virtuelles Humanistes*<sup>80</sup>, e às páginas que mantém nas plataformas SourceForge<sup>81</sup> e Slideshare<sup>82</sup>.

A pesquisa simples incide sobre os campos título, autor, data, edição e cota e permite articular refinadores relativos à coleção (*Fac-similés; Epistemon (textes); Manuscripts*), à categoria dos resultados, como informação (*Notice*), ou texto integral (*Plein texte*), e à biblioteca detentora do original (*Epistemon; Medic@; Gallica; Bibliothèque André-Desguine*). Na coleção de fac-similes, que corresponde às imagens de obras renascentistas digitalizadas pelos parceiros do projeto, é possível fazer a pesquisa por nome do autor da obra (*auteur*), do prefácio ou textos introdutórios (*auteur de pièce liminaire*), do tradutor (*traducteur*) ou do editor (*éditeur scientifique*).

A pesquisa avançada (Figura 26) efetua a busca nos elementos bibliográficos da obra e é assistida pela introdução de listas junto a determinados campos: índice de autores (*Index des auteurs*); lugar de edição (*Lieu d'édition*); lugar de impressão (*Lieu d'impression*); índice de livrarias (*Index des librairies*); índice de impressores (*Index des imprimeurs*); índice de formatos (*Index des formats*), segundo a classificação do tamanho em função da dobragem das folhas; língua (*Index des langues*); assunto (*Index des matières*) e localização do original (*Bibliothèque; Ville*). No campo data, é possível definir um período cronológico entre duas datas limite.

---

<sup>80</sup> Disponível em <https://bvh.hypotheses.org/>

<sup>81</sup> Disponível em <https://sourceforge.net/projects/bvh/?source=directory>

<sup>82</sup> Disponível em <https://www.slideshare.net/BVH-CESR>

Figura 26 – BVH: pesquisa avançada. Fonte: <http://www.bvh.univ-tours.fr/avancerech.asp> (31 jul. 2017)

Recherche avancée

Titre

Auteur  Index des auteurs ▼

Édition  Lieu d'édition ▼

Impression  Lieu d'impression - ▼

Libraire  Index des libraires ▼

Imprimeur  index des imprimeurs ▼

Date  < Période <

Format  Index des formats ▼

Langue  Index des Langues ▼

Matière  Index des matières ▼

Localisation  Bibliothèque ▼ Ville ▼

Résultats triés par

Opérateur intra rubriques

Opérateur extra rubriques

Titre ▼

Et ▼

Et ▼

Rechercher

A pesquisa pode ser assistida por listas (*Recherche par liste*) de autores e de temas, organizadas, alfabeticamente, por autor ou título e, cronologicamente, por data de publicação. Outra modalidade de pesquisa é a lista das últimas obras adicionadas (*Dernières numérisations*), correspondendo a um conjunto de 1.133 registos, os últimos dos quais foram colocadas em linha a 19 de maio de 2016<sup>83</sup>.

Os resultados da pesquisa são apresentados em lista com ligação aos registos bibliográficos das obras e a partir dos quais se acede ao respetivo conteúdo digitalizado.

<sup>83</sup> Segundo consulta efetuada a 31 de julho de 2017.

A versão digitalizada da obra apresenta-se numa página estruturada em três campos: a barra superior estática com o logótipo da BVH; uma barra vertical, à esquerda, com uma caixa de pesquisa simples, opções de navegação na obra, ligações à página de acolhimento, à lista de resultados e ao registo bibliográfico da obra; a reprodução digital da obra, no campo principal, encimada pela identificação das páginas (n.º e parte do livro, em função do respetivo sumário) e setas de navegação.

Estão disponíveis diferentes modalidades de navegação na obra:

- Navegação página a página (*Feuilletage*), cada uma das quais identificada pelo número e com indicação relativa à parte da obra em que se insere;
- Apresentação sequencial (*Chemin de fer*) das miniaturas (*thumbnail*) de todas as imagens digitalizadas da obra;
- Através do sumário da obra (*Sommaire*) que indica a parte da obra que está a ser visionada, referindo o número da página respetiva;
- Através das ilustrações (*Images*) existentes na obra, onde se inclui a iconografia, capitulares, emblemas, tarjas, vinhetas e outros elementos gráficos.

Por norma, a obra também está disponível em PDF, permitindo descarregá-lo na íntegra ou parcialmente. O PDF total possui três modalidades: a preto e branco de alta definição para impressão; a cores, com a transcrição do texto associada, não corrigida e não visível; a cores, só em formato imagem, para impressão.

A *Epistemon*, corpus de textos da Renascença, é uma base de dados textual, com a transcrição dos documentos para XML-TEI. A pesquisa é realizada através dos elementos da ficha bibliográfica ou através de palavra no texto, e é igualmente

assistida por listas de termos. Nem todos os textos estão completamente tratados, sendo que alguns estão disponíveis em HTML, com transcrição, mas sem a codificação em TEI. Depois de selecionar a obra pretendida, acede-se a uma página com informações acerca da obra, incluindo o registo bibliográfico, e, em seguida, à página de rosto exibida em modo imagem e texto. A navegação é feita página a página ou através dos verbetes do sumário.

A secção dos manuscritos disponibiliza documentos dos séculos XV e XVI, sobretudo, atos notariais, transcritos e codificados para TEI. A pesquisa simples faz-se por palavra ou sequência de caracteres e aceita o refinamento de critérios, como o nome do notário e da cidade, assistidos por lista de termos, assim como a data (dia, mês, ano) ou balizas temporais.

A secção de iconografia disponibiliza, de forma autónoma, as ilustrações das várias obras integradas na biblioteca. A pesquisa realiza-se através do Iconclass<sup>84</sup>. A referência à obra de origem mantém-se anexa à imagem. Nesta secção, são também disponibilizadas bases de dados de letras capitulares, de marcas dos impressores e de retratos, fazendo sempre a referência à obra de origem.

O processo está devidamente documentado, nomeadamente, no que se refere às opções de codificação em TEI e às políticas de transcrição.

Em 24 de abril de 2014, este projeto ganhou o prémio Succeed<sup>85</sup>, atribuído ao projeto de digitalização de documentos históricos.

---

<sup>84</sup> Sistema de classificação de imagens, disponível em <http://www.iconclass.org/>

<sup>85</sup> Vd. <http://succeed-project.eu/succeed-awards>

Não possui nenhum sistema que permita ao leitor ser informado acerca das novas obras adicionadas ou alteradas, o RSS funciona para notícias em geral, eventos, conferências, ofertas de emprego, etc. Também não permite a adição direta aos gestores de referências.

#### 4.2.5. Perseus Digital Library

##### **Perseus Digital Library 1987**

- Obras em domínio público;
- Disponibiliza as obras em XML, etiquetadas com TEI;
- Acesso: pesquisa e navegação;
- Pesquisa no conteúdo dos documentos.

Facilidade de utilização – média.

Utilidade – alta.

Performance – média.

O projeto Perseus Digital Library<sup>86</sup>, iniciado em 1987 e sediado no departamento de estudos clássicos (Department of the Classics) da Tufts University, nos Estados Unidos da América, e coordenado pelo editor-chefe Gregory Crane, é desenvolvido de forma colaborativa entre universidades europeias (Consiglio Nazionale delle Ricerche, Itália; University of Roma Tor Vergata, Itália; University of Leipzig, Alemanha) e americanas (Center for Hellenic Studies, College of the Holy Cross, Furman University, Tufts University, University of Massachusetts - Amherst, University of Missouri em Kansas City, dos Estados Unidos da América; Mount Allison University, Canadá). Além disso, congrega instituições e investigadores de diferentes especialidades, tais como e entre outras: línguas antigas, como latim, grego e árabe; análise de texto fragmentados;

---

<sup>86</sup> Disponível em <http://www.perseus.tufts.edu/hopper/>

edição multilingue; tratamentos de dados; edição digital. Constitui, por isso, um dos mais consistentes exemplos de projeto colaborativo e multidisciplinar em Humanidades e bibliotecas digitais.

Esta biblioteca digital desconstrói o livro físico, mas mantém a relação do texto com os capítulos ou secções ou com as linhas da edição impressa que lhe deu origem. No que se refere aos periódicos, mantém a lógica, perdendo-se a localização da notícia na página, os separadores e as variantes das fontes tipográficas. Quando foi constituída, tinha o objetivo de disponibilizar em linha o conteúdo de um conjunto de fontes primárias e secundárias que fundamentassem a história, a literatura e a cultura clássicas, em grego e latim, com a respetiva tradução para inglês. Posteriormente, a coleção foi aumentada e atualmente possui uma seção de arte e arqueologia (moedas, vasos, esculturas, edifícios) e outras coleções: textos em árabe; textos em alemão; documentos relativos ao século XIX americano; literatura inglesa da Renascença; poesia italiana do século XII à primeira metade do século XVI; e o periódico *Times Dispatch*, de Richmond<sup>87</sup>.

Na página de acolhimento (Figura 27), o cabeçalho inclui o logótipo, uma caixa de pesquisa simples e acesso a outras modalidades de pesquisa, incluindo uma lista de abreviaturas mais usadas, e uma barra de navegação, com os seguintes itens: página de acolhimento (*Home*); navegação por coleções (*Collections/Texts*); pesquisa no catálogo (*Perseus Catalog*); apresentação do projeto de investigação, incluindo o respetivo historial (*Research*); relatório financeiro (*Grants*); descrição da acessibilidade do projeto (*Open Source*); apresentação da equipa de investigação e das atividades e

---

<sup>87</sup> Disponível na Library of Congress, em <http://chroniclingamerica.loc.gov/lccn/sn84024738/>.

publicações relacionadas com o projeto, incluindo o acesso a eventuais ofertas de emprego ou investigação (*About*); e uma página de ajuda (*Help*). Este cabeçalho mantém-se persistente ao longo de todo o sítio eletrónico.

Figura 27 – Perseus: página de acolhimento. Fonte: <http://www.perseus.tufts.edu/hopper/> (31 jul. 2017)

The screenshot shows the Perseus Digital Library homepage. At the top, there is a navigation menu with links for Home, Collections/Texts, Perseus Catalog, Research, Grants, Open Source, About, and Help. A search bar is positioned in the top right corner. The main content area is divided into several sections. On the left, there is a 'Welcome to Perseus 4.0, also known as the Perseus Hopper.' section with a link to 'Read more on the Perseus version history'. Below this is the 'Perseus Updates' section, which contains three news items: 'January 9, 2017: Update on the new MA in Digital Tools for Premodern Studies at Tufts', 'December 22, 2016: A New Version of Perseus and Academic Partnerships', and 'December 21, 2016: Draft for a Request for Proposals for a new Perseus Digital Library'. The 'Release Announcements' section follows, with an entry for October 2013. On the right side, there are three main content boxes: 'Popular Texts' listing various classical works, 'Art and Archaeology' featuring images of the Temple of Aphaia, a silver obol, Satyr on an Attic red figure vase, and The Bartlett Head, and 'Exhibits'.

A página de acolhimento inclui a ligação ao Perseus 4.0, ou Perseus Hopper, à cronologia do projeto (*Perseus version history*), a um tutorial e às páginas do projeto nas redes sociais Facebook e Twitter. Apresenta, também, uma secção de notícias: atualizações (*Perseus Updates*), com ligação ao blogue Perseus Digital Library Updates<sup>88</sup>; e artigos relacionados (*Release Announcements*). À direita, apresenta caixas de destaques: textos mais consultados (*Popular Texts*); coleções de arte e arqueologia clássicas (*Art and Archaeology*); exposições virtuais (*Exhibits*); sítios eletrónicos com temáticas relacionadas (*Featured Sites*).

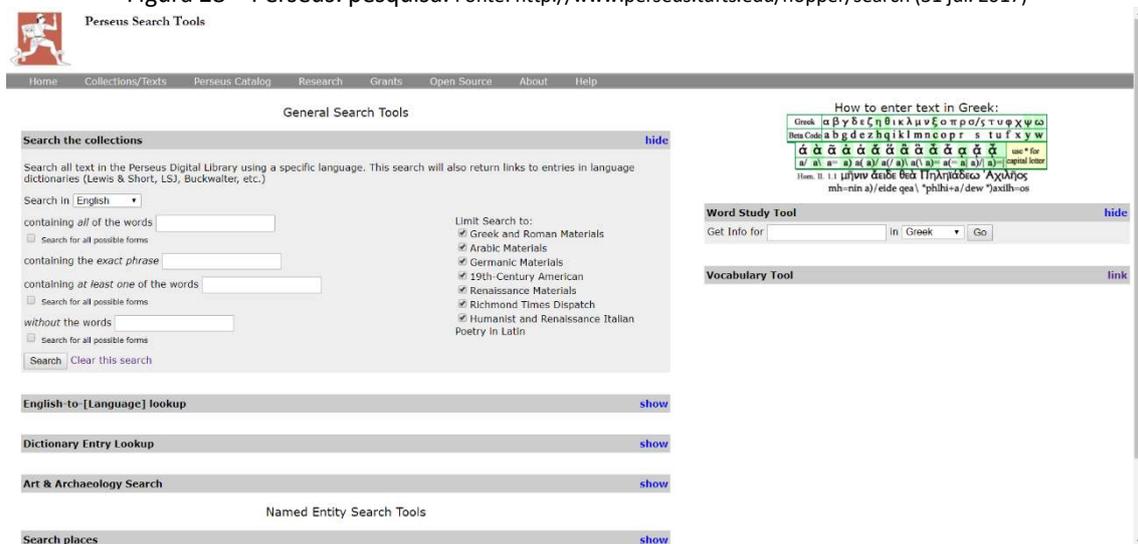
No registo inferior da página, uma breve descrição do projeto faculta a ligação aos sítios eletrónicos institucionais dos parceiros e patrocinadores.

<sup>88</sup> Disponível em <http://sites.tufts.edu/perseusupdates/>

Acede-se às obras por navegação na lista de autores, título, língua do documento original, edição, séries e assuntos, ou por pesquisa no catálogo. Este inclui as obras do projeto e de outras bibliotecas digitais. A interface é textual, não apresentando ícones para a navegação. O acesso às obras realiza-se através da navegação pelas páginas das várias coleções ou por pesquisa alargada ao conteúdo dos documentos.

A pesquisa avançada (*General Search Tools*) permite restringir a língua (inglês atual e arcaico, grego, latim, alemão e nórdico antigo) e pode ser limitada a determinada coleção: materiais greco-romanos (*Greek and Roman Materials*), árabes (*Arabic Materials*), germânicos (*German Materials*), novecentistas americanos (*19th-Century American*), renascentistas (*Renaissance Materials*), do jornal *Richmond Times-Dispatch* (*Richmond Times Dispatch*), e poesia em latim de autores italianos humanistas e renascentistas (*Humanist and Renaissance Italian Poetry in Latin*).

Figura 28 – Perseus: pesquisa. Fonte: <http://www.perseus.tufts.edu/hopper/search> (31 jul. 2017)



A pesquisa de termos em línguas antigas (grego, latim, árabe e vernáculas) (Figura 28) pode ser feita por grafia aproximada do termo (“words beginning with”; “words ending with”; “words containing”) nas línguas do projeto (*English-to-[Language] lookup*;

*Dictionary Entry Lookup*). Inclui, também, a pesquisa de artefactos (*Art and Archaeology Search*), lugares (*Search places*), pessoas (*Search people*) e datas (*Search dates*). Todas as modalidades de pesquisa incluem instruções de utilização.

Nos textos apresentados em resultado da pesquisa, as pessoas/personagens e os locais mencionados estão referenciadas através de hiperligações. Para os textos em versão bilingue, são indicadas as opções de tradução. As obras foram codificadas em TEI.

Todas as operações de pesquisa e recuperação da informação são executadas de forma amigável e transparente para o utilizador, o que torna o Perseus um dos projetos mais completos e exemplares no âmbito das Humanidades Digitais. “The rise of huge collections and new analytical methods is certainly important because it can provide new tools with which scholars can see what later ages had to say about Greek and Latin sources” (Crane, 2011, p. 19). Além disso, este projeto desenvolveu um conjunto significativo de competências adicionais, como a estruturação do conhecimento através de marcadores (*tag*), o multilinguismo, a aplicação extensiva do TEI-compliant XML, a otimização dos algoritmos de pesquisa e de reconhecimento de caracteres (OCR) (Bamman, Babeu, & Crane, 2010), contribuindo para a definição de um conjunto de boas práticas aplicáveis a projetos idênticos. Constitui um *corpus* de informação legível pelo homem ou pelo computador. Como a informação está estruturada com metadados normalizados, é possível criar algoritmos para realizar outras “leituras” sobre os textos e criar novo conhecimento.

A informação é fornecida em três camadas, ou categorias de acesso (Perseus, s.d.):

- Conhecimento legível: imagens digitalizadas de objetos, lugares, inscrições, páginas impressas, informações geográficas e outras representações digitais de objetos e espaços. Nesta camada, acede-se diretamente ao objeto digitalizado.
- Conhecimento acionável: registros de catálogo, artigos de enciclopédia, verbetes e outras fontes de informação estruturadas.
- Conhecimento gerado: novos conhecimentos gerados através da análise de sistemas de informação automatizados. A máquina de conhecimento acionável permite, por exemplo, ver uma entrada de dicionário ou encontrar traduções pré-existentes para uma citação em latim ou grego.

O projeto está acessível a partir da Universidade de Tufts, mas possui cópias (*mirrors*) nas Universidades de Oxford e Berlim.

Numa apreciação global, pode dizer-se que esta biblioteca digital serve os fins específicos relacionados com a colocação em linha de textos gregos e latinos e respetiva tradução, cumprindo os objetivos inicialmente traçados. Em contrapartida, como modelo a ser implementado nas bibliotecas digitais patrimoniais, é complexo e requer um grande investimento em termos quer financeiros, quer humanos. Provavelmente, não se justifica a aplicação do TEI a todas as obras, mas reconhece-se a importância da sua aplicação a obras que sejam consideradas fontes primárias para os vários domínios do conhecimento.

#### **4.2.6. Estudo comparativo**

Todas as bibliotecas digitais cumprem, com mais ou menos eficácia, os requisitos de pesquisa e navegação, de acordo com o conceito atual de biblioteca digital, implicando que os objetos digitais, que constituem a sua coleção, estejam organizados e

difundidos de forma a permitir a pesquisa e a recuperação da informação (Tabela 5). A pesquisa na Gallica é a que permite mais refinamentos, conduzindo a listas mais assertivas de resultados. Em contrapartida, a navegação na BVH parece mais eficiente, o que pode justificar-se pelo facto de o volume de obras ser menor.

Tabela 5 – Quadro resumo bibliotecas digitais

	<b>BND</b>	<b>Alma</b>	<b>Gallica</b>	<b>IA</b>	<b>BVH</b>	<b>Perseus</b>
Pesquisa	✓	✓	✓	✓	✓	✓
Navegação	✓	✓	✓	✓	✓	✓
Sinalética	✓	✓	✓	✓		
Ajuda	✓	✓	✓	✓	✓	✓
Alertas RSS			✓	✓	✓	
App			✓			
ePUB			✓	✓		
Personalização		✓	✓	✓		
Transferir ficheiros	✓	✓	✓	✓	✓	
Informação estatística			✓	✓		
Gestores de referências	✓	✓		✓		

A sinalética utilizada nas várias bibliotecas digitais é bastante uniforme, utilizando o mesmo tipo de símbolos, com idênticos significados ou funções (Tabela 6 – Símbolos e funções).

Tabela 6 – Símbolos e funções

Símbolo	Descrição	Função
	Lupa (antecedida de caixa de texto)	Pesquisa no conteúdo da obra
	Lupa com o sinal mais (+)	Ampliar a imagem
	Lupa com o sinal menos (-)	Diminuir a imagem
	Livro aberto	Consultar a obra, ou ler a obra em linha
	Matriz 3x3 em quadrícula	Aceder às miniaturas das imagens ( <i>thumbnail</i> ) das páginas da obra
	Letra "i" (de Informação, ou <i>Information</i> )	Obter informações acerca da obra, em regra, a ficha bibliográfica
	Seta para baixo	Transferir, ou descarregar, a obra
	Seta curva	Partilhar nas redes sociais

Neste aspeto, a BND constitui uma exceção, dado que utiliza sinalética pouco comum e explícita, como, por exemplo, a utilização do livro aberto para aceder às miniaturas das páginas, enquanto que usa binóculos para a pesquisa no conteúdo dos documentos, com a agravante de o símbolo ser cinzento escuro (RGB 97,97,97) sobre um fundo quase negro (RGB 19,14,11).

A ajuda disponibilizada nas várias bibliotecas digitais é díspar. As páginas de ajuda da Gallica, IA, BVH e Perseus contribuem para a eficácia da consulta, tanto nas primeiras abordagens, como em utilizações mais avançadas. Em contrapartida, a página de ajuda da Alma Mater é muito incipiente e revela-se de pouca utilidade. A ajuda contextual, na BND e na Gallica, descodifica a sinalética utilizada. Por conseguinte, no caso da BND, contribui para compensar a irregularidade referida na sinalética. No entanto e em geral, continua a ser necessário incrementar a informação nas páginas de ajuda,

incluindo tutoriais em formato vídeo, para uma exploração mais eficiente dos recursos disponibilizados, sobretudo, atendendo a que algumas das funcionalidades não se encontram imediatamente acessíveis. Ainda que o facto de todas as bibliotecas incluírem algum tipo de ajuda aponte para o reconhecimento consensual da sua importância, o caso particular das bibliotecas digitais portuguesas, BND e Alma Mater, permanece apartado dos restantes, no sentido em que não implementam páginas de ajuda efetiva que permita uma utilização adequada e eficaz, para além daquela que a intuição do utilizador permite.

Pode considerar-se que todas as bibliotecas possuem alertas, na medida em que, em regra, todas exibem, na página de acolhimento, as últimas obras adicionadas. Contudo, este sistema, por si só, é pouco fiável e só tem alguma eficácia se forem adicionadas poucas obras. A Gallica, o IA e a BVH possuem o sistema de alertas por RSS. Na Gallica e na BVH, o acesso ao RSS é feito na página de ajuda, enquanto, no IA, é feito na página da pesquisa avançada. A Gallica permite a personalização dos conteúdos acerca dos quais o utilizador pretende ser notificado, criando alertas para determinado assunto ou tipo de obra ou mesmo autor. O IA também autoriza a individualização dos alertas, de forma a que o utilizador possa ser avisado sempre que uma obra do seu domínio de interesse seja colocada em linha. O sistema de RSS do IA abrange os respetivos fóruns e blogue. No entanto, a construção dos alertas no IA personalizados implica algum conhecimento de construção das *query* (consultas), o que torna o processo pouco amigável e intuitivo para o utilizador comum. A Gallica não possibilita uma caracterização tão detalhada quanto a do IA, mas, em contrapartida e, em parte, por isso mesmo, torna-se mais fácil de operar. Na BVH, os alertas são mais generalistas que os anteriores, no sentido em que, além de noticiar as obras adicionadas ou

alteradas e de informar acerca da digitalização e colocação em linha, também reportam eventos relacionados com a temática da biblioteca, com ofertas de emprego e eventos.

A opção de criação de um espaço pessoal permite ao leitor elaborar uma biblioteca digital de acordo com as suas preferências ou com o domínio da investigação em curso, permitindo criar múltiplos perfis e bibliografias específicas e autónomas, tornando-se uma ferramenta particularmente útil em contexto académico. Esta função está disponível e operacional na Gallica e no IA. Também se encontra disponível na Alma Mater, mas, aqui, as funcionalidades são mais limitadas, sobretudo, no que respeita ao acesso, não permitindo a entrada na área pessoal com muita regularidade, e também por não possibilitar a criação de várias listas.

Das bibliotecas analisadas, a Gallica é a única a disponibilizar uma *app* para dispositivos móveis, na App Store (Apple) ou no Google Play (Android), alargando a utilização dos conteúdos para lá dos contextos académico ou científico e potenciando o uso para fins lúdicos ou didáticos, em sentido lato. No mesmo sentido, o formato ePUB, que é especialmente adaptado à leitura em dispositivos móveis, além de ser um formato aberto, é facultado na Gallica e no IA.

Todas as bibliotecas têm vindo a aplicar a transcrição dos documentos aos conteúdos digitalizados, mas esta não é uma prática sistemática e a qualidade da transcrição é bastante variável, quer em função da tecnologia utilizada, mais limitada nas primeiras edições digitais, quer pelos aspetos formais inerentes à obra, como a tipografia do texto e a qualidade e estado de conservação da obra. A Gallica indica o nível de correção, em percentagem, da conversão da imagem em texto. No entanto, apesar da

transcrição dos documentos proporcionar mais um elemento para a eficácia das bibliotecas digitais, não podemos esquecer

[...] (for instance, digital library projects where the transcription serves only to provide searchable full text to accompany a page image). But for projects in which the TEI functions as a representation of scholarly work, the accommodation of nuanced intellectual expression is clearly crucial. (Flanders, 2012, p. 70)

À exceção do Perseus, todas as bibliotecas digitais analisadas permitem efetuar a transferência das obras em formato PDF. O IA é a biblioteca digital que oferece uma maior variedade de formatos: ABBYY; DAISY; EPUB; TEXT; KINDLE; PDF; JP2 e TORRENT. A Gallica disponibiliza a transferência nos formatos PDF, JPG, ePUB e texto. A BND disponibiliza-a em PDF e JPG e a BVH e a Alma Mater apenas em PDF.

A informação estatística sobre os conteúdos e sobre os utilizadores é essencial para legitimar o investimento na biblioteca digital e para motivar a associação de novos parceiros, como bibliotecas, centros de estudo e, também, novos investidores, ao projeto. Torna-se, portanto, crucial a disponibilização em linha de dados atualizados, que evidenciem a dinâmica da biblioteca. O IA atualiza diariamente os seus dados estatísticos, tanto em relação aos conteúdos adicionados, como à sua utilização, mostrando a respetiva evolução ao longo do tempo. A Gallica disponibiliza dados estatísticos detalhados no relatório anual, apresentado pela BnF, o que significa, no limite, o desfasamento de um ano face à realidade. No entanto, no sítio eletrónico da biblioteca digital, são disponibilizados regularmente dados estatísticos relativos à colocação de obras digitalizadas em linha.

Os gestores de referências bibliográficas são ferramentas informáticas cada vez mais utilizadas pelos investigadores para organizarem e reutilizarem as referências bibliográficas das leituras efetuadas. Por esse motivo, é pertinente e relevante que as bibliotecas digitais sejam configuradas de forma a permitir a exportação direta dos elementos bibliográficos da obra para os gestores de referências. À exceção da Gallica, as bibliotecas estabelecem ligações a gestores de referências, mas esta ligação é geralmente complexa. A Biblioteca Nacional Digital, a Alma Mater e o IA permitem a exportação direta para o Mendley. A Alma Mater é a única a colocar a opção de exportação dos dados bibliográficos junto à obra, sendo, por isso, mais facilmente identificada. Na BND é necessário aceder ao catálogo da BN e exportar a informação em formato RIS. No IA, esta opção está disponível para as obras com o registo bibliográfico no WorldCat, ou através da pesquisa na Open Library, onde são disponibilizadas as referências bibliográficas em formato RDF, JSON, OPDS e, também, num código compatível com a Wikipedia.

As bibliotecas digitais portuguesas analisadas não investem de forma expressiva na divulgação (Tabela 7), não possuem blogue, nem boletim de notícias, nem uma página própria nas redes sociais. A BND é divulgada, de forma esporádica, na página da BN no Facebook. A Alma Mater utiliza um espaço de difusão na página do Facebook das Bibliotecas da Universidade de Coimbra<sup>89</sup>. As restantes bibliotecas analisadas, além de manterem página própria no Facebook, estão representadas em diversas redes sociais. A Gallica, possui blogue e tem páginas no Facebook e no Google+ e contas no Twitter, Pinterest e Flickr e, através, da BnF, assegura a sua presença no LinkedIn e no Youtube.

---

<sup>89</sup> Disponível em:

<https://www.facebook.com/Bibliotecas-da-Universidade-de-Coimbra-116613321697537/>

Tabela 7 – Divulgação nas bibliotecas digitais

	<b>BND</b>	<b>Alma</b>	<b>Gallica</b>	<b>IA</b>	<b>BVH</b>	<b>Perseus</b>
Blogue			✓	✓	✓	
FB	BNP	UC	✓	✓		✓
Twitter		UC	✓	✓		✓
Boletim			✓		✓	
Pinterest			✓	✓	✓	
Instagram		UC		✓	✓	
Linkedin		UC	BnF	✓	✓	
Flickr		UC	✓	✓	✓	
Youtube		UC	BnF	✓	✓	
Google+		UC	✓			

O IA possui blogue, boletim informativo e contas no Twitter, no Pinterest, no Flickr, no LinkedIn e no Youtube. A BVH tem blogue, boletim informativo e contas no Pinterest, no Instagram, no Flickr, no LinkedIn e no Youtube. O Perseus tem uma página no Facebook e uma conta no Twitter e faz a divulgação em congressos académicos.

Os blogues, além de divulgarem notícias acerca dos conteúdos disponibilizados em linha, abrem a discussão acerca das problemáticas relacionadas com a colocação em linha das obras digitalizadas, técnicas e soluções. Os boletins informativos seguem a mesma orientação. A presença nas redes sociais centra-se na divulgação das coleções e de eventos. O Youtube é essencialmente utilizado para divulgar tutoriais e vídeos demonstrativos da funcionalidade de equipamentos, ou para difundir os resultados alcançados.

As bibliotecas digitais e, em particular, as BDLA têm vindo a atualizar-se para acompanhar as evoluções tecnológicas e, com isso, melhorar a capacidade de resposta às crescentes exigências dos seus utilizadores. No entanto, é perceptível que as atuais bibliotecas digitais ainda não refletem, nem otimizam, a totalidade das funcionalidades

que poderiam ter. Por outro lado, a informação técnica associada à criação da biblioteca digital nem sempre está disponível ou atualizada.

A maioria das bibliotecas analisadas já implementou algum tipo de ajuda aos utilizadores e disponibiliza, pelo menos, a ajuda contextual interativa. No entanto, e mesmo que esta seja essencial, não é suficiente nalgumas operações mais específicas, sendo necessário fornecer uma orientação complementar, através de tutoriais ou páginas de ajuda mais completas e estruturadas para uma utilização avançada dos recursos da biblioteca. Da mesma forma, o sistema de avisos, ainda que seja uma prática cada vez mais comum, não foi implementado no Perseus e nas duas bibliotecas digitais portuguesas analisadas, enquanto que as restantes não exploram todas as suas potencialidades. Além de notificar as adições de novos conteúdos, o sistema deve facultar a hipótese da personalização e seleccionar, ou priorizar, os alertas em função das preferências e da investigação de cada utilizador.

Em geral, a pesquisa nas bibliotecas digitais apresenta fragilidades e pode ser melhorada. Os metadados associados às obras digitalizadas são exíguos e de pouca qualidade, o que aumenta o índice de resultados indesejados. Por outro lado, a maioria dos sistemas de pesquisa exige a palavra escrita na grafia exata, o que, no caso do livro antigo, em que são frequentes grafias diversas, abreviaturas e erros, condiciona os resultados, sem que os metadados consigam compensar esta situação. Excetua-se, neste aspeto, a pesquisa no Perseus que proporciona a pesquisa através de termos truncados. A pesquisa facetada apenas é utilizada na Alma Mater, mas a lista de termos sugerida mostra-se insuficiente para incrementar a consulta, não permitindo a conjugação com os elementos do registo bibliográfico da obra, ao limitar

a aplicação de refinadores à navegação e não à pesquisa. A web semântica poderá trazer melhorias substanciais à pesquisa, mas implica que, antes, se resolvam os problemas decorrentes da proliferação de ontologias.

Em todas as bibliotecas analisadas, o sistema de leitura em linha permite a pesquisa dentro do texto, desde que tenha sido efetuada a transcrição da obra, mas não possibilita a criação de hiperligações, ou qualquer outra relação dinâmica, entre obras diferentes, partes de obras ou autores dentro ou fora da respetiva biblioteca digital. Além disso, não viabiliza procedimentos de leitura e de anotação colaborativa ou a marcação de múltiplas páginas de interesse.

O Perseus e a BVH, de forma consistente, e o IA, ainda de forma incipiente, utilizam o TEI como sistema de marcação do texto, o que facilita a utilização e reutilização da informação. A marcação com o TEI permite que a obra seja processada em diferentes versões para dispositivos móveis, incluindo *ebooks readers*, ou para exportação noutros formatos, como o HTML ou PDF. Além disso, o conteúdo da obra pode ser processado digitalmente por outros aplicativos, efetuando estudos de contexto, contabilizando elementos existentes no texto, ou estabelecendo relações semânticas, ou outras entre eles, de forma a gerar novo conhecimento.

À exceção do Perseus, que apenas faculta a leitura em linha ou a exportação em XML, as restantes bibliotecas digitais permitem a transferência das obras digitalizadas, pelo menos em PDF, para os computadores pessoais, facilitando a respetiva leitura sem estar ligado à Internet. A disponibilização de conteúdos para leitura em dispositivos móveis ainda é pouco expressiva, e o mesmo acontece em relação ao formato ePUB.

As bibliotecas digitais são diferentes entre si, quer em termos formais, quer de conteúdo. A pesquisa, a apresentação dos resultados, as funcionalidades divergem de biblioteca para biblioteca e, mesmo, dentro da mesma biblioteca, onde as obras não são apresentadas de forma homogénea, variando a qualidade das imagens, e do texto, a existência de sumários ou de outros marcadores e a navegação no interior da obra. Se o denominador comum é o facto de o conteúdo das obras estar disponível, a forma de o disponibilizar varia, tal como também são diversificados os moldes da consulta em linha ou os formatos de transferência dos ficheiros. Outro aspeto comum é a prevalência do conteúdo sobre a materialidade da obra. Os aspetos materiais e formais são negligenciados, no sentido em que a digitalização inibe a perceção das dimensões, altura, largura, espessura e não respeita a integralidade das encadernações, das lombadas, dos cortes ou das sobrecapas.

Tabela 8 - Bibliotecas Digitais

	<b>BND</b>	<b>Alma</b>	<b>Gallica</b>	<b>IA</b>	<b>BVH</b>	<b>Perseus</b>
Facilidade de utilização	Média	Alta	Alta	Média	Média	Média
Utilidade	Alta	Alta	Alta	Alta	Alta	Alta
Performance	Baixa	Baixa	Alta	Média	Média	Média

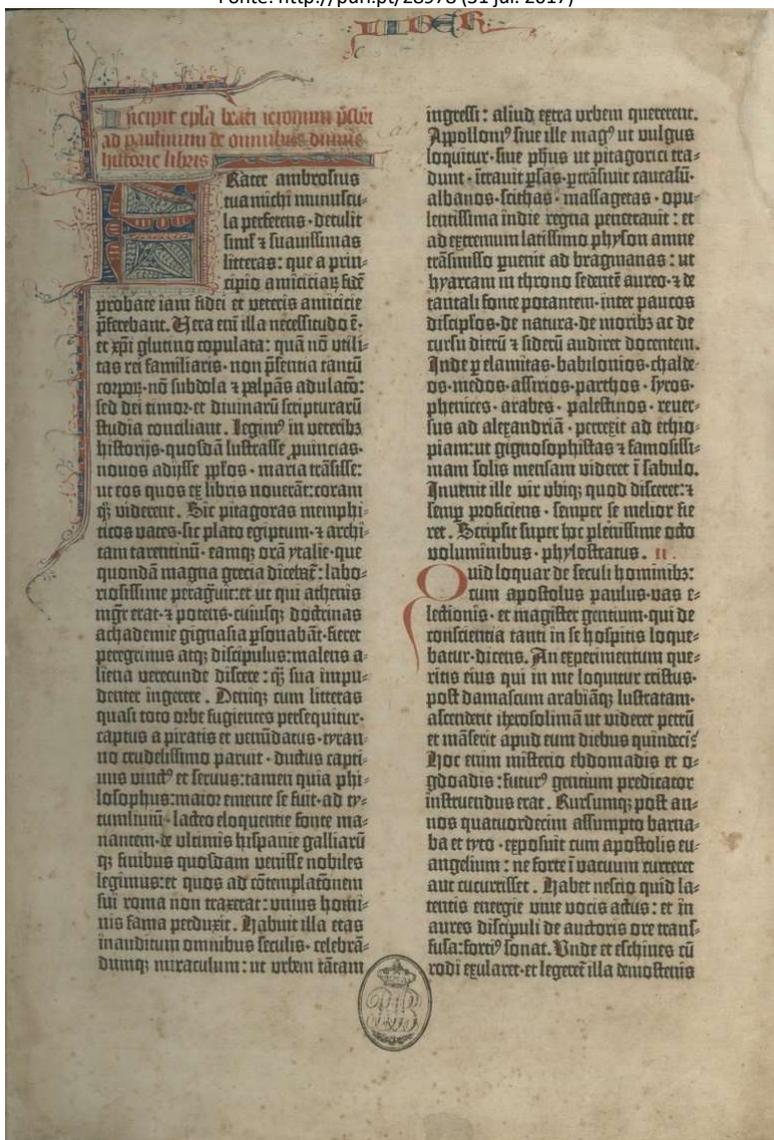
Todas as bibliotecas digitais analisadas disponibilizam conteúdos relevantes (Vd. Tabela 8), pelo que apresentam um elevado nível de utilidade. Entre as várias bibliotecas, a Alma Mater e a Gallica são as mais intuitivas e fáceis de utilizar. Em contrapartida, o desempenho (performance) que se afigura como o mais eficaz é o da Gallica.

## 5. Incunábulos e livro antigo

O primeiro livro impresso com caracteres móveis é a *Bíblia*, saída da oficina de Gutenberg, na Mogúncia, ca 1454-55, conhecida como a Bíblia das 42 linhas, de que existe um exemplar na Biblioteca Nacional de Portugal (INC. 305-306)<sup>90</sup> (Figura 29).

Figura 29 – *Bíblia*, dita das 42 linhas (Johann Gutenberg), ca 1454-1455: *incipit*.

Fonte: <http://purl.pt/28978> (31 jul. 2017)



<sup>90</sup> Disponível em <http://purl.pt/28978/>

Aos livros impressos com caracteres móveis entre 1450 e 1500, dá-se o nome de incunábulos (Faria & Pericão, 2008, p. 652). Do ponto de vista estrutural, as primeiras obras impressas não diferiam substancialmente dos manuscritos, dado que utilizavam o mesmo tipo de letra gótica e o texto era formatado em duas colunas, por vezes, com as capitais decoradas ou iluminadas, não possuem folha de rosto e iniciam no *incipit*. Em 1501, os livros ainda não se produziam de modo muito diferente, continuando a aparecer obras com características idênticas às dos incunábulos. A evolução do livro impresso foi gradual e irregular no espaço e no tempo. Na transição entre os séculos XVIII e XIX, ainda no contexto da Revolução Industrial, o aparecimento da litografia e da invenção da máquina plana, que produzia a folha de papel em contínuo, dando origem às rotativas, origina uma rutura no processo de fabrico do livro tal como se efetuara durante quatro séculos e meio.

### 5.1. Materialidade dos Incunábulos e livro antigo

**Incunábulos:** obras impressas com caracteres móveis entre 1450 (data de impressão da *Bíblia* de Gutenberg) e 1500.

**Livro antigo:** obras impressas com caracteres móveis entre 1501 (ano de publicação, por Aldo Manuzio, da obra *Vergilius [Bucolica...]*, integralmente impressa em itálico e formato *in-oitavo*, considerando-se o protótipo do livro moderno) e 1800.

Em **1801**, com a invenção da prensa rotativa para impressão em papel contínuo marca a industrialização da imprensa. Em rigor, o primeiro registo patenteado para a prensa rotativa surge em 1790 por William Nicholson, mas como as datas referidas são convencionais, optou-se por uma data que corresponde ao final século XVIII e princípio de XIX.

Para o estudo dos primeiros livros impressos, considera-se a janela temporal de 1450 a 1800 (Barbier, 2012; Febvre & Martin, 2000; McMurtrie, 1969), que engloba os incunábulos e o livro antigo, o qual se passa a designar genericamente por livro antigo.

Por definição, designa-se por livro antigo todo aquele que foi impresso entre 1501 e 1800 (Faria & Pericão, 2008, p. 764). Assim, todo o livro antigo que integra os acervos patrimoniais das bibliotecas encontra-se em domínio público e é possível disponibilizá-lo na íntegra na web.

O estudo dos incunábulos e do livro antigo faz-se em duas vertentes, das quais, uma é dirigida ao conteúdo das obras (Anselmo, 1996; Castillo Gómez, 2004; Martins, 2007), analisando quem são os autores, que obras foram impressas, por quem, em que línguas, em que países, enquanto outra foca a materialidade dos livros (Barbier, 2006, 2012; Febvre & Martin, 2000; McMurtrie, 1969), ou seja, que tipos de letras, ilustrações e partes constituintes, foram utilizados em cada época, bem como as características e as inovações que cada impressor veiculava nas obras que imprimia.

O acesso em linha ao livro antigo, assim como a outras obras digitalizadas, é realizado através de bibliotecas digitais. De resto, o livro antigo é presença obrigatória nas bibliotecas digitais patrimoniais, mas ainda não se encontrou um modelo que seja fácil de implementar e que sirva, tanto aos investigadores, como ao público em geral.

Os livros impressos entre 1450 e 1800 constituem os fundos antigos das bibliotecas, denominados (sobretudo no mundo anglo saxónico) por livros raros ou reservados. Como o nome indica, estas coleções têm acesso reservado e estão a despertar cada vez mais interesse, “o alargamento do espaço de investigação a países e continentes sem qualquer tradição no ramo, mostram à evidência que a História do Livro está cada vez mais presente na esfera das preocupações dominantes nas comunidades de Ciências Sociais e de Ciências Humanas” (Anselmo, 1996, p. 79), mas a história do livro

também é a “História das Ideias” (Anselmo, 1996, p. 81), motivo pelo qual o livro antigo continua a ser importante para as Humanidades.

As Humanidades Digitais, entendidas como a utilização do computador e das técnicas computacionais para rentabilizar o trabalho do homem, foram iniciadas com a informatização da obra de São Tomás de Aquino (vd. Cap. 2). Outros investigadores continuaram a dar uso ao computador para os projetos em Humanidades, mas continuaram a trabalhar com o livro: “By the 1960s, other researchers had begun to see the benefits of working with concordances” (Hockey, 2004).

A importância da utilização do computador para os trabalhos na área da linguística justificou a criação da revista *Computers and the Humanities* em 1966. “Humanists have been capturing, analysing, and presenting textual data in digital form for as long as there have been computers capable of processing alphanumeric symbols” (Deegan & Tanner, 2004). É sabido que as “Libraries provide access to original works, to correspondence and other commentary that helps contextualize those works, and to previous humanities commentaries” (Besser, 2004). As bibliotecas são essenciais para a investigação e a sua unidade basilar das bibliotecas são os livros. A “Biblioteca digital se haya convertido en el nombre genérico para designar este modelo de almacenamiento y difusión de documentos, ya sea a partir de la digitalización de objetos anteriores con diversos formatos (texto, audio, video, imágenes...)” (Lucía Megías, 2010, p. 373).

A digitalização facilitou a obtenção de cópias e a disseminação da informação, “We have seen that the primary aim for converting an image into a numeric or digital format is to facilitate copying, transmission and processing of the pictorial [or text]

**Aparar um livro** – cortar a cabeça (margem superior), pé (margem inferior) e margem (lateral) das folhas.

**Assinatura de caderno** – conjunto de letras, números e caracteres colocados na primeira página do caderno, ou em várias páginas, fora da mancha do texto, a fim de definir a sequência para a encadernação.

**Caderno** – conjunto de páginas de um livro, impressas numa única folha.

**Caldeirão ¶** – sinal em forma de C barrado, utilizado para destacar a informação.

**Corações ☞ ou camarão** – elemento tipográfico, utilizado para destacar a informação

**Desdobráveis** – folha de tamanho superior às restantes páginas da obra, que tem de ser dobrada.

**Elementos gráficos das obras:** Ilustrações, mapas, heráldica, esquemas existentes no interior das obras.

**Galerias** – escavação realizada por inseto, por norma num conjunto de folhas ou nas encadernações.

**Glosas** – Comentários, por norma em letra com corpo inferior ao restante texto.

**Lacunas** – espaço que ficou em branco.

(Vd. Faria & Pericão, 2008)

information” (Terras, 2008, p. 15). Os projetos à volta dos livros e das bibliotecas digitais têm vindo a conhecer grande incremento. “En los últimos años, los proyectos informáticos alrededor del libro y del texto, de su difusión en nuevos soportes y con originales posibilidades y herramientas se han multiplicado” (Lucía Megías, 2010, p. 372).

Os exemplares de incunábulos e livro antigo apresentam características materiais muito diversificadas, o que coloca questões, quer à digitalização, quer à colocação em linha. As definições para o formato, resolução, profundidade de cor e densidade das imagens digitais do livro têm vindo a estabelecer-se de forma empírica (Guerreiro, 2016; Guerreiro, Campos, & Lopes, 2009; Lee, 2001; Puglia, Reed, & Rhodes, 2004; Ross et al., 2002; Terras, 2008), mas, apesar disso, constituem um conjunto de regras aplicadas nas várias bibliotecas digitais, de forma mais ou menos coerente.

Tanto os incunábulos como o livro antigo são constituídos por conjunto de folhas unidas entre si, em sequência, e encadernadas em conjunto.

**Letras capituladas** – Letra maiúscula com corpo superior ao texto, por vezes com decoração.

**Margem interior** – Margem esquerda do reto e direita do verso, ou margem da costura do livro.

**Página de rosto ou folha de título** – primeira página da obra onde estão inscritos os elementos fundamentais da identificação.

**Página reto** – página da direita.

**Página verso** – página da esquerda.

**Pé de página** – parte inferior de uma página impressa.

**Reclamo** – primeira palavra impressa num caderno, é repetida no pé do caderno anterior, para facilitar a encadernação.

**Tarjas** – Cercadura, guarnição.

**Tinta ferrogálica** – feita à base de sulfato ferroso, muito durável, mas ao oxidar fica ácida e corrói o papel. Utilizada nas notas manuscritas.

(Vd. Faria & Pericão, 2008)

Ambos já possuíam dimensões definidas, que eram estipuladas a partir do tamanho da página o fólho, que varia entre 33 cm e +40 cm (Dias, 1994, p. 34), mas podendo ter alguma variação. Conforme as dobras a que a folha era sujeita, assim se define os formatos derivados:

- In-folio ou in-2º – página dobrada ao meio, cerca de 34 cm de altura, origina quatro páginas;
- In quarto ou In-4º – página dobrada duas vezes, cerca de 26 cm de altura, origina oito páginas;
- In oitavo ou In-8º – página dobrada três vezes, cerca de 18 cm de altura, origina dezasseis páginas.

Ao encadernar as obras, estas eram aparadas, com uma dupla finalidade: por um lado, para abrir os cadernos e, por outro, para a altura da obra se adaptar à estante do possuidor. Na

prática, isto significa que é possível encontrar a mesma obra, saída da mesma oficina de impressor e na mesma data, mas com dimensões diferentes.

As dimensões das obras também podem dificultar a digitalização e a visualização. Se a obra é muito grande, não se consegue ver a página inteira no ecrã e, simultaneamente,

ler o conteúdo, facto que pode tornar a navegação na obra confusa. As obras muito pequenas, com altura inferior ao in-8º, são difíceis de manusear e de digitalizar.

Uma das questões que se coloca à digitalização é a forma como é apresentada a sequência das páginas. Algumas obras não eram paginadas, ou apenas tinham numeração na página rosto (foliação) e não no verso, e havendo erros na encadernação da obra, estes tornam-se mais difíceis de detetar. Outras obras apenas indicam a sequência das páginas por assinaturas. Neste caso, é necessário estar na presença da obra para verificar se a sequência é a correta. Para facilitar a encadernação e o agrupamento sequencial das páginas das obras, os impressores começaram a colocar os reclusos. Neste caso, há que verificar, página a página, se a obra está corretamente encadernada. A apresentação do texto em duas colunas numeradas pode acarretar trabalho extra na verificação.

“Digitization, done successfully, can add value to both the end user and institution” (Terras, 2008, p. 101). Contudo, a digitalização e colocação em linha de ilustrações, mapas e desdobráveis evidenciam alguns problemas. Os desdobráveis não aparecem junto do texto que lhes corresponde, o mesmo acontecendo com ilustrações e mapas. A estes últimos, acresce o facto de quando ocupam duas páginas na íntegra dão pouca leitura junto da margem interior. Um outro tipo de ilustração, muito comum nas obras de astronomia e anatomia, é que possuem elementos móveis e/ou sobrepostos (*pop-up*). Por exemplo, a obra *Cosmographicus liber Petri Apiani mathematici studiose collectus*<sup>91</sup> ou *Kleiner welt spiegel, das ist, abbildung göttlicher schöpfung an dess menschen leib: mit beygesetzter schriftlicher Erklärung: so wo zu Gottes Weissheit: als*

---

<sup>91</sup> Disponível em <https://archive.org/stream/Cosmographicusl00Apia#page/n35/mode/2up>

*dess menschen selbst erkandtnuss dienend* de Remmelin, J., Remmelin, J., Kilian, L., Schultes, J., & Görlin, J. (1661)<sup>92</sup>, estes elementos móveis, têm funções didáticas, sendo necessário manuseá-los para o bom entendimento da obra. Uma solução implementada na obra *Dess Menschen Circkel und Lauff [...]*<sup>93</sup> foi apresentar as imagens separadas das várias peças e uma imagem de conjunto. Muito embora esta obra esteticamente seja muito interessante, a sua funcionalidade perdeu-se.

Em síntese, as características particulares do livro antigo que condicionam a sua digitalização e verificação e edição digital são:

- A paginação:
  - Obras sem numeração nos fólios;
  - Obras apenas com numeração nas páginas rosto (foliação);
  - Obras com as sequências das páginas dadas por assinaturas;
  - Obras com a paginação errada;
  - Obras com blocos de sequências em numeração árabe, romana ou sem numeração;
- Obras com o texto em duas colunas numeradas;
- Sequência da obra dada por reclamo;
- Obras com as margens, sobretudo a interior, demasiado pequenas ou truncadas;
- Obras cuja leitura depende do manuseamento de elementos móveis (*pop-up*) inseridos nas páginas;
- Obras com desdobráveis;

---

<sup>92</sup> Disponível em <http://www.columbia.edu/cgi-bin/cul/resolve?clio11497246>

<sup>93</sup> Disponível em <https://www.wdl.org/en/item/15131/>

**Encadernação digital:** Processo através do qual se agrupam as imagens digitais de uma obra, ordenando-as em conformidade com o original ou por outra ordem devidamente assinalada, bem como sumário ou outros apontadores de conteúdo e os respetivos metadados, de forma a obter uma representação digital coerente e fidedigna do original em suporte papel.

**Sumário:** Lista sequencial de títulos e subtítulos dos conteúdos (partes, capítulos, subcapítulos) pela ordem apresentada na obra e com indicação das páginas ou colunas onde iniciam.

**Sumário nível 1:** limitam-se à indicação da estrutura geral da obra.

**Sumário nível 2:** apresentam a estrutura da obra, em grandes divisões, geralmente correspondentes às partes ou livros.

**Sumário nível 3:** acompanham a estrutura da obra, com a indicação das partes, livros, capítulos, secções.

**Thumbnail:** miniatura da imagem, versão reduzida da imagem, cerca de 140px de altura.

- Obras com lombadas superiores a 10 cm;
- Encadernação da obra apertada, não permitindo grandes ângulos de abertura;
- Dimensão dos originais, demasiado grandes ou pequenos.

A aplicação dos programas de OCR para obter a transcrição do texto existente numa obra é tanto mais eficiente quanto menos problemas as imagens apresentarem. São fontes de erro: a transparência do papel, pois o texto do verso é identificado como caracter; o desgaste dos caracteres usados na impressão, dado que o texto se torna menos preciso; letras capituladas e muitas vezes decoradas; notas impressas nas margens laterais das obras, com um tipo de letra pequeno, dá pouca perceção das palavras; as glosas requerem um cuidado especial na marcação da área da imagem a transcrever; os livros usados com notas manuscritas e sublinhados introduzem ruído nas imagens, dificultando um resultado eficaz; lacunas e galerias. As manchas e a tinta ferrogália também provocam ruído nas imagens. Os caldeirões, corações e outros elementos

decorativos, podem ser problemáticos; porém, com os atuais programas de OCR, é possível personalizar os dicionários e atribuir uma sinalética específica a estes caracteres usados para identificar o parágrafo ou o início de capítulo.

A transcrição através de programas de OCR (*Optical Character Recognition*, ou Reconhecimento Ótico de Caracteres) é dificultada por:

- Textos em duas colunas;
- Textos a duas ou mais cores;
- Obras com papel pouco espesso, tornando-se transparente;
- Inclusão de letras capituladas, tarjas, caldeirões, corações e outros elementos decorativos ou iconográficos;
- Tabelas;
- Notas impressas na margem das obras;
- Livros usados, com o texto sublinhado e anotações manuscritas;
- Existência de lacunas e galerias;
- Borrões de tinta a tapar o texto ou manchas de tinta ferrogálica;
- Glosas;
- Manchas de humidade.

A digitalização e correspondente disponibilização em linha requer alguma atenção.

The primary source materials for humanists come in many data forms, and the digital capture of that data needs to be carefully considered in relation to its formats and the materials that form the substrates upon which it is carried.

[...] This is an interesting point, as many digitization projects aim to present

some kind of "true" representation of the original, and some materials can get closer to this goal than others. (Deegan & Tanner, 2004)

Numa tentativa de aproximar as versões digitais ao respetivo objeto físico, foram criadas versões digitais das obras em que se simula o folhear de página de um livro real, o que não traz nenhuma mais valia à versão digital; já a possibilidade de ver as miniaturas das páginas da obra (*thumbnail*) facilita a navegação e a localização dos elementos gráficos. Também a construção de sumários com as secções e partes das obras, ou a transcrição do texto, permitem uma localização mais facilitada do conteúdo. Existem vantagens na disponibilização de cópias digitais das obras, pelo menos na facilidade de acesso.

## 5.2. Os incunábulos e livro antigo na Biblioteca Nacional Digital

A BND<sup>94</sup> disponibilizava em linha, em julho de 2016, 3.979 obras referentes a livro antigo, no entanto, a atividade de digitalização e edição digital tem continuado e, um ano depois, em julho de 2017, contabilizava-se mais 897, perfazendo um total de 4.876 títulos disponíveis em linha<sup>95</sup> (Tabela 9).

Tabela 9 – Número de obras livro antigo na BND (jul. 2016; jul. 2017)

Data edição	N.º de obras	
	2016	2017
Séc. XV	107	122
Séc. XVI	1 074	1 124
Séc. XVII	1 429	1 830
Séc. XVIII	1 369	1 800
Total	3 979	4 876

---

<sup>94</sup> Disponível em <http://purl.pt>

<sup>95</sup> Dados atualizados em 30/07/2017

O universo analisado (Vd. Anexo XII) corresponde apenas às obras digitalizadas, obviamente muito inferior ao número total do acervo de livro antigo da BNP<sup>96</sup>.

No conjunto do acervo de livro antigo na BND, limitamos a amostra da análise ao conjunto de 265 obras publicadas entre 1451 e 1558, abrangendo os primeiros 100 anos da imprensa com caracteres móveis na BNP. A opção por estender o limite final do arco cronológico para lá de 1550, prende-se com a intenção de obter uma amostra mais representativa, dado que, entre 1551 e 1558, há 86 obras, o que corresponde a 32,4%, ou cerca de 1/3 da amostra.

No universo das obras digitalizadas e disponibilizadas em linha pela BND, confirma-se que as primeiras impressões seguiam o modelo dos manuscritos. Sem folha de rosto, umas obras continuam a iniciar com a fórmula dos manuscritos *incipit* (Figura 28), enquanto que outras surgem já com página de título, à qual vão sendo acrescentados elementos descritivos da obra (autor, título, data, impressor) e elementos decorativos ou com iconografia alusiva (vd. *Opuscula*, de Virgílio<sup>97</sup>), mas não de forma sistemática, coexistindo vários modelos. O colofão continua a ser o elemento privilegiado de recolha de informação sobre a obra. Há, nestas obras, aspetos particulares que influenciam a edição digital e devem ser levados em conta, caso a caso. Os capítulos aparecem evidenciados através de caldeirões ou corações (vd. *Manipulus curatorum*, de Guido de Monte Roteri<sup>98</sup>), o que perturba a leitura e a transcrição do texto,

---

<sup>96</sup> É difícil encontrar dados exatos relativos à existência de livro antigo na BNP. A título meramente informativo, refira-se que, segundo os dados disponíveis no sítio eletrónico da biblioteca, existem, na secção Reservados, 1.500 incunábulo e mais de 30.000 espécies de livro impresso. Fica fora destes números, o Fundo Geral, com cerca de 3 milhões de espécies a partir do século XVI e, portanto, onde haverá ainda um número substancial de livro antigo.

<sup>97</sup> Disponível em <http://purl.pt/23359>

<sup>98</sup> Disponível em <http://purl.pt/23037>

nomeadamente através da aplicação de programas de reconhecimento ótico de caracteres. As ilustrações e mapas em páginas duplas ou em desdobrável dificultam a leitura dos elementos existentes na margem interior (vd. *Cosmographiae uniuersalis*, de Sebastian Münster<sup>99</sup>), ou exigem cuidados especiais na planificação da folha desdobrada.

A data impressa na página de título pode não ser a data efetiva de impressão da obra, dado que as matrizes da portada em madeira, feitas em determinada data, eram posteriormente reutilizadas noutras obras (vd. *Primera parte de las Sentencias*, de Fr. Manuel de São Carlos, de 1554<sup>100</sup>, e *Os Lusíadas*, de Luís de Camões, de 1572<sup>101</sup>), o que gera dificuldades de catalogação e, por conseguinte, de acesso. Em contrapartida, a maioria das obras apresentam índices, tabuadas ou index (vd. *Cronica geral*, de Sabelico<sup>102</sup>), de grande utilidade para a edição digital.

A coleção de livro antigo é composta por monografias, ou livros, “una publicación impresa no periódica que consta como mínimo de 49 páginas, sin contar las de cubierta, editada en el país y puesta a disposición del público” (UNESCO, 1965, p. 150) e folhetos, ou seja, “la publicación impresa no periódica que consta de 5 a 48 páginas sin contar las de cubierta, impresa, editada en el país y puesta a disposición del público” (UNESCO, 1965, p. 150).

---

<sup>99</sup> Disponível em <http://purl.pt/13845/3/#/34>

<sup>100</sup> Disponível em <http://purl.pt/15188>

<sup>101</sup> Disponível em <http://purl.pt/1>

<sup>102</sup> Disponível em <http://purl.pt/14546>

No universo do livro antigo da BNP analisado, as monografias são a maioria, com 72%, em relação aos folhetos, com 28% (Tabela 10).

Tabela 10 – BND: livro antigo, por monografias e folhetos

<b>Tipo</b>	<b>N.º exemplares</b>	<b>%</b>
Folhetos	103	28%
Monografias	265	72%

No conjunto de monografias, a maioria é em português, 41,5% (42 obras); em seguida, as línguas mais representadas são o latim e o espanhol, com 34,7% e 15,8%, respetivamente (Tabela 11).

Tabela 11 – BND: livro antigo, por língua

<b>Língua</b>	<b>Número de obras</b>	<b>%</b>
Francês	8	3,0%
Hebreu	4	1,5%
Italiano	5	1,9%
Latim	92	34,7%
Português	110	41,5%
Espanhol	42	15,8%
Bilingues	4	1,5%

As línguas com menos exemplares são o francês, 3,0% (8 obras), o italiano, 1,9% (5 obras) e o hebreu, 1,5% (4). Há quatro obras bilingues em heb/lat, lat/spa, por/lat e spa/por. A identificação das línguas foi feita a partir da informação recolhida no campo 101\$c do registo UNIMARC. Registam-se quarenta e um locais de edição diferentes. A identificação dos locais foi feita através do campo 210\$a do registo UNIMARC.

No conjunto das mais representadas (Tabela 12), Lisboa surge em primeiro lugar, com 45,11%, seguida de Coimbra, com 17,29%, Paris, com 6,77%, Sevilha e Veneza, cada uma com 3,76%, e, por fim, Évora com 3,38%.

Tabela 12 – BND: livro antigo, por local de edição das obras

<b>Local de impressão</b>	<b>Número de obras</b>	<b>%</b>
Coimbra	46	17,29%
Évora	9	3,38%
Lisboa	120	45,11%
Paris	18	6,77%
Sevilha	10	3,76%
Veneza	10	3,76%
Outros	53	19,92%

A seguir a este grupo, há uma grande dispersão de locais de publicação (Tabela 13).

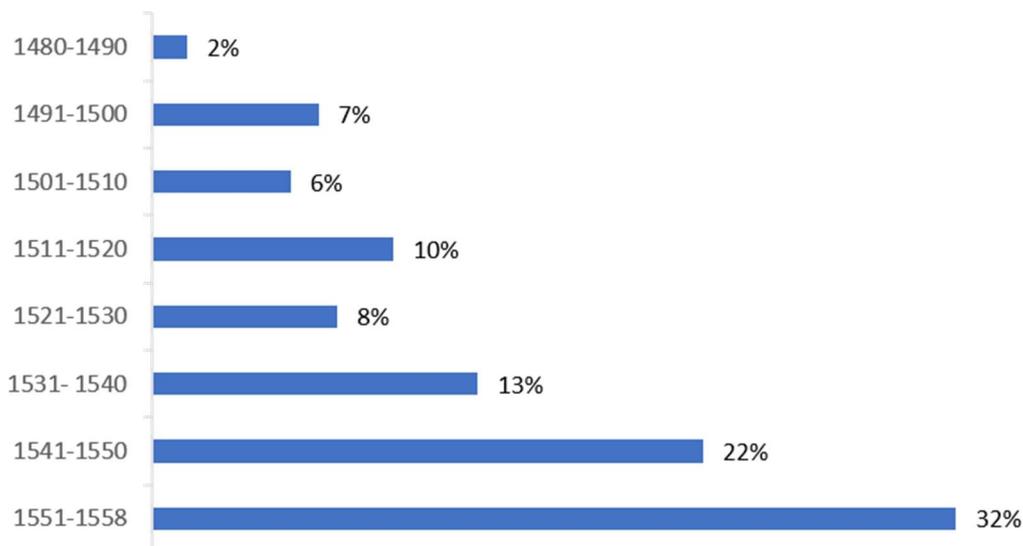
Neste conjunto, a representação dos vários locais de impressão é quase residual: Basileia possui 1,88% dos livros (5 exemplares); Braga com 1,50% (4 exemplares); Porto Roma e Salamanca com 1,13% (3 exemplares, cada); Ferrara, Leiria, Lyon, Saragoça, Setúbal e Valladolid, com 0,75% (2 obras, cada). Os restantes locais estão representados, cada um, por um exemplar. No total estão disponíveis cerca de 77,4% de obras impressas em Portugal. As obras sem o local de impressão (*sine loco*: s.l.) identificado são na totalidade 7 exemplares.

Tabela 13 – BND: livro antigo, por outros locais de edição

<b>Local Publicação</b>	<b>Número de obras</b>	<b>%</b>
Almeirim	1	0,38%
Anvers, Bélgica	1	0,38%
Barcelona	1	0,38%
Basileia	5	1,88%
Braga	4	1,50%
Burgos	1	0,38%
Chaves	1	0,38%
Évora, Lisboa	1	0,38%
Faro	1	0,38%
Ferrara	2	0,75%
Hispali	1	0,38%
Köln	1	0,38%
Leiria	2	0,75%
Leon	1	0,38%
Louanij	1	0,38%
Lyon	2	0,75%
Mainz	1	0,38%
Medina del Campo	1	0,38%
Mediolani	1	0,38%
Parma	1	0,38%
Porto	3	1,13%
Roma	3	1,13%
s.l.	7	2,63%
Salamanca	3	1,13%
Saragoça	1	0,38%
Setúbal	2	0,75%
Strasbourg	1	0,38%
Toledo	1	0,38%
Valladolid	2	0,75%
Zaragoça	1	0,38%

As obras analisadas concentram-se no arco cronológico entre 1480 e 1558 (Gráfico 1), dado que na BND, antes de 1480, há uma única obra de *ca* 1454-1455 (*Bíblia*, de Gutenberg).

Gráfico 1 – BND: livro antigo, por distribuição cronológica da data de edição (1480-1558)



A seguir a esta, há uma obra publicada em 1480 que se integrou no conjunto de 1480 a 1490, o único que compreende 11 anos. O último intervalo compreende os anos de 1551 a 1558, o último da amostra. As obras publicadas entre 1451 a 1500 correspondem a 9% (23) das obras analisadas. Na primeira metade do século XVI, os valores variam entre 7% (18 obras), na década de 1491 e 1500, e 22% (59), de 1541 a 1550. A década menos representada é a de 1501 a 1510, com 6% (15 obras). Os valores mais elevados registam-se nos 8 anos entre 1551 e 1558, com 19% (49), confirmando a tendência de crescimento das últimas décadas, sendo que em 1559 e 1650, anos já não contemplados na amostra, há, respetivamente, 12 e 19 exemplares.

A maioria das obras disponibilizadas (89%) não permite a pesquisa dentro do documento, porque não foi realizada a conversão das imagens digitalizadas em texto

(Tabela 14). Apenas 11% das obras foram transcritas, mas sem indicar o respetivo grau de erro.

Tabela 14 – BND: livro antigo, obras com e sem transcrição textual

<b>Transcrição do texto</b>	<b>%</b>	<b>Número de obras</b>
Sem	89,06%	236
Com	10,94%	29

Das obras com transcrição (Tabela 15), 62% são em português, 21% em espanhol e 17% em latim, não sendo a língua um impedimento para a aplicação dos programas OCR, nem para efetuar a respetiva transcrição dos documentos.

Tabela 15 – BND: livro antigo, obras a que foi realizada a transcrição, por idioma

<b>Língua</b>	<b>%</b>	<b>Número de obras</b>
Espanhol	21%	6
Latim	17%	5
Português	62%	18

A maioria das obras (72%) é disponibilizada em linha com sumário (Tabela 16), facilitando a navegação no respetivo conteúdo, havendo 28% sem sumário. No entanto, a qualidade destes sumários é muito variável, permitindo considerar três categorias, por ordem crescente de complexidade: nível 1, com 36%; nível 2, com 26%; e nível 3, com 28%.

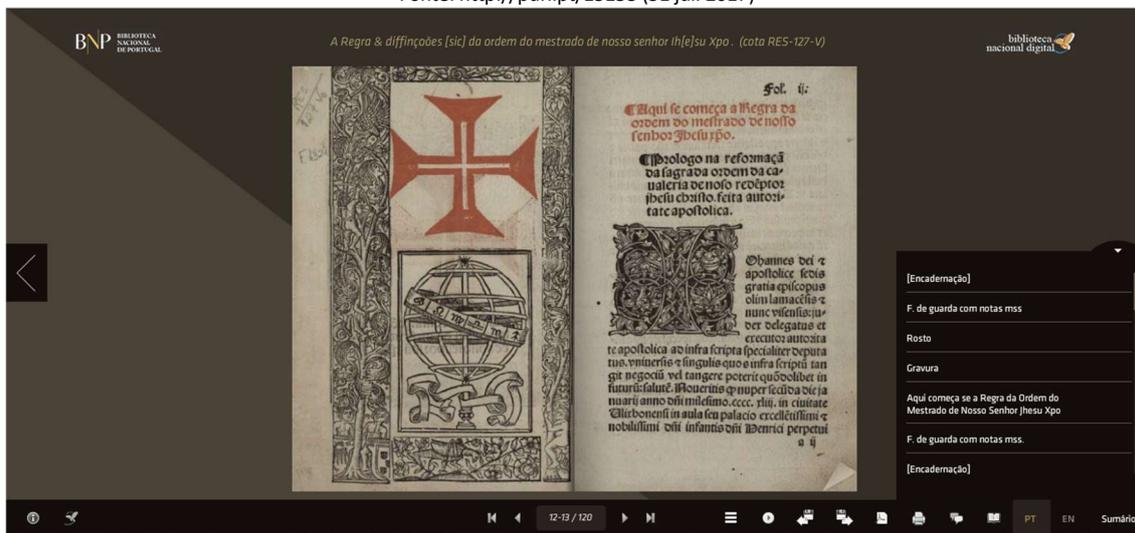
Tabela 16 – BND: livro antigo, obras com e sem sumário

<b>Sumário</b>	<b>%</b>	<b>Número de obras</b>
Sem sumário	28%	75
Com sumário	72%	190
Sumário nível 1	36%	97
Sumário nível 2	26%	66
Sumário nível 3	10%	27

Os sumários de nível 1 (vd. *A Regra & diffinções [sic]*, da Ordem de Cristo<sup>103</sup>, Figura 30), fornecem algumas informações esparsas, ou fazem o aglutinamento dos capítulos (*Vita Christi*, de Ludolfo de Saxónia<sup>104</sup>), sem identificar os conteúdos.

Figura 30 – *A Regra & diffinções [sic]*, da Ordem de Cristo: sumário de nível 1.

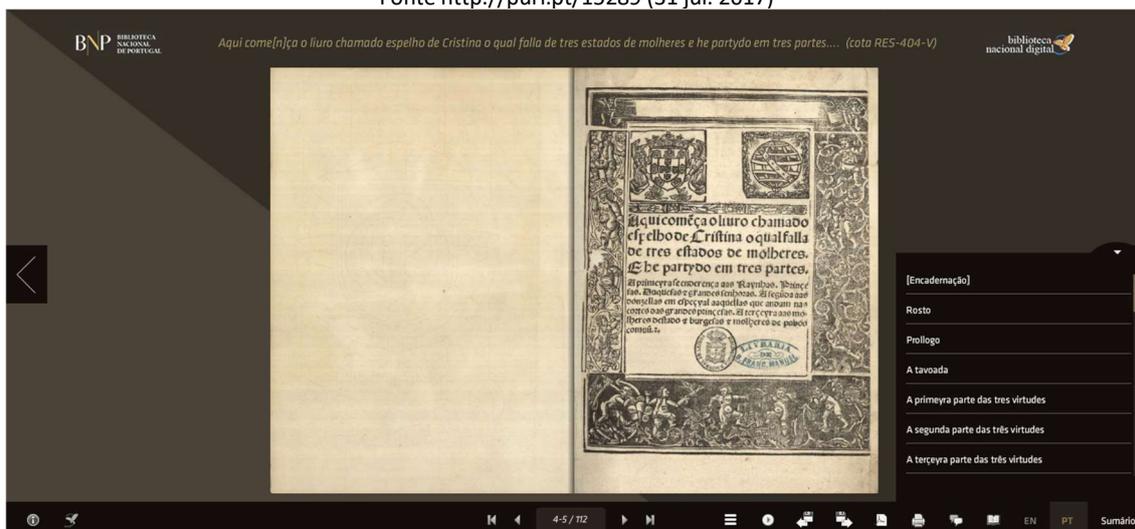
Fonte: <http://purl.pt/15153> (31 jul. 2017)



Os sumários de nível 2 (26%) (vd. *Aqui come[n]ça o liuro chamado espelho de Cristina*, de Christine de Pisan<sup>105</sup>, Figura 31).

Figura 31 – *Aqui come[n]ça o liuro chamado espelho de Cristina*: sumário de nível 2.

Fonte <http://purl.pt/15289> (31 jul. 2017)



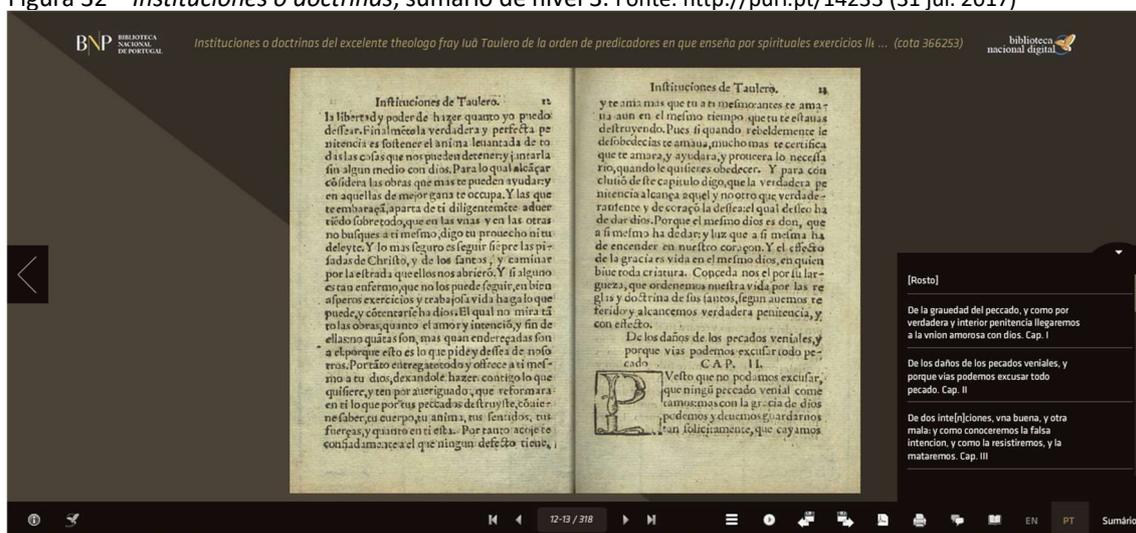
<sup>103</sup> Disponível em <http://purl.pt/15153>

<sup>104</sup> Disponível em <http://purl.pt/22010>

<sup>105</sup> Disponível em <http://purl.pt/15289>

Os sumários de nível 3 (10%) (vd. *Instituciones o doctrinas*, de Johann Tauler<sup>106</sup>, Figura 32), informam acerca do respetivo conteúdo.

Figura 32 – *Instituciones o doctrinas*, sumário de nível 3. Fonte: <http://purl.pt/14233> (31 jul. 2017)



Em regra, apenas é referido o conteúdo textual das obras. Nas obras que contenham ilustrações (vd. *Reg[ra] e Sstatut[us]*, da Ordem de Avis<sup>107</sup> ou *Ho liuro da regra & perfeçãõ da cõuersaçã dos monges*, de Santo Justiniano Lourenço<sup>108</sup>) ou mapas (vd. *Delle Nauigationi et Viaggi*, de Giovanni Battista Ramusio<sup>109</sup>), estes não são referidos no sumário.

O estilo dos sumários não é uniforme: uns fazem a transcrição do texto, respeitando a grafia da época, abrindo, ou não, as abreviaturas (vd. *A Regra & diffinçoões [sic]*, da Ordem de Cristo<sup>110</sup>, Figura 30; *Instituciones o doctrinas*, de Johann Tauler<sup>111</sup>, Figura 32); nas obras em latim (o *Missale monasticum secundum morem et ritum Casinensis*

<sup>106</sup> Disponível em <http://purl.pt/14233>

<sup>107</sup> Disponível em <http://purl.pt/15207>

<sup>108</sup> Disponível em <http://purl.pt/16678>

<sup>109</sup> Disponível em <http://purl.pt/21896>

<sup>110</sup> Disponível em <http://purl.pt/15153>

<sup>111</sup> Disponível em <http://purl.pt/14233>

*Congregationis*<sup>112</sup>), em particular nos livros religiosos, é frequente a tradução para português.

Não existe uma relação direta entre a existência, ou não, de sumário e a língua da obra (Tabela 17).

Tabela 17 – BND: livro antigo, obras com sumário, por língua

Língua	Número de obras	%
Latim	70	37%
Português	77	41%
Francês	5	3%
Espanhol	30	16%
Italiano	4	2%
Hebreu/Latim	1	1%

Embora a maioria das obras com sumário seja em português (por), com 41%, há obras com sumário em francês (fre), italiano (ita), latim (lat) e espanhol (spa) e uma obra bilingue em hebraico e latim (heb/lat), com sumário em latim (*Luah ha-diqduq = Tabula in grammaticen hebraeam*, de Nicolao Clenardo<sup>113</sup>)

As características físicas da obra têm impacto, quer na digitalização, quer nos procedimentos seguintes para a respetiva edição digital. Em particular, algumas características e sinais tipográficos, como as capitais ornamentadas, os caldeirões e os corações, os reclusos, as ornamentações e as notas marginais dificultam o reconhecimento ótico de caracteres, incrementando o nível de erro nas respetivas transcrições.

---

<sup>112</sup> Disponível em <http://purl.pt/24994>

<sup>113</sup> Disponível em <http://purl.pt/25773>

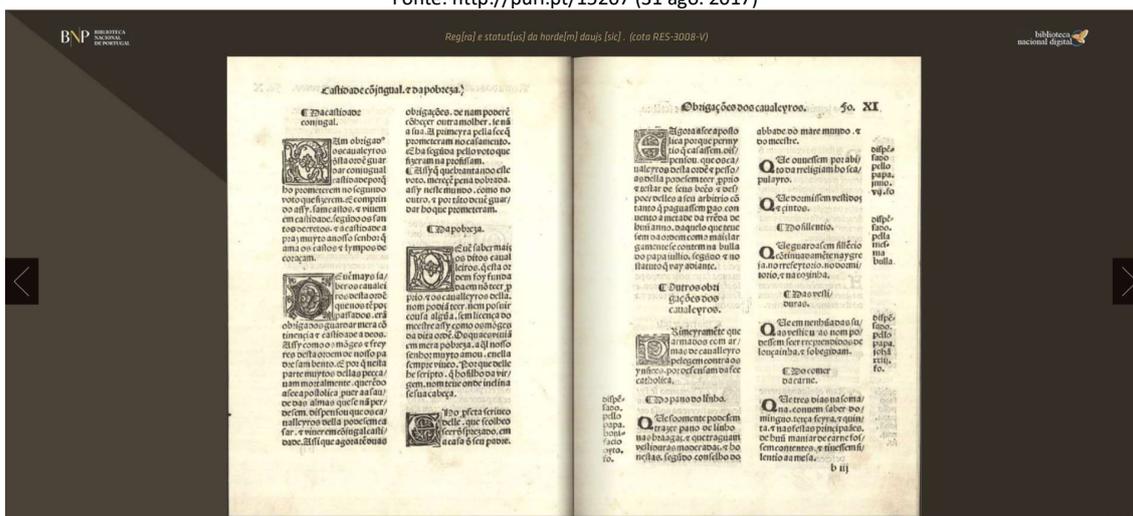
Os exemplares analisados em função das características físicas (Tabela 18) incluem: capitais (70,45%), caldeirões ou corações (52,27%), paginação por assinatura (91,29%), encabeçamento (60,23%) ou reclamo (19,32%).

Tabela 18 – BND: livro antigo, por características das obras

Elementos do livro	%	Número de obras
Capitais	70,45%	186
Caldeirões ou corações	52,27%	138
Assinatura	91,29%	241
Paginação	62,88%	166
Encabeçamento	60,23%	159
Reclamo	19,32%	51

Figura 33 – *Reg[ra] e statut[us]*, da *Ordem de Avis*: páginas com diferentes capitais e caldeirões.

Fonte: <http://purl.pt/15207> (31 ago. 2017)



A utilização das capitais, muito variadas inclusive na mesma obra (vd. *Reg[ra] e statut[us]*, da *Ordem de Avis*<sup>114</sup> (

Figura 33), de caldeirões (vd. id. *Ibid.*), corações (vd. *Tratado da sphaera*, de Pedro Nunes<sup>115</sup>, ou *De verboru[m] coniugatione*, André de Resende<sup>116</sup>), reclamos (vd.

<sup>114</sup> Disponível em <http://purl.pt/15207>

*Rhetorica*, de Ringelbergij<sup>117</sup>), ou encabeçamentos, podem aumentar o erro na versão textual das obras, pois condicionam a aplicação dos programas de OCR. Por outro lado, a numeração das páginas apenas por assinatura (vd. *Ho Preste loam das indias*, de Francisco Alvares<sup>118</sup>), ou a numeração dos fólhos na página de rosto, sem indicação no verso, dificultam a verificação das imagens digitalizadas.

Tabela 19 – BND: livro antigo, obras com transcrição, por tipo de letra

<b>Tipo de letra</b>	<b>Número de obras</b>
Romano	6
Gótico	4
Romano, itálico	3

O tipo de letra pode ser um condicionante para a obtenção de uma transcrição de qualidade, com poucos erros. No conjunto das obras analisadas há obras em romano, em romano e itálico e, mesmo, em gótico com transcrição (Tabela 19). Verifica-se que as obras que foram transcritas têm maioritariamente letra tipo romano (46%). O menor grupo é o das obras em romano e itálico (23%), mas deve atender-se ao facto de o itálico ser menos expressivo neste conjunto de obras, além de só ter surgido, e com restrições, em 1501. A percentagem de obras com letra gótica é, neste contexto, bastante elevada, com 31%, mas não se conhece o nível de erro que apresenta. Outra condicionante à obtenção de transcrições de qualidade, através de programas de OCR,

---

<sup>115</sup> Disponível em <http://purl.pt/14445>

<sup>116</sup> Disponível em <http://purl.pt/15169>

<sup>117</sup> Disponível em <http://purl.pt/23268>

<sup>118</sup> Disponível em <http://purl.pt/15320>

é a existências de glosas (vd. *Tractado da forma dos libelos*, de Gregório Caminha<sup>119</sup>) ou de texto rasurado (vd. *Curationum medicinalium*, de Amato Lusitano, pseud.<sup>120</sup>).

No conjunto das obras analisadas (Tabela 20), só 19% possuem ilustrações (vd. *Medidas del Romano*, Diego de Sagredo<sup>121</sup>).

Tabela 20 – BND: Livro antigo, obras por ilustração e mapas

	%	Número de obras
Ilustrações	19%	51
Mapas	1%	3

Três obras (1%) incluem mapas (vd. *Sei Libri Della Cosmografia Vniuersale*, de Sebastian Münster<sup>122</sup>). Por um lado, as obras com elementos iconográficos são poucas, mas, por outro, são importantes, pelo que justificam o trabalho acrescido para as refletir nos sumários.

Na amostra em análise, 77% são impressas em Portugal, enquanto 42% são em português. Ou seja, uma parte substantiva das obras impressas em Portugal são noutras línguas que não o português. Estão neste caso, por exemplo, os incunábulo em hebreu (vd. *[Perush Ha-Torah] = Comentário do Pentateuco*, de Moisés ben Nahman<sup>123</sup>) ou em latim, impressos em Braga, Coimbra, Évora, Leiria e Lisboa. Geralmente, versam temas religiosos, como a *Bíblia* ou tratados teológicos (vd. *Cronologia, seu ratio temporum*, de Nicolau Coelho de Amaral<sup>124</sup>), embora também se

---

<sup>119</sup> Disponível em <http://purl.pt/23247>

<sup>120</sup> Disponível em <http://purl.pt/26772>

<sup>121</sup> Disponível em <http://purl.pt/15281>

<sup>122</sup> Disponível em <http://purl.pt/21907>

<sup>123</sup> Disponível em <http://purl.pt/16559>

<sup>124</sup> Disponível em <http://purl.pt/23040>

registem obras laicas (vd. *Vrbis Olisiponis Descriptio*, de Damião de Góis<sup>125</sup>). A seguir ao português e latim, a língua que surge com maior frequência é o espanhol, maioritariamente em obras de temática religiosa (vd. *[Libro primero d'l Espejo...]*, de Francisco Monzon<sup>126</sup>), o que se pode explicar pela circulação dos livros impressos em espaço ibérico e porque esta era também uma língua frequente em Portugal, sobretudo em espaços eruditos. Em contrapartida, nesta amostra, não há obras em francês ou italiano, impressas em Portugal.

As obras sem sumário (28%) apenas dispõem do número sequencial da imagem digitalizada como auxiliar de navegação. Porém, mesmo as obras com sumário de nível 1 ou 2 (62%) contêm pouca informação e, portanto, não constituem um bom auxiliar de navegação na obra. Para melhorar a pesquisa no conteúdo, sobretudo em obras sem transcrição, seria necessário que todas as obras tivessem sumários de nível 1. Por outro lado, a apresentação do sumário, numa janela *pop-up* com dimensão fixa e com navegação por elevador lateral, dificulta a consulta e, sobretudo, a relação entre os apontadores e o conteúdo da obra, pelo que também seria conveniente alterar este aspeto. Além disso, os sumários não fazem referência aos elementos gráficos das obras (vd. *De medica matéria*, de Amato Lusitano, pseud.<sup>127</sup>, Figura 34), truncando, desta forma, uma parte substancial do seu conteúdo, pelo que seria de incluir referências às ilustrações (estampas, mapas, etc.) e, pelo menos nalguns casos, estabelecendo a ligação entre os elementos visuais e o texto correspondente.

---

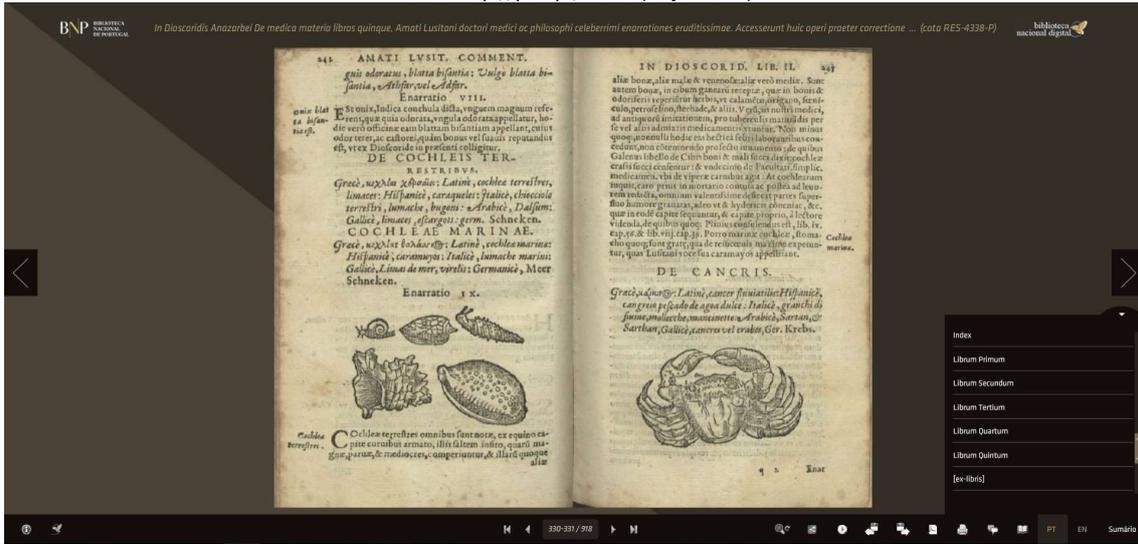
<sup>125</sup> Disponível em <http://purl.pt/23351>

<sup>126</sup> Disponível em <http://purl.pt/15319>

<sup>127</sup> Disponível em <http://purl.pt/22520>

Figura 34 – *De medica matéria*: obra sem indicação relativa às ilustrações.

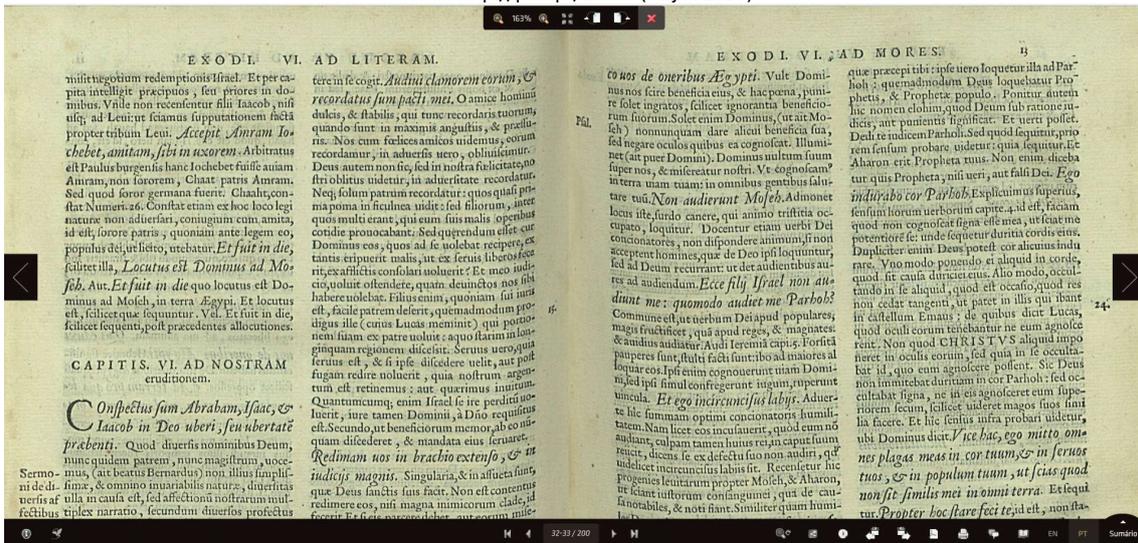
Fonte: <http://purl.pt/22520> (31 jul. 2017)



A leitura em linha das obras de grandes dimensões, superiores à Série Preta (Vd. Anexo XIII), é difícil, dado que exige a utilização de zoom para atingir um nível de legibilidade suficiente, mas nesta configuração, não estão disponíveis barras para mover a página, que fica truncada no ecrã (vd. *Reuerendi admodum...*, de Jerónimo de Azambuja<sup>128</sup>, Figura 35).

Figura 35 – *Reuerendi admodum*: visualização da imagem em zoom, sem barras de navegação.

Fonte: <http://purl.pt/17386> (31 jul. 2017)



<sup>128</sup> Disponível em <http://purl.pt/17386>

A disponibilização em linha de obras sem paginação nem foliação (vd. *Missale monasticum*<sup>129</sup>), ou com vários esquemas de numeração de página acarreta alguns problemas. Na BND, a opção mais comum é fazer uma numeração sequencial do número de imagens digitalizadas (*Delle Nauigationi et Viaggi*, de Giovanni Battista Ramusio<sup>130</sup>), sem atender a possíveis reinícios, por exemplo, após os conteúdos pré-textuais, o que provoca ambiguidades na referência de um trecho da obra. Porém, este procedimento não se aplica a todas as obras. Esta ambiguidade é maior quando há erros frequentes na paginação do original (vd. *Oraciones y exercícios de deuociõ*, de Frei Luís de Granada<sup>131</sup>), a qual é corrigida na versão digital, havendo, por isso, discrepâncias entre ambas. Nalguns casos, as páginas não numeradas no original estão identificadas na versão digital, com numeração romana (vd. *Liuro das obras*, de Garcia de Resende<sup>132</sup>). Ao transferir a versão digital de uma página em formato JPG, na denominação do ficheiro inclui a indicação do número da página do original, enquanto que na leitura em linha é identificado o número da imagem digital da obra que, na maioria dos casos, difere do número de página. A numeração por parágrafos existente nalgumas obras (vd. *L. Andr. Resendij Vincentius leuita et martyr*, de André de Resende<sup>133</sup>), não é refletida na edição digital.

Algumas obras são colocadas em linha a partir de ficheiros de imagem p/b, substituindo a cor de fundo por uma tonalidade neutra que procura simular a cor do

---

<sup>129</sup> Disponível em <http://purl.pt/24994>

<sup>130</sup> Disponível em <http://purl.pt/21896>

<sup>131</sup> Disponível em <http://purl.pt/23139>

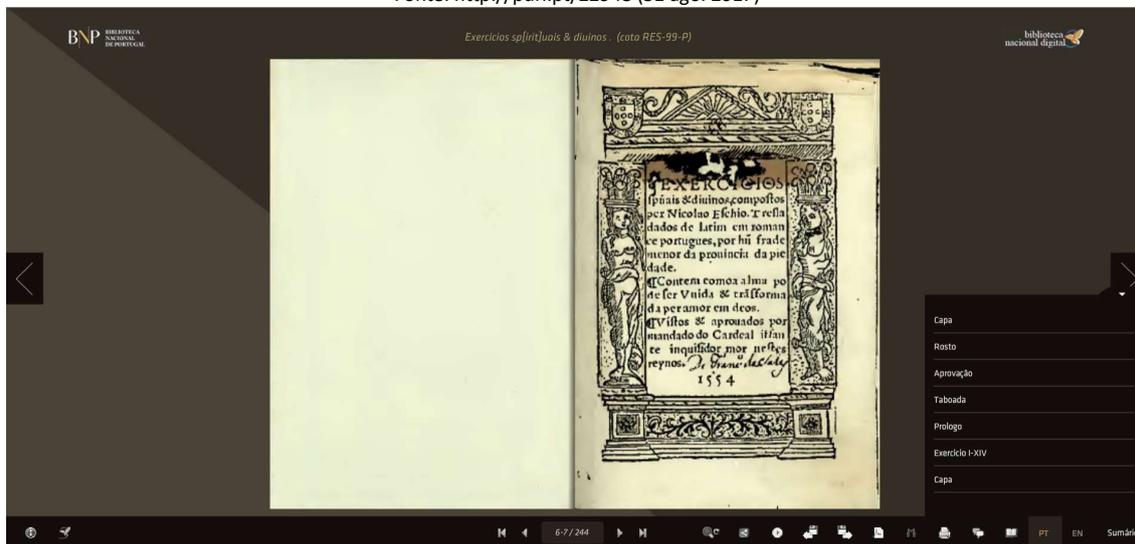
<sup>132</sup> Disponível em <http://purl.pt/15301>

<sup>133</sup> Disponível em <http://purl.pt/15168>

papel (vd. *Exercicios sp[irit]uais & diuinos*, de Nicolaus van Esch<sup>134</sup>, Figura 36), alterando as cores, nomeadamente o vermelho, e interferindo na leitura.

Figura 36 – *Exercicios sp[irit]uais & diuinos*: páginas com o fundo alterado.

Fonte: <http://purl.pt/22948> (31 ago. 2017)



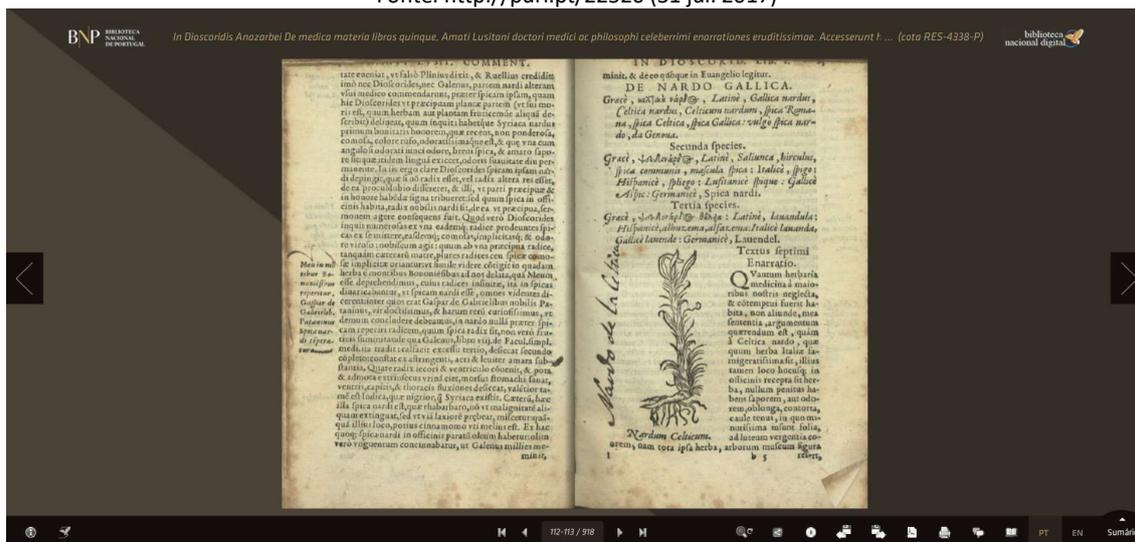
As obras inferiores a A6 (10,5x14,8 cm), geralmente com ângulos de leitura inferior a 180º, são difíceis de manusear e de colocar nos equipamentos de digitalização, sem forçar a encadernação. As obras de grandes dimensões, geralmente, muito pesadas são também difíceis de manusear. Este item não foi analisado, dado que os registos bibliográficos nem sempre referem o formato ou a altura do original e a perceção da sua dimensão é difícil na versão digital. No entanto, a observação das obras disponíveis permite concluir que a maior parte das obras da amostra se situam entre o In-oitavo (altura de 10 a 22 cm) e o In-fólio (altura varia de 33 a 40 cm) (Dias, 1994, p. 34). Na BNP, estes formatos refletem-se nas séries (Preta, Azul e Vermelha) que são mencionadas na cota, mas esta informação não é rigorosa (Vd. Anexo XIII).

---

<sup>134</sup> Disponível em <http://purl.pt/22948>

As obras digitalizadas com as margens truncadas colocam dúvidas em relação à sua integridade, não se percebendo se é o original que se encontra aparado sobre o texto, ou se este foi cortado no processo de digitalização (vd. *De medica matéria*, de Amato Lusitano, pseud.<sup>135</sup>, Figura 37).

Figura 37 – *De medica matéria*: obra com a margem superior aparada, truncando o texto.  
Fonte: <http://purl.pt/22520> (31 jul. 2017)



O processo de digitalização permite completar uma obra truncada com partes de outras proveniências, de forma a preservar a integridade do conteúdo, assinalando o facto em nota no registo bibliográfico (vd. *Este liuro he do começo da historea de nossa rede[n]çam*<sup>136</sup>, Figura 38).

<sup>135</sup> Disponível em <http://purl.pt/22520>

<sup>136</sup> Disponível em <http://purl.pt/15151>

Figura 38 – *Este liuro he do começo da historea de nossa rede[n]çam*: registo com nota relativa à digitalização a partir de dois originais. Fonte: <http://purl.pt/15151> (31 jul. 2017)

The screenshot shows a digital library record from the Biblioteca Nacional de Portugal. At the top, there are logos for 'BNP BIBLIOTECA NACIONAL DE PORTUGAL' and 'biblioteca nacional digital'. Below the header, the title of the record is 'BND > Este liuro he do começo da historea de nossa rede[n]çam que se fez pera consolação dos que nam...'. The main text describes the book as a Latin text by an author asking readers to pray for their souls, published in Lisbon in 1552. It includes the call number 'res-121-v' and a link to the full record. Under 'Cópias Digitais', there are three options: a public copy, a 39.9 MB PDF copy, and an 88.7 MB image copy. A note mentions the digitalization source as the Faculty of Letters at the University of Lisbon. At the bottom, there are social media sharing buttons for 'Gosto', 'Tweeter', and 'Partilhar', along with a PURL link and a timestamp.

Subjacente a estes fatores, a condição que mais influi no processo de digitalização do livro antigo decorre do seu valor patrimonial e da possível fragilidade do estado de conservação, impondo cuidados especiais no seu manuseamento e a criação de um ambiente adequado à sua preservação e segurança.

### 5.3. As edições digitais de Aldo Manuzio

Aldo Manuzio (ca 1450-1515) ou, em latim, Aldus Manutius, nascido precisamente na altura em que Gutenberg inventava a imprensa, foi um editor e impressor e tipógrafo italiano, cujo principal propósito era que textos escritos ao longo da história, sobretudo os clássicos que ainda não estavam publicados (McMurtrie, 1927, p. 124), fossem lidos pelo maior número de pessoas, tornando-os acessíveis a todos. “If the German, Johann Gutenberg, invented the process of printing by movable types, a Frenchman, Nicholas Jenson, raised it to its highest artistic level, while Aldus benefits could best be disseminated to the greatest number” (McMurtrie, 1927, pp. 117–118). Em certa medida, passados cinco séculos sobre a sua morte, cujo centenário se

comemorou em 2015, o seu objetivo foi amplamente cumprido, com as suas obras representadas na maior parte bibliotecas digitais com acervos patrimoniais de livro antigo. As obras produzidas por Aldo Manuzio constituem um universo próprio, pautado pela qualidade e pelo rigor reconhecido às coleções Aldinas.

### **5.3.1. Aldo Manuzio**

Aldo Pio Manuzio nasceu em Bassiano, *ca* 1450, e morreu em Veneza, a 6 de fevereiro de 1515. Estudou em Roma e Ferrara, onde contactou com os clássicos gregos e, em 1490, fixou-se em Veneza, onde fundou uma casa tipográfica (Manni, 2012; McMurtrie, 1969, p. 193). Embora esta seja conhecida como tipografia, era mais do que isso, no sentido em que “he had elaborated a careful scheme of action which he set in operation by assembling a group of the capable Greek scholars as well as the technical assistant’s necessary for equipping an ample printing establishment” (McMurtrie, 1927, p. 124). Além disso, reuniu um grupo de especialistas em grego antigo para que elaborassem uma revisão crítica dos textos encontrados em versões incompletas e dispersas. Ao juntar peritos e tipógrafos, Aldo Manuzio garantia a qualidade das obras, quer em relação ao conteúdo intelectual, quer em relação à forma, e definiu a função de disseminação do conhecimento atribuída ao editor.

Manuzio deu um forte impulso à divulgação da cultura e do pensamento clássico, recuperando e publicando as obras antigas, bem como os respetivos comentários académicos, disponibilizando-os “in editions characterised by impeccable philological care and an innovative design that improves readability and ease of use of the reader” (Crisci, 2017, p. 5), juntamente com a edição de autores humanistas como Dante,

Petrarca, Bembo e Erasmus. A edição destas obras era divulgada através de catálogos, o que era, à época, uma estratégia de promoção também inovadora.

Decorrente do objetivo de divulgação das obras, Manuzio preocupou-se igualmente com a sua acessibilidade e, por conseguinte, com a necessidade de baixar os custos de produção.

A new phase was initiated when publishers began to produce books for a more widely scattered public that was for the most part not known to them personally. Aldo Manuzio took great risks at this endeavour. [...] Manuzio was the first producer of books to arouse the interest of a large number of readers scattered throughout Europe. (Nuovo, 2013, p. 118)

A impressão de textos gregos com tipos móveis exigiu um considerável investimento de capital e competências específicas em caligrafia e metalurgia.

Além de conseguir apoio financeiro para as suas experiências, Manuzio teve a colaboração de Franceso Griffo, um gravador bolonhês, para desenvolver caracteres móveis gregos (Crisci, 2017). Mediante a utilização de contrações e ligaturas, Manuzio, com o apoio técnico de Griffo para a sua execução e usando a escrita cursiva do escriba Emanuele Rusota, como modelo, criou o tipo cursivo grego, o que lhe permitiu imprimir os textos nessa língua, com maior facilidade e economia de custos (Crisci, 2017). Este tipo foi usado pela primeira vez em 1495.

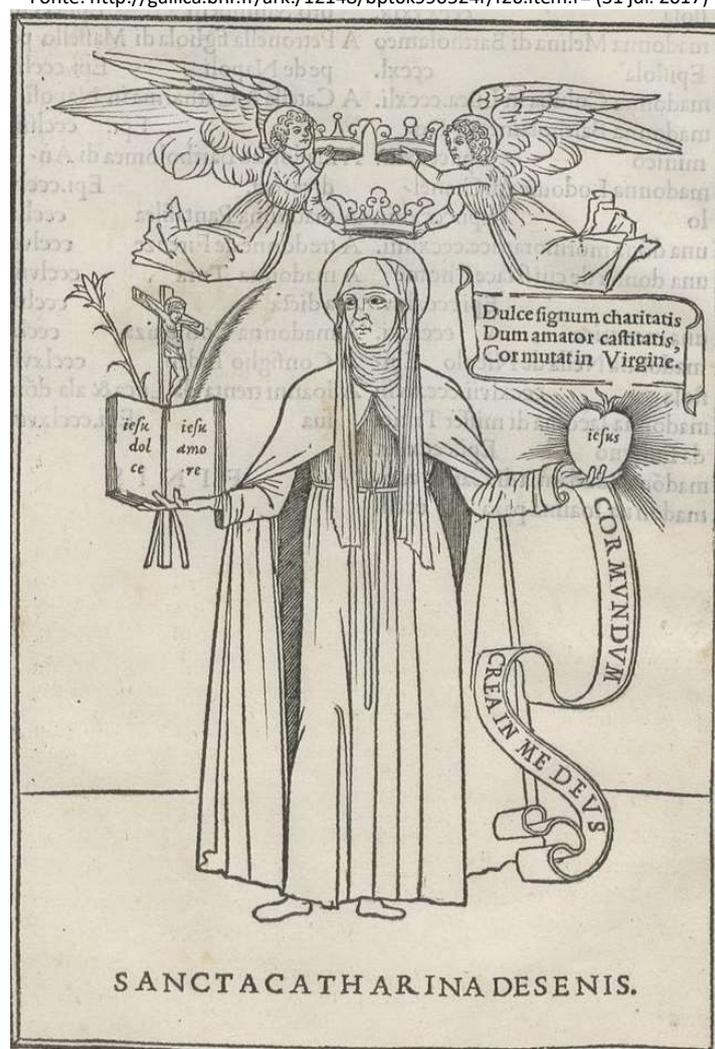
A redução do custo de produção passava, também, por diminuir o custo dos materiais utilizados, ou seja, do suporte físico do livro, o papel. Nesse sentido, Manuzio reduziu o tamanho dos livros, passando a imprimir no formato “oitavo”, ou in-8º, com uma

altura aproximada de cerca de 18 cm (Dias, 1994, p. 34). Por seu turno, a redução de tamanho das folhas implicou a diminuição do corpo de letra.

Para diminuir o corpo da letra e garantir a legibilidade, criou o itálico, permitindo-lhe uma economia de espaço na impressão de textos latinos e vernaculares, idêntica à que havia conseguido com as ligaturas nos textos em grego (McMurtrie, 1927, p. 127).

O tamanho condensado do tipo permitiu reduzir o tamanho da página e, com isso, diminuir o tamanho dos livros e baixar o consumo de papel, subtraindo os custos de impressão e tornando as obras mais acessíveis, melhorando, ao mesmo tempo, as condições de manuseio e portabilidade.

Figura 39 – *Epistole devotissime de sancta Catharina da Siena*, s.p.: Santa Catarina de Siena (il.)  
Fonte: <http://gallica.bnf.fr/ark:/12148/bpt6k990524r/f26.item.r=> (31 jul. 2017)

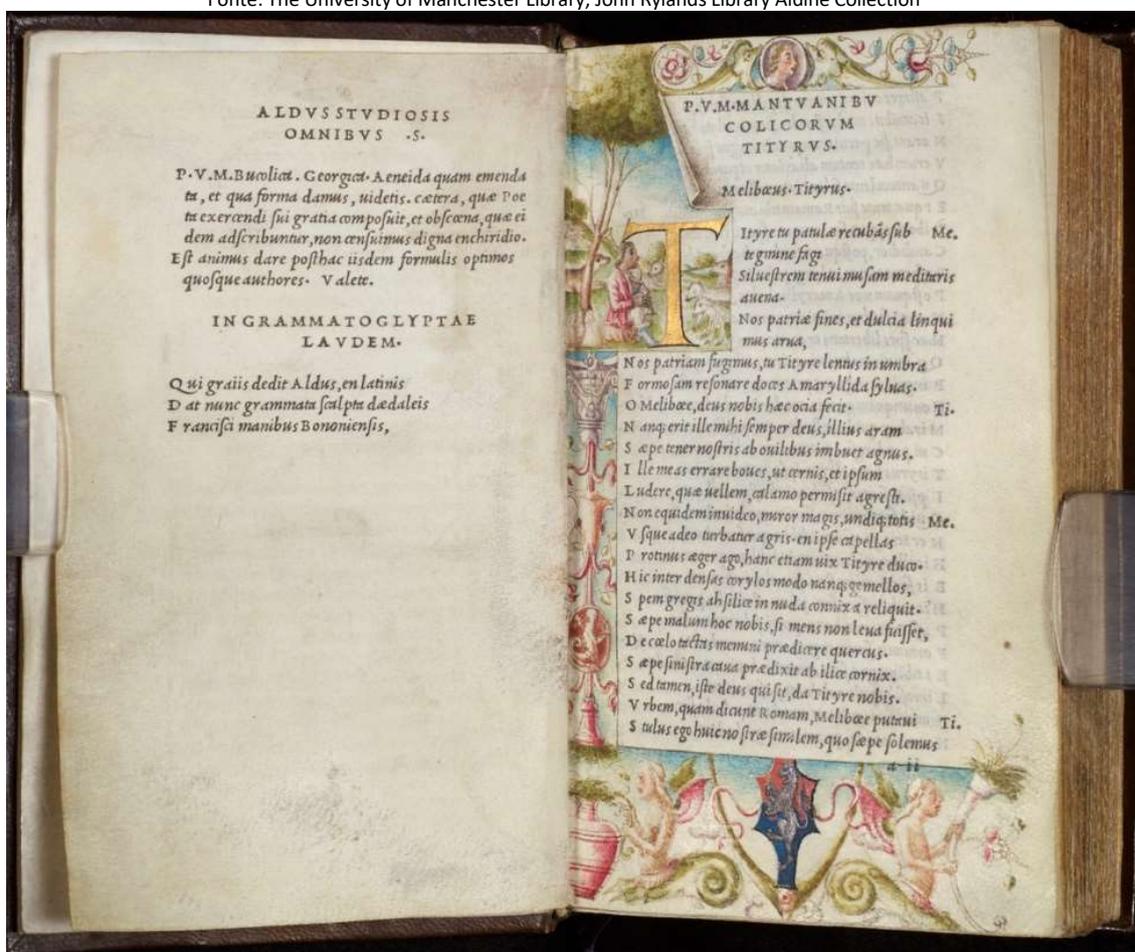


O itálico foi utilizado pela primeira vez em 1500, numa ilustração, com a representação de Santa Catarina, na edição das suas *Epistole devotissime* (Figura 39), em três frases (“*iesu dulce*”; “*iesu amore*”; e “*iesu*”) no interior do livro e do coração que a santa segura em cada uma das mãos.

Em 1501, foi publicada a coletânea das obras de Virgílio, [*Bucolica, Georgica, Aeneida quam emendata*] (Figura 40), com 228 páginas, a primeira obra impressa inteiramente em itálico, com caracteres abertos por Francesco Griffo, e no formato in-8º.

Figura 40 – Vergilius. [*Bucolica...*].

Fonte: The University of Manchester Library, John Rylands Library Aldine Collection<sup>137</sup>



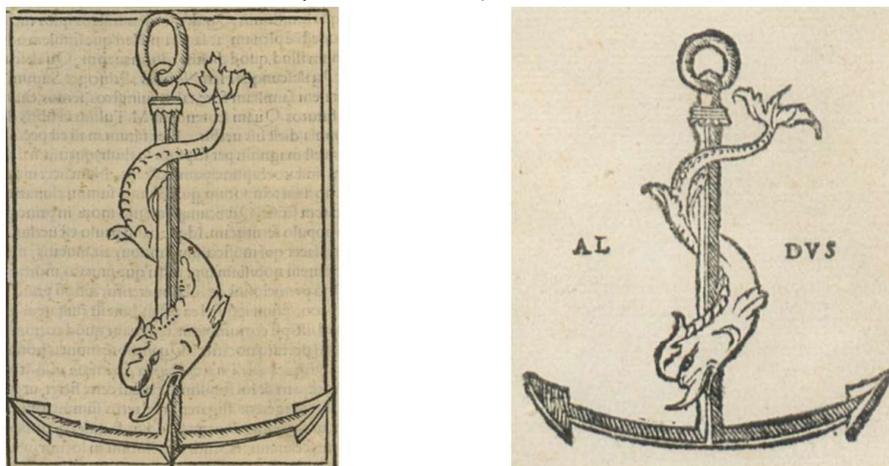
<sup>137</sup> Disponível em <http://luna.manchester.ac.uk/luna/servlet/detail/Manchester~91~1~408746~128632?qvq=q%3AVirgil%27s%2BBucolica&mi=3&trs=4>

Aldo Manuzio e Francesco Griffo criaram também o tipo romano, em caracteres redondos, usado pela primeira vez na impressão da obra *De Aetna dialogus*, do cardeal Pietro Bembo, em 1495 (Figura 48).

Os melhoramentos e as criações tipográficas conseguidos por Manuzio e Griffo resultam dos avanços que conseguiram implantar no equipamento, nos materiais e no fabrico dos tipos e na formação dos artífices. Porém, pelo avanço que representavam numa época de forte concorrência neste campo, muitas das suas invenções, nomeadamente os tipos de letra e formatos da obra foram copiados, o que motivou o aceso protesto de Aldo Manuzio e a tentativa de os registar (McMurtrie, 1969, p. 200).

Criou, por isso, uma forma de identificar as suas obras, acrescentando-lhes a sua marca de impressor (Figura 41), uma âncora com o golfinho enrolado, representando o adágio latino “*Festina lente*”, que significa “apressa-te devagar” (Fletcher, 1988).

Figura 41 – Marcas do impressor Aldo Manuzio: à esq., *Aristotelis de Natura animalium lib. IX* (1513); à dir., *Plutarchi Opuscula LXXXVII* (1509). Fonte: Gallica (31 jul. 2017)



Outro dos contributos mais significativos de Aldo Manuzio foi a sistematização da pontuação, utilizada nas suas impressões pela primeira vez na forma moderna: o ponto, como fecho de período; o ponto e vírgula, para encerrar frases no período; e a vírgula, o apóstrofo e os acentos (Barolini, 1992, p. XVII).

Ainda que os números não sejam consensuais, estima-se que tenha produzido 132 livros impressos durante os seus vinte anos de atividade (Bühler, 1950, p. 210), mas a imprensa aldina manteve-se através dos seus herdeiros. A Manuzio, que é considerado o primeiro editor moderno, ficou a dever-se uma série de inovações com impacto até aos nossos dias (Crisci, 2017; Fletcher, 1988): o uso de catálogos das obras; a impressão em formato in-8º; a quantidade de tipos e, entre eles, sobretudo, o itálico.

### 5.3.2. Representação nas várias coleções digitais

A obra de Aldo Manuzio está representada em quatro das seis bibliotecas digitais estudadas (Tabela 21): um exemplar, na BND; dezasseis, na Gallica; três, na BVH; cinquenta e um, no IA. Há, também, uma obra de Manuzio no Perseus, onde o conteúdo prevalece, em detrimento da materialidade da obra, pelo que esta não é aqui considerada.

Tabela 21 – Obras Aldo Manuzio nas bibliotecas digitais

<b>Biblioteca</b>	<b>N.º de obras</b>
BND	1
Gallica	16
Bibliothèques Virtuelles Humanistes	3
Internet Archive	51

A opção de pesquisa na BND só é possível através do catálogo da BN. Efetuando a pesquisa com os termos “Manuzio, Aldo” em autor e com a data de produção definida entre 1450 e 1515, limitando a coleção a “Biblioteca Nacional Digital”, foram obtidos dois resultados (Figura 42), um dos quais extrínseco às condições da pesquisa.

Figura 42 – BNP: resultado da pesquisa para “Manuzio, Aldo” (31 jul. 2017)

The screenshot shows the search results for 'Manuzio, Aldo' on the BNP website. The search bar contains 'Manuzio, Aldo' and the results are limited to 'Biblioteca Nacional Digital'. Two results are listed:

1. Quincti Horatii Flacci Poemata omnia. Adjectus est in fine Aldus Manutius De metris Horatianis. Ulyssipone : apud Petrum Craesbeeck Typographum Regium, 1629. (Thumbnail of the book cover)
2. Institutiones graecae grammaticae. Venezia : Aldo Manuzio. 1497. (Thumbnail of a page from the book)

O primeiro resultado aponta para uma obra, onde o termo “Manutius” aparece nos campos 200\$c do registo UNIMARC, relativo ao título e menção de responsabilidade, e 702, relativo à responsabilidade secundária, mas a data “1629” não se enquadra no arco cronológico solicitado.

A obra de Manuzio na BND é a *Institutiones graecae grammaticae*, de Urbano Bolzanio, impressa em 1497-98<sup>138</sup> (Figura 43).

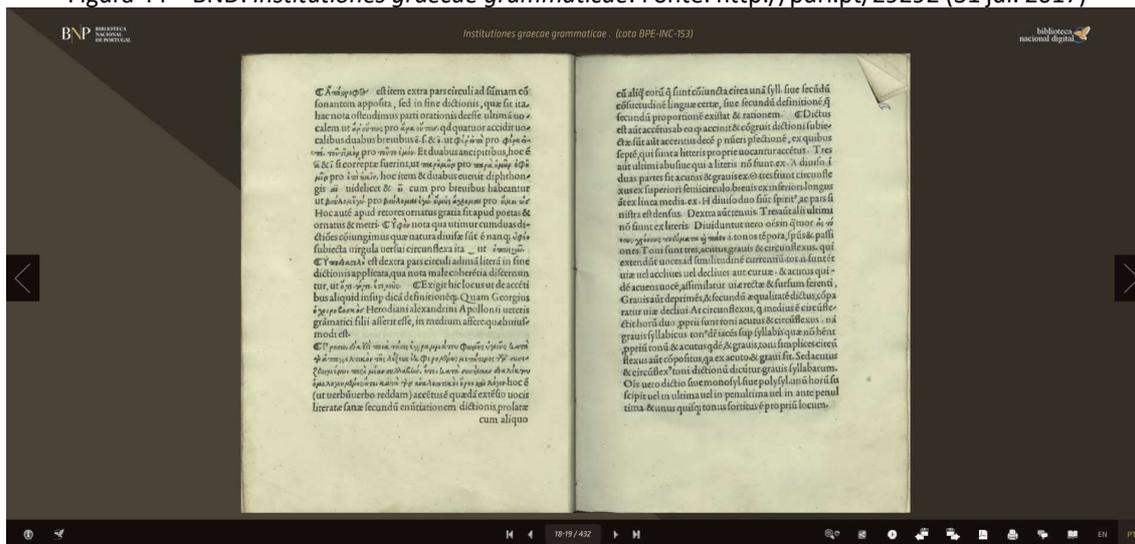
Figura 43 – BND: *Institutiones graecae grammaticae*. Fonte: <http://purl.pt/29292> (31 jul. 2017)

The screenshot shows the detailed record for 'Institutiones graecae grammaticae' in the BND. The record includes the following information:

- BND > Institutiones graecae grammaticae, Venezia, Janeiro 1497/98**
- BOLZANIO, Urbano, ca 1443-1524**
- Institutiones graecae grammaticae. - Venezia : Aldo Manuzio Janeiro 1497/98. - [214] f. ; 4º**
- Cota do exemplar digitalizado: bpe-inc-153
- Ver registo completo
- Cópias Digitais**
  - Cópia pública
  - Cópia pública, 22.6 MB, 1 ficheiro
  - Cópia pública, 667.5 MB
- Thumbnail of a page from the book
- Social media buttons: Gosto 0, Tweetar, Partilhar
- URL: PURL 29292
- Última atualização em 2016-12-06T10:27:20

<sup>138</sup> Disponível em <http://purl.pt/29292>

Figura 44 – BND: *Institutiones graecae grammaticae*. Fonte: <http://purl.pt/29292> (31 jul. 2017)



A obra é disponibilizada nas versões *flash*, PDF e JPG. A obra é bilingue em grego e latim (Figura 44), foi impressa em in-4º (com cerca de 22 cm de altura), tem paginação por assinaturas e apresenta reclamos, além de capitulares ou o respetivo espaço em branco e um esquema.

Na leitura em linha, a obra é apresentada em página dupla. Não tem sumário e, não tendo sido transcrita, não permite a pesquisa no documento. Embora seja legível, é possível efetuar *zoom*. Na versão em *flash*, o *zoom* apresenta variações pré-definidas, não permitindo um aumento progressivo. O botão para ajustar a obra ao ecrã não funciona. A numeração apresentada segue o número sequencial da imagem digitalizada. A versão em JPG é idêntica, com a vantagem de que, nesta versão, o botão para ajustar a obra ao ecrã funciona. Depois de iniciada a leitura, é difícil distinguir as duas versões. O ficheiro PDF foi comprimido, resultando num arquivo com pouca qualidade para ser utilizado ou processado, não ficando imediatamente disponível para ser trabalhado. Esta obra também está disponível no IA.

O tratamento aplicado a esta obra é dos mais rudimentares da BND, sem algumas das funcionalidades que esta permite, limitando-se a disponibilizar uma versão digital fac-similada. A navegação na obra é feita página a página ou avançando para um determinado número da digitalização. Sem índice ou quaisquer outros apontadores, para lá do facto de poder ser consultada em linha, não apresenta vantagens face à consulta do original.

A Gallica possui 16 obras atribuídas à oficina de Aldo Manuzio. No sítio eletrónico da Gallica, a pesquisa avançada por autor e data (Figura 45), feita em termos idênticos à realizada na BN, resulta mais assertiva e correta.

Figura 45 – Gallica: pesquisa avançada, por autor e data (31 jul. 2017)

The screenshot shows the Gallica advanced search interface. At the top, there is a search bar with the text 'Rechercher...' and a magnifying glass icon. Below the search bar, there are several options and filters. On the left, there is a dropdown menu for 'Trier vos résultats par'. In the center, there are two buttons: 'Lancer la recherche' and 'Effacer la recherche'. On the right, there are checkboxes for 'Recherche exacte' and 'Liste de résultats : Regroupée par titre' (checked) and 'Au numéro'. Below these, there is a section for search criteria. It includes a dropdown for 'Auteur / Contributeur' with the value 'Manuzio, Aldo'. There are also dropdowns for 'Titre' and 'Table des matières / Légendes'. Below this, there is a section for 'Par année d'édition' with a range from '1450' to '1519'. There are also checkboxes for 'En mode texte', 'En 3D', and 'En version EPUB' (with sub-options 'Tous les EPUB' and 'EPUB accessibles'). Finally, there is a section for 'Par langue' with checkboxes for 'Toutes', 'Français', 'Anglais', 'Allemand', 'Espagnol', 'Italien', 'Chinois', 'Grec', and 'Latin'.

A lista dos resultados integra 16 obras (Anexo XIV), sendo 13 pertencentes à coleção da BnF e três provenientes das coleções da e|rara.ch<sup>139</sup>, da Rosalis<sup>140</sup> e da Tolosana<sup>141</sup>. A partir da lista de resultados, é possível aceder ao registo bibliográfico da obra e, deste, à respetiva versão em formato digital. No entanto, a múltipla proveniência, a

<sup>139</sup> Biblioteca que disponibiliza obras dos séculos XV a XVII de bibliotecas da Suíça. Disponível em <http://www.erara.ch>

<sup>140</sup> Biblioteca municipal de Toulouse. Disponível em <http://rosalis.bibliotheque.toulouse.fr/>

<sup>141</sup> Biblioteca universitária. Disponível em <http://tolosana.univ-toulouse.fr/>

partir de quatro sistemas não uniformizados, implica diferentes modalidades de consulta.

A biblioteca e|rara apresenta o registo bibliográfico com a página de título, permitindo as seguintes opções de visualização: sumário da obra (*Content*); elenco de miniaturas (*Thumbnails*); navegação na obra, página a página, acompanhada de caixas de informação relativa ao título da obra, à identificação do conteúdo (localização na obra) e número de página (*Page view*); opção de écran inteiro; número de página; botões de navegação; botões de ajuste da imagem à largura e à altura do écran, de rodar à direita e à esquerda (90º) e *zoom*. Ao mover o elevador, na barra à direita, a imagem da obra ocupa o écran inteiro, perdendo-se o acesso à barra de navegação superior.

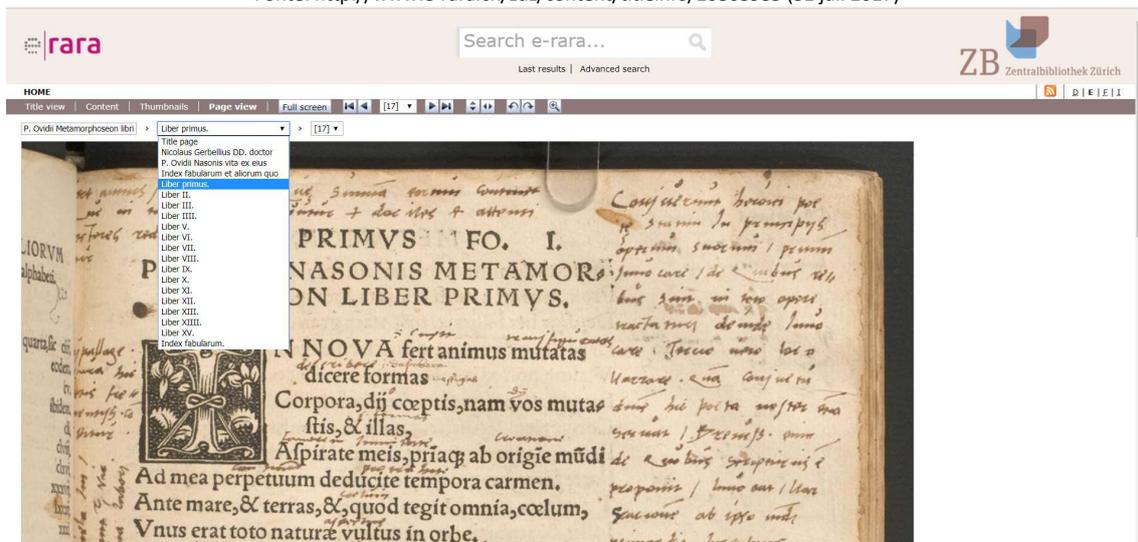
Na Rosalis, a versão *flash* tem incorporada uma animação para simular o folhear, permite a navegação por miniaturas (*Vignettes*, ou *thumbnails*), a marcação de páginas, a partilha total ou parcial nas redes sociais (Facebook, Twitter, LinkedIn, MySpace, Tumblr), num blogue associado à conta ou por correio eletrónico, a transferência do documento (página ou obra) em PDF, ligar ou desligar o som (neste caso, apenas para a passagem das páginas), visualização em écran inteiro e aplicação de *zoom*. A obra disponibilizada não tem paginação e o número que aparece na barra superior corresponde ao número sequencial da imagem digitalizada. Não tem sumário ou qualquer outro apontador que facilite a navegação na obra.

Na Tolosana, a barra à esquerda do registo bibliográfico apresenta as seguintes opções: transferir (*Télécharger*), folhear (*Feuilleter*), disponibilizar nas redes sociais (*Partager*) ou enviar uma mensagem para a biblioteca (*Une remarque?*). A leitura da obra em linha faz-se a partir do botão *Feuilleter*, abrindo a imagem noutra janela,

simulando o folhear da obra e permitindo *zoom* ao movimentar o cursor sobre a imagem. Uma barra horizontal sob a imagem permite navegar na obra, mas sem fornecer qualquer informação relativa à localização da imagem na obra, à paginação ou número de imagem digitalizada, ou ao conteúdo.

No conjunto das obras disponibilizadas na Gallica, há 10 com as imagens a cores. Apenas uma obra tem sumário (*P. Ovidii Metamorphoseon libri XV*, Figura 46). Nenhuma das obras possui versão em texto. As obras da BnF indicam o número da página do original. O maior obstáculo à consulta reside no facto de obrigar o utilizador a adaptar-se a diferentes configurações e modos de navegação.

Figura 46 – Gallica / e|rara: *P. Ovidii Metamorphoseon libri XV*.  
Fonte: <http://www.e-rara.ch/zuz/content/titleinfo/10868985> (31 jul. 2017)

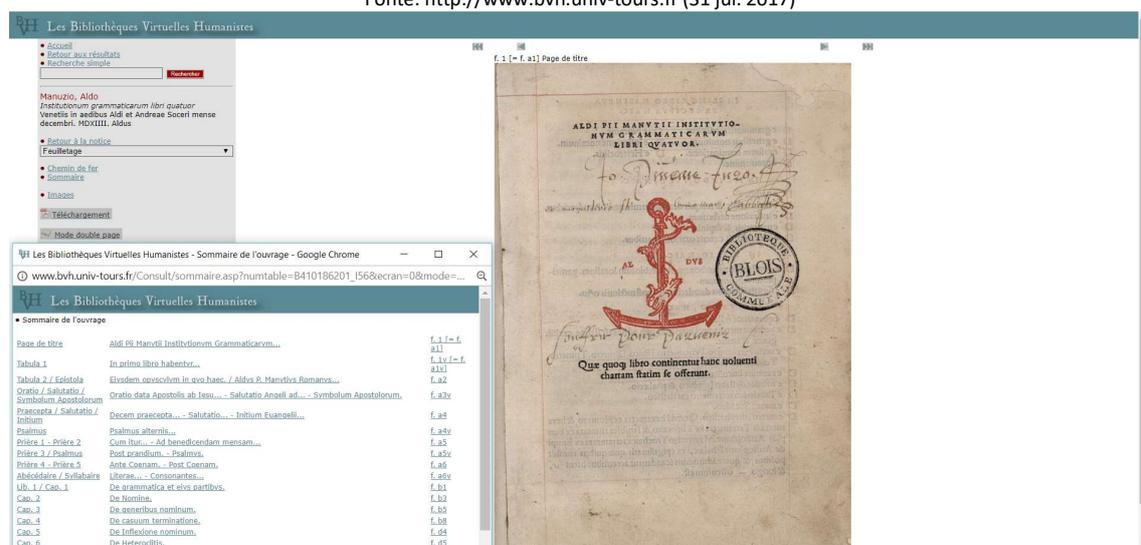


Na BVH, a pesquisa pode realizar-se a partir da página de acolhimento ou na coleção *Fac-similés*, sendo a única biblioteca que, na pesquisa avançada, permite fazer a pesquisa por impressor. No entanto, os resultados obtidos não são coerentes: a pesquisa simples devolve um resultado (*Institutionum grammaticarum libri quatuor*); a pesquisa avançada fornece dois registos, acrescentando, à anterior, a obra *Institutionis oratoriae libri XII*, de Quintiliano. Dado que apenas a primeira obra regista o termo

“Manuzio, Aldo” como autoridade, depreende-se que a pesquisa simples não abrange o campo do impressor. A pesquisa simples por “Aldo Manuzio” recupera uma lista de termos relativos a Manuzio, como autor (*Auteur*), comentador (*Comentateur*), editor (*Éditeur scientifique*), impressor (*Imprimeur*) e livreiro (*Libraire*). Neste caso, ao fazer a pesquisa por “Manuzio, Aldo, Imprimeur” são devolvidos cinco resultados, o que indicia uma falta de coerência na definição dos metadados. Incluindo, na pesquisa, a opção texto (*Plein-texte*), obtém-se uma lista com cinco documentos em PDF com a transcrição do texto associada.

A obra *Institutionum grammaticarum libri quatuor* (Figura 47) é apresentada a cores e tem sumário. É possível transferir o ficheiro PDF. A navegação na obra é realizada através dos botões de avanço e retrocesso. Por cima da imagem da obra, mantém-se visível a informação relativa à respetiva secção e página. Não permite aumentar a imagem no ecrã, apesar de, neste caso concreto, se conseguir ler a obra. Possui uma secção com as ilustrações contidas na obra.

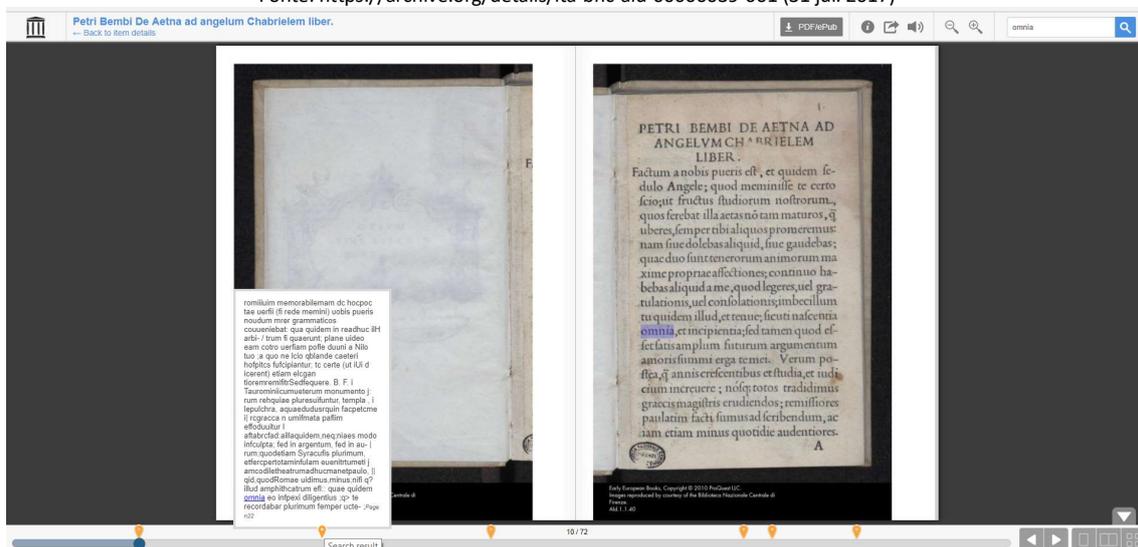
Figura 47 – BVH: *Institutionum grammaticarum libri quatuor*.  
 Fonte: <http://www.bvh.univ-tours.fr> (31 jul. 2017)



No IA, na pesquisa avançada, o campo data (*Date range*), tabelado, não permite circunscrever um intervalo de datas anterior a 1800. Ao lançar a pesquisa no campo de autoridade (*Creator*) com a expressão "Manuzio, Aldo", é apresentada uma lista com 337 obras; com a expressão "Manuzio, Aldo, 1450-1515", a lista obtida contém 9 resultados. Numa biblioteca colaborativa como o IA, esta discrepância pode justificar-se pela multiplicidade de contribuidores e critérios de catalogação e metacodificação. Ordenando por ordem crescente os resultados por data de criação (impressão), identificam-se 51 obras anteriores a 1515 (Anexo XV). Tal como registámos na Gallica, as múltiplas proveniências provocam diferentes configurações que, aqui, se verificam nos formatos para transferência e na qualidade da digitalização e dos metadados. Em contrapartida, a visualização da obra e a leitura em linha (Figura 48) são sempre efetuadas no mesmo leitor, mantendo uma configuração uniforme.

Figura 48 – IA: *Petri Bembi de Aetna dialogus*.

Fonte: <https://archive.org/details/ita-bnc-ald-00000039-001> (31 jul. 2017)



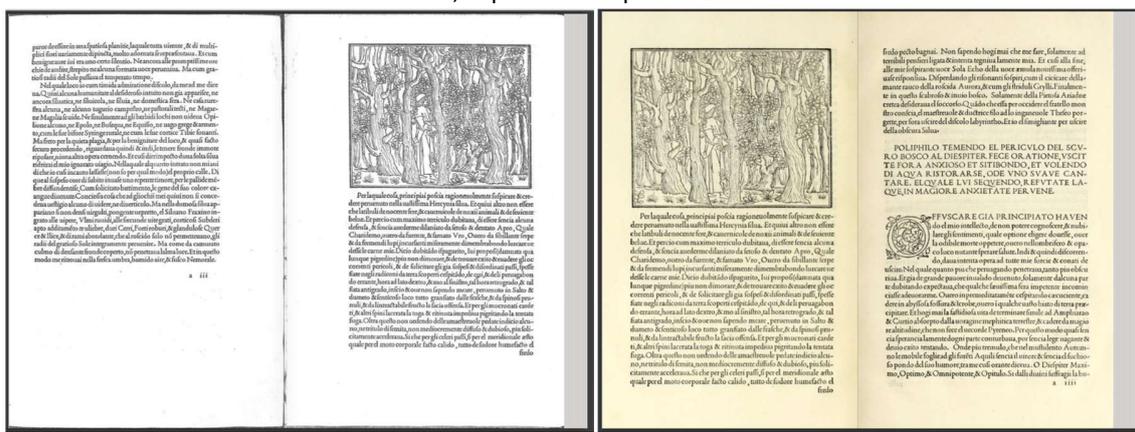
A navegação nas páginas é realizada através do número da imagem digitalizada. As obras estão disponíveis em diferentes formatos de imagem e de texto, permitindo a pesquisa por palavra no texto. O termo ou expressão de pesquisa é assinalado no

corpo do texto e, na barra inferior, surgem marcadores indicando as várias ocorrências. Ao passar o curso sobre qualquer um dos marcadores (não necessariamente da página em visualização), abre uma janela com a palavra inserida no respetivo contexto. As obras, mesmo as que são em língua grega, possuem transcrição, mas não há informação relativa ao índice de erro, não garantindo que a pesquisa recupere todos os termos ou expressões solicitados e existentes na obra.

O volume de recursos disponíveis no IA é permeável à duplicação de obras. No caso das obras da imprensa aldina, encontram-se, por exemplo, dois exemplares da *Hyperotomachia Poliphili*, publicada em 1499, cópias digitais dos originais existentes na Biblioteca de la Universidad de Sevilla e na Boston Public Library.

A análise comparativa destas obras permite observar algumas diferenças. A obra da Biblioteca de Sevilla, a p/b, tem um total de 464 imagens; a de Boston, a cores, tem 466 imagens. A paginação que, no original, é por assinaturas, nas cópias digitais a numeração segue a ordem sequencial da digitalização.

Figura 49 – IA: *Hyperotomachia Poliphili*: à esq. cópia do exemplar de Sevilla; à direita, cópia do exemplar de Boston



Na Figura 49, estão representadas: à esquerda, as páginas com assinatura aaiii//[aiiiiv], do exemplar da Biblioteca de Sevilla, correspondendo à imagem digital

8/464; à direita, as páginas com assinatura [aaiiav]//aaiiii, da Biblioteca de Boston, correspondendo à imagem digital 18/464. A discrepância de números entre o total de imagens de cada uma deve-se ao facto de a obra de Boston incluir 10 imagens de partes pré-textuais que não aparecem na de Sevilha.

Ambas as imagens parecem reproduzir a obra aberta. Contudo, apenas a de Boston está na sequência correta, dado que a de Sevilha apresenta a página *recto* na posição de *verso* e a página *verso* que lhe corresponde é visível na posição de *recto*. A capa da obra não foi digitalizada, iniciando pelas duas guardas, enquanto o sistema assume que o início da apresentação da obra se faz com uma única imagem. Ao editar digitalmente a obra e ao juntar as imagens para simular o livro aberto, foi gerado o erro que afeta toda a obra e que não é facilmente detetável pelo utilizador.

Esta ocorrência fica registada como um eventual inconveniente da consulta em suporte digital e que poderá justificar, nalguns casos, a necessidade de requerer o cotejo com o original.

A obra *Institutiones graecæ grammaticæ* está disponível na BND<sup>142</sup> e no IA<sup>143</sup>, sendo uma duplicação inevitável, atendendo ao volume de obras em linha. No IA, a forma como a obra é apresentada traz vantagens à utilização, dado que é possível trabalhar com facilidade nos ficheiros de texto ou mesmo com os ficheiros de metadados. Ao contrário do IA, a obra da BND é apresentada a cores.

Foram analisadas um total de 75 obras da imprensa aldina disponibilizadas nas bibliotecas BND, Gallica (BnF, e|rara, Rosalis e Tolosana), BVH e IA (Tabela 22).

---

<sup>142</sup> Disponível em <http://purl.pt/29292>

<sup>143</sup> Disponível em <https://archive.org/details/A336161>

Tabela 22 – Características das obras de Aldo Manúsio, por biblioteca digital

Biblioteca	Total	Sumário	Trans. <sup>144</sup>	Cor	p/b	Formatos transferência	Leitura em linha	Paginação original
BND	1			1		PDF/JPG	1	
Gallica	13							
BnF	13			7	6	PDF/JPG	13	2
e rara	1	1		1		PDF	1	
Rosalis	1			1		PDF	1	
Tolosana	1			1		PDF	1	
BVH	7	2		2	5	PDF	2	2
IA	51		51	34	17	Vários	51	

Nesta amostra, a maior parte das obras não tem sumário, nem outros apontadores para a navegação no conteúdo da obra. Excetuam-se a BVH, biblioteca ligada a um centro de estudos em Humanidades, e a e|rara que tem os investigadores como público-alvo. O IA permite a pesquisa por palavra no conteúdo textual da obra, mas é a única biblioteca a disponibilizar os ficheiros em formato texto, pelo que as restantes não facultam esta pesquisa. A maioria das obras disponibilizadas são a cores (47), enquanto 28 estão a p/b.

Todas as bibliotecas disponibilizam as obras em formato PDF, a BND e a Gallica também disponibilizam os ficheiros JPG e o IA fornece as obras num variado conjunto de formatos texto e imagem. Sendo que todas as bibliotecas permitem a leitura em linha, a BVH tem também obras em formato PDF pesquisável, com transcrição do texto associado, apenas disponíveis para transferência. Nesta amostra, apenas a BnF e a BVH refletem a paginação do original.

---

<sup>144</sup> Transcrição.

A primeira dificuldade para efetuar este levantamento foi a própria pesquisa da obra de Aldo Manuzio, dado que, depois da sua morte (1515), a oficina de impressão continuou em atividade com o filho e o neto. Nas bibliotecas digitais consultadas não há um registo de autoridade comum. O nome aparece em diferentes idiomas e grafias: em latim, Aldus Pius Manutius, ou Albus Manutius; em italiano, Aldo Pio Manuzio; em inglês, Aldus Manutius the Elder; em francês, Manuzio, Aldo. Em português, pode aparecer como Manúcio ou Manuzio. Surge, também, como Manucci. A terminologia não normalizada configura-se como o primeiro obstáculo à consulta em linha e, por conseguinte, à utilização das bibliotecas digitais. “Search technology makes information readily available that may previously have been public in principle, but impossible to find in practice” (Witten et al., 2009, p. 37). A pesquisa por determinado termo não abrange os termos equivalentes. Embora haja projetos, como o sistema Computer Says Yes (Bainbridge, Twidale, & Nichols, 2011) para a variação dos nomes das autoridades, a investigação neste campo ainda não apresenta resultados satisfatórios.

A segunda dificuldade relaciona-se com a recuperação da informação. Por um lado, a lista de resultados inclui dados não solicitados, sem obedecer às condições definidas pelos termos de pesquisa, por outro lado, é difícil ordenar os resultados de forma a obter as obras pretendidas. Esta matéria relaciona-se com a qualidade dos metadados, ou seja, a conversão dos registos bibliográficos em metadados para a obra digital, no sentido em que formatos como o Marc foram desenhados em função de uma pesquisa rígida por campos e a pesquisa em ambiente digital é dinâmica (Fenlon et al, 2014). A investigação aponta no sentido da implementação da pesquisa semântica, em detrimento da lexical, por ser mais maleável e adequada à complexidade da linguagem

natural. A rede Concept-in-Context (Hinze et al., 2015) que tem vindo a ser desenvolvida com este objetivo permite algumas expectativas de otimização de pesquisa e recuperação da informação, mas aponta para a necessidade de conversão das obras em formato imagem para formato texto.

A principal lacuna detetada relaciona-se com o acesso à informação intelectual da obra, ou seja, com a pesquisa no texto. Apenas o IA permite esta pesquisa, mas sem indicar o nível de erro. Fazendo pesquisas por palavras visíveis no texto e não detetadas na pesquisa supõe-se que esta margem de erro da transcrição não seja irrisória. Além disso, mesmo nos documentos transcritos, a informação não se encontra estruturada, apesar da investigação e dos projetos que têm vindo a ser desenvolvidos nesse sentido (Budig et al., 2016; Choudhury et al., 2006).

A dificuldade no acesso ao conteúdo das obras não incide só no texto, mas também no conteúdo gráfico das obras. Apenas a BVH possui uma secção com as ilustrações contidas na obra. Na investigação, são também conhecidos alguns avanços neste âmbito (Agosti et al., 2005; Audenaert & Furuta, 2009, 2010), com o objetivo de elaborar anotações das imagens e ilustrações dos textos, atribuindo-lhes autorias, descrevendo a respetiva iconografia e estabelecendo relações.

## **6. Fatores de acessibilidade da informação: levantamento das necessidades da investigação em Ciências Sociais e Humanas**

Para determinar quais os fatores de acessibilidade da informação, foram utilizados o grupo focal e o inquérito por questionário em duas rondas. Com o grupo focal pretendeu-se auscultar a opinião dos investigadores acerca das necessidades sentidas na utilização das bibliotecas digitais. Posteriormente, o resultado foi aferido com o inquérito por questionário. Com esta técnica de recolha de dados, pretendemos obter respostas não influenciadas pelos outros membros do painel.

### **6.1. Grupo focal**

Os participantes no grupo focal foram selecionados em função da investigação em Humanidades (Ciências Sociais e Humanas), cobrindo as seguintes áreas: História e História da Arte; Linguística; Comunicação Cultural; Estudos da Cultura; Antropologia Cultural e Etnografia; Museologia e Património. Todos os participantes têm o grau de Doutor ou são doutorandos, tendo sido recrutados em várias instituições do ensino superior, em Lisboa: Universidade de Lisboa (1); Universidade Nova de Lisboa (1); Universidade Católica Portuguesa (2); Universidade Europeia – Laureate International Universities (2). A investigadora que reuniu o grupo assumiu as funções de moderadora e redatora.

O guião (ANEXO IV) foi seguido de forma semiestruturada, permitindo total liberdade na exposição das ideias. Foi realizada a transcrição do grupo focal (Anexo V) a partir da gravação sonora efetuada.

**Pergunta 1** - Ao aceder a uma biblioteca digital de livro antigo, o que costuma encontrar?

Em resposta a esta pergunta, foi dada uma definição de biblioteca digital como “a transposição do conhecimento para o digital” (R5), e, em referência às grandes bibliotecas mundiais, foi dito que “a tendência é cada vez maior de anexarem objetos, museológicos, toda a qualidade de artefactos, livros, etc., com toda a informação, não só com a legenda, mas com toda a informação” (R1).

Um dos participantes reformulou a questão em função das expectativas, referindo que não se trata tanto do que se costuma encontrar, mas, sobretudo do que se espera encontrar, acrescentando: “Quando faço uma consulta na internet ou na biblioteca é o que eu espero, encontrar o livro” (R1).

A maior parte dos comentários incidiu sobre as funcionalidades existente no acesso ao conteúdo. No que respeita aos sumários, foi considerado que “o índice [sumário] é fundamental” (R5), mas, em contrapartida, foi referido que, os espólios digitalizados possuem poucos pontos de acesso e que, em geral, “não têm índices, ou [estão] mal associados às imagens” (R3). A este respeito, foram citados os casos da BND e Alma Mater que terão perdido sumários que anteriormente estavam disponíveis: “Aqueles sumários maravilhosos que antigamente a Biblioteca Nacional e a de Coimbra tinham e que deixaram de ter. Isto desgosta-me porque eu preciso de saber em que ponto do livro estou, qual é o assunto que está antes e o assunto que está depois” (R1). Outro participante, lembrou que “Havia um índice do texto e um índice de imagens” (R3), o que permitia consultar apenas a componente iconográfica da obra. No que respeita à

consulta, foi dito que o reconhecimento ótico de caracteres pode compensar a inexistência do sumário, no sentido em que “com o OCR, pode fazer-se a pesquisa dentro do conteúdo do livro” (R1), embora sublinhando que a maior parte das obras que encontra estão em formato imagem. Por seu turno, a língua original dos documentos também pode constituir um impedimento à consulta. Como referiu um dos participantes: “Faço muitas pesquisas em latim e aflijo-me porque não sei qual é aquela tradução” (R2).

Ainda no âmbito das funcionalidades, foi referida a possibilidade de transferência dos documentos – “Além de ler *on-line*, descarrego os livros” (R1) –, sendo que, neste caso, “A velocidade [de transferência] também [é fundamental]” (R5).

No âmbito desta pergunta, foi levantada a questão do pagamento dos serviços, embora a moderadora tenha elucidado que o acesso às bibliotecas digitais é gratuito e que “Existem aquelas que são pagas, mas são de editoras” (M). Contudo, um dos investigadores referiu que nos Estados Unidos da América se pagava, citando um exemplo: “Era uma biblioteca que tinha documentos sobre a expansão. Eles cobravam uma taxa simbólica [...]. uma pequena taxa” (R5).

Em síntese, aquilo que se encontra numa biblioteca digital de livro antigo são as obras digitalizadas e disponibilizadas, na íntegra e em acesso livre. Algumas obras, em formato imagem, têm sumários que auxiliam a localização das partes que constituem a obra e, por conseguinte, facilitam o acesso ao conteúdo. Nas obras em formato texto, é possível fazer a pesquisa por palavra e localizar a conteúdo pretendido. O mesmo se aplica aos elementos iconográficos, considerando que a imagem faz parte intrínseca da obra, em parceria com o texto e, não apenas, em complementaridade, necessitando

de sumários ou apontadores específicos. Ainda neste âmbito, foi referida a necessidade de traduzir os conteúdos de forma automática, o que recolheu o consenso do grupo, embora esta funcionalidade não seja um tema recorrente na literatura consultada. A perceção que os investigadores têm da biblioteca digital passa, por conseguinte, pela forma como esta permite a pesquisa e a recuperação da informação.

**Pergunta 2** – Quais as vantagens das Bibliotecas Digitais de Livro Antigo face às bibliotecas que possuem apenas obras em papel?

Esta questão revelou-se consensual e unânime: “Só tem vantagens. Permite-nos ter acesso às bibliotecas todas do mundo” (R2)

A principal vantagem prende-se com a acessibilidade, isto é, com “a possibilidade de fazer a pesquisa alargada em todas as bibliotecas e em qualquer tempo e lugar” (R1). Salientou-se, ainda, a vantagem da biblioteca digital enquanto reprodução e a possibilidade de usar a cópia como divulgação, citando alguns exemplos de exposições museológicas. “Numa exposição, o facto de os documentos poderem ser digitalizados permite fazer trabalhos completamente diferentes, imaginando que está ali o documento original” (R5). Como notou um dos participantes, esta é uma “questão mais sociológica. Essa divulgação abrange mais grupos sociais, é menos classista. Pode ir ao encontro de pessoas que se calhar entraram ali sem saber o porquê... com o olhar menos educado em relação ao documento...” (R6).

A outra vantagem referida prende-se com a conservação dos originais: “Tem todas as vantagens, até a preservação. Poupa os livros” (R3)

Em síntese, as vantagens oferecidas pelas Bibliotecas Digitais de Livro Antigo, face às bibliotecas que possuem apenas obras em papel, são inequívocas: do ponto de vista patrimonial e das bibliotecas detentoras do original, em termos de preservação; do ponto de vista do utilizador, em termos de acessibilidade que, através delas, se alarga a públicos que habitualmente não frequentam as bibliotecas.

**Pergunta 3** – Quais os inconvenientes que assinala na utilização da Biblioteca Digital de Livro Antigo?

A resposta a esta questão não obteve o consenso da anterior e causou alguma estranheza. No entanto, sobretudo os investigadores mais relacionados com a temática do livro, acabaram por referir que “Há coisas num livro que no digital às quais não temos acesso: páginas de dimensão diferentes... as marcas de água... não temos a perspetiva do tamanho” (R3). Em contrapartida, quando um dos participantes referiu a ausência do “cheiro” (R6), retorquiu-lhe que “Isso é mais sentimental” (R3). Direcionando este assunto para o estudo do livro, justificou que esta questão “não é só uma desvantagem” (R3), por impedir a investigação neste âmbito, salientando que “é mesmo uma impossibilidade, [porque] quem estuda questões físicas tem de ver o objeto ao vivo” (R3), além de que “a encadernação, as texturas... isso perde-se...” (R3). Por outro lado, um investigador mencionou que “a investigação torna-se isolada” (R1), recorrendo a um exemplo da sua experiência pessoal: “Quando vou à BN acabo por encontrar outras pessoas. Às vezes torna-se interessante. Trocamos ideias, sabemos o que os outros estão a fazer. No digital, não é possível” (R1).

As desvantagens da biblioteca digital de livro antigo prendem-se, essencialmente, com um campo específico da investigação: o estudo e a história do livro, enquanto objeto, não em relação ao conteúdo. Por outro lado, é mais abrangente em relação à socialização e à permuta intelectual entre investigadores.

**Pergunta 4** – Quais as vantagens para o ensino?

Para o ensino, as bibliotecas digitais têm “as mesmas vantagens” (R6) que para os utilizadores em geral. A acessibilidade é, aqui, subjacente, no que se refere à utilização dos conteúdos para a elaboração do material didático, referindo-se que a “lógica das bibliotecas digitais permite criar dispositivos mais didáticos e interessantes” (R6). A utilização efetua-se de diversas formas: “Às vezes recomendo [obras], outras, descarrego com mais facilidade. Disponibilizo os acessos” (R2). A vantagem decorrente prende-se com a adequação aos hábitos do aluno, no sentido em que “Para o aluno de hoje em dia é mais atrativo do que deslocar-se à biblioteca que passou a ser um sítio maldito” (R6).

Para o caso específico do ensino das línguas foi referida a “possibilidade de ouvir o livro na própria língua” (R4).

Para o ensino, a vantagem principal é a possibilidade de criação de materiais didáticos, recorrendo à reprodução ou à divulgação do acesso a fontes documentais.

**Pergunta 5** – Quais as vantagens para a investigação?

No que se refere à investigação, a principal vantagem referida foi a economia de meios, diminuindo os custos inerentes, e de tempo: “O que eu gosto nas bibliotecas digitais é que me baixaram o custo da investigação. E o tempo. Ficou mais barato, porque já não preciso de me deslocar às bibliotecas. Deixei de precisar de pedir coisas, fotocópias e pagá-las. E levo menos tempo a reunir a informação” (R4). Por outro lado, a disponibilização das obras permite um acesso mais alargado às fontes e abrange “Públicos mais generalistas. Não só os que têm possibilidade de ir aos EUA, ou para aqueles que tiveram uma bolsa” (R5).

Para a investigação, a vantagem principal reside na facilidade de acesso às fontes que, no caso das disciplinas que trabalham com o livro antigo, são muitas vezes exemplares únicos ou raros, dispersos em bibliotecas geograficamente distantes e onde o acesso pode ser condicionado ou reservado.

**Pergunta 6** – Que funcionalidades seriam necessárias para tornar a utilização das Bibliotecas Digitais de Livro Antigo mais eficaz?

As funcionalidades necessárias para otimizar as bibliotecas digitais prendem-se com as perceções enunciadas na pergunta 1, a partir das quais se formularam as hipóteses de desenvolvimento.

A funcionalidade principal prende-se com a pesquisa: por um lado, por “sofisticar os motores de busca” (R4), mas de forma transparente para o utilizador, ou seja, “que essa complexidade fosse do lado de lá” (R3). No mesmo sentido, um dos participantes referiu que “o grande desafio vai ser tornar os processos mais acessíveis” (R2). No caso

particular da pesquisa nos motores de busca, foi referido que será útil que as bibliotecas disponibilizem os seus catálogos digitais na web, permitindo um acesso direto aos respetivos conteúdos: “Gostava de encontrar os livros na pesquisa do *google*, sem ter de ir a uma biblioteca específica” (R4). Um dos participantes referiu a pesquisa de obras em contexto, ao afirmar que “Gostava que, quando faço a pesquisa para um livro, ele me indicasse outros livros semelhantes que houvesse na mesma biblioteca digital” (R1), estabelecendo um paralelismo com a procura de obras em bibliotecas com consulta em livre acesso: “Quando vou a uma biblioteca analógica, física, por exemplo, numa biblioteca universitária, vou à estante e vejo o livro e os livros próximos que lá estão” (R1).

Em articulação com a pesquisa, a recuperação da informação constitui outro aspeto a desenvolver, nomeadamente, na pesquisa no documento: “Gostava de conseguir encontrar a parte do texto que eu pretendo. Pesquisa por palavra, dentro do documento” (R1).

Ainda que as bibliotecas digitais tendem a privilegiar o conteúdo textual, os investigadores, neste grupo, em particular, os que estudam história da arte, mostraram-se particularmente interessados no conteúdo iconográfico: “Gostava de saber o assunto da imagem que lá está, as ilustrações, as gravuras, gostava de as identificar. Eu procuro sobretudo imagens... os livros que eu procuro têm ilustrações. As bibliotecas devem promover esse conhecimento” (R2). No entanto, também em relação ao texto, um dos participantes afirmou que gostaria que “O formato disponibilizado fosse acessível, por exemplo para fazer uma citação, dentro do documento. Sem ter de estar a escrever o texto” (R4).

Neste âmbito e atendendo a que “quando se fala em livro antigo, a maior parte deles são em latim” (R2) e que “hoje, o conhecimento em latim é limitadíssimo” (R2), um investigador considerou que “falta aqui um trabalho de tradução que as bibliotecas podiam promover” (R2), acrescentando que não se trata apenas “da tradução do latim, [mas] de outras línguas... árabe, chinês” (R2). No mesmo sentido, um outro participante confirmou que algumas das suas pesquisas devolvem resultados em línguas que não domina e que “se houvesse essa possibilidade [tradução], queria ter o acesso ao conteúdo desse livro” (R5). Foi salientado o facto de o acesso à informação, atualmente, passar pelo desconhecimento das línguas e que este é um fator que limita a investigação, quer por parte dos investigadores portugueses em relação a outras culturas – “A cultura asiática... está praticamente tudo vedado, tudo naquelas línguas” (R6) –, mas também “o contrário também para os investigadores lá fora” (R2). Como solução, propôs que “podia haver uma colaboração com as universidades” (R2).

Ainda no âmbito das funcionalidades, foi referida a necessidade de incrementar os serviços de alerta, refletindo a prática corrente das aquisições em linha: “Depois, que me desse a informação também das outras obras que viessem a aparecer, como quando se está na Amazon. Faço uma compra e, passado tempo, sugerem outros. Também podem fazê-lo para bibliotecas. Queria que aparecesse um alerta: isto já está disponível” (R1). No mesmo sentido, sugeriu ainda a possibilidade de criação de perfis personalizados, que segmentassem a informação em função dos respetivos interesses: “Deviam ter perfis de utilizador, é fácil criá-los” (R1).

Outras questões, surgidas ao longo da discussão, foram a autenticidade e legibilidade dos conteúdos e a gratuitidade dos serviços. A fim de preservar a reprodução fidedigna

do original, um dos participantes afirmou: “Não quero carimbos digitais nos livros, podem estar sobre a informação, podem perturbar a leitura” (R1), ao que outro acrescentou “Não quero nada que comprometa o conteúdo” (R5). Esta questão dos carimbos digitais prende-se, também, com a questão dos direitos. Como foi referido em relação à BNP, “Se estiverem a pôr o livro antigo cá fora, podem cobrar direitos de utilização” (R3). O que conduz à questão da gratuitidade de acesso às bibliotecas digitais.

Este terá sido um dos pontos mais discutidos. Enquanto um afirmava que “O património e a sua utilização é de todos” (R5), outros objetavam que “Isto tem custos. É preciso é que isto cresça. Sou a favor do utilizador pagador” (R2) e “não choca que [alguns serviços] sejam cobrados. Uma espécie de taxa moderadora que ajudasse a manter a biblioteca”, ou que “as pessoas preferem pagar e ter um serviço de qualidade a não o ter” (R5). Afirmaram a possibilidade de pagamento, como contrapartida a serviços de qualidade: “Mesmo que tenha de pagar alguns serviços, por exemplo, se eu quiser imagens de qualidade, posso pagar” (R1).

Todo este desenvolvimento tem subjacente uma expectativa de melhoria: “Aumentem os espólios digitalizados” (R3).

Em regra, os investigadores do grupo pretendem uma biblioteca digital com conteúdos estruturados e com auxiliares de navegação, tanto para o conteúdo textual, como para o conteúdo iconográfico ou visual das obras. Os mecanismos de pesquisa devem ser melhorados de forma a que a recuperação da informação seja mais eficaz, propondo algumas vias de desenvolvimento: disponibilizar os catálogos nos motores de busca; sofisticar os motores de busca das bibliotecas digitais; incrementar o acesso em

contexto. As obras em formato texto são uma exigência, a que se acrescenta a respetiva tradução. Insere-se, aqui, o conceito de que a informação está disponível, mas que o seu acesso ainda se mantém bloqueado pela barreira da língua.

A criação de perfis de utilizador e de sistemas de alerta personalizados permitiriam, às bibliotecas, um maior conhecimento acerca do utilizador, adequando-lhes a oferta, e, ao utilizador, manter-se atualizado acerca dos conteúdos e dos serviços disponibilizados pela biblioteca.

As bibliotecas são tendencialmente gratuitas, mas os investigadores admitem o pagamento de serviços excecionais de qualidade.

Subjacente a todas as propostas de melhoria, está o incremento dos conteúdos digitais, no caso, do livro antigo.

## **6.2. Inquérito por questionário**

Em função dos dados obtidos na reunião do grupo focal foi estruturado um inquérito (Vd. Anexo VII). O inquérito foi respondido em linha. Foram obtidas 19 respostas, com uma taxa média de conclusão de 95%. O tempo médio de resposta foi de 8 minutos.

A maioria dos respondentes (9) tem idades compreendidas entre os 40 e 49 anos; houve 4 respondentes no intervalo etário entre 50 e 59 anos e outros 4 com mais de 60 anos. Abaixo dos 40 anos, houve 1 respondente no intervalo entre 30 e 39, não se registando nenhum na casa dos 20 anos.

No que respeita ao género, 11 respondentes são mulheres e 7 são homens, tendo havido 1 respondente que não assinalou nenhuma opção. A maioria dos investigadores (9) são doutorados e 4 têm pós-doutoramento; houve, ainda, 5 respondentes com

mestrado. Um dos respondentes não assinalou o grau académico. A área de investigação (Gráfico 2) mais representada foi História e Arqueologia (7), seguida de Línguas e Literatura (5) e Arte, na especialidade de História da Arte (3). Outras áreas referidas são: Ciências da Comunicação; Humanidades Digitais e Humanística Digital; Filosofia, Ética e Religião; e Sociologia (Antropologia e Etnologia); Ciência da Informação; Biblioteconomia; Estudos de Cultura; e Museologia e Património.

Gráfico 2 – Respondentes, por área de investigação (ronda 1)



Dada a dimensão da amostra, foi efetuada uma análise qualitativa dos dados recolhidos, verificando as tendências das respostas.

**Pergunta 1** - Ao aceder a uma biblioteca digital de livro antigo, o que costuma encontrar?

- 1.1. Informação organizada, de forma validada e credível
- 1.2. Obras disponibilizadas em acesso livre
- 1.3. Obras disponibilizadas na íntegra

1.4. Obras disponibilizadas com pontos de acesso explícitos, por exemplo, sumários

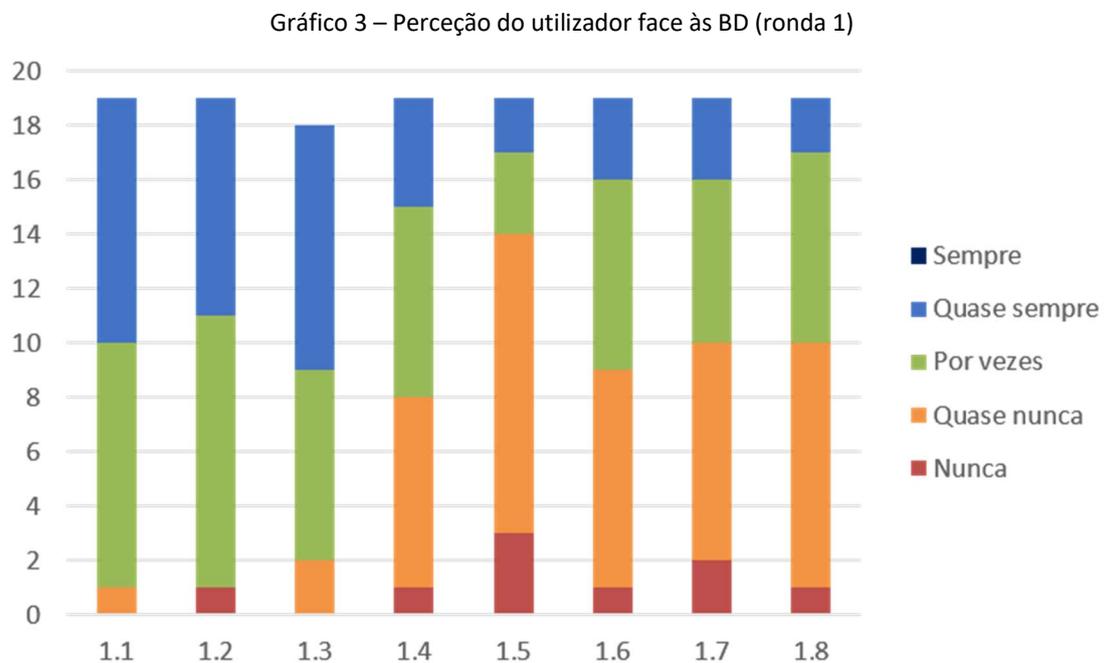
1.5. Identificação dos elementos gráficos das obras

1.6. Pesquisa nos elementos textuais dos documentos (OCR)

1.7. Exportação direta das referências bibliográficas para os gestores de referências bibliográficas

1.8. Partilha das obras digitalizadas nas redes sociais

Os resultados obtidos estão sintetizados no Gráfico 3.



Um respondente não assinalou nenhuma opção na resposta 1.3. A opção “Sempre” não foi assinalada em nenhuma das respostas.

Este conjunto de perguntas reflete a forma como os investigadores percecionam as bibliotecas digitais que utilizam.

As respostas mais consensuais indicam que a maior parte dos investigadores reconhece que a informação se encontra quase sempre organizada, de forma validada

e credível (1.1) e que costuma encontrar as obras disponibilizadas em acesso livre (1.2) e na íntegra (1.3). Apenas um afirma nunca ter encontrado as obras em acesso livre.

Em relação à pergunta 1.1, os respondentes consideram que a informação está, quase sempre (9), ou por vezes (9), bem organizada. Em relação às perguntas 1.2 e 1.3, encontram quase sempre as obras em acesso livre (8) e na íntegra (9); ou encontram-nas quase sempre em acesso livre (8) e na íntegra (7).

Por outro lado, a identificação dos elementos gráficos das obras (1.5), nunca (3), ou quase nunca (11) é encontrada.

No que respeita à pesquisa ou navegação das obras, a tendência das respostas não é muito definida. O número de respondentes que encontram, quase sempre (4), ou por vezes (7), pontos de acesso ao conteúdo das obras (1.4), como, por exemplo, sumários, é ligeiramente superior aos que nunca (1) ou quase nunca (7), os encontram. A diferença é ainda menos expressiva entre os que encontram quase sempre (3), ou por vezes (7), a funcionalidade de pesquisa no texto (1.6) e os que nunca (1) ou quase nunca (8), a encontram.

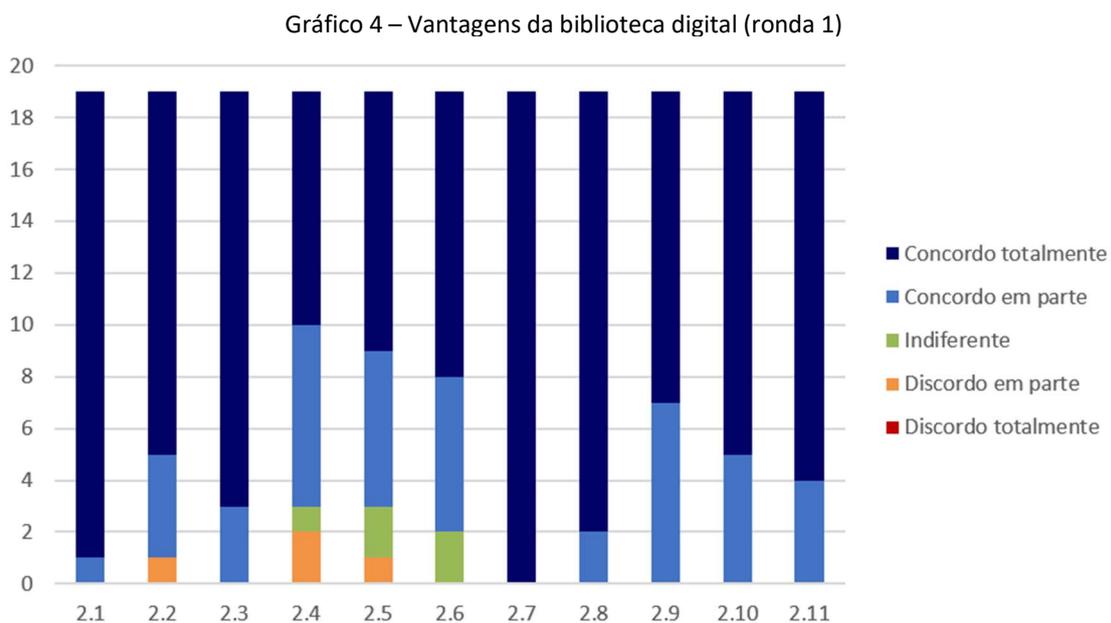
A distribuição de respostas é equilibrada entre os que encontram, ou não, as funcionalidades de exportação direta das referências bibliográficas para os gestores de referências bibliográficas (1.7) ou de partilha das obras digitalizadas nas redes sociais (1.8).

**Pergunta 2** - Quais as vantagens da biblioteca digital de livro antigo face às bibliotecas físicas?

2.1. Acesso às obras em qualquer lugar e em qualquer horário

- 2.2. Maior facilidade de leitura por pessoas com limitações físicas
- 2.3. Preservação dos documentos originais
- 2.4. Maior legibilidade do texto
- 2.5. Preservação do conteúdo intelectual da obra
- 2.6. Acesso a fontes documentais no contexto da atividade letiva
- 2.7. Diminuição de custos na investigação decorrente das deslocações e das reproduções
- 2.8. Otimização do processo de investigação pela redução do tempo na recolha dos dados
- 2.9. Novas leituras e análises através da aplicação de novas ferramentas (contagem de palavras, estudos de contexto, georreferenciação, etc.)
- 2.10. Alargamento da informação, como geradora de oportunidades
- 2.11. Divulgação da cultura e do conhecimento

Os resultados obtidos estão sintetizados no Gráfico 4.



Todas as respostas tiveram opções assinaladas. A opção “Discordo totalmente” não foi assinalada em nenhuma das respostas.

Este conjunto de perguntas incide sobre as vantagens que os investigadores atribuem às bibliotecas digitais.

O grau de concordância dos respondentes em todas as respostas foi elevado, pelo que o reconhecimento das vantagens trazidas pelas bibliotecas digitais é consensual.

A concordância total foi unânime quanto ao benefício das bibliotecas digitais no que respeita à diminuição dos custos da investigação decorrentes das deslocações e das reproduções (2.7).

À exceção de um respondente, que apenas concorda parcialmente, todos os restantes concordam totalmente que o acesso às obras em qualquer lugar e em qualquer horário (2.1) é vantajoso. Também uma larga maioria de respondentes (18) concorda totalmente que o processo de investigação é otimizado pela redução do tempo na recolha de dados (2.8).

Também no que respeita ao contributo das bibliotecas digitais para a preservação dos documentos originais (2.3), houve um elevado índice de concordância total (16).

Houve, ainda, três aspetos em que a concordância acerca das vantagens das bibliotecas digitais foi generalizada, embora com maior variação entre os respondentes que concordam totalmente ou parcialmente: divulgação da cultura e conhecimento (2.11); alargamento da informação como geradora de oportunidades (2.10); e novas leituras e análises através da aplicação de novas ferramentas digitais (2.9). A cada uma destas respostas, a concordância total foi global em, respetivamente, 15, 14 e 12 respostas.

No que respeita ao proveito retirado do acesso às fontes documentais para a atividade letiva, a concordância também foi ampla, mas houve respondentes que se afirmaram indiferentes (2).

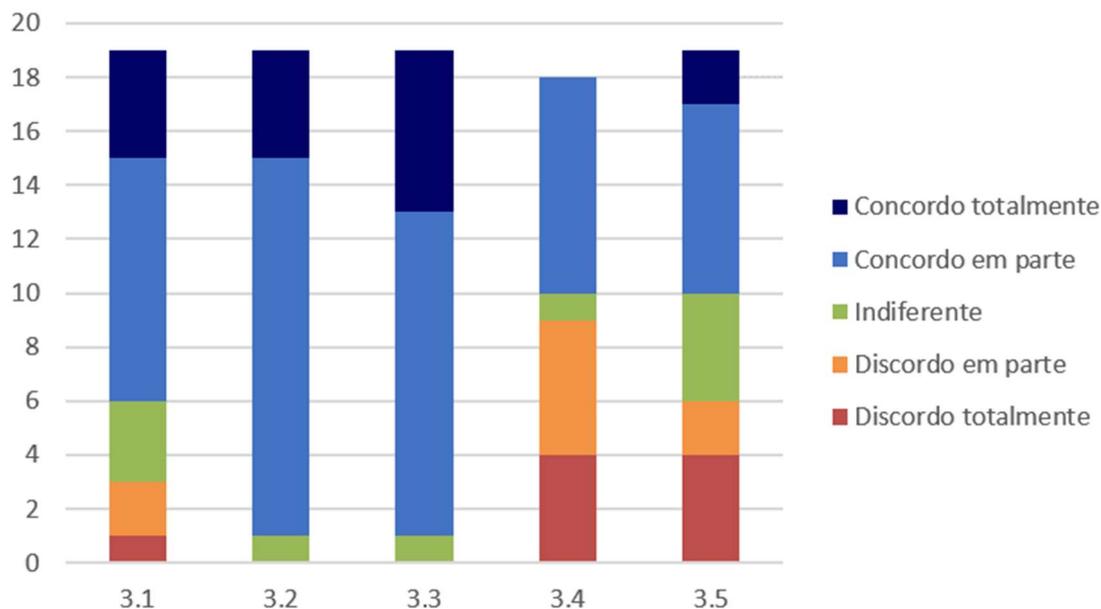
Apenas se registou discordância parcial no que respeita à maior facilidade de leitura por pessoas com limitações físicas (2.2), à maior legibilidade do texto (2.4) e à preservação do conteúdo intelectual da obra.

**Pergunta 3** – Quais as desvantagens das bibliotecas digitais de livro antigo em relação às bibliotecas físicas?

- 3.1. Falta de contacto físico com livro
- 3.2. Diminuição do acesso aos aspetos materiais da obra
- 3.3. Elementos móveis das obras, não são representáveis digitalmente na sua dinâmica
- 3.4. Dificuldade no processamento da informação, decorrente do excesso de resultados
- 3.5. Impossibilidade de socialização e a troca de ideias entre pares

Os resultados obtidos estão sintetizados no Gráfico 5.

Gráfico 5 – Desvantagens da biblioteca digital (ronda 1)



Um respondente não assinalou nenhuma opção na resposta 3.4.

Este conjunto de perguntas incide sobre as desvantagens que os investigadores atribuem às bibliotecas digitais.

As principais desvantagens indicadas são o facto de os elementos móveis da obra não serem representáveis digitalmente na sua dinâmica original (3.3) e a diminuição do acesso aos aspetos materiais da obra (3.2). Em cada uma destas respostas, houve um respondente que se afirmou indiferente, mas não houve qualquer discordância.

A maioria dos respondentes (13) concorda, quer totalmente (4), quer parcialmente (9), que uma das desvantagens das bibliotecas digitais é a ausência de contacto com o livro físico (3.1), tendo havido respondentes (3) que se afirmaram indiferentes.

Em relação à dificuldade no processamento da informação decorrente do excesso de resultados (3.4), houve uma discrepância entre os que concordam parcialmente (3) e os que discordam total (4) ou parcialmente (5). Sem que tenha havido nenhuma concordância total, um respondente mostrou-se indiferente.

A maior dispersão nas respostas verificou-se em relação à impossibilidade de socialização e de troca de ideias entre pares (3.5), tendo sido também aqui que se registou o maior número de respondentes indiferentes (4).

**Pergunta 4** – Quais as suas expectativas de mudança/otimização nas bibliotecas digitais de livro antigo?

- 4.1. Pesquisa no conteúdo dos documentos a partir dos motores de busca da web
- 4.2. Possibilidade de pesquisa no conteúdo textual dos documentos
- 4.3. Possibilidade de pesquisa das ilustrações pelo respetivo tema
- 4.4. Pesquisa das ilustrações autónoma em relação ao texto
- 4.5. Ligação entre as ilustrações e o texto que lhe corresponde
- 4.6. Exportação das obras em múltiplos formatos, nomeadamente, para dispositivos móveis
- 4.7. Criação de auxiliares à navegação (ajudas interativas e tutoriais)
- 4.8. Criação de pontos de acesso ao conteúdo das obras (por exemplo, sumários, índices)
- 4.9. Visualização da referência à parte da obra onde o leitor se encontra
- 4.10. Possibilidade de criação de perfis de utilizador
- 4.11. Possibilidade de criação de várias listas em cada perfil de utilizador
- 4.12. Possibilidade de partilha de listas pessoais
- 4.13. Possibilidade de personalização do sistema de alertas
- 4.14. Possibilidade de navegação, a partir de uma obra, em obras semelhantes

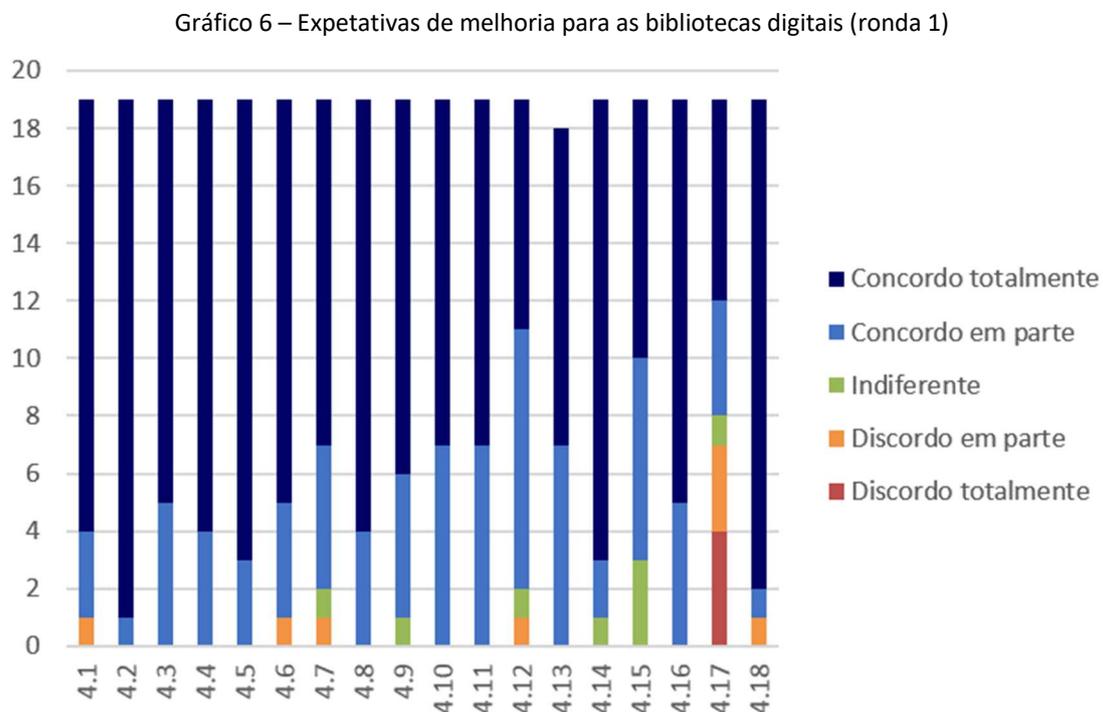
4.15. Possibilidade de tradução automática dos documentos em linha

4.16. Possibilidade de melhorar o reconhecimento de caracteres de forma colaborativa

4.17. Acesso a obras digitais de qualidade mediante pagamento de taxas moderadoras

4.18. Incremento da colocação em linha de conteúdos em português

Os resultados obtidos estão sintetizados no Gráfico 6.



Um respondente não assinalou nenhuma opção na resposta 4.13.

Este conjunto de perguntas reflete as expetativas dos investigadores em relação ao desenvolvimento, ou otimização, das bibliotecas digitais.

Os níveis mais elevados de expetativas dizem respeito à possibilidade de pesquisa no conteúdo textual dos documentos (4.2) e ao incremento da coleção em linha de

conteúdos em português (4.18). Em relação a este último aspeto, apenas um respondente discorda em parte.

A concordância, total ou parcial, foi unânime acerca dos seguintes pontos, apresentados por ordem decrescente de concordância total: a ligação entre as ilustrações e o texto que lhe corresponde (4.5); pesquisa das ilustrações, autónoma em relação ao texto (4.4); criação de pontos de acesso aos conteúdos das obras, como, sumários e índices (4.8); possibilidade de pesquisa das ilustrações pelo respetivo tema (4.3); possibilidade de melhorar o reconhecimento de caracteres de forma colaborativa (4.16); possibilidade de criação de perfis de utilizador (4.10); possibilidade de criação de várias listas em cada perfil de utilizador (4.11); possibilidade de personalização do sistema de alerta (4.13).

Houve aspetos em relação aos quais não houve discordância, mas onde alguns respondentes se afirmaram indiferentes: visualização da referência à parte da obra onde o leitor se encontra (4.9); possibilidade de navegação, a partir de uma obra, em obras semelhantes (4.14); possibilidade de tradução automática dos documentos (4.15).

Houve um registo de discordância parcial em cada um dos seguintes aspetos: pesquisa no conteúdo dos documentos a partir dos motores de busca da web (4.1); exportação das obras em múltiplos formatos, nomeadamente, para dispositivos móveis (4.6); incremento da colocação em linha de conteúdos em português (4.18); criação de auxiliares à navegação, como as ajudas interativas e tutoriais (4.7); possibilidade de partilha das listas pessoais (4.12). Em cada um dos dois últimos aspetos, houve também o registo de uma resposta indiferente.

O acesso a obras digitais de qualidade, mediante pagamento de taxas moderadoras, foi o ponto com maior dispersão de resultados, embora tendencialmente positivos: 12 concordam totalmente e 4, parcialmente, contra 4 que discordam totalmente e 3, parcialmente. Um respondente afirmou-se indiferente.

Para a segunda ronda, o questionário foi alterado em função dos resultados (Vd. Anexo VIII), tendo sido retiradas algumas perguntas que se revelaram menos relevantes e as restantes foram ordenadas por ordem de decrescente de concordância.

O inquérito foi feito em linha, nos mesmos moldes que o anterior. Foram obtidas 17 respostas, com uma taxa de conclusão de 100%. O convite para participar foi enviado ao mesmo conjunto de especialistas, mas, sendo anónimo, não é possível confirmar em que medida os indivíduos desta amostra coincidem com a da primeira ronda. O tempo médio de resposta foi de 5 minutos, portanto, menos 3 minutos que o anterior.

A maioria dos respondentes situa-se no intervalo etário entre os 40 e 59 anos: 6, com idades compreendidas entre os 40 e 49 anos; 5, com idades entre 50 e 59 anos. Houve, ainda 4 respondentes com idades acima dos 60 anos e 2, no intervalo entre 30 e 39. No que respeita ao género, 11 respondentes são mulheres e 6 são homens. A maioria dos investigadores (9) são doutorados e 5 têm pós-doutoramento, tendo havido 3 respondentes com mestrado.

Gráfico 7 – Respondentes, por área de investigação



A área de investigação (Gráfico 7) mais representada foi História e Arqueologia (6), seguida de Línguas e Literatura e Estudos da Cultura, cada uma assinalada por 4 respondentes, e de Arte, na especialidade de História da Arte, assinalada por 3. As áreas de Ciências da Comunicação e Informação, Humanidades Digitais e Humanística Digital e Museologia e Património foram, cada uma, assinaladas por 2 respondentes. Foram, ainda, referidas as áreas de Filosofia, Ética e Religião e de Sociologia (Antropologia e Etnologia). A caracterização desta amostra é, em termos genéricos, idêntica à da anterior. Dada a sua dimensão, foi também efetuada uma análise qualitativa dos dados recolhidos, verificando as tendências das respostas.

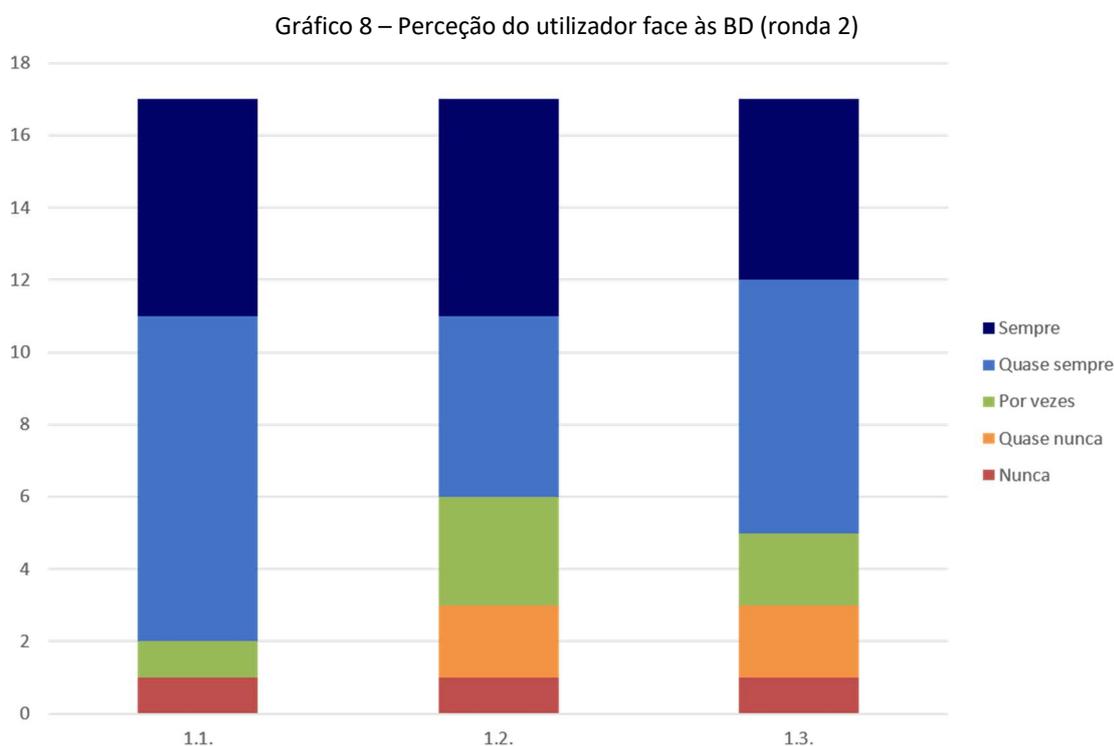
**Pergunta 1** - Ao aceder a uma biblioteca digital de livro antigo, o que costuma encontrar?

1.1. Informação organizada, de forma validada e credível

1.2. Obras disponibilizadas em acesso livre

1.3. Obras disponibilizadas na íntegra

Os resultados obtidos estão sintetizados no Gráfico 8.



Em cada uma das perguntas, a opção “nunca” foi escolhida uma vez. A resposta mais consensual foi a pergunta 1.1, onde a grande maioria dos respondentes consideram que a informação está sempre (6), ou quase sempre (9) organizada, de forma validada e credível e apenas 1 respondente optou pela opção “por vezes”.

Na pergunta 1.2, o grau de concordância foi ligeiramente inferior, dado que, embora a maioria dos respondentes tenham considerado que as obras estão sempre (6) em acesso livre, houve menos respondentes a achar que está quase sempre (5), apenas

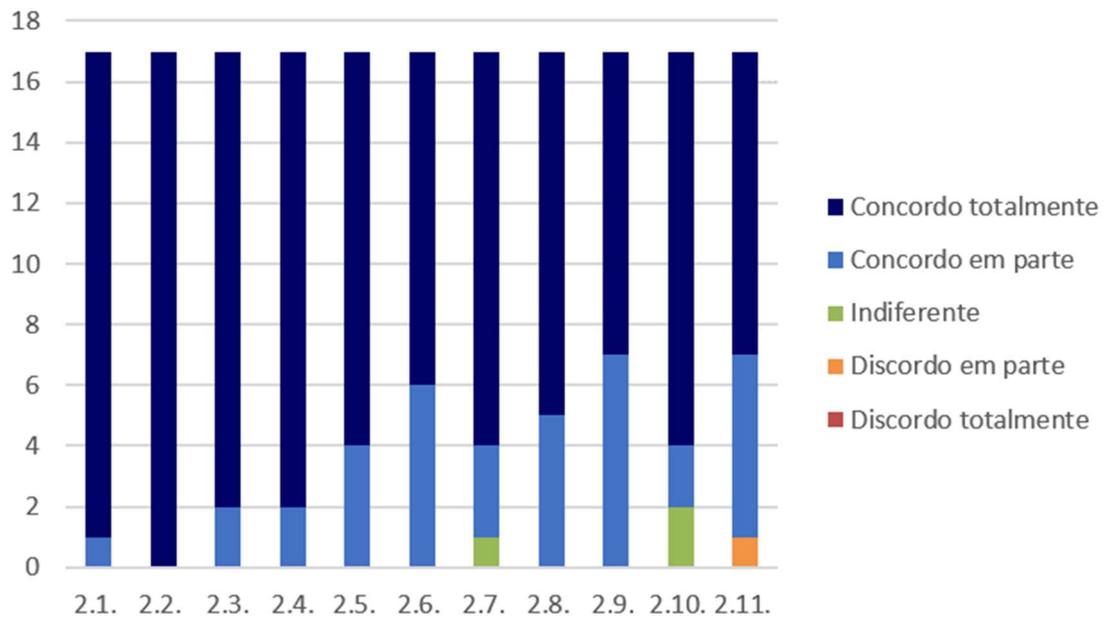
por vezes (3) ou quase nunca (2). O consenso foi idêntico na pergunta 1.3, tendo 5 respondentes considerado que as obras estão sempre (5) ou quase sempre disponibilizadas na íntegra, mas havendo quem refira quem só as encontra por vezes (2), ou quase nunca (2).

**Pergunta 2** - Quais as vantagens da biblioteca digital de livro antigo face às bibliotecas físicas?

- 2.1. Diminuição de custos na investigação decorrente das deslocações e das reproduções
- 2.2. Acesso às obras em qualquer lugar e em qualquer horário
- 2.3. Otimização do processo de investigação pela redução do tempo na recolha dos dados
- 2.4. Preservação dos documentos originais
- 2.5. Divulgação da cultura e do conhecimento
- 2.6. Alargamento da informação, como geradora de oportunidades
- 2.7. Novas leituras e análises através da aplicação de novas ferramentas (contagem de palavras, estudos de contexto, georreferenciação, etc.)
- 2.8. Maior facilidade de leitura por pessoas com limitações físicas
- 2.9. Acesso a fontes documentais no contexto da atividade letiva
- 2.10. Preservação do conteúdo intelectual da obra
- 2.11. Maior legibilidade do texto

Os resultados obtidos estão sintetizados no Gráfico 9.

Gráfico 9 – Vantagens da biblioteca digital (ronda 2)



A opção “Discordo totalmente” não foi assinalada em nenhuma das respostas.

Este conjunto de perguntas incide sobre as vantagens que os investigadores atribuem às bibliotecas digitais. Em todas as perguntas, a opção “Concordo totalmente” foi a mais escolhida, tendo sido o índice 10/17 o mais baixo.

O grau de concordância dos respondentes em todas as respostas foi muito elevado, pelo que o reconhecimento das vantagens trazidas pelas bibliotecas digitais é consensual. A concordância total foi unânime quanto ao benefício das bibliotecas digitais no que respeita ao acesso às obras em qualquer lugar e em qualquer horário (2.2) e quase unânime em relação à diminuição dos custos da investigação decorrentes das deslocações e das reproduções (2.1), onde apenas 1 respondente que concordou parcialmente, e em relação à otimização do processo de investigação (2.3) e à preservação dos documentos originais (2.4), em cada uma das quais, houve 2 respondentes a assinalar a opção “Quase sempre”.

A concordância também foi sempre positiva nas questões 2.5, 2.6., 2.8 e 2.9, onde os respondentes concordam totalmente que os seguintes aspetos são vantajosos: a divulgação da cultura e do conhecimento (13); a maior facilidade de leitura por pessoas com limitações físicas (12); o acesso a fontes documentais no contexto da atividade letiva (10); o alargamento da informação como geradora de oportunidades (11).

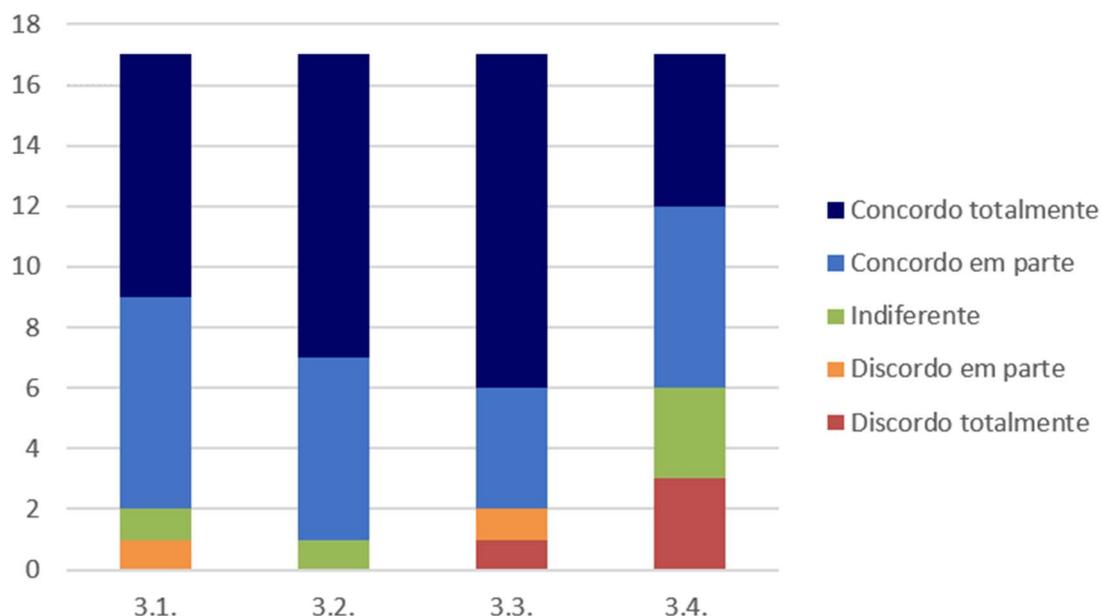
A opção “Discordo em parte” foi assinalada uma única vez na resposta 2.11., relativa à maior legibilidade do texto. Houve também respondentes que se afirmaram indiferentes em relação a novas leituras e análises através da aplicação de novas ferramentas (contagem de palavras, estudos de contexto, georreferenciação, etc.) (1) e à preservação do conteúdo intelectual da obra (2).

**Pergunta 3** – Quais as desvantagens das bibliotecas digitais de livro antigo em relação às bibliotecas físicas?

- 3.1. Elementos móveis das obras, não são representáveis digitalmente na sua dinâmica
- 3.2. Diminuição do acesso aos aspetos materiais da obra
- 3.3. Falta de contacto físico com livro
- 3.4. Impossibilidade de socialização e a troca de ideias entre pares

Os resultados obtidos estão sintetizados no Gráfico 10.

Gráfico 10 – Desvantagens da biblioteca digital (ronda 2)



Este conjunto de perguntas incide sobre as desvantagens que os investigadores atribuem às bibliotecas digitais. A opção “Concordo totalmente” foi a mais escolhida nas perguntas 3.2 e 3.3. O índice mais baixo verificou-se na pergunta 3.4, relativa à impossibilidade de socialização e de troca de ideias entre pares, onde houve maior dispersão nas respostas: 5 concordaram totalmente; 6 concordaram parcialmente; 3 manifestaram-se indiferentes; 3 discordaram totalmente.

A única pergunta sem discordância foi a 3.2, relativa à diminuição do acesso aos aspetos materiais da obra, onde apenas 1 respondente se mostrou indiferente.

A pergunta onde a concordância total (11) foi maior, foi a 3.3, relativa à falta de contacto físico com o livro, havendo, no entanto, 1 respondente a discordar totalmente e 1 parcialmente.

Relativamente à impossibilidade de representar digitalmente os elementos móveis da obra na sua dinâmica, a concordância foi total (8) ou parcial (7), mas houve 1 respondente indiferente e 1 a discordar parcialmente.

A área de investigação (Gráfico 7) mais representada foi História e Arqueologia (6), seguida de Línguas e Literatura e Estudos da Cultura, cada uma assinalada por 4 respondentes, e de Arte, na especialidade de História da Arte, assinalada por 3. As áreas de Ciências da Comunicação e Informação, Humanidades Digitais e Humanística Digital e Museologia e Património foram, cada uma, assinaladas por 2 respondentes. Foram, ainda, referidas as áreas de Filosofia, Ética e Religião e de Sociologia (Antropologia e Etnologia).

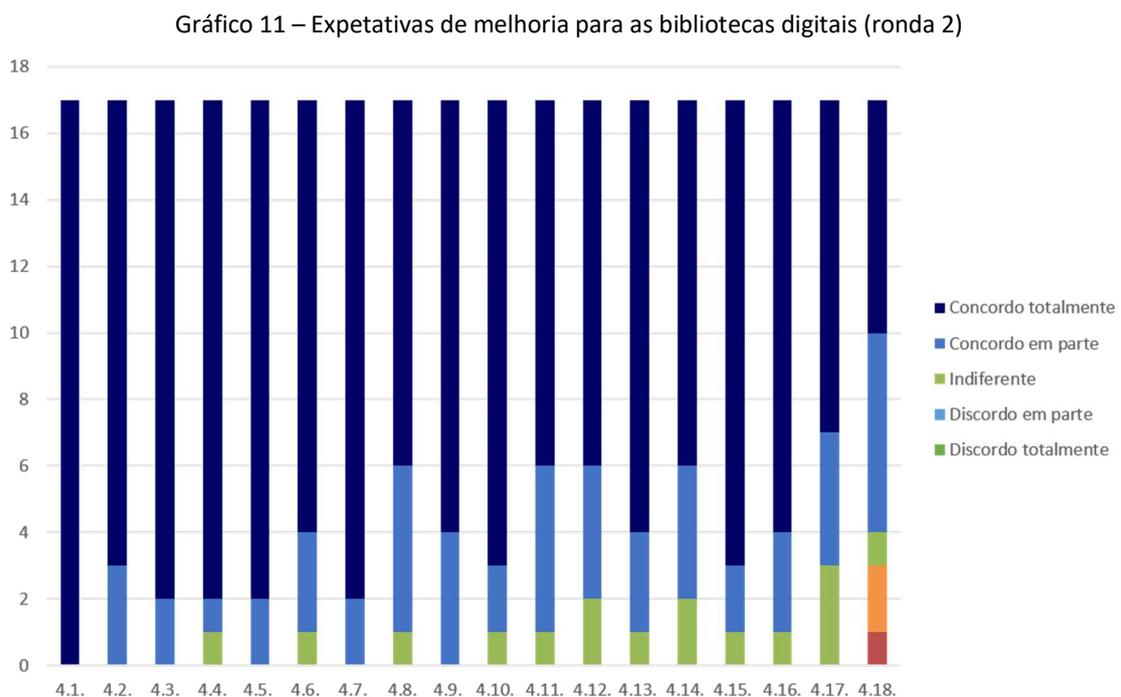
A caracterização desta amostra é, em termos genéricos, idêntica à da anterior. Dada a sua dimensão, foi também efetuada uma análise qualitativa dos dados recolhidos, verificando as tendências das respostas.

**Pergunta 4** – Quais as suas expectativas de mudança/otimização nas bibliotecas digitais de livro antigo?

- 4.1. Possibilidade de pesquisa no conteúdo textual dos documentos
- 4.2. Ligação entre as ilustrações e o texto que lhe corresponde
- 4.3. Possibilidade de navegação, a partir de uma obra, em obras semelhantes
- 4.4. Pesquisa das ilustrações autónoma em relação ao texto
- 4.5. Criação de pontos de acesso ao conteúdo das obras (por exemplo, sumários, índices)
- 4.6. Incremento da colocação em linha de conteúdos em português
- 4.7. Possibilidade de pesquisa das ilustrações pelo respetivo tema
- 4.8. Possibilidade de melhorar o reconhecimento de caracteres de forma colaborativa

- 4.9. Pesquisa no conteúdo dos documentos a partir dos motores de busca da web
- 4.10. Visualização da referência à parte da obra onde o leitor se encontra
- 4.11. Possibilidade de criação de perfis de utilizador
- 4.12. Possibilidade de criação de várias listas em cada perfil de utilizador
- 4.13. Exportação das obras em múltiplos formatos, nomeadamente, para dispositivos móveis
- 4.14. Possibilidade de personalização do sistema de alertas
- 4.15. Criação de auxiliares à navegação (ajudas interativas e tutoriais)
- 4.16. Possibilidade de tradução automática dos documentos em linha
- 4.17. Possibilidade de partilha de listas pessoais
- 4.18. Acesso a obras digitais de qualidade mediante pagamento de taxas moderadoras

Os resultados obtidos estão sintetizados no Gráfico 11.



Este conjunto de perguntas reflete as expectativas dos investigadores em relação ao desenvolvimento, ou otimização, das bibliotecas digitais.

A concordância total foi unânime em relação à possibilidade de pesquisa no conteúdo textual dos documentos (4.1). O índice mais baixo de concordância verificou-se na pergunta 4.18, relativa ao pagamento de uma taxa para acesso a obras digitais de qualidade, onde houve também maior dispersão nas respostas: 7 concordaram totalmente; 6 concordaram parcialmente; 1 manifestou-se indiferente; 2 discordaram parcialmente; 1 discordou totalmente.

A concordância, total ou parcial, foi unânime acerca dos seguintes pontos, apresentados por ordem decrescente de concordância total: possibilidade de navegação, a partir de uma obra, em obras diferentes (4.3); criação de pontos de acesso aos conteúdos das obras, como, sumários e índices (4.5); possibilidade de pesquisa das ilustrações pelo respetivo tema (4.7); ligação entre as ilustrações e o texto que lhe corresponde (4.2); pesquisa no conteúdo dos documentos, a partir dos motores de busca da web (4.9).

Os níveis de concordância, genericamente superiores em relação à primeira ronda, foram reorganizados por ordem diferente, tendo havido manifestações de neutralidade onde na ronda anterior a concordância fora sempre positiva: pesquisa das ilustrações, autónoma em relação ao texto (4.4); possibilidade de melhorar o reconhecimento de caracteres de forma colaborativa (4.8); possibilidade de criação de perfis de utilizador (4.11); possibilidade de criação de várias listas em cada perfil de utilizador (4.12); possibilidade de personalização do sistema de alerta (4.14).

Outras perguntas sem discordância e onde houve registo de neutralidade foram: incremento da colocação em linha de conteúdos em português (4.6); possibilidade de criação de várias listas em cada perfil de utilizador (4.12); exportação das obras em múltiplos formatos, nomeadamente, para dispositivos móveis (4.13); criação de auxiliares à navegação (ajudas interativas e tutoriais) (4.15); possibilidade de tradução automática dos documentos em linha (4.16); possibilidade de partilha de listas pessoais (4.17).

O grau de concordância aumentou em todas as baterias de perguntas. Dado que a amostra é reduzida, os dados não são tratados estatisticamente. Para obter o grau de concordância, foi atribuída uma ponderação a cada uma das opções de resposta, conforme explícito na Tabela 23; para obter o grau de concordância por respondente, o grau de concordância obtido foi dividido pelo número de respondentes em cada amostra.

Tabela 23 – Fator de ponderação das respostas

<b>Opções de resposta</b>		<b>Ponderação</b>
Nunca	Discordo totalmente	1
Quase nunca	Discordo em parte	2
Por vezes	Indiferente	3
Quase sempre	Concordo em parte	4
Sempre	Concordo totalmente	5

Nas perguntas relativas à perceção das bibliotecas digitais pelos utilizadores (1), a média de concordância em função do número de respondentes aumentou de 2,8 (valor máximo: 5), na primeira ronda, para 3,9 (valor máximo: 5), na segunda, mantendo-se a ordem de concordância entre as perguntas. Os investigadores da

amostra reconhecem que as bibliotecas digitais disponibilizam a informação de forma organizada, de forma validada e credível, e encontram as obras na íntegra e em acesso livre.

Nas perguntas relativas às vantagens das bibliotecas digitais, reconhecidas pelos utilizadores (2), a média de concordância por número de respondentes manteve-se em 4,7 (valor máximo: 5), onde se detetou um ligeiro aumento de 0,07, embora a ordem de preferência não tenha sido exatamente a mesma. Genericamente, as principais vantagens das bibliotecas digitais são, em primeiro lugar, a ubiquidade de acesso, ou seja, a possibilidade de consulta em toda a parte e ao mesmo tempo, e a otimização dos recursos derivada da redução do tempo na recolha de dados e, conseqüentemente, a diminuição dos custos inerentes à investigação.

Nas perguntas relativas às desvantagens das bibliotecas digitais, reconhecidas pelos utilizadores (3), a média de concordância por número de respondentes aumentou de 3,6 (valor máximo: 5), na primeira ronda, para 4,3 (valor máximo: 5), na segunda. As principais desvantagens apontadas são, por um lado, a diminuição do acesso aos aspetos materiais da obra e, por outro lado, com a falta de contacto físico com o livro.

Nas perguntas relativas às expetativas de desenvolvimento e melhoria que os investigadores têm acerca das bibliotecas digitais (4), a média de concordância por número de respondentes aumentou de 4,6 (valor máximo: 5), na primeira ronda, para 4,7 (valor máximo: 5), na segunda. As principais expetativas de desenvolvimento das bibliotecas digitais relacionam-se com a localização dos conteúdos textuais e iconográficos na totalidade da obra: a possibilidade de pesquisa no conteúdo textual dos documentos; a pesquisa das ilustrações pelo respetivo tema e autónoma em

relação ao texto; a ligação entre as ilustrações e o texto que lhe corresponde; a criação de pontos de acesso ao conteúdo das obras (por exemplo, sumários, índices); a pesquisa no conteúdo dos documentos a partir dos motores de busca da web. Para melhorar a pesquisa no conteúdo textual dos documentos, afigura-se, como hipótese, a possibilidade de melhorar o reconhecimento de caracteres de forma colaborativa.

Outro aspeto a melhorar que reuniu grande consenso entre os respondentes foi a necessidade de aumentar os conteúdos em língua portuguesa. Por outro lado, um dos resultados mais relevantes prende-se com a necessidade de traduzir, quer as línguas mortas, como o latim e o grego antigo, quer as línguas orientais, para poder incrementar o acesso a fontes que estão disponíveis em linha, mas cujo conteúdo não é compreensível para grande parte dos investigadores. Ou seja, fica aqui implícito que, sendo a disponibilização em linha um dado adquirido, as expectativas de melhoria das bibliotecas digitais de livro antigo prendem-se agora com a criação de ferramentas que permitam tornar inteligível o respetivo conteúdo intelectual.

## 7. Conclusões

Os incunábulos e o livro antigo, que têm estado resguardados em bibliotecas geralmente com acesso reservado ou limitado a alguns utilizadores, estão a ser submetidos a uma transferência de suporte, do papel para o digital, passando a ser disponibilizados em linha através das bibliotecas digitais, dada a reconhecida importância como fonte para a investigação na área das Humanidades. Apesar das melhorias efetuadas nas últimas duas décadas, o processo ainda se mantém muito próximo da mera transferência de suportes, sem adicionar vantagens, ou aproveitar a totalidade dos benefícios que a tecnologia permite.

As principais limitações registadas referem-se à incipiente edição do conteúdo textual e iconográfico dos documentos digitalizados. As bibliotecas digitais que realizam a transcrição dos textos, ou não efetuam a sua marcação, ou os ficheiros disponibilizados não têm qualidade suficiente para extrair o texto e marcá-lo. As ilustrações inseridas nos documentos, na maioria dos casos, não são referidas; só em casos residuais é feito algum tratamento específico à iconografia. Os investigadores em Humanidades ultrapassaram o preconceito que os separava da tecnologia digital e reconhecem que as bibliotecas digitais constituem um recurso fundamental para a investigação. No entanto, identificam algumas limitações no acesso aos respetivos conteúdos, considerando que os modelos de disponibilização são demasiado básicos, sem acrescentar valor à consulta do texto impresso, nomeadamente, a falta de pontos de acesso aos diversos componentes estruturais da obra e a elementos complementares do texto. Cumprindo a missão inerente às Humanidades Digitais, norteadas por princípios de articulação entre os dois domínios do conhecimento,

sugere-se que, para lá da colaboração já habitual entre bibliotecários e informáticos, também a comunidade académica seja chamada a intervir para que a seleção dos conteúdos seja a mais representativa e a sua organização mais eficaz.

### **7.1. Contexto da investigação**

Esta investigação centrou-se, por um lado, na análise dos modelos de disponibilização de incunábulos e livro antigo em bibliotecas digitais constituídas por coleções digitalizadas de documentos e, por outro, na forma como estes são utilizados pelos investigadores na área das Humanidades, com o objetivo de contribuir com propostas para a otimização do acesso aos conteúdos nas bibliotecas digitais.

Com esse intuito foi efetuado, em primeiro lugar, uma revisão da literatura que incidiu no estudo das Humanidades Digitais e das bibliotecas digitais para compreender o estado da investigação, procurando descrever os respetivos problemas e identificar as soluções que se encontram disponíveis, ou em teste, no âmbito de programas experimentais.

Em seguida, foi efetuada a observação direta de bibliotecas digitais, selecionadas, no universo das que disponibilizam incunábulos e livro antigo em linha, em função da respetiva relevância para a investigação em Humanidades. No decurso da observação, os serviços e as funcionalidades de cada biblioteca foram registados e sistematizados numa grelha de avaliação.

O estudo dos incunábulos e do livro antigo foi organizado em três componentes sequenciais: definição das tipologias a partir da literatura consultada; análise do conjunto das obras publicadas entre 1450 e 1558, disponibilizadas em linha na BND,

usando uma grelha de avaliação, para registrar os dados relativos às características físicas e particularidades formais, passíveis de interferir no processo de transferência de suporte; análise da edição digital da obra tipográfica de Aldo Manuzio num estudo transversal às várias bibliotecas digitais analisadas, com o objetivo de avaliar a forma como objetos de características idênticas são disponibilizadas em linha nas várias bibliotecas.

Por último, o estudo centrou-se na pesquisa, através de questionários, aplicados a investigadores em Humanidades, com o objetivo de identificar as suas opiniões em relação às bibliotecas digitais. Nesse sentido, foi realizado um grupo focal para identificar as expectativas de utilização das BDLA, bem como as respetivas vantagens e desvantagens e as expectativas de melhoria. Em função da análise dos resultados do grupo focal, foi desenhado um inquérito por questionário, efetuado em duas rondas, para validar as opiniões recolhidas.

## **7.2. Objetivos e conclusões**

Verificou-se, ao longo das várias fases da investigação, que o estudo dos incunábulos e do livro antigo em suporte digital continua pertinente, dado que ainda não se encontram fixadas modalidades suficientemente eficientes para a acessibilidade destas obras em linha.

O primeiro objetivo enunciado neste trabalho foi analisar um conjunto representativo de bibliotecas digitais que disponibilizam incunábulos e livro antigo, o que permitiu extrair algumas ilações.

Confirma-se que uma biblioteca, seja em espaço físico ou virtual, tem subjacentes as mesmas tarefas de seleção, organização, manutenção, preservação e difusão das coleções. Assim, é possível comparar a navegação numa biblioteca digital à deambulação por entre as estantes na sala de leitura de uma biblioteca material: não sendo o método mais eficaz para encontrar uma obra específica, permite traçar uma ideia geral acerca da temática da biblioteca e dos respetivos conteúdos. Da mesma forma, a navegação em bibliotecas digitais com um volume considerável de obras pode ser uma atividade frustrante, que dificilmente conduz a resultados positivos, embora possa revelar-se útil nas primeiras consultas com carácter exploratório.

Nas bibliotecas digitais, a pesquisa pode ser feita através dos elementos bibliográficos das obras, sendo tanto mais assertiva quanto o sistema de informação permita a utilização de refinadores, ou na medida em que as pesquisas avançadas possam combinar os elementos bibliográficos com a pesquisa nos conteúdos e nos metadados do objeto digital.

A sinalética utilizada nas bibliotecas digitais tende a ser uniforme, permitindo uma interpretação quase intuitiva dos símbolos utilizados. A facilidade de acesso depende do esforço em utilizar os mesmos símbolos, com idêntica formalização e significado.

A maioria das bibliotecas digitais disponibiliza modalidades de ajuda quer para orientar os utilizadores nas primeiras utilizações, quer em operações mais elaboradas. Contudo, diferem tanto em relação à forma como a ajuda é facultada, como quanto ao nível de auxílio prestado: ajuda contextual, que descodifica a sinalética utilizada; página de F.A.Q., que permite a consulta de algumas questões previamente definidas pelo administrador; página de ajuda, com exemplos que permitem a exploração das

várias funcionalidades. Esta última página não está disponível em todas as bibliotecas digitais com a totalidade das funcionalidades existentes, apesar de se reconhecer que este procedimento seria fundamental para uma utilização mais eficiente. Nenhuma das bibliotecas analisadas facultava tutoriais em vídeo.

Dado que a constituição de bibliotecas digitais é um processo em aberto, com coleções em contínuo crescimento, seria vantajoso que os utilizadores fossem notificados relativamente a novas obras disponibilizadas em linha ou a alterações efetuadas. As bibliotecas digitais destacam as novidades onde, de forma mais ou menos exaustiva, são apresentadas as novas obras adicionadas, mas, regra geral, sem indicar a data de atualização da página. O sistema de alertas RSS, implementado nalgumas bibliotecas, permite a criação de alertas personalizados que correspondam aos interesses de cada utilizador, possibilitando que, em função do perfil definido, este seja mantido ao corrente das atualizações. Nalguns casos, a criação de um espaço pessoal permite guardar as referências bibliográficas em diversas listas.

A transcrição das obras não é uma prática sistemática em todas as bibliotecas digitais, ou em todas as obras das bibliotecas que o fazem, além de que a marcação e o nível de correção não são uniformes. Não obstante, a vantagem da transcrição, ao permitir efetuar pesquisas no conteúdo da obra, afigura-se consensual. Um dos problemas mais referidos relaciona-se com as diferentes grafias, dado que, em geral, o termo da pesquisa apenas recupera as palavras com a mesma grafia, ignorando as grafias semelhantes ou alternativas. Além disso, e em geral, nas bibliotecas também não são disponibilizados as normas e os critérios de transcrição e de transliteração, embora

estejam a ser implementadas soluções para efetuar a pesquisa através de palavras truncadas.

A marcação das obras, nomeadamente, em TEI, permite gerar cópias em múltiplos formatos, como HTML, PDF ou ePUB, entre outros, com a vantagem de facultar conteúdos pesquisáveis. Além disso, a marcação das obras possibilita o processamento informático com ferramentas específicas para a análise do conteúdo, estudos de contexto, contabilização de elementos no texto, de forma a gerar novo conhecimento.

A maioria das bibliotecas digitais analisadas permite a transferência das obras em formato PDF. No entanto, as imagens, mesmo quando possuem uma resolução suficiente para a legibilidade do texto, nem sempre suportam a aplicação de programas de OCR e a realização da respetiva transcrição.

As ferramentas para exportação dos registos bibliográficos são particularmente úteis no contexto da investigação, mas, apesar de já se encontrarem implementadas nas bibliotecas digitais, o acesso pode ser dificultado por não estarem devidamente assinaladas ou não se encontrarem na página da respetiva obra.

As obras digitalizadas, no seio da mesma biblioteca, não são uniformes, registando-se variações consideráveis na qualidade das imagens e da transcrição, nos formatos disponibilizados e na qualidade dos sumários. Por outro lado, a materialidade da obra não é considerada, sendo que a digitalização dificulta a perceção das dimensões reais das obras, da espessura, das lombadas, dos cortes ou das sobrecapas, assim como das filigranas e avergoados do papel.

Apesar de ser possível a introdução de apontadores nas obras, esta marcação é individual e não permite anotações individuais ou colaborativas, nem a criação de relações entre as obras.

A informação estatística sobre as obras disponibilizadas e os utilizadores não é prática corrente em Portugal, embora se considere a pertinência da disponibilização destes dados e da sua atualização periódica. Também, a política de divulgação das bibliotecas digitais, sobretudo das portuguesas, é pouco efetiva. Apesar de terem sítios eletrónicos próprios ou de estarem presentes nas principais redes sociais, não desenvolvem ações de formação ou de divulgação.

O segundo objetivo desta investigação era a caracterização dos incunábulos e do livro antigo através das obras disponibilizados em linha na Biblioteca Nacional Digital.

Foi possível confirmar que as primeiras obras impressas seguiam o modelo dos manuscritos, sem folha de título, as obras continuavam a iniciar com a fórmula dos manuscritos, *incipit*. Com o tempo surge uma folha com o título a anteceder a obra, a qual vai sendo progressivamente enriquecida, de forma pouco sistemática, com elementos descritivos, contendo dados bibliográficos (autor, título, data, impressor) e elementos decorativos ou com iconografia alusiva. Para o período analisado, o colofão continua a ser a principal fonte de informação sobre a obra.

A maioria dos impressos analisados na BND corresponde à tipologia monografia, seguida pelos folhetos. A língua dominante é o português, seguida pelo latim e pelo espanhol. As línguas menos representadas são o francês, o hebreu, e o italiano, registando-se obras bilingues, em hebreu/latim, latim/espanhol, português/latim e espanhol/português (uma de cada). Verifica-se a grande proliferação de locais de

edição, nos quais Lisboa surge em lugar cimeiro, seguida da cidade universitária de Coimbra e de Évora que, na altura, gozava de grande dinamismo cultural. Observa-se também a presença de locais estrangeiros, como Paris, Sevilha e Veneza.

No âmbito das especificidades do livro antigo, a maioria das obras disponibilizadas em linha na BND possuem capitais, caldeirões ou corações, paginação por assinaturas, encabeçamentos ou reclusos. As capitais podem ser mais ou menos ornamentadas, mas apresentam grande variedade de modelos, sendo possível identificar capitais diferentes para a mesma letra dentro da mesma obra. Também os reclusos e encabeçamentos podem aumentar o erro na versão textual das obras, pois requerem mais atenção na aplicação dos programas de OCR. Em contrapartida, os caldeirões e corações são muito semelhantes entre si, o que facilita a aplicação de algoritmos para a respetiva identificação.

As primeiras obras são em gótico, cuja transcrição e reconhecimento é mais difícil. Este tipo continua a ser utilizado sobretudo em obras de carácter religioso, enquanto se começa a implantar o romano, com grande legibilidade, e o itálico, cujo reconhecimento também é mais fácil. No entanto, continuam a ser frequentes obras com mais de um tipo de letra e de diferentes tamanhos, o que dificulta o reconhecimento automático.

As obras com ilustrações presentes na amostra são residuais.

No que se refere à dimensão, as obras situam-se entre o In-oitavo (altura de 10 a 22 cm) e o In-fólio (altura varia de 33 a 40 cm), mas a maioria dos registos bibliográficos não facultam a indicação das dimensões reais da obra, o que é, potencialmente, perturbador do fluxo de trabalho de digitalização.

O terceiro objetivo era avaliar as características específicas dos incunábulos e livro antigo e a forma como afetam a edição digital das obras, desde a digitalização à disponibilização em linha.

As diferentes paginações das obras, sem numeração nos fólios ou apenas com numeração nas páginas reto (foliação), com blocos de sequências em numeração árabe, romana ou sem numeração, com as sequências das páginas dadas por assinaturas ou por reclamo, ou com erros na paginação, constituem um dos principais obstáculos à correta verificação da versão digital da obra, conferindo eventuais falhas na digitalização.

A encadernação das obras, por vezes, demasiado apertada, dificultando a respetiva abertura, ou as obras demasiado aparadas, com margens muito pequenas, podem perturbar a leitura integral do texto. Verifica-se, também, que a digitalização pode truncar a imagem da página e compromete, sobretudo, a leitura dos cabeçalhos e das assinaturas.

Algumas obras apresentam elementos móveis, nomeadamente as de Astronomia e de Medicina, que apresentam esquemas dinâmicos, cujas propriedades são difíceis de transferir para o suporte digital. Nestes casos, para lá da leitura à distância, a obra em formato digital não apresenta vantagens e, pelo contrário, a leitura é mais limitada. No conjunto das obras analisadas, não foram encontradas soluções adequadas a estas situações, apresentando apenas uma imagem estática do elemento numa determinada posição fixa. Em nenhuma dos casos observados, foi introduzida uma imagem em movimento que simulasse a leitura da obra analógica.

No caso dos desdobráveis, a digitalização também não apresenta as funcionalidades da obra original. Os desdobráveis destinavam-se a ser mantidos abertos durante a consulta da obra; por isso, a parte que corresponde à página da obra e ao verso é mantida em branco. Porém, na versão digital as zonas em branco costumam ser eliminadas e a imagem do desdobrável é inserida na sequência regular da paginação. Também neste caso e no conjunto das obras analisadas, não foi encontrada qualquer solução que permitisse manter a opção de visualização do desdobrável durante a leitura do texto correspondente.

Da mesma forma, a versão digital dificulta a observação dos elementos iconográficos ou cartográficos apresentados em duas páginas, quer porque o original tem a encadernação muito apertada, impedindo a visualização do conjunto, quer porque a digitalização se faz página a página e as imagens são, posteriormente, editadas para formar uma imagem única, sem resolver o problema de concordância e apresentando deslocamentos, sobreposições ou redundâncias.

Os elementos textuais que remetem para o conteúdo da obra, como índices, tabuadas ou índices, são apenas reproduzidos em formato digital, sem interação que permita remeter diretamente para o respetivo conteúdo.

A dimensão dos originais, demasiado grandes ou pequenos, também condiciona a digitalização. A digitalização de obras com lombadas superiores a 10 cm, mesmo quando efetuada em digitalizadores com compensador de lombada, podem originar imagens com uma curvatura acentuada junto à margem interna, o que dificulta a leitura da página em ecrã e a transcrição do texto com programas de OCR.

De resto, são diversos os fatores que introduzem ruído nas imagens e, por conseguinte, comprometem a eficácia da aplicação destes programas para realizar a

transcrição de texto: a transparência do papel, trazendo a interferência do texto existente no verso da folha; o desgaste dos caracteres usados na impressão, tornando a grafia imprecisa; as letras capituladas com diferentes decorações, dificultando a definição de um padrão de identificação; as notas impressas nas margens laterais das obras, com um tipo de letra demasiado pequeno, tornando os caracteres indefinidos; as glosas, criando caixas de texto cuja representação é difícil de reproduzir. Também provocam interferências indesejáveis, os aspetos derivados do uso ou do estado de conservação da obra: notas manuscritas e texto sublinhado; lacunas, galerias, manchas de humidade, sujidade e oxidação da tinta ferrogálica, muito usada nas notas manuscritas. Do mesmo modo, a disposição do texto em colunas, a formatação do texto com vários tipos de letras e a utilização de duas ou mais cores, perturbam a transcrição do texto através de programas de OCR. As obras em mais de uma língua também apresentam problemas consideráveis na transcrição automatizada do texto, dado que, em regra, os programas apenas utilizam um dicionário.

A maioria das obras disponibilizadas em linha apresenta um sumário que facilita a navegação no respetivo conteúdo. No entanto, os sumários têm uma qualidade variável, sendo que, umas vezes, refletem a estrutura completa da obra, e noutras, apenas indicam, de forma pouco precisa, as partes principais da obra ou conjuntos de capítulos. Da mesma forma, o estilo dos sumários não é uniforme. Por vezes, é feita a transcrição do texto, respeitando a grafia da época, mas a abertura das abreviaturas não é uniforme. Nas obras em latim, é frequente fazer a tradução das partes das obras e capítulos para português. Além disso, em regra, os sumários apresentam o conteúdo textual das obras, sem referir eventuais ilustrações ou outros elementos gráficos.

As obras sem sumário apenas apresentam, como auxiliar de navegação, o número sequencial da imagem digitalizada, o qual, na maioria dos casos, não corresponde à paginação do original.

Por conseguinte e como apreciação geral, regista-se uma elevada falta de uniformização na colocação dos conteúdos em linha e na divulgação dos critérios utilizados para tal.

O quarto objetivo era identificar as funcionalidades que os investigadores em Humanidades gostariam de ver implementadas nas bibliotecas digitais.

Para compreender a perceção dos investigadores face às bibliotecas digitais, foi realizado um grupo focal. A sua análise permitiu elaborar um conjunto de questões que foram estruturadas num inquérito por questionário, em duas rondas, junto de especialistas para validar as conclusões do grupo focal.

Na ótica dos investigadores que participaram no grupo focal, a biblioteca digital fornece informação organizada, validada e credível e em acesso livre, e as obras encontram-se digitalizadas na íntegra. Consideraram que a navegação no interior da obra deve ser auxiliada, através da indicação de pontos de acesso específicos, como os livros, partes, capítulos, índices, folha de título, erratas, indicação dos cadernos, colofão, etc. Ou seja, a estrutura da obra e as suas diversas componentes devem ser registadas no sumário que acompanha a obra digital. O mesmo foi referido em relação à identificação das ilustrações (elementos iconográficos, heráldicos, cartográficos), tabelas, desdobráveis ou outros elementos. Sublinharam, ainda, que os textos devem ser transcritos para permitir a pesquisa no conteúdo textual dos documentos. A

exportação direta das referências bibliográficas para os gestores de referências bibliográficas e a partilha das obras, incluindo as redes sociais, também foi referida.

As duas rondas, do inquérito por questionário, confirmaram, com um elevado grau de concordância, as sugestões apontadas no grupo focal.

São manifestamente sublinhadas, por todos os investigadores, as vantagens das bibliotecas digitais face às suas congéneres físicas, em particular, por permitirem o acesso às obras em qualquer lugar e em qualquer horário. Além disso, também referiram outros atributos relevantes: a maior facilidade de leitura por pessoas com limitações físicas; a conservação dos documentos originais, que podem deixar de ser manuseados para consulta; a maior legibilidade do texto, que pode ser aumentado consoante a vontade do leitor; a difusão do conteúdo intelectual da obra; o acesso a fontes documentais no contexto da atividade letiva; a otimização do processo de investigação pela redução do tempo na recolha dos dados; a diminuição de custos na investigação decorrente das deslocações e das reproduções; a possibilidade de novas leituras e análises através da aplicação de ferramentas digitais; o alargamento da informação, como geradora de oportunidades e de divulgação da cultura e do conhecimento.

No entanto, também foram referidas algumas desvantagens, ou limitações, das bibliotecas digitais, como a diminuição ou, mesmo, a inibição de acesso aos aspetos materiais da obra e, nomeadamente, o facto dos elementos móveis incluídos nas publicações não serem digitalmente representados na sua dinâmica. Para além destes aspetos, os investigadores referiram, igualmente, a falta de contacto físico com livro, a dificuldade no processamento da informação decorrente do excesso de resultados

obtidos nas pesquisas e a impossibilidade de socialização e de troca de ideias entre pares.

Neste processo, foram apresentadas algumas sugestões de melhoria para as bibliotecas digitais, as quais podem ser sistematizadas em três vertentes: acesso às obras em linha; personalização dos acessos; serviços adicionais.

No acesso às obras em linha foi referido que seria benéfico se a obra pudesse ser acedida a partir da pesquisa de excertos nos motores de busca, o que implicaria que todas as obras digitalizadas fossem transcritas. Além disso, as obras deveriam possuir pontos de acesso ao conteúdo das obras (por exemplo, sumários, índices) e permitir a pesquisa no conteúdo textual dos documentos, o que, mais uma vez, exigiria a transcrição das obras.

No que se refere ao conteúdo gráfico, foi referida a possibilidade de pesquisa das ilustrações pelo respetivo tema, acarretando um trabalho de investigação adicional, eventualmente, em colaboração com a academia, para a descrição dos elementos iconográficos e heráldicos. Os investigadores defenderam, ainda, que a pesquisa das ilustrações devia ser autónoma em relação ao texto, por exemplo, através de um sumário ou base de ilustrações, e que fosse fomentada a ligação entre as ilustrações e o texto que lhe corresponde, tendo em vista a compreensão global do conteúdo da obra.

Atendendo a que, na leitura em linha, o leitor facilmente perde a referência à parte da obra onde se encontra, foi considerada a necessidade de visualizar o ponto de leitura em relação com o todo, por exemplo, mantendo presente a indicação das respetivas páginas ou parte da obra.

Em relação à personalização dos acessos foi referido que, à semelhança do que já existe nalgumas bibliotecas, seria interessante a possibilidade de criação de perfis de utilizador e, dentro de cada perfil, permitir a criação e a gestão de várias listas e, mesmo, a sua partilha. Além disso, os sistemas de alerta seriam mais úteis se permitissem a personalização, emitindo alertas em função de uma tipologia de obra, coleção, autor, etc., associada às preferências definidas em cada perfil.

Foram, ainda, referidos alguns serviços adicionais que as bibliotecas digitais deveriam implementar. Considerando que há um universo de informação disponível em linha, mas cujo conteúdo se mantém inacessível por barreiras linguísticas, foi apontada a necessidade de implementar a tradução automática dos documentos em linha. Os investigadores sugeriram, ainda, que a transcrição dos textos fosse melhorada através de trabalho colaborativo envolvendo os leitores, em geral, e a academia, em particular. Outros aspetos a melhorar são a possibilidade de navegação entre obras semelhantes, a exportação das obras em múltiplos formatos e a criação de auxiliares à navegação, para uma utilização mais racional das bibliotecas digitais.

O quinto objetivo era propor um conjunto de diretivas que adequassem as bibliotecas digitais à investigação em Humanidades.

Considerando que a pesquisa é habitualmente referida como um dos pontos fracos das bibliotecas digitais, sugere-se a implementação de modalidades de pesquisa das obras por autor, título ou extratos de texto, através dos motores de busca da web, de forma direta, à semelhança do que acontece na pesquisa filtrada em livros e artigos, com a criação de um filtro ou separador que limitasse a pesquisa ao universo das bibliotecas digitais. No mesmo sentido, sugere-se a organização de coleções digitais, autorizando

a pesquisa das obras em contexto específico. O objetivo é tornar a pesquisa mais complexa do que existe atualmente para obter resultados mais adequados, mantendo a simplicidade de utilização.

Dado que a pesquisa nos OPAC pode não devolver o conteúdo pretendido, seria vantajoso o seu enriquecimento através do cruzamento dos elementos das fichas bibliográficas com os metadados do objeto digital, com o conteúdo da obra, com as listas facetadas ou de assuntos, ou através da pesquisa em cada um destes elementos.

No que se refere aos elementos iconográficos das obras, sugere-se uma pesquisa através de descritores normalizados, como, por exemplo, o ICONCLASS, ou através do reconhecimento da imagem.

Também a possibilidade de pesquisa por termos truncados seria proveitosa. Dada a existência de diferentes grafias, derivadas a erros cometidos por quem faz a pesquisa ou por quem adicionou a informação, a pesquisa por termos truncados permitiria obter uma lista de termos que se adequem, devidamente ordenada por similitude.

Para que o acesso em linha às obras seja eficiente é essencial utilizar formatos normalizados nas várias fases do processo e endereços persistentes para as obras.

Ao digitalizar as obras, um dos aspetos cruciais é a numeração das imagens. Por conseguinte, a designação individual das imagens digitais deve incluir uma referência que permita associá-la à obra e à respetiva sequência, além da indicação das várias componentes, como capa, guarda, folha de título, página em branco, página manuscrita, ilustração (iconografia, heráldica, cartografia), desdobrável, guarda, ou a página expressa na obra. Deve, ainda, ser definida a forma de numerar as obras sem

paginação ou com a foliação dada por assinaturas, e proceder em conformidade e de forma coerente em toda a coleção.

Propõe-se que os conteúdos transcritos indiquem a respetiva taxa de concordância da transcrição com o original, referindo igualmente o programa e as regras de transcrição utilizadas. O mesmo se sugere para a marcação em TEI. Os textos devem ser transcritos e marcados, considerando-se que será vantajosa a exibição simultânea do texto transcrito e da respetiva imagem digital.

Prevê-se a criação de índices para cada obra e para coleções de obras. Os sumários devem refletir toda a estrutura da obra, quer para o conteúdo textual, quer para as ilustrações. Em complemento, na leitura em linha, deve ser exibida a localização da página na respetiva estrutura da obra.

Sobretudo para a investigação e para o ensino, seria útil a criação de perfis individuais que permitissem aceder a funções complementares, como a criação e a partilha de bibliografias específicas, ou o acesso a imagens das obras numa resolução maior. Os investigadores também poderiam colaborar com a biblioteca digital através da transcrição e estruturação das obras de forma colaborativa. A marcação e a anotação das obras de forma individual ou coletiva, incluindo as ilustrações, poderá conduzir a um estudo mais profundo das obras. A construção de produtos digitais que ajudem a compreender e a divulgar as obras, poderia ser um trabalho desenvolvido entre as bibliotecas digitais e a academia, com benefícios para a população em geral.

A criação de alguns serviços adicionais afigura-se, igualmente, profícua: a tradução dos conteúdos das obras; a exportação direta dos dados bibliográficos das obras para os gestores de referências; a disponibilização de páginas de ajuda, com pequenos

tutoriais; a construção de alertas personalizáveis sem implicar conhecimentos de programação.

A disponibilização das normas e procedimentos seguidos na construção da biblioteca digital contribuiria para que os utilizadores percecionassem as propriedades das mesmas, facilitando uma utilização mais arguta e eficaz. Da mesma forma se defende a divulgação dos dados estatísticos relativos à disponibilização das obras em linha e dos seus leitores, bem como a respetiva atualização periódica, divulgando toda a sua produção, como uma estratégia essencial para cativar investidores e leitores.

Ao construir ou ao modificar o aspeto de uma biblioteca digital, a sinalética adotada deve seguir a tendências verificada nas outras bibliotecas e na web, selecionando símbolos com funções definidas e facilmente identificáveis, mantendo, obviamente, essas funções.

O utilizador das bibliotecas digitais, na atualidade, pretende que os acessos e os conteúdos lhe sejam facultados à sua medida (*tailoring access to content*). As bibliotecas digitais, além de divulgadoras do conhecimento, também são produtoras de novos conteúdos digitais, criados a partir dos seus acervos e de forma a corresponder às necessidades dos seus utilizadores. Colocar um livro em linha não se resume à mudança de suporte. A transferência de suporte do papel para o digital implica um trabalho moroso, metuculoso e rigoroso, para que a obra no novo suporte seja devidamente processada, acrescentando-lhe novas funcionalidades, de forma a permitir novas leituras e a gerar novo conhecimento.

### **7.3. Limitações do estudo**

As principais limitações enfrentadas ao longo deste trabalho derivam da carência de dados relativos às bibliotecas digitais portuguesas. Na revisão da literatura efetuada, foram encontrados poucos estudos acerca da realidade das bibliotecas digitais em Portugal e dos seus utilizadores, bem como dados estatísticos atualizados acerca das obras disponibilizadas em linha. Também não se encontraram dados relativos às normas e procedimentos seguidos, pelo que estes não puderam ser referidos no trabalho.

### **7.4. Linhas futuras de investigação**

No universo das bibliotecas digitais aplicadas à otimização das coleções em linha as vias de investigação a seguir correspondem ao alargamento deste estudo a outras áreas complementares e afins.

- Este estudo procurou definir quais as exigências dos investigadores em relação às bibliotecas digitais, pelo que a pesquisa deve ser ampliada de forma a abranger as preferências dos utilizadores, avaliando o seu grau de satisfação e motivação para a utilização das bibliotecas digitais.
- Considerando a evolução das Humanidades Digitais, a investigação deve manter-se ativa na observação e análise das exigências apresentadas pelos investigadores, para que as bibliotecas digitais lhes correspondam de forma adequada.

Assim, a investigação deve prosseguir no sentido de:

- Definir um conjunto de procedimentos para a edição digital, incluindo a elaboração dos sumários com os conteúdos textuais e iconográficos das obras.
- Criar instrumentos que permitam a avaliação dos desenhos (*layouts*) das bibliotecas digitais e das obras disponibilizadas em linha.
- Fazer um estudo extensivo da aplicação do TEI a conteúdos em português, conjuntamente com a aplicação do ALTO (Analyzed Layout and Text Object), tendo em vista a descrição do aspeto da obra física em meio digital.
- Estender o estudo a outras espécies bibliográficas como os periódicos, iconografia e cartografia.

Os incunábulos e o livro antigo possuem particularidades que exigem ações específicas. Contudo, os resultados deste estudo são replicáveis e os procedimentos inerentes a estas obras podem ser generalizados a outras tipologias monográficas.

A evolução das bibliotecas digitais passa pela articulação e colaboração entre bibliotecários, informáticos e investigadores. No entanto, o papel do bibliotecário continua basilar, competindo-lhe definir o rumo a seguir para que a biblioteca digital corresponda às expectativas dos utilizadores e se mantenha integrada na missão e nos objetivos da instituição, enquanto mediadora do conhecimento.

## Referências bibliográficas

- Aalberg, T., Papatheodorou, C., Dobрева, M., Tsakonas, G., & Farrugia, C. J. (Eds.). (2013). Research and advanced technology for digital libraries. In *International Conference on Theory and Practice of Digital Libraries, TPDL 2013, Valletta, Malta, September 22-26, 2013*. Springer. <http://doi.org/10.1007/978-3-642-40501-3>
- ADHO - Alliance of Digital Humanities Organizations. (s.d.). Acedido em 4 mar. 2014, em <http://adho.org/>
- Agosti, M., Ferro, N., & Orio, N. (2005). Annotating illuminated manuscripts: An effective tool for research and education. In *Proceedings of the 5th ACM/IEEE-CS Joint Conference on Digital Libraries* (pp. 121–130). New York, NY: ACM. <http://doi.org/10.1145/1065385.1065412>
- AHDig. (s.d.). Acedido em 13 ago. 2017, em <https://ahdig.org/>
- Al-Hajj, Y. A. A., & Küster, M. W. (2013). The text-image-link-editor: A tool for linking facsimiles and transcriptions, and image annotations. *Literary and Linguistic Computing*, 28(2), 190–198. <http://dx.doi.org/10.1093/lc/fqs067>
- Ampudia de Haro, F., Serafim, J., Cobra, J., Faria, L., Roque, M. I. R., Ramos, M., & Carvalho, P., Costa, R. (2016). *Investigação em ciências sociais: Guia prático do estudante*. Lisboa: Pactor.
- Anselmo, A. (1996). Fronteiras da história do livro. *Revista da Faculdade de Ciências Sociais e Humanas*, 32, 79–82. Acedido em <http://run.unl.pt/handle/10362/6829>
- Audenaert, N., & Furuta, R. (2009). Annotated facsimile editions: Defining macro-level structure for image-based electronic editions. *Literary and Linguistic Computing*, 24(2), 143. <http://doi.org/10.1093/lc/fqp008>
- Audenaert, N., & Furuta, R. (2010). What humanists want: How scholars use source materials. In *Proceedings of the 10th Annual Joint Conference on Digital libraries* (pp. 283–292). New York, NY: ACM. <http://doi.org/10.1145/1816123.1816166>
- Bailey Jr, C. W. (2008). Institutional repositories, tout de suite. *Digital Scholarship*, 10,

- s.p.. Acedido em <http://digital-scholarship.org/ts/irtoutsuite.pdf>
- Bainbridge, D., Thompson, J., & Witten, I. H. (2003). Assembling and enriching digital library collections. *Proceedings of the ACM/IEEE Joint Conference on Digital Libraries, 2003–January* (pp. 323–334). New York, NY: ACM.  
<http://doi.org/10.1109/JCDL.2003.1204885>
- Bainbridge, D., Twidale, M. B., & Nichols, D. M. (2011). That’s “é” not “p” “?” or “□”: A user-driven context-aware approach to erroneous metadata in digital libraries. In *Proceedings of the 11th Annual International ACM/IEEE Joint Conference on Digital Libraries* (pp. 39–48). New York, NY: ACM.  
<http://doi.org/10.1145/1998076.1998084>
- Bamman, D., Babeu, A., & Crane, G. (2010). *Multilingual alignment and projection*. Medford. Acedido em <http://www.perseus.tufts.edu/publications/jcdl27-bamman.pdf>
- Barbier, F. (2006). *L’Europe de Gutenberg: Le livre et l’invention de la modernité occidentale (XIII-XVI siècle)*. Paris: Belin.
- Barbier, F. (2012). *Histoire du livre en Occident* (3rd ed.). Paris: Armand Colin.
- Barbier, F. (2013). *Histoire des bibliothèques: D’Alexandrie aux bibliothèques virtuelles*. Paris: Colin.
- Barolini, H. (1992). *Aldus & his dream book: An Illustrated Essay*. [S.l.]. Italica Press.
- Barton, M. R., & Waters, M. M. (2004). *Creating an institutional repository: LEADIRS workbook*. Cambridge: MIT Libraries. Acedido em <http://hdl.handle.net/1721.1/26698>
- Bell, J. (2004). *Como realizar um projecto de investigação: Um guia para a pesquisa em ciências sociais e da educação* (3ª ed.). Lisboa: Gradiva.
- Berry, D. M. (2012). Introduction. In D. M. Berry (Ed.), *Understanding digital humanities* (p. 1–20). Houndmills Basingstoke Hampshire; New York: Palgrave Macmillan. <http://doi.org/10.1057/9780230371934>
- Besser, H. (2002). The next stage: Moving from isolated digital collections to interoperable digital libraries. *First Monday*, 7(6), s.p.. Acedido em

<http://www.firstmonday.org/ojs/index.php/fm/article/view/958/879>

Besser, H. (2004). The past, present, and future of digital libraries: Why are digital libraries important to humanists? What is a library ? In S. Schreibman, R. Siemens, & J. Unsworth (Eds.), *A companion to digital humanities* (pp. 557–575). Oxford: Blackwell.

Bia, A., & Pedreño, A. (2001). The Miguel de Cervantes digital library: the Hispanic voice on the web. *Literary and Linguistic Computing*, 16(2), 161–177.  
<http://dx.doi.org/10.1093/lc/16.2.161>

Blandford, A., & Gow, J. (2006). Digital Libraries in the context of users' broader activities. *D-Lib Magazine*, 12(7/8), 8–10.

Bloch, R. H., & Hesse, C. A. (2015). *Future libraries*. London: University of California Press.

BnF. (2016). *Innovation numérique à la BnF*. Retrieved from  
[http://www.bnf.fr/fr/professionnels/innovation\\_numerique.html](http://www.bnf.fr/fr/professionnels/innovation_numerique.html)

BnF. (2017). *Rapport d'activité 2016*. Retrieved from  
<http://webapp.bnf.fr/rapport/index.html>

Bogdan, R., & Biklen, S. (1994). *Investigação qualitativa em educação: Uma introdução à teoria e aos métodos*. Porto: Porto.

Borges, M. M. (2003). Biblioteca digital: Materialização e utopia. *Revista da Faculdade de Letras: Ciências e Técnicas do Património*, 2, 653–664. Acedido em  
<http://hdl.handle.net/10316/243>

Borges, M. M. (2006). A esfera: Comunicação académica e novos media. *Universo*, 1–801. Acedido em <http://estudogeral.sib.uc.pt/handle/10316/8557>

Borgman, C. L. (2009). The digital future is now: A call to action for the Humanities. *DHQ: Digital Humanities Quarterly*, 3(4), 1–30. Acedido em  
<http://www.digitalhumanities.org/dhq/vol/3/4/000077/000077.html>

Borgmann, C. L. (2000). *From Gutenberg to the global information infrastructure*. Massachusetts: MIT Press.

- Buchanan, G., Bainbridge, D., Don, K. J., & Witten, I. H. (2005). A new framework for building digital library collections. In *Proceedings of the 5th ACM/IEEE-CS Joint Conference on Digital Libraries* (pp. 23–31). New York, NY: ACM.  
<http://doi.org/10.1145/1065385.1065392>
- Buchanan, G., & Hinze, A. (2005). A generic alerting service for digital libraries. In *Proceedings of the 5th ACM/IEEE-CS Joint Conference on Digital Libraries* (pp. 131–140). New York, NY: ACM. <http://doi.org/10.1145/1065385.1065414>
- Budig, B., Dijk, T. C. Van, & Kirchner, F. (2016). Glyph miner: A system for efficiently extracting glyphs from early prints in the context of OCR. In *Proceedings of the 16th ACM/IEEE-CS on Joint Conference on Digital Libraries* (pp. 31–34). New York, NY: ACM. <http://doi.org/10.1145/2910896.2910915>
- Bühler, C. F. (1950). Aldus Manutius: The first five hundred years. *The Papers of the Bibliographical Society of America*, 44(3), 205–215. Acedido em <http://www.jstor.org/stable/24298605>
- Burdick, A., Drucker, J., Lunenfeld, P., Presner, T., & Schnap, J. (2012). *Digital humanities*. Cambridge, MA: MIT Press.
- Busa, R. A. (1980). The annals of humanities computing: The index thomisticus. *Computers and the Humanities*, 14, 83–90.
- Busa, R. A. (2004). Foreword: Perspectives on the digital humanities. In S. Schreibman, R. Siemens, & J. Unsworth (Eds.), *A companion to digital humanities* (pp. XVI–XXII). Oxford: Blackwell.
- Campos, F. M. G. de. (2005). A Biblioteca Nacional e a memória digital do património bibliográfico português: A experiência da Biblioteca Nacional Digital. *Cadernos BAD*, 2, 19–32. Acedido em <https://www.bad.pt/publicacoes/index.php/cadernos/article/view/811/810>
- Campos, F. M. G. de, Cordeiro, M. I., Galvão, R. M., & Study Group on the Functional Requirements for Bibliographic Records. (2008). *Requisitos funcionais dos registos bibliográficos*. Lisboa: Biblioteca Nacional de Portugal. Acedido em <http://purl.pt/23509>

- Candela, L., Castelli, D., Ferro, N., Ioannidis, Y., Koutrika, G., Meghini, C., ... Schuldt, H. (2007). *The DELOS digital library reference model foundations for digital libraries*. Acedido em [http://www.delos.info/files/pdf/ReferenceModel/DELOS\\_DLReferenceModel\\_096.pdf](http://www.delos.info/files/pdf/ReferenceModel/DELOS_DLReferenceModel_096.pdf)
- Carmo, H., & Ferreira, M. M. (1998). *Metodologia da investigação: Guia para auto-aprendizagem*. Lisboa: Universidade Aberta.
- Cassella, M. (2010). Institutional repositories: An internal and external perspective of the IRs value for the researchers' communities. *LIBER Quarterly*, 20(July), 210–225. <http://doi.org/URN:NBN:NL:UI:10-1-113593>
- Castillo Gómez, A. (2004). *Das tabuinhas ao hipertexto: Uma viagem na história da cultura escrita*. Lisboa: Biblioteca Nacional.
- Charpin, D. (2008). *Lire et écrire à Babylone*. Paris: Presses Universitaires de France.
- Choudhury, G. S., DiLauro, T., Ferguson, R., Droettboom, M., & Fujinaga, I. (2006). Document recognition for a million books. *D-Lib Magazine*, 12(3), 31–41. <http://doi.org/10.1045/march2006-choudhury>
- Choudhury, G. S., & Seaman, D. (2008). The virtual library. In R. Siemens, & S. Schreibman, *A Companion to digital literary studies* (pp. 534–546). Oxford: Wiley-Blackwell.
- Clarkson, E. C., Navathe, S. B., & Foley, J. D. (2009). Generalized formal models for faceted user interfaces. In *Proceedings of the 9th ACM/IEEE-CS Joint Conference on Digital Libraries* (pp. 125–134). New York, NY: ACM. <http://doi.org/10.1145/1555400.1555422>
- Cole, T. W., Han, M.-J. K., & Vannoy, J. A. (2012). Descriptive metadata, iconclass, and digitized emblem literature. In *Proceedings of the 12th ACM/IEEE-CS Joint Conference on Digital Libraries* (pp. 111–120). New York, NY: ACM. <http://doi.org/10.1145/2232817.2232839>
- Collins, E., & Jubb, M. (2012). How do Researchers in the Humanities Use Information Resources ? *Liber Quarterly*, 21(2), 176–187. Acedido em

- <http://liber.library.uu.nl/index.php/lq/article/view/8017>
- Coutinho, C. (2011). *Metodologia de investigação em ciências sociais e humanas: Teoria e prática*. Coimbra: Almedina.
- Crane, G. (2011). *White paper report*. Medford. Acedido em <http://hdl.handle.net/10427/75558>
- Crane, G., Smith, D. A., & Wulfman, C. E. (2001). Building a hypertextual digital library in the Humanities: A case study on London. In *Proceedings of the 1st ACM/IEEE-CS Joint Conference on Digital Libraries* (pp. 426–434). New York, NY: ACM. <http://doi.org/10.1145/379437.379756>
- Crane, G., & Wulfman, C. (2003). Towards a cultural heritage digital library. In *Proceedings of the 3rd ACM/IEEE-CS Joint Conference on Digital Libraries* (pp. 75–86). Washington, DC: IEEE Computer Society. Acedido em <http://dl.acm.org/citation.cfm?id=827140.827150>
- Creswell, J. (1994). *Research design: Qualitative & quantitative approaches*. Thousand Oaks, CA: Sage Publications.
- Crisci, F. (2017). The world of Aldus Manutius (1494-1515), a Renaissance publishing venture: Materiality, cultural entrepreneurship, and institutional dynamics in markets. In *Sinergie-SIMA, Annual Conference, Sinergie Italian Journal of Management - "Value co-creation: management challenges for business and society", Napoli, 15-16 giugno*. Napoli. Acedido em [http://criticalmanagement.uniud.it/fileadmin/user\\_upload/Crisci\\_2017b.pdf](http://criticalmanagement.uniud.it/fileadmin/user_upload/Crisci_2017b.pdf)
- Dacos, M. (2013). La stratégie du sauna finlandais. Acedido em 3 fev. 2016, em <http://bn.hypotheses.org/11138>
- Dacos, M., & Mounier, P. (2014). *Humanités numériques: État des lieux et positionnement de la recherche française dans le contexte international*. Paris: Institut Français. Acedido em <http://www.institutfrancais.com/fr/actualités/humanites-numeriques>
- Dalkey, N., & Helmer, O. (1963). An experimental application of the DELPHI method to the use of experts. *Management Science*, 9(3), 458–467.

<http://doi.org/10.1287/mnsc.9.3.458>

- Dawei, W., & Yigang, S. (2010). The National Digital Library Project. *D-Lib Magazine*, 16(5/6), 1–8. <http://doi.org/10.1045/may2010>
- Deegan, M., & Tanner, S. (2004). Conversion of primary sources. In S. Schreibman, R. Siemens, & J. Unsworth (Eds.), *A companion to digital humanities* (pp. 488–504). Oxford: Wiley-Blackwell.
- Denscombe, M. (2007). *The good research guide: For small-scale social research projects*. London: Open University Press.
- Denscombe, M. (2014). *The good research guide for small-scale social research projects* (3th ed.). Maidenhead, Berkshire: Open University Press.
- Denzin, N. K., & Lincoln, Y. S. (2008). *Strategies of qualitative inquiry*. Thousand Oaks, CA: Sage Publications.
- Dias, J. (1994). *Iniciação à bibliofilia*. Lisboa: Pró-Associação Portuguesa de Alfarrabistas.
- Digibís. (s.d.). *Datasets*. Acedido em <http://enumeratedataplatform.digibis.com/>
- Droettboom, M., Fujinaga, I., MacMillan, K., Chouhury, G. S., DiLauro, T., Patton, M., & Anderson, T. (2002). Using the Gamera Framework for the recognition of cultural heritage materials. In *Proceedings of the 2Nd ACM/IEEE-CS Joint Conference on Digital Libraries* (pp. 11–17). New York, NY: ACM.  
<http://doi.org/10.1145/544220.544223>
- Dylla, K., Frischer, B., Mueller, P., Ulmer, A., & Haegler, S. (2009). *Rome reborn 2.0 : A Case study of virtual city reconstruction using procedural modeling techniques*. Acedido em [http://www.romereborn.virginia.edu/rome\\_reborn\\_2\\_documents/papers/Dylla2\\_Frischer\\_Rome\\_Reborn.pdf](http://www.romereborn.virginia.edu/rome_reborn_2_documents/papers/Dylla2_Frischer_Rome_Reborn.pdf)
- Ernst-Gerlach, A., & Fuhr, N. (2007). Retrieval in text collections with historic spelling using linguistic and spelling variants. In *Proceedings of the 7th ACM/IEEE-CS Joint Conference on Digital Libraries* (pp. 333–341). New York, NY: ACM.  
<http://doi.org/10.1145/1255175.1255242>

- Europeana collections*. (2008). Acedido em 22 fev. 2016, em <http://www.europeana.eu/portal/>
- Evans, L., & Rees, S. (2012). An interpretation of digital humanities. In D. M. Berry (Ed.), *Understanding digital humanities* (pp. 21–41). Houndmills, Basingstoke, Hampshire: Palgrave Macmillan.
- Faria, M. I., & Pericão, M. da G. (2008). *Dicionário do livro: Da escrita ao livro electrónico*. Coimbra: Almedina.
- Febvre, L., & Martin, H.-J. (2000). *O aparecimento do livro*. Lisboa: Fundação Calouste Gulbenkian.
- Feng, S., & Manmatha, R. (2006). A hierarchical, HMM-based automatic evaluation of OCR accuracy for a digital library of books. In *Proceedings of the 6th ACM/IEEE-CS Joint Conference on Digital Libraries* (pp. 109–118). New York, NY: ACM. <http://doi.org/10.1145/1141753.1141776>
- Fenlon, K., Fallaw, C., Cole, T., & Han, M.-J. (2014). A preliminary evaluation of HathiTrust metadata: Assessing the sufficiency of legacy records. In *Proceedings of the 14th ACM/IEEE-CS Joint Conference on Digital Libraries* (pp. 317–320). Piscataway, NJ: IEEE Press. Acedido em <http://dl.acm.org/citation.cfm?id=2740769.2740824>
- Ferreira, C. A. S. (2011). *Preservação da informação digital: Uma perspectiva orientada para as bibliotecas*. Dissertação de Mestrado, Universidade de Coimbra, Coimbra. Acedido em <https://estudogeral.sib.uc.pt/jspui/handle/10316/15001>
- Fitzpatrick, K. (2012). The Humanities, done digitally. In M. K. Gold (Ed.), *Debates in the Digital Humanities* (p. 504). Minneapolis; London: University of Minnesota Press. Acedido em <http://dhdebates.gc.cuny.edu/debates/text/30>
- Flanders, J. (2012). Collaboration and dissent: Challenges of collaborative standards for Digital Humanities. In W. Mccarty & M. Deegan (Eds.), *Collaborative research in the Digital Humanities* (pp. 67–80). Farnham: ASGATE.
- Fletcher, H. G. (1988). *New Aldine studies: Documentary essays on the life and works of Aldus Manutius*. San Francisco: Rosenthal.

- Flick, U. (2009). *An introduction to qualitative research* (4th ed.). Los Angeles, CA; London: Sage Publications.
- Flick, U., Kardorff, E. von, & Steinke, I. (Eds.). (2004). *A companion to qualitative research*. London; Thousand Oaks, CA: Sage Publications.
- Fox, E. A. (1993). *Source book on digital libraries*. Blacksburg, VA: Dept. of Computer Science, Virginia Polytechnic Institute and State University.
- Freitas, C. V. de. (2010). *A autenticidade dos objectos digitais*. Dissertação de Mestrado, Universidade de Coimbra, Coimbra. Acedido em <https://estudogeral.sib.uc.pt/handle/10316/15266>
- Fuhr, N., Tsakonas, G., Aalberg, T., Agosti, M., Hansen, P., Kapidakis, S., ... Sølvsberg, I. (2007). Evaluation of digital libraries. *International Journal on Digital Libraries*, 8(1), 21–38. <http://doi.org/10.1007/s00799-007-0011-z>
- Galina Russell, I. (2011a). ¿Qué Son las Humanidades Digitales? *Revista Digital Universitaria*, 12(7), s.p.. Acedido em <http://www.revista.unam.mx/vol.12/num7/art68/>
- Galina Russell, I. (2011b). El papel de las bibliotecas en las humanidades digitales. In *World Library and Information Congress : 77th IFLA General Conference and Assembly*. San Juan: IFLA. Acedido em [http://www.paginaspersonales.unam.mx/files/638/Publica\\_20110923192631.pdf](http://www.paginaspersonales.unam.mx/files/638/Publica_20110923192631.pdf)
- Gao, L., Tang, Z., Lin, X., Liu, Y., Qiu, R., & Wang, Y. (2011). Structure extraction from PDF-based book documents. *Proceedings of the 11th ACM/IEEE-CS Joint Conference on Digital Libraries* (pp. 11–20). <http://doi.org/10.1145/1998076.1998079>
- Gardiner, E., & Musto, R. G. (2015). *The Digital Humanities: A primer for students and scholars*. New York, NY: Cambridge University Press.
- Garrison, W. A. (2001). Retrieval issues for the Colorado digitization project's heritage database. *D-Lib Magazine*, 7(10), 1–11. <http://doi.org/10.1045/october2001-garrison>
- Gee, D., & West, J. (2008). *About Open Library*. Acedido em

<https://openlibrary.org/help/faq/about>

Glaser, B. G., & Strauss, A. L. (2006). *The discovery of grounded theory: Strategies for qualitative research*. Hawthorne, N.Y: Aldine de Gruyter.

Gold, M. K. (Ed.). (2012). *Debates in the digital humanities. Digital scholarship in the Humanities*. Minneapolis: University of Minnesota Press.

<http://doi.org/10.5749/j.ctttv8hq>

Gonçalves, M. F., & Banza, A. P. (2013). In limine. In M. F. Gonçalves & A. P. Banza (Eds.), *Património textual e Humanidades Digitais: da antiga à nova filologia*.

Évora: CIDEHUS. Acedido em <http://dspace.uevora.pt/rdpc/handle/10174/10468>

Gong, X., Ke, W., Zhang, Y., & Broussard, R. (2013). Interactive search result clustering: A study of user behavior and retrieval effectiveness. In *Proceedings of the 13th ACM/IEEE-CS Joint Conference on Digital Libraries* (pp. 167–170). New York, NY: ACM. <http://doi.org/10.1145/2467696.2467726>

Gooding, P., Terras, M., & Warwick, C. (2013). The myth of the new: Mass digitization, distant reading, and the future of the book. *Literary and Linguistic Computing*, 28(4), 629–639. Acedido em <http://dx.doi.org/10.1093/lc/fqt051>

Gorman, G. E., & Clayton, P. (2005). *Qualitative research for the information professional: A practical handbook* (2nd ed.). London: Facet.

Gray, D. E. (2004). *Doing research in the real world*. London, Thousand Oaks, CA: Sage Publications.

Greenhalgh, M. (2004). Art history. In S. Schreibman, R. G. Siemens, & J. Unsworth (Eds.), *A companion to digital humanities* (p. 31–45). Oxford: Wiley-Blackwell.

Guerreiro, D. M., & Borbinha, J. L. (2014). Humanidades Digitais: Novos desafios e oportunidades [Ciências da Informação]. *Cadernos BAD*, (1), 63–78. Retrieved from <http://www.bad.pt/publicacoes/index.php/cadernos/article/view/1060>

Guerreiro, D. M. G. (2016). Bibliotecas digitais: Divulgação e preservação. *Cidade Solidária*, (35), 130–137.

Guerreiro, D. M. G., Campos, F. M. G. de, & Lopes, P. F. (2009). *Repositório digital de património cultural móvel: Uma aplicação a objectos do culto católico*.

Dissertação de Mestrado, ISCTE, Lisboa. Acedido em  
<http://hdl.handle.net/10071/1829>

Günther, H. (2003). *Como elaborar um questionário*. Brasília: UnB, Laboratório de Psicologia Ambiental. Acedido em [www.psi-ambiental.net/pdf/01Questionario.pdf](http://www.psi-ambiental.net/pdf/01Questionario.pdf)

Günther, H. (2006). Pesquisa qualitativa versus pesquisa quantitativa: Esta é a questão?. *Revista Psicologia: Teoria e Pesquisa*, 22(2), 201–210. Acedido em <http://www.scielo.br/pdf/ptp/v22n2/a10v22n2>

Hagedorn, K., Kargela, M., Noh, Y., & Newman, D. (2011). A new way to find: Testing the use of clustering topics in digital libraries. *D-Lib Magazine*, 17(9/10), 1–9. <http://doi.org/10.1045/september2011>

Harris, M., Levene, M., Zhang, D., & Levene, D. (2014). The anatomy of a search and mining system for Digital Humanities. In *Proceedings of the 14th ACM/IEEE-CS Joint Conference on Digital Libraries* (pp. 165–168). Piscataway, NJ: IEEE Press. Acedido em <http://dl.acm.org/citation.cfm?id=2740769.2740796>

Hart, C. (1998). *Doing a literature review*. Newbury Park, CA: Sage Publications.

Hayles, N. K. (2012). How we think: Transforming power and digital technologies. In D. M. Berry (Ed.), *Understanding digital humanities*. Houndmills, Basingstoke, Hampshire: Palgrave Macmillan.

Hill, M. M., & Hill, A. (2002). *Investigação por questionário*. Lisboa: Sílabo.

Hinze, A., Taube-Schock, C., Bainbridge, D., Matamua, R., & Downie, J. S. (2015). Improving access to large-scale digital libraries through semantic-enhanced search and disambiguation. In *Proceedings of the 15th ACM/IEEE-CS Joint Conference on Digital Libraries* (pp. 147–156). New York, NY: ACM. <http://doi.org/10.1145/2756406.2756920>

Hockey, S. (2004). The history of humanities computing. In S. Schreibman, R. G. Siemens, & J. Unsworth (Eds.), *A companion to digital humanities* (pp. 3–19). Oxford: Wiley-Blackwell.

Hoque, Z., Parker, L. D., Covaleski, M. A., & Haynes, K. (Eds.). (2017). *The Routledge*

*companion to qualitative accounting research methods*. Abingdon, Oxon; New York, NY: Routledge.

Huggett, M., & Rasmussen, E. (2012). Dynamic Online Views of Meta-indexes. In *Proceedings of the 12th ACM/IEEE-CS Joint Conference on Digital Libraries* (pp. 233–236). New York, NY: ACM. <http://doi.org/10.1145/2232817.2232860>

Huot, R. (2002). *Métodos quantitativos para as ciências humanas*. Lisboa: Instituto Piaget.

IFLA, & UNESCO. (2011). Manifiesto de la Federación Internacional de Asociaciones de Bibliotecarios y Bibliotecas (IFLA) sobre las bibliotecas digitales. In *UNESCO, Conferencia General, 36.ª reunión*. Paris: UNESCO, IFLA. Retrieved from <http://www.ifla.org/ES/publications/manifiesto-de-las-ifla-unesco-sobre-las-bibliotecas-digitales>

Jett, J., Nurmikko-Fuller, T., Cole, T. W., Page, K. R., & Downie, J. S. (2016). Enhancing scholarly use of digital libraries: A comparative survey and review of bibliographic metadata ontologies. In *Proceedings of the 16th ACM/IEEE-CS on Joint Conference on Digital Libraries* (pp. 35–44). New York, NY: ACM. <http://doi.org/10.1145/2910896.2910903>

Jorgensen, D. L. (1998). *Participant observation: A methodology for human studies*. London: Sage Publications.

Kader, C. C. C., & Richter, M. G. (2013). Linguística de corpus: Possibilidades e avanços. *Instrumento*, 15(1), 13–23.

Kalinichenko, L. A., Korenkov, V. V., Shirikov, V. P., Sissakian, A. N., & Sunturenko, O. V. (2003). Digital Libraries: Advanced methods and technologies, digital collections. *D-Lib Magazine*, 9(1), 1–7.

Kichuk, D. (2007). Metamorphosis: Remediation in early English books online (EEBO). *Literary and Linguistic Computing*, 22(3), 291–303. <http://dx.doi.org/10.1093/lc/fqm018>

Kirschenbaum, M. G. (2012). What is digital humanities and what's it doing in English departments. In M. K. Gold (Ed.), *Debates in the digital humanities* (pp. 3–11).

- Minneapolis, MN: University of Minnesota Press.
- Knijff, J. van der. (2011). JPEG 2000 for long term preservation: JP2 as a preservation format. *D-Lib Magazine*, 17(5/6), 1–14. <http://doi.org/10.1045/may2011>
- Koller, G. (2012). Digital history: Ready made. Acedido em 6 jan. 2013, em <http://wethink.hypotheses.org/701>
- Kott, K., Dunn, J., Halbert, M., Johnston, L., Milewicz, L., & Shreeves, S. (2006). Digital Library Federation (DLF) Aquifer project. *D-Lib Magazine*, 12(5), 1–10. Acedido em <http://www.dlib.org/dlib/may06/kott/05kott.html>
- Krueger, R. A. (2002). Designing and conducting focus group interviews. Acedido em <http://www.eiu.edu/ihec/Krueger-FocusGroupInterviews.pdf>
- Krueger, R. A., & Casey, M. A. (2009). *Focus groups: A practical guide for applied research* (4th ed.). Los Angeles, CA [etc.]: Sage Publications.
- Kucsma, J., Reiss, K., & Sidman, A. (2010). Using Omeka to build digital collections: The METRO case study. *D-Lib Magazine*, 16(3/4), 1–11. <http://doi.org/10.1045/march2010-kucsma>
- Kules, B., Capra, R., Banta, M., & Sierra, T. (2009). What do exploratory searchers look at in a faceted search interface? *JCDL '09 Proceedings of the 9th ACM/IEEE-CS Joint Conference on Digital Libraries* (pp. 313–322). New York, NY: ACM. <http://doi.org/10.1145/1555400.1555452>
- Lawless, S., Conlan, O., & Hampson, C. (2016). Tailoring access to content. In S. Schreibman, R. Siemens, & J. Unsworth (Eds.), *A New Companion to Digital Humanities* (p. 586). Chichester, West Sussex, UK: Wiley-Blackwell.
- Lee, S. (2001). *Digital imaging: A practical handbook*. London: Facet.
- Léon, J. (2006). A Lingüística de Corpus: história, problemas, legitimidade. *Filologia E Lingüística Portuguesa*, 8, 51–81.
- Lesk, M. (2005). *Understanding Digital Libraries* (2nd ed.). San Francisco: Elsevier.
- Lucía Megías, J. M. (2010). De las bibliotecas digitales a las plataformas de conocimiento (notas sobre el futuro del texto en la era digital). In M. Arbor Aldea

- & A. Fernández Guiadanes (Eds.), *Estudos de edición crítica e lírica galego-portuguesa* (pp. 369–401). Santiago de Compostela: Universidade de Santiago de Compostela. Acedido em <http://eprints.ucm.es/10767/>
- Manifeste des Digital Humanities*. (2010). *THATCamp Paris 2010*. Acedido em <http://tcp.hypotheses.org/497>
- Manni, D. M. (2012). *Vita di Aldo Pio Manuzio*. Roma: LiberLiber. Acedido em <http://www.eclap.eu/119661>
- Martins, J. V. de P. (2007). *Histórias de livros para a história do livro*. Lisboa: Fundação Calouste Gulbenkian.
- Maxwell, J. A. (2013). *Qualitative research design: An interactive approach*. Los Angeles, CA [etc.]: Sage Publications.
- Mayernik, M. S., Wallis, J. C., & Borgman, C. L. (2013). Unearthing the infrastructure: Humans and sensors in field-based scientific research. *Computer Supported Cooperative Work*, 22(1), 65–101. <http://doi.org/10.1007/s10606-012-9178-y>
- McKenna, G. (coord.). (s.d.). *ENUMERATE*. Acedido em 27 fev. 2016, em <http://www.enumerate.eu/>
- McMurtrie, D. (1927). *The golden book: The story of fine books and bookmaking: Past & present*. Chicago: Pascal Covici. Acedido em <https://archive.org/details/goldenbookhistor00mcmu>
- McMurtrie, D. (1969). *O livro: Impressão e fabrico*. Lisboa: Fundação Calouste Gulbenkian.
- Miguéis, A., & Fiolhais, C. (2014). Recursos digitais em livre acesso na Universidade de Coimbra: Estudo Geral e Alma Mater. *Revista Recis*, 8(2), 231–242. <http://doi.org/10.3395/recis.v8i2.937.pt>
- Moore, N. (2006). *How to do research: A practical guide to designing and managing research projects* (3rd. ed.). London: Facet.
- Murray, S. (2009). *The library: An illustrated history*. New York, NY; Chicago, IL: Skyhorse; ALA Editions.

- Nascimento, M. F. B. do. (2000). O corpus de referência do português contemporâneo e os projectos de investigação do Centro de Linguística da Universidade de Lisboa sobre variedades do português falado e escrito. In E. Gärtner, C. Hundt, & A. Schönberger (Eds.), *Estudos de gramática portuguesa* (pp. 185–200). Frankfurt am Main: TFM.
- Nauta, G. J., Heuvel, W. van den, & DEN Foundation. (2015). *Survey report on digitisation in European cultural heritage institutions 2015*. Acedido em <http://www.enumerate.eu/fileadmin/ENUMERATE/documents/ENUMERATE-Digitisation-Survey-2014.pdf>
- Neves, B. P., & Borges, M. M. (2011). A gestão de conteúdos digitais em bibliotecas universitárias: O caso do portal da biblioteca da FLUC. In *Límites, fronteras y espacios comunes: encuentros y desencuentros en las Ciencias de la Información. V Encuentro Ibérico EDICIC 2011, Badajoz, 17-19 November*. Acedido em <http://eprints.rclis.org/18757/>
- Newman, D., Noh, Y., Talley, E., Karimi, S., & Baldwin, T. (2010). Evaluating topic models for digital libraries categories and subject descriptors. *JCDL'10- Proceedings of the 10th annual joint conference on Digital libraries, Gold Coast, Queensland, Australia, 21-25 June* (pp. 215–224). <http://doi.org/10.1145/1816123.1816156>
- Noonan, D. W., McCrory, A., & Black, E. L. (2010). PDF/A: A viable addition to the preservation toolkit. *D-Lib Magazine*, 16(11/12), 1–13. <http://doi.org/10.1045/november2010>
- Nuovo, A. (2013). *The book trade in the Italian Renaissance*. Leiden: Brill.
- Padilla, T. (2016). Humanities data in the library: Integrity, form, access. *D-Lib Magazine*, 22(3/4), 1–6. <http://doi.org/10.1045/march2016>
- Patrício, H. S. (2015). Evolução do modelo de informação da Biblioteca Nacional Digital. In *12.º Congresso Congressos Nacionais de Bibliotecários, Arquivistas e Documentalistas: Ligar. Transformar. Criar Valor. Évora - Universidade de Évora, 21, 22 e 23 de outubro de 2015* (pp. 1–13). Acedido em <http://www.bad.pt/publicacoes/index.php/congressosbad/article/view/1391>

- Pedrosa, G., & Borbinha, J. L. (2010). ContentE: Uma ferramenta para estruturar obras digitalizadas. In *Atas de congresso: Congresso Nacional de BAD: Políticas de Informação na Sociedade em Rede, Guimarães, 7, 8 e 9 de Abril de 2010*. Acedido em <https://www.bad.pt/publicacoes/index.php/congressosbad/article/view/236/237>
- Pedrosa, G., Edmundo, J., Borbinha, J., & Freire, N. (2012). Agregação de conteúdos para a The European Library e EUROPEANA. In *Atas de congresso: Congresso Nacional de BAD: Integração, Acesso e Valor Social*, Lisboa, Fundação Calouste Gulbenkian, 18, 19 e 20 outubro de 2012. Acedido em <http://www.bad.pt/publicacoes/index.php/congressosbad/article/view/293>
- Pickard, A. J. (2007). *Research methods in information*. London: Facet.
- Puglia, S., Reed, J., & Rhodes, E. (2004). *Technical guidelines for digitizing archival materials for electronic access: Creation of production master files – raster images*. Washington. Acedido em <http://www.archives.gov/preservation/technical/guidelines.pdf>
- Quivy, R., & Campenhoudt, L. van. (2008). *Manual de investigação em ciências sociais* (5.ª ed.). Lisboa: Gradiva.
- Rahnemoonfar, M., & Plale, B. (2013). Automatic performance evaluation of dewarping methods in large scale digitization of historical documents. In *Proceedings of the 13th ACM/IEEE-CS Joint Conference on Digital Libraries* (pp. 331–334). New York, NY: ACM. <http://doi.org/10.1145/2467696.2467744>
- Ramel, J.-Y., Sidère, N., & Rayar, F. (2013). Interactive layout analysis, content extraction, and transcription of historical printed books using pattern redundancy analysis. *Literary and Linguistic Computing*, 28(2), 301–314. <http://doi.org/10.1093/lc/fqs077>
- Reeves, T. C., Apedoe, X., & Woo, Y. H. (2005). *Evaluating digital libraries: A user-friendly guide*. Acedido em 15 abr. 2013, em <http://www.dpc.ucar.edu/projects/evalbook/index.html>
- Religion: Sacred electronics. (1956). *Time*, 68(27).

- Reto, L., & Nunes, F. (1999). Métodos como estratégia de pesquisa: Problemas tipo numa investigação. *Revista Portuguesa de Gestão, 1*, 21–31. Acedido em <http://repositorio-iul.iscte.pt/handle/10071/1386>
- Rieder, B., & Rohle, T. (2012). Digital methods: Five challenges. In D. M. Berry (Ed.), *Understanding digital humanities* (p. 67–84). Houndmills, Basingstoke, Hampshire: Palgrave Macmillan.
- Rodrigues, M. E. P., Amante, M. J., Pais, C., Segurado, T., & Lopes, S. (2014). Avaliação de repositórios institucionais: Análise comparativa. *Cadernos BAD, (2)*, 15–28. Acedido em <http://www.bad.pt/publicacoes/index.php/cadernos/article/view/1174>
- Romero Frías, E., & Del-Barrio-García, S. (2014). Una visión de las humanidades digitales a través de sus centros. *El Profesional de la Información, 23(5)*, 485–492. <http://doi.org/10.3145/epi.2014.sep.05>
- Ross, S., Anderson, I., Duffy, C., Economou, M., Gow, A., McKinney, P., ... National Initiative for a Networked. (2002). The NINCH guide to good practice in the digital representation and management of cultural heritage materials. <http://doi.org/10.1515/MFIR.2002.131>
- Rydberg-Cox, J. A. (2006). *Digital libraries and the challenges of digital humanities*. Oxford: Chandos.
- Saint-Georges, P. de. (2011). Pesquisa e crítica das fontes de documentação nos domínios económicos, social e político. In L. Albarello, F. Digneffe, J.-P. Hiernaux, C. Maroy, D. Ruquoy, & P. de Saint-Georges (Eds.), *Práticas e métodos de investigação em ciências sociais* (pp. 15–47). Lisboa: Gradiva.
- Saldaña, J. (2009). *The coding manual for qualitative researchers*. London: Sage Publications.
- Sánchez, A. (1995). *Cumbre: corpus lingüístico del español contemporáneo: Fundamentos, metodología y aplicaciones*. Madrid: Sociedad General Española de Librería.
- Sanderson, R., Albritton, B., Schwemmer, R., & Sompel, H. V. (2011). SharedCanvas: A

- collaborative model for medieval manuscript layout dissemination. In *Proceedings of the 11th Annual International ACM/IEEE Joint Conference on Digital Libraries* (Vol. 1104, pp. 175–184). New York, NY: ACM.  
<http://doi.org/10.1145/1998076.1998111>
- Sardinha, T. B. (2000). Lingüística de corpus: Histórico e problemática. *DELTA: Documentação de Estudos em Lingüística Teórica e Aplicada*, 16(2), 323–367.  
<http://dx.doi.org/10.1590/S0102-44502000000200005>
- Schilit, B. N., & Kolak, O. (2008). Exploring a digital library through key ideas. In *Proceedings of the 8th ACM/IEEE-CS Joint Conference on Digital Libraries* (pp. 177–186). New York, NY: ACM. <http://doi.org/10.1145/1378889.1378920>
- Schleiermacher, F., & Tice, T. N. (2011). *Brief outline of theology as a field of study: Revised translation of the 1811 and 1830 editions*. Louisville, KY: Westminster John Knox Press.
- Schöneberg, H., Schmidt, H.-G., & Höhn, W. (2013). A scalable, distributed and dynamic workflow system for digitization processes. In *Proceedings of the 13th ACM/IEEE-CS Joint Conference on Digital Libraries* (pp. 359–362). New York, NY: ACM.  
<http://doi.org/10.1145/2467696.2467729>
- Schreibman, S., Siemens, R., & Unsworth, J. (Eds.). (2004a). *A companion to digital humanities*. Oxford: Wiley-Blackwell.
- Schreibman, S., Siemens, R., & Unsworth, J. (2004b). The digital humanities and humanities computing: An introduction. In S. Schreibman, R. G. Siemens, & J. Unsworth (Eds.), *A companion to digital humanities* (pp. XXIII–XXVII). Oxford: Wiley-Blackwell.
- Schreibman, S., Siemens, R., & Unsworth, J. (Eds.). (2016). *A new companion to digital humanities*. Oxford: Wiley-Blackwell.
- Shnayder, E. (2010). A data model for spatial history. *Spatial History Lab*. Acedido em <http://www.stanford.edu/group/spatialhistory/cgi-bin/site/pub.php?id=23>
- Siemens, L., Cunningham, R., Duff, W., & Warwick, C. (2011). A tale of two cities: Implications of the similarities and differences in collaborative approaches within

- the digital libraries and digital humanities communities. *Literary and Linguistic Computing*, 26(3), 335–348. Acedido em <http://dx.doi.org/10.1093/lc/fqr028>
- Siemens, R., & Schreibman, S. (Eds.). (2008). *A companion to digital literary studies*. Oxford: Wiley-Blackwell.
- Sousa, M. C. P. (2012). A filologia digital em língua portuguesa: Alguns caminhos. In A. P. Banza & M. F. Gonçalves (Eds.), *Património textual e humanidades digitais: Da antiga à nova filologia*. Évora: CIDEHUS.
- Stewart, D. W., & Shamdasani, P. N. (2015). *Focus groups: Theory and practice*. Thousand Oaks, CA: Sage Publications.
- Stroeker, N., Vogels, R., & Panteia. (2012). *ENUMERATE: Survey report on digitisation in European cultural heritage institutions 2012*. Acedido em <http://www.enumerate.eu/fileadmin/ENUMERATE/documents/ENUMERATE-Digitisation-Survey-2014.pdf>
- Stroeker, N., Vogels, R., & Panteia. (2014). *ENUMERATE: Survey report on digitisation in European cultural heritage institutions 2014*. Acedido em <http://www.enumerate.eu/fileadmin/ENUMERATE/documents/ENUMERATE-Digitisation-Survey-2014.pdf>
- Strötgen, J., & Gertz, M. (2012). Event-centric search and exploration in document collections. In *Proceedings of the 12th ACM/IEEE-CS Joint Conference on Digital Libraries* (pp. 223–232). New York, NY: ACM. <http://doi.org/10.1145/2232817.2232859>
- Sukovic, S. (2002). Beyond the scriptorium: The role of the library in text encoding. *D-Lib Magazine*, 8(1), 1–6. Acedido em <http://www.dlib.org/dlib/january02/sukovic/01sukovic.html>
- Tanner, S., Muñoz, T., & Ros, P. H. (2009). Measuring mass text digitization quality and usefulness. *D-Lib Magazine*, 15(7/8), 1–12. <http://doi.org/10.1045/july2009-munoz>
- Terras, M. (2012). *Infographic: Quantifying digital humanities*. Acedido em 4 fev. 2016, em <https://blogs.ucl.ac.uk/dh/2012/01/20/infographic-quantifying-digital->

humanities/

Terras, M. (2008). *Digital images for the information professional*. Aldershot: Ashgate.

Terras, M., Nyhan, J., & Vanhoutte, E. (2013). *Defining digital humanities*. Aldershot: Ashgate.

UNESCO. (1965). Recomendacion sobre la normalizacion internacional de las estadísticas relativas a la edicion de libros. In *Actas de la Conferencia General, 13a reunión, Paris, 1964* (pp. 149–153). Paris: UNESCO. Acedido em <http://unesdoc.unesco.org/images/0011/001145/114581s.pdf#page=153>

Vinopal, J., & McCormick, M. (2013). Supporting digital scholarship in research libraries: Scalability and sustainability. *Journal of Library Administration, 53*(1), 27–42. <http://doi.org/10.1080/01930826.2013.756689>

Warwick, C., Galina, I., Rimmer, J., Terras, M., Blandford, A., Gow, J., & Buchanan, G. (2009). Documentation and the users of digital resources in the humanities. *Journal of Documentation, 65*(1), 1–39. <http://doi.org/10.1108/00220410910926112>

What is DH? (s.d.). Acedido em 14 fev. 2014 em <http://cdh.ucla.edu/about/what-is.html>

White, R. (2010). What is spatial history ? Acedido em <http://www.stanford.edu/group/spatialhistory/cgi-bin/site/pub.php?id=29>

Witten, I. H., Bainbridge, D., & Nichols, D. M. (2009). *How to build a digital library* (2th ed.). Burlington, MA: Morgan Kaufmann.

Xie, H. I. (2006). Evaluation of digital libraries: Criteria and problems from users' perspectives. *Library & Information Science Research, 28*(3), 433–452. <http://doi.org/10.1016/j.lisr.2006.06.002>

Yin, R. (2001). *Estudo de caso: planejamento e métodos*. Porto Alegre: Bookman.

Yin, R. (2003). *Estudo de caso* (2.<sup>a</sup> ed.). Porto Alegre: Bookman.

## Sítios institucionais

Digibís. (s.d.). Datasets. Retrieved from <http://enumeratedataplatfom.digibis.com/>

McKenna, G. (coord.). (s.d.). ENUMERATE. Retrieved February 27, 2016, from <http://www.enumerate.eu/>

Europeana Collections. (2008). Retrieved February 27, 2016, from <http://www.europeana.eu/portal/>

Library of Congress. (s.d.). Digital Collections and Services: Access to print, pictorial and audio-visual collections and other digital services. Retrieved February 27, 2016, from <https://www.loc.gov/library/libarch-digital.html>

OpenEdition: four platforms for electronic resources in the humanities and social sciences: OpenEdition Books, Revues.org, Hypotheses, Calenda. (s.d.). Retrieved February 24, 2016, from <https://www.openedition.org/>

Stanford University. (s.d.). Spatial History Project - Chinese Canadian Immigrant Pipeline, 1912-1923. Retrieved February 23, 2016, from <https://web.stanford.edu/group/spatialhistory/cgi-bin/site/viz.php?id=393>

Stanford University. (s.d.). Spatial History Project - Holocaust Geographies Collaborative. Retrieved February 23, 2016, from <https://web.stanford.edu/group/spatialhistory/cgi-bin/site/project.php?id=1015>

Stanford University. (s.d.). Spatial History Project. Retrieved February 23, 2016, from <http://web.stanford.edu/group/spatialhistory/cgi-bin/site/index.php>

ToolDatabase. (s.d.). Retrieved February 23, 2016, from <https://wiki.digitalmethods.net/Dmi/ToolDatabase>

Visualization Tools. (s.d.). Retrieved February 23, 2016, from <http://keshif.me/demo/VisTools>

DiRT Directory (Digital research tools). (s.d.). Retrieved February 23, 2016, from <http://dirtdirectory.org/>

Network for Digital Methods in the Arts and Humanities (NeDiMAH). (s.d.). Retrieved from <http://www.nedimah.eu/>

TAPoR - Text Analysis Portal for Research. (s.d.). Retrieved February 23, 2016, from <http://www.tapor.ca/>

Associação das Humanidades Digitais (AHDig). (s.d.). Retrieved from <http://ahdig.org/>

dh + lib. (s.d.). Retrieved from <http://acrl.ala.org/dh/>

HASTAC. (s.d.). Retrieved from <https://www.hastac.org/>

Red de Humanidades Digitales (Red-HD). (s.d.). Retrieved from <http://www.humanidadesdigitales.net/>

Global Outlook::Digital Humanities. (s.d.). Retrieved February 16, 2016, from <http://www.globaloutlookdh.org/>

4humanities. (s.d.). Retrieved from <http://4humanities.org/>

Digital Services Infrastructure for Social Sciences and Humanities (DASISH). (s.d.). Retrieved from <http://dasish.eu/>

Council of European Social Science Data Archives (CESSDA). (s.d.). Retrieved February 14, 2016, from <http://cessda.net/>

Common Language Resources and Technology Infrastructure - CLARIN-ERIC. (s.d.). Retrieved February 14, 2016, from <http://clarin.eu/>

DARIAH - Digital Research Infrastructure for the Arts and Humanities. (s.d.). Retrieved March 4, 2014, from <http://dariah.eu/>

NeDiMAH - Network for Digital Methods in the Arts and Humanities. (s.d.). Retrieved March 4, 2014, from <http://nedimah.eu/>

JADH - Japanese Association for Digital Humanities. (s.d.). Retrieved March 4, 2014, from <http://www.jadh.org/>

aaDH - Australian Association for Digital Humanities. (s.d.). Retrieved March 4, 2014, from <http://aa-dh.org/>

centerNet. (s.d.). Retrieved March 4, 2014, from <http://dhcenternet.org/>

CSDH/SCHN - Canadian Society for Digital Humanities / Société canadienne des humanités numériques. (s.d.). Retrieved March 4, 2014, from <http://csdh-schn.org/>

ACH - Association for Computers and the Humanities. (s.d.). Retrieved March 4, 2014, from <http://ach.org/>

EADH - The European Association for Digital Humanities. (s.d.). Retrieved March 4, 2014, from <http://eadh.org/>

ADHO - Alliance of Digital Humanities Organizations. (s.d.). Retrieved March 4, 2014, from <http://adho.org/>

What Is DH? (s.d.). Retrieved February 14, 2014, from  
<http://cdh.ucla.edu/about/what-is.html>

UC. (s.d.). DigLitWebAdeArquivos. Retrieved September 15, 2012, from  
<http://www.ci.uc.pt/diglit/DigLitWebAdeArquivos.html>

CenterNet. (s.d.). centerNet | An international network of digital humanities centers.  
Retrieved September 20, 2012, from <http://digitalhumanities.org/centernet/>



## **Glossário**

### **Acesso livre**

Obras disponibilizadas na web sem custos para o utilizador. Disponibilização dos conteúdos em linha sem restrições.

### **Ajuda interativa**

Funcionalidade de ajuda, obtida ao passar o cursor sobre um ícone ou texto, abrindo uma janela de texto com indicações sobre a ação associada.

### **Anotações colaborativas**

Adição de comentários marcações e destaques, em linha de forma sincronizada, por vários utilizadores.

### **Aparar um livro**

Cortar as margens de um livro, geralmente com guilhotina, de forma a ficar com a dimensão pretendida.

### **ARK**

Archival Resource Key, endereço persistente de recursos eletrónicos.

### **Assinatura de caderno**

Conjunto de letras, números e caracteres colocados na primeira página do caderno, ou em várias páginas, fora da mancha do texto, a fim de definir a sequência para a encadernação.

**Autoridade**

Forma autorizada, pela entidade nacional (em Portugal a BNP é a Agência Bibliográfica Nacional e Organismo de Normalização Sectorial), de referir o responsável por uma obra seja ele individual ou coletivo.

**Bibframe**

Bibliographic Framework Initiative. Fornece uma base para a descrição bibliográfica na web, fundamentado em técnicas Linked Data.

**Biblioteca**

Local onde se faculta o acesso à informação através de documentos, como monografias, panfletos, periódicos, iconografia e cartografia, em suporte papel ou digital, e disponibilizados de forma organizada, com critérios definidos para o acesso e gestão da coleção.

**Biblioteca digital**

Biblioteca constituída por um conjunto de objetos digitais, devidamente descritos, estruturados e disponibilizados ao público em vários formatos. Assegura a gestão e a preservação desses recursos, bem como a sua disseminação de forma consistente e persistente no tempo. Disponibiliza serviços que auxiliem os na localização e utilização dos recursos digitais.

### **Biblioteca digital de livro antigo**

O mesmo que biblioteca digital, com a particularidade do objeto ser o livro antigo. Esta espécie bibliográfica encontra-se em domínio público e para ser disponibilizada é necessário ser digitalizada e encadernada digitalmente.

### **BIBO**

Bibliographic Ontology Specification. É uma ontologia criada para descrever objetos bibliográficos como livros ou revistas.

### **Cadernos**

Conjunto de páginas de um livro, impressas numa única folha.

### **Caldeirões**

Sinal em forma de C barrado, utilizado para destacar a informação.

### **Códice**

Ou *codex* livro manuscrito, normalmente em pergaminho.

### **Coleção personalizada**

Coleção de obras digitais organizada pelo utilizador de bibliotecas digitais, constituída em função de critérios particulares, com a possibilidade de ser partilhada com outros utilizadores, nomeadamente, através das redes sociais ou de sistemas de gestão de aprendizagem (learning management system).

**ContentE**

Programa desenvolvido na BND para a edição digital e associação de metadados às imagens das obras digitalizadas.

**Corações**

Ou camarão, elemento tipográfico, utilizado para destacar a informação.

**Data mining**

Prospecção de dados ou mineração de dados pesquisa de dados com valor. Processo de extração de informação pertinente a partir de grandes massas documentais, através da detecção de padrões, regras de associação, etc.

**Desdobráveis**

Folha de tamanho superior às restantes páginas da obra, que tem de ser dobrada.

**Dewey Decimal Classification**

Sistema de classificação para assuntos em bibliotecas. Começou a ser desenvolvido em 1873, por Melvil Dewey. Organiza os materiais da biblioteca por assunto.

**Digitalização na íntegra**

Digitalização de todo o conteúdo da obra, no caso das monografias, de capa a capa.

**DjVu**

Formato para armazenagem e distribuição de imagens na web. Vantajoso para documentos digitalizados que contenham imagem a cores, linhas e texto, pode substituir o PDF.

**DSpace**

Programa em código aberto, utilizado para criar e gerir repositórios académicos.

**Dublin Core**

Esquema de metadados para a descrição dos objetos digitais.

**Elementos gráficos das obras**

Ilustrações, mapas, heráldica, esquemas existentes no interior das obras.

**Em linha**

O mesmo que *on line*.

**Encadernação digital**

Processo através do qual se agrupam as imagens digitais de uma obra, ordenando-as em conformidade com o original ou por outra ordem devidamente assinalada, bem como sumário ou outros apontadores de conteúdo e os respetivos metadados, de forma a obter uma representação digital coerente e fidedigna do original em suporte papel.

**ePUB (ou e-Pub)**

Abreviatura de Electronic Publication - Publicação Eletrónica. É um formato de arquivo digital para ebooks, livros digitais.

**Estante digital**

Coleção individual e personalizada de obras em formato digital, acedida através de identificação e autenticação, no sítio eletrónico da biblioteca digital.

**F.A.Q.**

Frequently Asked Questions. Perguntas mais frequentes.

**FaBiO**

FRBR-aligned Bibliographic Ontology. Ontologia estruturada de acordo com os FRBR

**Facetas**

Termos pré-definidos, devidamente estruturados, para a pesquisa e recuperação da informação. Utilizam-se em biblioteconomia desde 1933 (ver Ranganathan “Colon classification”). Classifica cada elemento de informação ao longo de múltiplas dimensões explícitas, denominadas facetas, permitindo que as classificações sejam acedidas e ordenadas de várias maneiras.

**Fedora**

Repositório digital, desenvolvido na Universidade da Virgínia, baseado no Flexible Extensible Digital Object and Repository Architecture (Fedora).

**Folheto**

Publicação não-periódica impressa com pelo menos 5 páginas e no máximo 48, excluindo-se as páginas de capa, publicada em determinado país e disponibilizada ao público.

**FRBR**

Functional Requirements for Bibliographic Records, requisitos funcionais dos registos bibliográficos. É um modelo conceptual de entidade relação.

## **Galerias**

Escavação realizada por inseto, por norma num conjunto de folhas ou nas encadernações.

## **Gamera**

Kit de ferramentas, que combinam a edição de imagem com o reconhecimento ótico de caracteres e a respetiva transcrição do texto, produzindo documentos estruturados.

## **Geographic information systems (GIS)**

Ou sistema de informação geográfica (SIG). É um sistema de informação elaborado para capturar, armazenar, manipular, analisar, gerenciar e apresentar dados espaciais ou geográficos.

## **Gestores bibliográficos**

Programas (ex. Mendeley, Zotero, Endnote) que permitem criar uma base de dados de referências bibliográficas, para utilização pessoal, com a possibilidade de referenciar citações e criar bibliografias, de forma automática, em trabalhos académicos ou de investigação.

## **Glosas**

Comentário, por norma em letra com corpo inferior ao restante texto.

## **GNU Public License**

Licença aplicada ao software livre. Concede ao utilizador tem o direito executar, copiar, distribuir, estudar, mudar e melhorar os programas.

### **GreaseMonkey**

É uma extensão para o navegador de web Mozilla Firefox, que pode modificar a exibição das páginas de site previamente selecionados pelo utilizador

### **Greenstone**

É um conjunto de ferramentas de informática para a construção manutenção e distribuição de coleções de bibliotecas digitais na web ou em CD-ROM.

### **Hipertexto**

Texto em formato digital que permite a ligação imediata a outros conteúdos, textuais ou não.

### **Humanidades Digitais**

Área do conhecimento que articula as Humanidades com as tecnologias digitais. A utilização das tecnologias digitais na investigação, no ensino e na divulgação das Humanidades, proporciona novas formas de olhar para os problemas e novos métodos, ferramentas e estruturas de suporte para o ensino e investigação. Além de facilitar e tornar mais aliciantes as tarefas de estudantes e investigadores, as Humanidades Digitais permitem divulgar a importância das Humanidades junto do grande público e angariar financiamento.

### **Hyrex 1**

É um mecanismo de recuperação de hipertexto para XML.

## **ICONCLASS**

ICONCLASS é um sistema de classificação para iconografia. É amplamente aceita para a descrição e recuperação de temas com imagens (obras de arte, ilustrações de livros, reproduções, fotografias, etc.).

## **Incunábulos**

Livro impresso com caracteres móveis, no ocidente, produzido entre 1450 e 1500.

## **in-oitavo**

Tamanho do papel para os incunábulos e livro antigo. Página dobrada três vezes, cerca de 18 cm de altura, origina dezasseis páginas.

## **JSON**

JavaScript Object Notation, é um formato de padrão aberto que utiliza texto legível por humanos para transmitir objetos de dados consistindo de pares atributo-valor.

## **Lacunas**

Espaço que ficou em branco.

## **Legibilidade do texto**

Qualidade do texto que permite distinguir e reconhecer os caracteres e, portanto, realizar a respetiva leitura.

## **Letras capituladas**

É uma letra no início da obra, ou de uma seção, de maior dimensão que o restante corpo do texto, podendo ou não ser decorada ou iluminada.

### **Linked Open Data**

Traduzindo literalmente: dados abertos ligados entre si. Modelo para a estruturação normalizada da web de dados abertos, formando assim uma rede de dados mais coerente e eficiente.

### **Livro**

Publicação não-periódica, impressa, com pelo menos 49 páginas, excluindo-se as páginas de capa, publicada em determinado país e disponibilizada ao público.

### **Livro antigo**

Livro impresso com caracteres móveis, no ocidente, produzido entre 1501 e 1800.

### **Livro de emblemas**

Livros de temática civil ou religiosa, ilustrados com gravuras, muito populares entre os séculos XVI a XVIII.

### **MADSRDF**

Metadata Authority Description Schema in RDF

### **Marcação de textos**

Aplicar etiquetas normalizadas (esquema TEI) ao conteúdo dos documentos, indicando parágrafos, frases, sinais gráficos, etc.

### **Margem interior**

Margem esquerda do reto e direita do verso margem da costura do livro.

### **Metacodificação**

Ação de associar os metadados

## **Metadados ou metainformação**

Literalmente, dados acerca de outros dados. Refere-se aos dados utilizados para ajudar a identificar, descrever e localizar os recursos, nomeadamente da web. Inteligíveis por computador

## **METS**

Metadata Encoding and Transmission Standard. O METS foi especialmente criado e desenvolvido para realizar a recuperação e preservação digital dos dados dos objetos digitais. É mantido pelo Network Development and MARC Standards Office, da Library of Congress, e pela Digital Library Federation.

## **Miscelâneas**

Obras que foram encadernadas juntas. Espécie bibliográfica que contem uma coleção artificialmente organizada de várias unidades bibliográficas, com ou sem afinidade temática, com a mesma encadernação.

## **MODS**

Metadata Object Description Schema. Descrição bibliográfica em XML, baseada no formato MARC.

## **MODSRDF**

RDF ontologia para o MODS. MODS é um esquema em XML para os elementos bibliográficos.

## **Monografia**

Termo técnico para livro.

## **MySQL**

O MySQL é um sistema de gerenciamento de banco de dados (SGBD), que utiliza a linguagem SQL (Linguagem de Consulta Estruturada, do inglês Structured Query Language) como interface. WIKI

## **Nado digital**

É criado com ferramentas ou processos digitais

## **Navegação ou pesquisa facetada**

Navegação ou pesquisa recorrendo às facetas.

## **n-gram statistical language model (SLM)**

É um modelo matemático que representa a distribuição probabilística de sequências de palavras ou caracteres encontrados em corpos de texto de linguagem natural.

## **Novas ferramentas**

(contagem de palavras, estudos de contexto, georreferenciação, etc.). Programas informáticos que permitem novas formas de trabalhar um texto

Ferramenta de contagem de palavras – Programa informático que permite contar as palavras de um texto e visualizar a sua representatividade.

Ferramenta para estudos de contexto – Programa informático que permite relacionar palavras ou conceitos dentro de um texto ou entre textos, criando, por ex., redes de influência.

Ferramenta de georreferenciação – Programa informático que permite associar uma localização geográfica à designação de um lugar e construir mapas dinâmicos.

## **OAI**

Open Archive Initiative, desenvolve e promove padrões de interoperabilidade que visam facilitar a disseminação eficiente de conteúdo.

## **OAI-PMH**

Open Archives Initiative Protocol for Metadata Harvesting, é um protocolo desenvolvido pela Open Archives Initiative que define um mecanismo para recolha e registos de metadados nos repositórios

## **Objeto digital**

Unidade de informação na biblioteca digital. Ficheiro, ou conjuntos de ficheiros digitais, representativos de uma realidade. No caso particular da biblioteca digital de livro antigo, é o ficheiro, ou o conjunto de ficheiros digitais, que reproduzem os documentos textuais.

## **Obras em domínio público**

Obra que não está abrangida pelo direito de autor. Por norma nos países signatários da Convenção de Berna, uma obra entra em domínio público setenta anos após o falecimento de seu autor, deixando de haver restrições à utilização da obra.

## **OCR**

Optical Character Recognition ou reconhecimento ótico de caracteres; tecnologia utilizada para converter as imagens dos caracteres existentes nas obras digitalizadas num arquivo de texto editável, ou seja, para transcrever os documentos.

## **Ontologias**

É um modelo de dados que representa um conjunto de conceitos dentro de um domínio e os relacionamentos entre eles. Uma ontologia é utilizada para realizar inferência sobre os objetos do domínio.

## **OPAC**

Online Public Access Catalog, catálogo informatizado de uma biblioteca ou bibliotecas em linha e disponível ao público.

## **Open-source software**

Programas informáticos cujo código fonte é disponibilizado e licenciado com uma licença de código aberto. É permitido o estudar, modificar e distribuir o programa de graça para qualquer entidade e para qualquer finalidade.

## **Página de rosto ou folha de título**

Primeira página da obra onde estão inscritos os elementos fundamentais da identificação.

## **Página de título**

Também designada por página de rosto.

## **Página reto**

Página da direita de um livro.

## **Página verso**

Página esquerda de um livro

### **Pattern Redundancy Analysis**

Reconhecimento de padrões é uma área da ciência cujo objetivo é a classificação de objetos dentro de um número de categorias ou classes

### **PDF**

Portable Document Format é um formato utilizado para representar documentos de maneira independente do aplicativo, do computador ou sistema operativo onde foi produzido.

### **Perfil de cor ICC**

Perfil de cor ICC (International Color Consortium) é a informação, embebida no próprio ficheiro de imagem, que descreve a forma como o equipamento gere a cor, a gama de cores com que o equipamento funciona e quais as aproximações realizadas

### **Perfil de utilizador**

Registo eletrónico de dados pessoais, que pode permitir aceder à informação de forma personalizada, em função dos interesses definidos pelo utilizador, e/ou criar um espaço pessoal para gestão de conteúdos.

### **Pergaminho**

Pele de animal, geralmente de cabra, carneiro, cordeiro, ovelha, ou vitela preparada para nela se escrever.

### **Pointwise mutual information (PMI)**

O modelo de classificação que fornece previsões relativamente boas sobre a classificação humana.

## **PURL**

Presistente Uniform Resource Locator, é um endereço de recursos uniforme, ou seja, um identificador de recurso uniforme ou um URI baseado em localização, ou seja o utilizador vê sempre o mesmo endereço.

## **RDF**

Resource Description Framework, modelo simples de dados, com uma semântica formal.

## **Reclamo**

Primeira palavra impressa num caderno, é repetida no pé do caderno anterior, para facilitar a encadernação.

## **Reconhecimento ótico de caracteres (OCR)**

Tecnologia utilizada para converter os caracteres existentes em imagens digitalizadas de texto, num ficheiro de texto editável.

## **Referências bibliográficas**

Conjunto de dados (elementos) sobre uma obra que permitam a sua identificação exclusiva. Consideram-se, entre outros, os seguintes elementos: autoria, título, n.º de edição, local, editora, data, código de barra, n.º de páginas, tradução, ISBN, etc.

## **RGB**

Red, Green e Blue, sistema de cores aditivas baseado no vermelho, verde e azul.

## **Rolo pergaminho**

Ou *volumen*, pergaminho enrolado em forma de rolo, geralmente continha uma obra.

## **schema.org**

É uma iniciativa lançada pelos operadores dos motores de busca do mundiais "criar e suportar um conjunto comum de esquemas para dados estruturados e marcação em páginas da web.

## **Sistema de alerta**

Sistema tecnológico utilizado para informar o subscritor acerca de uma determinada mudança; neste caso particular, informação relativa a alteração ou adição de obras numa biblioteca digital.

## **SKOS**

Simple Knowledge Organization System, recomendação do W3C destinada a representação de tesouros, esquemas de classificação, taxonomias, sistemas de título de assunto ou qualquer outro tipo de vocabulário estruturado e controlado.

## **SQL database**

Structured Query Language é a linguagem de pesquisa declarativa padrão para as bases de dados relacionais.

## **Stemming**

Método para reduzir uma palavra à sua raiz. São utilizados algoritmos com este método, nos sistemas de recuperação da informação, com o objetivo de melhorar a pesquisa.

## **Suffix tree**

É uma árvore comprimida, contendo todos os sufixos de um dado texto, assim como suas chaves e posições no texto.

### **Sumário nível 1**

Apontadores que permitem a navegação na obra através das partes constituintes.

Limitam-se à indicação da estrutura geral da obra.

### **Sumário nível 2**

Apontadores que permitem a navegação na obra através das partes constituintes, apresentam a estrutura da obra, em grandes divisões, geralmente correspondentes às partes ou livros

### **Sumário nível 3**

Apontadores que permitem a navegação na obra através das partes constituintes, acompanham a estrutura da obra, com a indicação das partes, livros, capítulos, secções.

### **Sumário obra digital**

Lista sequencial de títulos e subtítulos dos conteúdos (partes, capítulos, subcapítulos) pela ordem apresentada na obra e com indicação das páginas ou colunas onde iniciam.

Permite a navegação na obra.

### **Tabuinhas**

Placas de argila com escrita cuneiforme.

### **Tarjas**

Elementos tipográficos decorativos. Cercadura, orla.

### **Taxonomia**

A conceção, nomeação e classificação dos grupos de objetos, ideia, organismos, etc.

## **TEI**

Text Encoding Initiative, é um consórcio que desenvolve e mantém uma norma para a representação dos textos em formato digital.

## **Text clustering**

É uma técnica para fazer agrupamento automático de dados segundo o seu grau de semelhança

## **Texto marcado**

Texto que foi estruturado de acordo com uma norma como por exemplo o TEI

## **Thumbnail**

Miniatura da imagem, versão reduzida da imagem, cerca de 140px de altura.

## **Tinta ferrogálica**

Tinta feita à base de sulfato ferroso, muito durável, mas ao oxidar fica ácida e corrói o papel. Utilizada nas notas manuscritas.

## **Tinta repassada**

As palavras escritas no verso da folha são lidas no reverso.

## **TOPIC MODEL**

É um modelo estatístico para descobrir os "tópicos" abstratos que ocorrem numa coleção de documentos. A modelagem de tópicos é uma ferramenta de mineração de texto usada para a descoberta de estruturas semânticas escondidas nos corpos dos textos.

**Tutorial**

Ferramenta de aprendizagem (vídeo, texto, imagem, etc.) detalhando, a par e passo o funcionamento de algo e, neste caso particular, as funcionalidades associadas a uma biblioteca digital.

**URL**

Uniform Resource Locator, endereço eletrónico, tanto pode ser utilizado na web como em redes.

***Volumen***

O mesmo que rolo

**W3C**

World Wide Web Consortium (W3C) é a principal organização de normalização para o World Wide Web.

**Word-sense discrimination**

Relacionado com a identificação automática dos sentidos de uma palavra, isto é, os seus significados.

**WorldCat**

World Catalog, catálogo mundial em português) é um catálogo em linha gerido pelo Online Computer Library Center (OCLC) e considerado o maior catálogo em linha do mundo.

## **XML**

eXtensible Markup Language é uma linguagem de marcação que define um conjunto de regras para codificar documentos em formato que seja legível por humanos e legível por máquina. É uma recomendação da W3C para gerar linguagens de notação para necessidades especiais.



## **Anexos**



## I. Grelha observação livro antigo na BND

Itens a considerar na observação do livro antigo

		Elementos	Obra
Identificação da obra	Com base nos elementos do registo bibliográfico da obra	Endereço eletrónico	
		Cota	
		Século	
		Data	
		Língua 101 \$c (do UNIMARC)	
		Local de impressão 210\$a (do UNIMARC)	
		Título da obra 200\$a (do UNIMARC)	
Estruturação dos conteúdos	Disponibilização em linha	Tema	
		Obra com transcrição textual	
		Sumário	
Características das obras	Formas de iniciar a obra	Início da obra	
		Título/Portada	
		Título folha única	
		Título a 2 cores	
		Falso Título, Anterosto	
	Características do texto impresso	Texto em colunas	
		Tipo de letra	
		Capitais	
		Caldeirões ou corações	
		Capitulares iluminadas	

		Iluminado	
	Ilustrações	Ilustrações	
		Mapas	
	Modo de paginar a obra	Assinaturas	
		Paginação	
		Frente	
		Frente e verso	
		Início paginação	
	Outros elementos	Reclamo	
		Cortina	
		Cabeça de capítulo ou encabeçamento	
		Índice ou tabuada ou índice	
		Errata	
		Localização da errata	
		Marca de impressor	
		Invocação	
		Colofão	
		Notas na margem, impressas	
		Cadernos enunciados no fim (Registrum)	
		Observações	

## II. Grelha de observação das bibliotecas digitais com livro antigo

Itens a considerar na observação das bibliotecas digitais:

Identificação		Nome
		Instituição
		Pais
		Endereço eletrónico
		Data de inicio
Conteúdos	Tipologia das coleções digitais	Livros (texto)
		Imagem
		Vídeo
		Som
		Outros
Funcionalidades	Acesso à BD	Pesquisa
		Elementos da ficha bibliográfica
		No conteúdo da obra
		Temática por assuntos
		Página dinâmica
		Página estática
	Permite a navegação por	Autor
		Título
		Data de impressão
		Assuntos
		Na totalidade das obras
Tipologia das obras		
Formatos de transferência	JPG	
	Tiff	
	PDF	
	PDF com transcrição	
	GIF	
	PNG	
	ePub	
Dejavu		

Funcionalidades das obras em linha		Rotação de imagem
		Obras com Verbetes/sumários
		Pesquisa temática/por assuntos
		<i>Zoom</i>
		Menu vertical
		Menu horizontal
		Outro
Ajudas		Tutoriais
		Ajuda interativa
Alertas		Página de novidades
		Alertas primeira página
		RSS
		Alertas personalizáveis
Referências bibliográficas		Gestores Referências
Perfis pessoais		Gestão de listas de referências
Disseminação		Personalizada / envio por mail
		Redes sociais
Dados técnicos		Informação técnica
		Estatísticas obras
		Estatísticas de uso
		Data
Facilidade de utilização		
Utilidade		
Performance		

### III. Sondagem grupo focal

No âmbito do programa de PhD na Universidade de Évora, estou a fazer um levantamento acerca das mudanças que são necessárias implementar nas bibliotecas digitais com acervos digitalizados.

As bibliotecas digitais apresentam alguns problemas. Por vezes, é difícil encontrar a obra pretendida e, quando se encontra, nem sempre a pesquisa corresponde às expectativas. O que se poderia mudar?

Gostava de contar com a vossa opinião.

Selecione uma ou mais opções ou adicione outras.

**Muito obrigada pela sua colaboração.**

#### **Opções:**

- A pesquisa das obras digitalizadas deve realizar-se apenas no catálogo em linha
- Os textos em linha devem ser pesquisáveis pelos motores de busca
- As obras devem ser disponibilizadas só em formato texto
- As obras devem ser disponibilizadas só em formato imagem
- As obras devem ser disponibilizadas em formato imagem e texto
- No catálogo deveriam ser criadas hiperligações a bibliografias sobre a obra digitalizada
- Criação de apontadores, na obra digitalizada, para os capítulos, partes, etc.
- Obras muito ilustradas, navegação por texto e por imagem
- Pesquisa nas imagens
- Exportação direta para os gestores de referências (Mendeley, Zotero, etc.)
- 

Opção acrescentada por um respondente:

- As obras devem ser apresentadas em formato acessível (norma w3c)

Figura 50 - Sondagem no Facebook, grupo "Bibliotecas Digitais", 21 jun. 2013



Figura 51 - Sondagem no blogue "Bibliotecas para as Humanidades Digitais", 21 jun. 2013

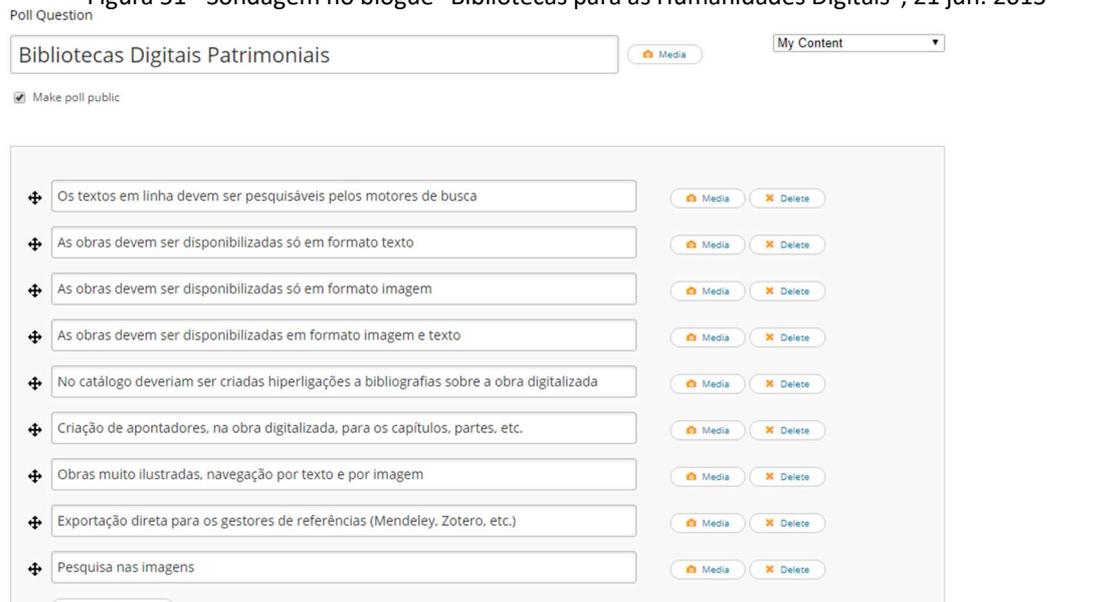
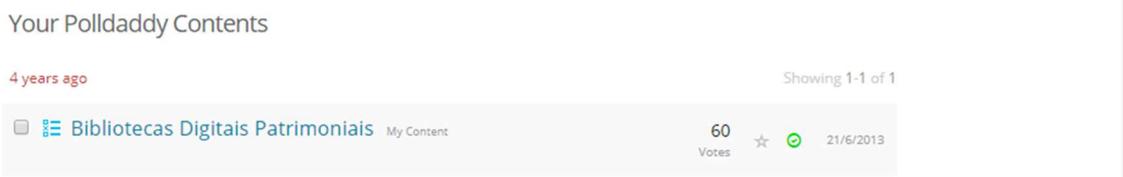


Figura 52 - Sondagem no blogue "Bibliotecas Digitais para as Humanidades": índice de votação



#### IV. Guião do grupo focal

O presente grupo focal integra-se na investigação conducente à obtenção do grau de doutor em Ciências da Informação e da Documentação

Título da tese: Bibliotecas e Humanidades Digitais: novos desafios e oportunidades

Nome: Dália Maria Godinho Guerreiro

Orientadores: Prof. Doutor José Luís Borbinha e Prof. Doutor José António Calixto

Filiação: Universidade de Évora / CIDEHUS

Bolsa de SFRH/BD/82229/2011

Objetivo:

- Identificar as necessidades sentidas pelos investigadores em Humanidades em relação às bibliotecas digitais:
  - Identificar as expetativas de utilização das Biblioteca Digital de Livro Antigo
  - Identificar as vantagens e desvantagens das Biblioteca Digital de Livro Antigo
  - Identificar as expectativas de mudança/otimização

Data: 22/04/2017, às 18:00

Local: Lisboa (residência particular)

Elemento	Área
1	História, História de Arte, Museologia
2	Antropologia, Património
3	Estudos da Cultura
4	Linguística, História do Livro
5	Comunicação Cultural
6	Filosofia

### **Introdução:**

- Apresentação do tema da tese;
  - Justificação das razões de seleção do grupo: investigadores de diversos domínios do conhecimento (História, História da Arte, Linguística, Filosofia, Antropologia, Estudos da Cultura, Comunicação Cultural, Património, Museologia e História do Livro);
  - Duração prevista: cerca de 1 hora
  - Pedido de autorização para a gravação das sessões presenciais;
  - Pedido de autorização para posterior divulgação dos dados obtidos;
  - É garantido o anonimato de todos os participantes.
  - Os dados recolhidos só serão utilizados para o fim enunciado.
- 
1. Quando consulta uma biblioteca digital de livro antigo, o que encontra?
  2. Quais as vantagens das Biblioteca Digital de Livro Antigo face às bibliotecas que possuem apenas obras em papel?
  3. Quais os inconvenientes que assinala na utilização da Biblioteca Digital de Livro Antigo?
  4. Quais as vantagens para o ensino?
  5. Quais as vantagens para investigação?
  6. Que funcionalidades seriam necessárias para tornar a utilização das Biblioteca Digital de Livro Antigo mais eficaz?

**Muito obrigada pela sua colaboração.**

## V. Transcrição do grupo focal

[Apresentação dos elementos do grupo]

**Moderador (M)** – No âmbito da minha investigação, a biblioteca digital é uma coleção de obras digitalizadas de livro antigo, devidamente organizada e estruturada e disponível 24 sobre 24 h, 7 dias por semana (twenty-four seven - 24 hours a day, 7 days a week; all the time) e sem custo para o utilizador. Tem de ter mecanismos de pesquisa e recuperação da informação, para, basicamente não andar à toa numa lista de obras. O livro antigo, que é o que me interessa particularmente, é considerado, grosso modo, os publicados entre 1450 e 1800, mais ou menos. Os objetivos são identificar as expectativas, vantagens e desvantagens da utilização das bibliotecas digitais e ver o que nós queremos mudar.

**A minha primeira questão é: quando se consulta uma biblioteca digital, sobretudo de livro antigo, o que se espera encontrar?**

**R1** – Não é bem o que encontro. É o que espero encontrar facilmente: o livro. Quando faço uma consulta na internet ou na biblioteca é o que eu espero, encontrar o livro. Depois, que o livro esteja em formato imagem, mas com o OCR [com transcrição textual], pode fazer-se a pesquisa dentro do conteúdo do livro. O OCR significa que há um programa que vai ler o livro e, portanto, como se estivesse em formato de texto e que vai ter ao *sítio*.

Também faz jeito encontrar aqueles sumários maravilhosos que antigamente a Biblioteca Nacional e a de Coimbra tinham e que deixaram de ter. Isto desgosta-me porque eu preciso de saber em que ponto do livro estou, qual é o assunto que está antes e o assunto que está depois. Também gostava que, quando faço a pesquisa para

um livro, ele me indicasse outros livros semelhantes que houvesse na mesma biblioteca digital. Quando vou a uma biblioteca analógica, física, por exemplo, numa biblioteca universitária, vou à estante e vejo o livro e os livros próximos que lá estão. Outra coisa é quando têm imagem. Gostava de saber o assunto da imagem que lá está, as ilustrações, as gravuras, gostava de as identificar.

**R4** – Gostava de encontrar os livros na pesquisa do *google*, sem ter que ir a uma biblioteca específica e que, nessa pesquisa, o formato disponibilizado fosse acessível, por exemplo para fazer uma citação, dentro do documento. sem ter de estar a escrever o texto...

**R1** – Não é só isso. Gostava de conseguir encontrar a parte do texto que eu pretendo. Pesquisa por palavra, dentro do documento.

**R1** – Além de ler *on-line*, descarrego os livros.

**R2** – Acrescentaria duas coisas: quando se fala em livro antigo, a maior parte deles são em latim. Existem centenas, não sei se milhares de livros *on-line*, disponíveis para *download* gratuito, mas porque estão em latim... Falta aqui um trabalho de tradução que as bibliotecas podiam promover. Hoje, o conhecimento em latim é limitadíssimo. Quem estuda este período está, obviamente, limitado ao acesso da produção de conhecimento, às fontes, pelo menos.

A segunda questão é a das imagens. Eu procuro sobretudo imagens... os livros que eu procuro têm ilustrações. As bibliotecas devem promover esse conhecimento. Encaminhar, de imagem em imagem... tenho sempre dificuldade: num livro de 800 páginas, por exemplo, talvez um OCR que nos podia ajudar. Isso é uma loucura...

**R3** – Não. Não é loucura. Havia um índice do texto e um índice de imagens. Eu posso consultar o índice das imagens, só a parte das imagens.

**R5** – O índice é fundamental. A velocidade também.

**R4** – O acesso em termos de custo? É gratuito?

**M** – Sim. Existem aquelas que são pagas, mas são de editoras.

**R1** – Por mim, quero ter acesso livre e gratuito. Mesmo que tenha de pagar alguns serviços, por exemplo, se eu quiser imagens de qualidade, posso pagar. O resto...

**R5** – Não só. Nos EUA tinha que se pagar. Era uma biblioteca que tinha documentos sobre a expansão. Eles cobravam uma taxa simbólica.

**R2** – Eu não me importava de pagar.

**R5** – Eles cobravam uma pequena taxa. Quase o custo do envio.

**R2** – Não choca que sejam cobrados. Uma espécie de taxa moderadora que ajudasse a manter a biblioteca.

**R4** – O que eu gosto nas bibliotecas digitais é que me baixaram o custo da investigação. E o tempo. Ficou mais barato, porque já não preciso de me deslocar às bibliotecas. Deixei de precisar de pedir coisas, fotocópias e pagá-las. E levo menos tempo a reunir a informação.

**R5** – As pessoas preferem pagar e ter um serviço de qualidade a não o ter.

**R2** – Até no sentido do abuso...

**R6** – Vamos imaginar que em antropologia eu quero ter acesso ao espólio das reservas do Museu. Esta nova informação poder-me-ia facilitar isso?

**R1** – Sim. Hoje em dia, já. Nas grandes bibliotecas digitais mundiais, a tendência é cada vez maior de anexarem objetos, museológicos, toda a qualidade de artefactos, livros, etc, com toda a informação, não só com a legenda, mas com toda a informação.

**R4** – O acesso ao Matriz que eu tenho em casa. Isto engloba o conceito de biblioteca digital?

**R1** – Sim. Por exemplo, a Europeana que é a europeia engloba todas as coleções (?)

**R5** – A biblioteca digital é a transposição do conhecimento para o digital.

**R2** – Faço muitas pesquisas em latim e aflijo-me porque não sei qual é aquela tradução. Eu quero mais valias numa biblioteca digital. Podia haver uma colaboração com as universidades. O latim, muitas vezes, já vem das fontes dos árabes e dos gregos...

**M** – **Quais as vantagens da biblioteca digital face a uma biblioteca só com obras em papel?**

**R2** – Só tem vantagens. Permite-nos ter acesso às bibliotecas todas do mundo.

**R3** – Tem todas as vantagens, até a preservação. Poupa os livros.

**M** – Mas o livro antigo apresenta poucos problemas. O papel é muito bom. Os manuscritos, esses sim, podem ter problemas. Aliás, se os compararmos com os de final do séc. XIX, não tem nada a ver...

**R1** – A principal vantagem é a possibilidade de fazer a pesquisa alargada em todas as bibliotecas e em qualquer tempo e lugar.

**M** – **Não existem desvantagens?**

**R5** – Desvantagens?

[Silêncio]

**R3** – Sim. Há coisas num livro que no digital às quais não temos acesso: páginas de dimensão diferentes... as marcas de água... não temos a perspetiva do tamanho.

**R6** – O cheiro...

**R3** – Isso é mais sentimental!

**R1** – Pode haver páginas diferentes, o formato das letras, mesmo na mesma edição há diferença a nível das letras, nos caracteres.

**R3** – A encadernação, as texturas...isso perde-se....

**R6** – Mas isso é sentimental...

**R1** – Para a história do livro não é.

**R3** – A textura não é. Vemos a data do livro através das marcas das filigranas e das marcas de água....

**R2** – Ele não é disponibilizado sem estar datado?

**R1** – Isso é a ficha bibliográfica

**R3** – Às vezes conseguimos datar pelas letras e pelo conteúdo, quando é manuscrito.

Há ali margens de 100 anos.

[Conversa sobre a história e os aspetos materiais do livro]

**M** – Voltando, ainda, às desvantagens das bibliotecas digitais do livro antigo

**R1** – Houve a tentativa de fazer a memória do livro. Há essa desvantagem, mas não é.

Às vezes há pormenores que não são perceptíveis a olho nu, mas fazemos uma ampliação outra informação que às vezes não conseguimos ver

**R4** – O conteúdo intelectual passa.

**R3** – Aqui não é só uma desvantagem, é mesmo uma impossibilidade, quem estuda questões físicas tem de ver o objeto ao vivo.

**R1** – Pode haver outra coisa. A investigação trona-se isolada. Quando vou à BN acabo por encontrar outras pessoas. Às vezes torna-se interessante. Trocamos ideias, sabemos o que os outros estão a fazer. No digital, não é possível.

**M** – **Consideram que as bibliotecas digitais são vantajosas. E nas vossas atividades, tanto no ensino como na investigação? Usam os livros digitalizados nas aulas? No ensino, qual é a vantagem?**

**R6** – As mesmas vantagens.

**R2** – Às vezes recomendo, outras, descarrego com mais facilidade. Disponibilizo os acessos.

**R6** – Para o aluno de hoje em dia é mais atrativo do que deslocar-se à biblioteca que passou a ser um sítio maldito, uma chatice. Não pela curiosidade que o livro devia facultar, mas para fazer trabalhos. Há uma mesa, é sossegado, podem estar juntos.

**R3** – Tem vantagens até para os mais pequenos, descarrega-se os livros digitais que são lidos. Alguém lê ao mesmo tempo que se passam as folhas.

**R4** – Mesmo para as línguas... tem a possibilidade de ouvir o livro na própria língua.

**R5** – Numa exposição, o facto de os documentos poderem ser digitalizados permite fazer trabalhos completamente diferentes, imaginando que está ali o documento original, A parte da divulgação. Expor um documento original, as pessoas passam e não conseguem ler o que está lá, mas tanto é possível trabalhá-los graficamente como utilizar dispositivos mais legíveis e didáticos. É interessante a parte sensorial, a cor do

pergaminho, por exemplo, mas em termos de extração dali...é zero. Assim o documento pode ser trabalhado sem perder ligação ao próprio documento.

**R1** – A exposição sobre a guerra, da Gulbenkian, é feita com reproduções....

**R5** – A “ciência do tempo dos descobrimentos” é um exemplo disto, desta lógica das bibliotecas digitais permite criar dispositivos mais didáticos e interessantes

**R4** – E sem custos a vários níveis: preservação, transporte com segurança...

**R6** – E a questão mais sociológica. Essa divulgação abrange mais grupos sociais, é menos classista. Pode ir ao encontro de pessoas que se calhar entraram ali sem saber o porquê... com o olhar menos educado em relação ao documento...

**R4** – Públicos mais generalistas. Não só os que têm possibilidade de ir aos EUA, ou para aqueles que tiveram uma bolsa. O ato da divulgação. Até no formato remetente ao original...

**M** – **Última questão: o que é que gostariam de ver implementado nas bibliotecas digitais, o que se devia mudar, o que se devia fazer. Se consultarmos uma Gallica, um Internet Archive, temos um manancial de obras novas que são disponibilizadas em linha. O que se poderia pedir que fosse mais eficiente?**

**R1** – Encontrar informação. Se estou à procura de alguma coisa, onde encontro livros? Depois, que me desse a informação também das outras obras que viessem a aparecer, como quando se está na Amazon. Faço uma compra e, passado tempo, sugerem outros. Também podem fazê-lo para bibliotecas. Também que aparecesse um alerta: isto já está disponível...

**R4** – Sofisticar os motores de busca

**R3** – Que essa complexidade fosse do lado de lá.

**R2** – O grande desafio vai ser tornar os processos mais acessíveis: palavra chave, índice, a questão da tradução do latim, de outras línguas... árabe, chinês. Os japoneses têm milhares de coisas na área das imagens... mas vedado por causa da língua. O trabalho da gestão. Globalizar. **R4** – É isso. Criar mecanismos mais sofisticados.

**R5** – Se faço uma pesquisa vou ter a *sítios* o alemão, ao holandês. O que faço? Vou ao tradutor automático. Se houvesse essa possibilidade eu queria ter o acesso ao conteúdo desse livro.

**R6** – A cultura asiática...está praticamente tudo vedado, tudo naquelas línguas.

**R2** – E o contrário também para os investigadores lá fora.

**R3** – Aumentem os espólios digitalizados. São raríssimos na Biblioteca Nacional. Quando há, não tem índices ou mal associados às imagens. O Matriz é insuficiente e com erros gravíssimos.

**R1** – Quem for para a Biblioteca Nacional isso deixou de existir.

**R5** – Nos grandes museus está tudo digitalizado e nas duas línguas. O Matriz tem erros gravíssimos, é insuficiente.

**R2** – Não foram atualizados.

**R1** – Eu quero manusear a imagem, se for tridimensional eu quero vê-la

**R5** – Havia uma mentalidade de não dar acesso. Este movimento de livre acesso a tudo vai acabar por contaminar o nosso país. Há sítios públicos, em Portugal, que é preciso contar com a boa disposição do funcionário.

**R1** – Não quero carimbos digitais nos livros, podem estar sobre a informação, podem perturbar a leitura.

**R2** – Isto tem custos. É preciso é que isto cresça. Sou a favor do utilizador pagador.

**R5** – Não quero nada que comprometa o conteúdo. Posso pagar uma utilização ou uma impressão de qualidade. No Museu Britânico, quando estou a descarregar uma imagem, perguntam qual o objetivo. O património e a sua utilização é de todos.

**R4** – É serviço público.

[Conversa sobre museus]

**M** – Voltando ao nosso tema, estávamos a falar de melhorias nas bibliotecas digitais e, agora, devendo ou não ser pagas...

**R3** – A Biblioteca Nacional... Se estiverem a pôr o livro antigo cá fora, podem cobrar direitos de utilização.

**R1** – Deviam ter perfis de utilizador, é fácil criá-los. Terias acesso a imagem de melhor qualidade.

**R5** – Por exemplo, os arquivos da RTP. Cobravam direitos absurdos.

**R1** – Agora já mudou. Os arquivos estão quase todos, os mais relevantes, *on line*.

[Conversa sobre arquivos RTP]

**R2** – A torre do Tombo tem muitos problemas. Não se consegue navegar. Para descarregar documentos com 300 páginas, leva um dia... as imagens levam muito tempo.

**R4** – Estamos muito atrasados. Tem de haver uma mudança de mentalidades. A Europa está a pressionar o país.

**R2** – Não estamos a falar só de Lisboa...

**R4** – Cartórios, batismos, nascimentos, etc. Está muita coisa. Os paroquiais, depende. Muitos nem tratados, quanto mais digitalizados...arrumados em caixinhas...

**R2** – Lisboa está quase todos. Vejam *o site e-tombo* ....

**M** – Antes de concluirmos, vou tentar fazer um pequeno resumo do que foi dito, a ver se todos concordam. Se tomei as notas corretas, ou qualquer coisa assim.

Em relação à primeira pergunta, que era sobre o que encontram numa biblioteca digital, pode dizer-se que são livros digitalizados e disponibilizados, na íntegra, alguns estão em formato texto, alguns têm sumário. Portanto, a primeira coisa que pretendem é encontrar o livro com facilidade, o que tem a ver com a pesquisa e recuperação da informação. Depois querem que as obras estejam em formato texto, que consigam pesquisar no interior dos documentos. Se não estiverem em formato texto, querem auxiliares de navegação, como a ligação a sumários e índices de gravuras. Referiram que deveria haver um sistema que permitisse a tradução das obras que estão em linha. Disseram que queriam o acesso integral e gratuito. Houve quem referisse que não se importava de pagar uma pequena taxa se isso significasse um serviço de qualidade. Disseram ainda que a biblioteca pode ser definida como o repositório da produção intelectual humana, que abrange, não só os livros, mas também os artefactos e documentação de arquivo.

Em relação às vantagens, houve logo alguém que disse que só tem vantagens. A primeira é que permite a preservação das obras. Temos muitas bibliotecas à disposição, sem sairmos do sítio e em qualquer altura, porque estamos em casa. O que disseram também, que me parece desvantagem é que não permite o acesso à materialidade da obra, com encadernação da obra, marcas de água, perspectiva do tamanho que é muito importante. Foram falando das vantagens e desvantagens.

Também falaram que a investigação se torna isolada. Como grande vantagem temos que o digital permite manusear a imagem para ver detalhes que escapam a olho nu. Além disso, o conteúdo intelectual da obra é preservado.

Na atividade letiva, dizem que as vantagens são as mesmas. Dizem que disponibilizam aos alunos o acesso às obras. Mesmo para as línguas, os leitores de PDF permitem a consulta em modo de leitura que é bom para ouvirem o livro na própria língua. Além disso, permite a divulgação das obras, por exemplo, em exposições. Acho que em relação a isto, não disseram assim mais nada. Ah, sim. Em relação aos custos, todos pareceram concordar que as bibliotecas diminuam os gastos com as deslocações e, provavelmente, também com as reproduções, mas não anotei isso.

Em relação às expectativas, também foram dizendo algumas coisas na primeira pergunta. Sugeriram a sofisticação dos motores de busca, a criação de sistemas de alerta, dizendo que queriam ser informados de acordo com os interesses mostrados em consultas anteriores. Falaram da tradução, não só do latim, mas de outras, mesmo das orientais. Falaram de aumentar os espólios digitados, sobretudo, em português. Da existência, ou antes, de não gostarem dos carimbos nos livros, nem de nada que comprometa o conteúdo. De perfis para terem acesso a imagens, ou antes, a conteúdos de qualidade.

Acho que é tudo. Não sei se concordam. Há algum ponto que não reflita a vossa opinião. Não sei se querem acrescentar alguma coisa. Façam favor.

[Anuência geral]

Agradecendo a vossa colaboração, dou por terminada a sessão.



## VI. Grupo focal resultados estruturados nível 2

### Grupo focal – resultados

#### 1. Quando consulta uma biblioteca digital de livro antigo, o que encontra?

Definição de BD	<ul style="list-style-type: none"><li>– Nas grandes bibliotecas digitais mundiais, a tendência é cada vez maior de anexarem objetos, museológicos, toda a qualidade de artefactos, livros, etc., com toda a informação, não só com a legenda, mas com toda a informação. (R1)</li><li>– A biblioteca digital é a transposição do conhecimento para o digital. (R5)</li></ul>
Pesquisa na BD	<ul style="list-style-type: none"><li>– Quando faço uma consulta na internet ou na biblioteca é o que eu espero, encontrar o livro. (R1)</li></ul>
Funcionalidades no acesso ao conteúdo	<ul style="list-style-type: none"><li>– Que o livro esteja em formato imagem, mas com o OCR [com transcrição textual], pode fazer-se a pesquisa dentro do conteúdo do livro (R1)</li><li>– Aqueles sumários maravilhosos que antigamente a Biblioteca Nacional e a de Coimbra tinham e que deixaram de ter. Isto desgosta-me porque eu preciso de saber em que ponto do livro estou, qual é o assunto que está antes e o assunto que está depois. (R1)</li><li>– Havia um índice do texto e um índice de imagens. Eu posso consultar o índice das imagens, só a parte das imagens. (R3)</li><li>– O índice é fundamental. (R5)</li><li>– Quando há, não tem índices ou mal associados às imagens. (R3)</li><li>– Faço muitas pesquisas em latim e aflijo-me porque não sei qual é aquela tradução. (R2)</li></ul>
Transferência	<ul style="list-style-type: none"><li>– Além de ler <i>on-line</i>, descarrego os livros. (R1)</li><li>– A velocidade [de transferência] também [é fundamental]. (R5)</li></ul>

Pagamento de serviços	<ul style="list-style-type: none"> <li>– [O acesso é gratuito]. Existem aquelas que são pagas, mas são de editoras. (M)</li> <li>– Nos EUA tinha que se pagar. Era uma biblioteca que tinha documentos sobre a expansão. Eles cobravam uma taxa simbólica [...]. Eles cobravam uma pequena taxa. Quase o custo do envio. (R5)</li> </ul>
-----------------------	--

2. Quais as vantagens das Bibliotecas Digitais de Livro Antigo face às bibliotecas que possuem apenas obras em papel?

Geral	– Só tem vantagens. Permite-nos ter acesso às bibliotecas todas do mundo. (R2)
Acessibilidade	<ul style="list-style-type: none"> <li>– A principal vantagem é a possibilidade de fazer a pesquisa alargada em todas as bibliotecas e em qualquer tempo. (R1)</li> <li>– Numa exposição, o facto de os documentos poderem ser digitalizados permite fazer trabalhos completamente diferentes, imaginando que está ali o documento original, A parte da divulgação [...]. Assim, o documento pode ser trabalhado sem perder ligação ao próprio documento (R5)</li> <li>– E a questão mais sociológica. Essa divulgação abrange mais grupos sociais, é menos classista. Pode ir ao encontro de pessoas que se calhar entraram ali sem saber o porquê... com o olhar menos educado em relação ao documento... (R6)</li> </ul>
Preservação	– Tem todas as vantagens, até a preservação. Poupa os livros. (R3)

3. Quais os inconvenientes que assinala na utilização da Biblioteca Digital de Livro Antigo?

Perda de contacto com	– Há coisas num livro que no digital às quais não temos acesso: páginas de dimensão diferentes... as marcas de água...não
-----------------------	---

a materialidade do livro	temos a perspetiva do tamanho. (R3) – O cheiro... (R6)
Estudo do livro	– Aqui não é só uma desvantagem, é mesmo uma impossibilidade, quem estuda questões físicas tem de ver o objeto ao vivo. (R3) – Pode haver páginas diferentes, o formato das letras, mesmo na mesma edição há diferença a nível das letras, nos caracteres. (R1) – A encadernação, as texturas...isso perde-se... (R3)
Socialização	– A investigação torna-se isolada. Quando vou à BN acabo por encontrar outras pessoas. Às vezes torna-se interessante. Trocamos ideias, sabemos o que os outros estão a fazer. No digital, não é possível. (R1)

#### 4. Quais as vantagens para o ensino?

Geral	– As mesmas vantagens. (R6)
Material didático	– Às vezes recomendo, outras, descarrego com mais facilidade. Disponibilizo os acessos. (R2) – [...] desta lógica das bibliotecas digitais permite criar dispositivos mais didáticos e interessantes (R5)
Adequação aos hábitos dos alunos	– Para o aluno de hoje em dia é mais atrativo do que deslocar-se à biblioteca que passou a ser um sítio maldito, uma chatice. (R6)
Ensino da língua	– [...] descarrega-se os livros digitais que são lidos. Alguém lê ao mesmo tempo que se passam as folhas (R3) – Mesmo para as línguas... tem a possibilidade de ouvir o livro na própria língua (R4)

#### 5. Quais as vantagens para a investigação?

Diminuição de	– O que eu gosto nas bibliotecas digitais é que me baixaram o
---------------	---

custos/tempo	custo da investigação. E o tempo. Ficou mais barato, porque já não preciso de me deslocar às bibliotecas. Deixei de precisar de pedir coisas, fotocópias e pagá-las. E levo menos tempo a reunir a informação. (R4)
Maior acesso às fontes	– Públicos mais generalistas. Não só os que têm possibilidade de ir aos EUA, ou para aqueles que tiveram uma bolsa. O ato da divulgação. Até no formato remetente ao original... (R5)

6. Que funcionalidades seriam necessárias para tornar a utilização das Biblioteca Digital de Livro Antigo mais eficaz?

Pesquisa	<ul style="list-style-type: none"> <li>– Gostava de encontrar os livros na pesquisa do <i>google</i>, sem ter que ir a uma biblioteca específica e que, nessa pesquisa, o formato disponibilizado fosse acessível, por exemplo para fazer uma citação, dentro do documento. Sem ter de estar a escrever o texto... (R4)</li> <li>– Sofisticar os motores de busca (R4)</li> <li>– Que essa complexidade fosse do lado de lá (R3)</li> <li>– O grande desafio vai ser tornar os processos mais acessíveis: palavra chave, índice, [...]. Os japoneses têm milhares de coisas na área das imagens... mas vedado por causa da língua. (R2)</li> </ul>
Acesso em contexto	– Gostava que, quando faço a pesquisa para um livro, ele me indicasse outros livros semelhantes que houvesse na mesma biblioteca digital. Quando vou a uma biblioteca analógica, física, por exemplo, numa biblioteca universitária, vou à estante e vejo o livro e os livros próximos que lá estão. (R1)
Pesquisa no documento	– Gostava de conseguir encontrar a parte do texto que eu pretendo. Pesquisa por palavra, dentro do documento (R1)
Acesso a conteúdos textuais	– O formato disponibilizado fosse acessível, por exemplo para fazer uma citação, dentro do documento. Sem ter de estar a escrever o texto... (R4)

<p>Acesso a conteúdos não textuais</p>	<ul style="list-style-type: none"> <li>– Quando têm imagem. Gostava de saber o assunto da imagem que lá está, as ilustrações, as gravuras, gostava de as identificar. Eu procuro sobretudo imagens... os livros que eu procuro têm ilustrações. As bibliotecas devem promover esse conhecimento. Encaminhar, de imagem em imagem... tenho sempre dificuldade. (R2)</li> <li>– Eu quero manusear a imagem, se for tridimensional eu quero vê-la (R1)</li> </ul>
<p>Tradução dos conteúdos</p>	<ul style="list-style-type: none"> <li>– Quando se fala em livro antigo, a maior parte deles são em latim. Existem centenas, não sei se milhares de livros <i>on-line</i>, disponíveis para <i>download</i> gratuito, mas porque estão em latim... Falta aqui um trabalho de tradução que as bibliotecas podiam promover. Hoje, o conhecimento em latim é limitadíssimo. Quem estuda este período está, obviamente, limitado ao acesso da produção de conhecimento, às fontes, pelo menos. (R2)</li> <li>– Eu quero mais valias numa biblioteca digital. Podia haver uma colaboração com as universidades. O latim, muitas vezes, já vem das fontes dos árabes e dos gregos... (R2)</li> <li>– [...] a questão da tradução do latim, de outras línguas...árabe, chinês. (R2)</li> <li>– Se faço uma pesquisa vou ter a <i>sítios</i> o alemão, ao holandês. O que faço? Vou ao tradutor automático. Se houvesse essa possibilidade eu queria ter o acesso ao conteúdo desse livro. (R5)</li> <li>– A cultura asiática... está praticamente tudo vedado, tudo naquelas línguas (R6)</li> <li>– E o contrário também para os investigadores lá fora (R2)</li> </ul>
<p>Serviço de alertas</p>	<ul style="list-style-type: none"> <li>– Encontrar informação. Se estou à procura de alguma coisa, onde encontro livros? Depois, que me desse a informação também das outras obras que viessem a aparecer, como</li> </ul>

	quando se está na Amazon. Faço uma compra e, passado tempo, sugerem outros. Também podem fazê-lo para bibliotecas. Queria que aparecesse um alerta: isto já está disponível... (R1)
Perfis de utilizador	– Deviam ter perfis de utilizador, é fácil criá-los. (R1)
Autenticidade / Legibilidade  Direitos	– Não quero carimbos digitais nos livros, podem estar sobre a informação, podem perturbar a leitura (R1) – Não quero nada que comprometa o conteúdo. Posso pagar uma utilização ou uma impressão de qualidade. (R5) – A Biblioteca Nacional... Se estiverem a pôr o livro antigo cá fora, podem cobrar direitos de utilização (R3)
Gratuidade / serviços pagos	– Mesmo que tenha de pagar alguns serviços, por exemplo, se eu quiser imagens de qualidade, posso pagar. (R1) – No Museu Britânico, quando estou a descarregar uma imagem, perguntam qual o objetivo. O património e a sua utilização é de todos (R5) – Eu não me importava de pagar. (R2) – Não choca que [alguns serviços] sejam cobrados. Uma espécie de taxa moderadora que ajudasse a manter a biblioteca. (R2) – As pessoas preferem pagar e ter um serviço de qualidade a não o ter. (R5) – Isto tem custos. É preciso é que isto cresça. Sou a favor do utilizador pagador. (R2)
Incremento dos conteúdos	– Aumentem os espólios digitalizados. São raríssimos na Biblioteca Nacional. (R3)

## VII. Inquérito por questionário ronda 1

### Introdução

O presente questionário integra-se na investigação conducente à obtenção do grau de doutor em Ciências da Informação e da Documentação, na Universidade de Évora/CIDEHUS, com o título “Bibliotecas e Humanidades Digitais: Novos desafios e oportunidades”, realizada por Dália Maria Godinho Guerreiro e tendo, como orientadores, os Professores Doutor José Luís Borbinha e Doutor José António Calixto. Esta investigação é financiada pela FCT através da bolsa de formação avançada, SFRH/BD/82229/2011.

Objetivo: compreender a utilização das bibliotecas digitais pelos investigadores em Humanidades:

- Identificar as modalidades de utilização das bibliotecas digitais de livro antigo;
- Identificar as vantagens e desvantagens das bibliotecas digitais de livro antigo;
- Identificar as expectativas de mudança/otimização das bibliotecas digitais de livro antigo.

Com este questionário, pretendemos validar os resultados do grupo focal e os dados coletados na literatura consultada.

É garantido o anonimato de todos os participantes.

Os dados recolhidos só serão utilizados para o fim enunciado.

Muito obrigada pela sua colaboração.

Tempo calculado de resposta: 7min

## Questões

<b>1. Ao aceder a uma biblioteca digital de livro antigo, o que costuma encontrar?</b>	Nunca	Quase nunca	Por vezes	Quase sempre	Sempre
--	-------	-------------	-----------	--------------	--------

1.1. Informação organizada, de forma validada e credível					
1.2. Obras disponibilizadas em acesso livre					
1.3. Obras disponibilizadas na íntegra					
1.4. Obras disponibilizadas com pontos de acesso explícitos, por exemplo, sumários					
1.5. Identificação dos elementos gráficos das obras					
1.6. Pesquisa nos elementos textuais dos documentos (OCR)					
1.7. Exportação direta das referências bibliográficas para os gestores de referências bibliográficas					
1.8. Partilha das obras digitalizadas nas redes sociais					

<b>2. Quais as vantagens da biblioteca digital de livro antigo face às bibliotecas físicas?</b>	Discordo totalmente	Discordo em parte	Indiferente	Concordo em parte	Concordo totalmente
---	---------------------	-------------------	-------------	-------------------	---------------------

2.1. Acesso às obras em qualquer lugar e em qualquer horário					
2.2. Maior facilidade de leitura por pessoas com limitações físicas					
2.3. Preservação dos documentos originais					

2.4. Maior legibilidade do texto					
2.5. Preservação do conteúdo intelectual da obra					
2.6. Acesso a fontes documentais no contexto da atividade letiva					
2.7. Diminuição de custos na investigação decorrente das deslocações e das reproduções					
2.8. Otimização do processo de investigação pela redução do tempo na recolha dos dados					
2.9. Novas leituras e análises através da aplicação de novas ferramentas digitais*					
2.10. Alargamento da informação, como geradora de oportunidades					
2.11. Divulgação da cultura e do conhecimento					

\* Por exemplo, contagem de palavras, estudos de contexto, georreferenciação, etc.

<b>3. Quais as desvantagens das bibliotecas digitais de livro antigo em relação às bibliotecas físicas?</b>	Discordo totalmente	Discordo em parte	Indiferente	Concordo em parte	Concordo totalmente
3.1. Falta de contacto físico com livro					
3.2. Diminuição do acesso aos aspetos materiais da obra					
3.3. Elementos móveis das obras, não são representáveis digitalmente na sua dinâmica					
3.4. Dificuldade no processamento da informação, decorrente do excesso de resultados					
3.5. Impossibilidade de socialização e a troca de ideias entre pares					

4. Quais as suas expectativas de mudança/otimização nas bibliotecas digitais de livro antigo?	Discordo totalmente	Discordo em parte	Indiferente	Concordo em parte	Concordo totalmente
4.1. Pesquisa no conteúdo dos documentos a partir dos motores de busca da web					
4.2. Possibilidade de pesquisa no conteúdo textual dos documentos					
4.3. Possibilidade de pesquisa das ilustrações pelo respetivo tema					
4.4. Pesquisa das ilustrações autónoma em relação ao texto					
4.5. Ligação entre as ilustrações e o texto que lhe corresponde					
4.6. Exportação das obras em múltiplos formatos, nomeadamente, para dispositivos móveis					
4.7. Criação de auxiliares à navegação (ajudas interativas e tutoriais)					
4.8. Criação de pontos de acesso ao conteúdo das obras (por exemplo, sumários, índices)					
4.9. Visualização da referência à parte da obra onde o leitor se encontra					
4.10. Possibilidade de criação de perfis de utilizador					
4.11. Possibilidade de criação de várias listas em cada perfil de utilizador					
4.12. Possibilidade de partilha de listas pessoais					
4.13. Possibilidade de personalização do sistema de alertas					
4.14. Possibilidade de navegação, a partir de uma obra, em obras semelhantes					
4.15. Possibilidade de tradução automática dos documentos em linha					

4.16. Possibilidade de melhorar o reconhecimento de caracteres de forma colaborativa					
4.17. Acesso a obras digitais de qualidade mediante pagamento de taxas moderadoras					
4.18. Incremento da colocação em linha de conteúdos em português					

### Dados sociodemográficos

Idade    20-29 anos    30-39 anos    40-49 anos    50-59 anos    + 60 anos

Género    Masculino                    Feminino

Formação académica    Mestrado                    Doutoramento                    Pós-Doutoramento

Área de investigação \*

    Sociologia (antropologia e etnologia)                   

    Ciências da comunicação                   

    História e arqueologia                   

    Línguas e literatura                   

    Filosofia, ética e religião                   

    Arte (história da arte)                   

\* Nomenclatura a partir da Classificação de Domínios Científicos e Tecnológicos, 2007 (FOS)



## VIII. Inquérito por questionário ronda 2

### Introdução

O presente questionário corresponde à segunda ronda do Inquérito por Questionário e integra-se na investigação conducente à obtenção do grau de doutor em Ciências da Informação e da Documentação, na Universidade de Évora/CIDEHUS, com o título “Bibliotecas e Humanidades Digitais: Novos desafios e oportunidades”, realizada por Dália Maria Godinho Guerreiro e tendo, como orientadores, os Professores Doutor José Luís Borbinha e Doutor José António Calixto.

Esta investigação é financiada pela FCT através da bolsa de formação avançada, SFRH/BD/82229/2011.

Síntese dos resultados obtidos na primeira ronda:

- A maior parte dos investigadores reconhece que a informação se encontra quase sempre organizada, de forma validada e credível e que costuma encontrar as obras disponibilizadas em acesso livre e na íntegra.
- O reconhecimento das vantagens trazidas pelas bibliotecas digitais é consensual, com um elevado índice de concordância total em todas as questões.
- A diminuição do acesso aos aspetos materiais da obra é a principal desvantagem apontada às bibliotecas digitais
- Relativamente ao desenvolvimento das bibliotecas digitais, os níveis mais elevados de expectativas dizem respeito à possibilidade de pesquisa no conteúdo textual dos documentos e ao incremento da coleção em linha de conteúdos em português.

Depois de analisados os resultados da primeira ronda, procedeu-se à hierarquização das respostas e reordenou-se o questionário em função do consenso obtido.

É garantido o anonimato de todos os participantes.

Os dados recolhidos só serão utilizados para o fim enunciado.

O tempo estimado para conclusão deste questionário é de 6 min.

Muito obrigada pela sua colaboração.

Para prosseguir no inquérito, por favor, clique em OK e em seguida em Próx

### Questões

<b>1. Ao aceder a uma biblioteca digital de livro antigo, o que costuma encontrar?</b>	Nunca	Quase nunca	Por vezes	Quase sempre	Sempre
--	-------	-------------	-----------	--------------	--------

1.1. Informação organizada, de forma validada e credível					
1.2. Obras disponibilizadas em acesso livre					
1.3. Obras disponibilizadas na íntegra					

<b>2. Quais as vantagens da biblioteca digital de livro antigo face às bibliotecas físicas?</b>	Discordo totalmente	Discordo em parte	Indiferente	Concordo em parte	Concordo totalmente
---	---------------------	-------------------	-------------	-------------------	---------------------

2.12. Diminuição de custos na investigação decorrente das deslocações e das reproduções					
2.13. Acesso às obras em qualquer lugar e em qualquer horário					
2.14. Otimização do processo de investigação pela redução do tempo na recolha dos dados					
2.15. Preservação dos documentos originais					
2.16. Divulgação da cultura e do conhecimento					

2.17. Alargamento da informação, como geradora de oportunidades					
2.18. Novas leituras e análises através da aplicação de novas ferramentas*					
2.19. Maior facilidade de leitura por pessoas com limitações físicas					
2.20. Acesso a fontes documentais no contexto da atividade letiva					
2.21. Preservação do conteúdo intelectual da obra					
2.22. Maior legibilidade do texto					

\* Por exemplo, contagem de palavras, estudos de contexto, georreferenciação, etc.

<b>3. Quais as desvantagens das bibliotecas digitais de livro antigo em relação às bibliotecas físicas?</b>	Discordo totalmente	Discordo em parte	Indiferente	Concordo em parte	Concordo totalmente
---	---------------------	-------------------	-------------	-------------------	---------------------

3.1. Elementos móveis das obras, não são representáveis digitalmente na sua dinâmica					
3.2. Diminuição do acesso aos aspetos materiais da obra					
3.3. Falta de contacto físico com livro					
3.4. Impossibilidade de socialização e a troca de ideias entre pares					

<b>4. Quais as suas expectativas de mudança/otimização nas bibliotecas digitais de livro antigo?</b>	Discordo totalmente	Discordo em parte	Indiferente	Concordo em parte	Concordo totalmente
--	---------------------	-------------------	-------------	-------------------	---------------------

4.1. Possibilidade de pesquisa no conteúdo textual dos documentos					
---	--	--	--	--	--

4.2. Ligação entre as ilustrações e o texto que lhe corresponde					
4.3. Possibilidade de navegação, a partir de uma obra, em obras semelhantes					
4.4. Pesquisa das ilustrações autónoma em relação ao texto					
4.5. Criação de pontos de acesso ao conteúdo das obras (por exemplo, sumários, índices)					
4.6. Incremento da colocação em linha de conteúdos em português					
4.7. Possibilidade de pesquisa das ilustrações pelo respetivo tema					
4.8. Possibilidade de melhorar o reconhecimento de caracteres de forma colaborativa					
4.9. Pesquisa no conteúdo dos documentos a partir dos motores de busca da web					
4.10. Visualização da referência à parte da obra onde o leitor se encontra					
4.11. Possibilidade de criação de perfis de utilizador					
4.12. Possibilidade de criação de várias listas em cada perfil de utilizador					
4.13. Exportação das obras em múltiplos formatos, nomeadamente, para dispositivos móveis					
4.14. Possibilidade de personalização do sistema de alertas					
4.15. Criação de auxiliares à navegação (ajudas interativas e tutoriais)					
4.16. Possibilidade de tradução automática dos documentos em linha					
4.17. Possibilidade de partilha de listas pessoais					

4.18. Acesso a obras digitais de qualidade mediante pagamento de taxas moderadoras					
--	--	--	--	--	--

**Dados sociodemográficos**

Idade    30-39 anos    40-49 anos    50-59 anos    + 60 anos  
                                                           

Género    Masculino    Feminino  
                   

Formação académica    Mestrado    Doutoramento    Pós-Doutoramento  
                                       

Área de investigação    Arte (história da arte)                      
Ciências da comunicação e da informação                      
Estudos de cultura  
Filosofia, ética e religião                      
História e arqueologia                      
Humanidades digitais / Humanística Digital  
Línguas e literatura                      
Museologia e património  
Sociologia (antropologia e etnologia)



## IX. Glossário do Inquérito por questionário

Livro antigo – Livro impresso com caracteres móveis entre 1501 e 1800.

Acesso livre – Disponibilização dos conteúdos em linha sem restrições.

Digitalização na íntegra – Digitalização de todo o conteúdo da obra, de capa a capa.

Sumário – Lista sequencial de títulos e subtítulos dos conteúdos (partes, capítulos, subcapítulos) pela ordem apresentada na obra e com indicação das páginas ou colunas onde iniciam.

Elementos gráficos das obras – Ilustrações, mapas, heráldica, esquemas existentes no interior do documento.

OCR (Optical Character Recognition) – Reconhecimento ótico de caracteres; tecnologia utilizada para converter as imagens dos caracteres existentes nas obras digitalizadas num arquivo de texto editável, ou seja, para transcrever os documentos.

Gestores de referências bibliográficas – Programas (ex. Mendeley, Zotero, Endnote) que permitem criar uma base de dados de referências bibliográficas, com a possibilidade de referenciar citações e criar bibliografias, de forma automática, em trabalhos académicos ou de investigação.

Legibilidade do texto – Qualidade do texto que permite distinguir e reconhecer os caracteres e, portanto, realizar a respetiva leitura.

Novas ferramentas (contagem de palavras, estudos de contexto, georreferenciação, etc.) – Programas informáticos que permitem novas formas de trabalhar um texto

Ferramenta de contagem de palavras – Programa informático que permite contar as palavras de um texto e visualizar a sua representatividade.

Ferramenta para estudos de contexto – Programa informático que permite relacionar palavras ou conceitos dentro de um texto ou entre textos, criando, por ex., redes de influência.

Ferramenta de georreferenciação – Programa informático que permite associar uma localização geográfica à designação de um lugar e construir mapas dinâmicos.

Ajuda interativa – Funcionalidade de ajuda, obtida ao passar o cursor sobre um ícone ou texto, abrindo uma janela de texto com indicações sobre a ação associada.

Tutorial – Ferramenta de aprendizagem (vídeo, texto, imagem, etc.) detalhando, a par e passo o funcionamento de algo e, neste caso particular, as funcionalidades associadas a uma biblioteca digital.

Perfil de utilizador – Registo eletrónico de dados pessoais, que pode permitir aceder à informação de forma personalizada, em função dos interesses definidos pelo utilizador, e/ou criar um espaço pessoal para gestão de conteúdos.

Sistema de alerta – Sistema tecnológico utilizado para informar o subscritor acerca de uma determinada mudança; neste caso particular, informação relativa a alteração ou adição de obras numa biblioteca digital.

Disponível em: <http://bdh.hypotheses.org/category/glossario>

## **X. Convite à participação no inquérito por questionário, ronda 1 e 2**

### **Ronda 1**

No âmbito da tese de doutoramento que tenho vindo a desenvolver sobre bibliotecas digitais e considerando a sua competência e conhecimentos acerca deste tema, gostaria que integrasse um grupo para validar as hipóteses que apresento.

Caso esteja disponível, peço-lhe que responda ao questionário disponível em: <https://pt.surveymonkey.com/r/BibDigLivreAnt2>

O tempo de resposta estimado é de 5 minutos.

Agradeço, desde já, a sua colaboração, com os mais cordiais cumprimentos,

### **Ronda 2**

Estando a realizar a segunda ronda do Inquérito por Questionário, venho pedir-lhe, novamente, que responda ao novo questionário disponível em:

<https://pt.surveymonkey.com/r/BibDigLivreAnt2>

Nesta ronda, alguns termos do questionário estão associados a um glossário.

Dado que este questionário é feito a partir do anterior, eliminando as questões que se revelaram menos importantes, o tempo de resposta estimado é de 5 minutos.

Atendendo ao elevado nível de concordância obtido na primeira ronda, prevê-se que esta seja a última consulta.

Agradeço, desde já e muito reconhecidamente a sua colaboração, com os mais cordiais cumprimentos,



## **XI. Normas ISO**

ISO 10196:2003

Document imaging applications -- Recommendations for the creation of original documents

ISO 10918-2:1995

Information technology -- Digital compression and coding of continuous-tone still images: Compliance testing

ISO 10918-4:1999

Information technology -- Digital compression and coding of continuous-tone still images: Registration of JPEG profiles, SPIFF profiles, SPIFF tags, SPIFF colour spaces, APPn markers, SPIFF compression types and Registration Authorities (REGAUT)

ISO 10918-5:2013

Information technology -- Digital compression and coding of continuous-tone still images: JPEG File Interchange Format (JFIF)

ISO 12033:2001

Electronic imaging -- Guidance for selection of document image compression methods

ISO 12233:2000

Photography - Electronic still-picture cameras – Resolution measurements

ISO 12234-1:2012

Electronic still-picture imaging -- Removable memory -- Part 1: Basic removable-memory model

ISO 12234-2:2001

Electronic still-picture imaging -- Removable memory -- Part 2: TIFF/EP image data format

ISO 12234-3:2016

Electronic still picture imaging -- Removable memory -- Part 3: XMP for digital photography

ISO 12639:2004

Graphic technology -- Prepress digital data exchange -- Tag image file format for image technology (TIFF/IT)

ISO 12641:1997

Graphic technology - Prepress digital data exchange - Colour targets for input scanner calibration

ISO 14289-1:2014

Document management applications -- Electronic document file format enhancement for accessibility -- Part 1: Use of ISO 32000-1 (PDF/UA-1)

ISO 15076-1:2005

Image technology colour management -- Architecture, profile format and data structure -- Part 1: Based on ICC.1:2004-10

ISO 15444-11:2007

Information technology -- JPEG 2000 image coding system: Wireless

ISO 15444-9:2005

Information technology -- JPEG 2000 image coding system: Interactivity tools, APIs and protocols

ISO 15739:2017

Photography -- Electronic still-picture imaging -- Noise measurements

ISO 15740:2013

Photography -- Electronic still picture imaging -- Picture transfer protocol (PTP) for digital still photography devices

ISO 15801:2004

Electronic imaging -- Information stored electronically -- Recommendations for trustworthiness and reliability

ISO 15930-6:2003

Graphic technology -- Prepress digital data exchange using PDF -- Part 6: Complete exchange of printing data suitable for colour-managed workflows using PDF 1.4 (PDF/X-3)

ISO 16067-1:2003

Photography - Electronic scanners for photographic images - Spatial resolution measurements - Part 1: Scanners for reflective media

ISO 16067-2:2004

Photography - Electronic scanners for photographic images - Spatial resolution measurements -- Part 2: Film scanners

ISO 17321-1:2006

Graphic technology and photography -- Colour characterisation of digital still cameras (DSCs) -- Part 1: Stimuli, metrology and test procedures

ISO 18916:2007

Imaging materials -- Processed imaging materials -- Photographic activity test for enclosure materials

ISO 18921:2008

Imaging materials -- Compact discs (CD-ROM) -- Method for estimating the life expectancy based on the effects of temperature and relative humidity

ISO 18923:2000

Imaging materials -- Polyester-base magnetic tape -- Storage practices

ISO 18925:2013

Imaging materials -- Optical disc media -- Storage practices

ISO 18926:2012

Imaging materials -- Information stored on magneto-optical (MO) discs -- Method for estimating the life expectancy based on the effects of temperature and relative humidity

ISO 18927:2013

Imaging materials -- Recordable compact disc systems -- Method for estimating the life expectancy based on the effects of temperature and relative humidity

ISO 18930:2011

Imaging materials -- Pictorial colour reflection prints -- Methods for evaluating image stability under outdoor conditions

ISO 18933:2012

Imaging materials -- Magnetic tape -- Care and handling practices for extended usage

ISO 18934:2011

Imaging materials -- Multiple media archives -- Storage environment

ISO 18938:2014

Imaging materials -- Optical discs -- Care and handling for extended storage

ISO 18943:2014

Imaging materials -- Magnetic hard drives used for image storage -- Care and handling

ISO 19005-1:2005

Document management -- Electronic document file format for long-term preservation  
-- Part 1: Use of PDF 1.4 (PDF/A-1)

Preservação de gestão documental: formatos de documentos eletrônicos para  
preservação de longa duração, uso de PDF (versão 1.4)

ISO 19263-1:2017

Photography -- Archiving systems -- Part 1: Best practices for digital image capture of  
cultural heritage material

ISO 19264-1:2017

Photography -- Archiving systems -- Image quality analysis -- Part 1: Reflective originals

ISO 19566-1:2016

Information technology -- JPEG Systems -- Part 1: Packaging of information using  
codestreams and file formats

ISO 21550 Sept. 2004

Photography - Electronic scanners for photographic images - Dynamic range  
measurements

ISO 21550:2004

Photography -- Electronic scanners for photographic images -- Dynamic range  
measurements

ISO 22028-1:2016

Photography and graphic technology -- Extended colour encodings for digital image  
storage, manipulation and interchange -- Part 1: Architecture and requirements

ISO 22028-2:2013

Photography and graphic technology -- Extended colour encodings for digital image storage, manipulation and interchange -- Part 2: Reference output medium metric RGB colour image encoding (ROMM RGB)

ISO 22028-3:2006

Photography and graphic technology -- Extended colour encodings for digital image storage, manipulation and interchange -- Part 3: Reference input medium metric RGB colour image encoding (RIMM RGB)

ISO 22028-3:2012

Photography and graphic technology -- Extended colour encodings for digital image storage, manipulation and interchange -- Part 3: Reference input medium metric RGB colour image encoding (RIMM RGB)

ISO 22028-4:2012

Photography and graphic technology -- Extended colour encodings for digital image storage, manipulation and interchange -- Part 4: European Colour Initiative RGB colour image encoding [eciRGB (2008)]

ISO 26324:2012

Information and documentation -- Digital object identifier system

ISO 32000-2:2017

Document management -- Portable document format -- Part 2: PDF 2.0

ISO 18948

[Under development]

Imaging Materials -- Photo Books -- Test Methods for Permanence and Durability

ISO 18945

[Under development]

Imaging materials -- Pictorial color reflection prints -- Comparison of image degradation observed between ISO 18930 accelerated weathering test method and outdoor exposure

ISO 15444-1:2016

(ISO 15444-1 a ISO 15444-14)

Information technology -- JPEG 2000 image coding system: Core coding system



## XII. Lista do Livro antigo na BND

### Amostra

Endereço	Cota	Séc.	Data
<a href="http://purl.pt/28978/">http://purl.pt/28978/</a>	INC. 305 E INC. 305	15	1451
<a href="http://purl.pt/26240">http://purl.pt/26240</a>	INC. 462	15	1480
<a href="http://purl.pt/26131">http://purl.pt/26131</a>	ALC. 71//3	15	1486
<a href="http://purl.pt/22008">http://purl.pt/22008</a>	INC. 1484	15	1489
<a href="http://purl.pt/16559">http://purl.pt/16559</a>	INC. 551	15	1489
<a href="http://purl.pt/16560">http://purl.pt/16560</a>	INC. 1412	15	1492
<a href="http://purl.pt/22003">http://purl.pt/22003</a>	INC. 1413	15	1492
<a href="http://purl.pt/24139">http://purl.pt/24139</a>	INC. 1438	15	1493
<a href="http://purl.pt/24408">http://purl.pt/24408</a>	INC. 1359	15	1493
<a href="http://purl.pt/22000">http://purl.pt/22000</a>	INC. 94	15	1494
<a href="http://purl.pt/22004">http://purl.pt/22004</a>	INC. 1414 E INC. 1415	15	1494
<a href="http://purl.pt/22011">http://purl.pt/22011</a>	INC. 1596	15	1494
<a href="http://purl.pt/22010">http://purl.pt/22010</a>	INC. 1541 INC. 1542 INC. 1543 INC. 1544		151495
<a href="http://purl.pt/22001">http://purl.pt/22001</a>	INC. 187	15	1496
<a href="http://purl.pt/22002">http://purl.pt/22002</a>	INC. 571	15	1496
<a href="http://purl.pt/21816">http://purl.pt/21816</a>	INC. 533	15	1497
<a href="http://purl.pt/22005">http://purl.pt/22005</a>	INC. 1425	15	1497
<a href="http://purl.pt/22006">http://purl.pt/22006</a>	INC. 1426	15	1497
<a href="http://purl.pt/22007">http://purl.pt/22007</a>	INC. 1427	15	1497
<a href="http://purl.pt/24024">http://purl.pt/24024</a>	INC. 579	15	1497
<a href="http://purl.pt/24277">http://purl.pt/24277</a>	INC. 1468	15	1498
<a href="http://purl.pt/14483">http://purl.pt/14483</a>	INC. 510	15	1500
<a href="http://purl.pt/23230">http://purl.pt/23230</a>	RES. 3837 P.	15	1500
<a href="http://purl.pt/6931">http://purl.pt/6931</a>	RES. 218 10 V.	16	1501
<a href="http://purl.pt/14703">http://purl.pt/14703</a>	RES. 431 V.	16	1502
<a href="http://purl.pt/15164">http://purl.pt/15164</a>	RES. 149 V.	16	1502
<a href="http://purl.pt/24276">http://purl.pt/24276</a>	RES. 329 V.	16	1502

<a href="http://purl.pt/24646">http://purl.pt/24646</a>	RES. 323 V.	16	1502
<a href="http://purl.pt/17368">http://purl.pt/17368</a>	RES. 287 V.	16	1503
<a href="http://purl.pt/24630">http://purl.pt/24630</a>	RES. 322 V.	16	1503
<a href="http://purl.pt/14885">http://purl.pt/14885</a>	RES. 163 A.	16	1504
<a href="http://purl.pt/15153">http://purl.pt/15153</a>	RES. 127 V.	16	1504
<a href="http://purl.pt/15024">http://purl.pt/15024</a>	RES. 75 2 V.	16	1505
<a href="http://purl.pt/15274">http://purl.pt/15274</a>	D.S. XVI -3	16	1508
<a href="http://purl.pt/14702">http://purl.pt/14702</a>	RES. 94 A.	16	1509
<a href="http://purl.pt/14794">http://purl.pt/14794</a>	RES. 93 A.	16	1509
<a href="http://purl.pt/14879">http://purl.pt/14879</a>	RES. 155 A.	16	1509
<a href="http://purl.pt/22528">http://purl.pt/22528</a>	RES. 1572 V.	16	1509
<a href="http://purl.pt/14876">http://purl.pt/14876</a>	RES. 2342 A.	16	1512
<a href="http://purl.pt/14911">http://purl.pt/14911</a>	RES. 60 1 A.	16	1512
<a href="http://purl.pt/21815">http://purl.pt/21815</a>	RES. 327 V.	16	1512
<a href="http://purl.pt/24632">http://purl.pt/24632</a>	RES. 325 V.	16	1512
<a href="http://purl.pt/12097">http://purl.pt/12097</a>	RES. 157 A.	16	1513
<a href="http://purl.pt/24484">http://purl.pt/24484</a>	C.A. 152 V.	16	1513
<a href="http://purl.pt/14708">http://purl.pt/14708</a>	RES. 68 A.	16	1514
<a href="http://purl.pt/15049">http://purl.pt/15049</a>	RES. 176 A.	16	1515
<a href="http://purl.pt/24994">http://purl.pt/24994</a>	RES. 1876 P.	16	1515
<a href="http://purl.pt/12096">http://purl.pt/12096</a>	RES. 111 A.	16	1516
<a href="http://purl.pt/14668">http://purl.pt/14668</a>	RES. 92 A.	16	1516
<a href="http://purl.pt/15207">http://purl.pt/15207</a>	RES. 3008 V.	16	1516
<a href="http://purl.pt/24338">http://purl.pt/24338</a>	RES. 6753 P.	16	1517
<a href="http://purl.pt/12100">http://purl.pt/12100</a>	RES. 217 2 A.	16	1518
<a href="http://purl.pt/14548">http://purl.pt/14548</a>	RES. 3392 V.	16	1518
<a href="http://purl.pt/14728">http://purl.pt/14728</a>	RES. 3393//1 V.	16	1518
<a href="http://purl.pt/15166">http://purl.pt/15166</a>	RES. 152 V.	16	1518
<a href="http://purl.pt/15289">http://purl.pt/15289</a>	RES. 404 V.	16	1518
<a href="http://purl.pt/24626">http://purl.pt/24626</a>	RES. 1849 P.	16	1518
<a href="http://purl.pt/24631">http://purl.pt/24631</a>	RES. 5781 P.	16	1518
<a href="http://purl.pt/24647">http://purl.pt/24647</a>	RES. 326 V.	16	1518

<a href="http://purl.pt/24093">http://purl.pt/24093</a>	RES. 1063 P.	16	1519
<a href="http://purl.pt/6955">http://purl.pt/6955</a>	RES. 218 18 V.	16	1520
<a href="http://purl.pt/12099">http://purl.pt/12099</a>	RES. 217//1 A.	16	1520
<a href="http://purl.pt/14469">http://purl.pt/14469</a>	RES. 124 V.	16	1520
<a href="http://purl.pt/25776">http://purl.pt/25776</a>	RES. 1032 A.	16	1520
<a href="http://purl.pt/109">http://purl.pt/109</a>	RES. 91 P.	16	1521
<a href="http://purl.pt/149">http://purl.pt/149</a>	RES. 3106 P.	16	1521
<a href="http://purl.pt/12182">http://purl.pt/12182</a>	RES. 70 A.	16	1521
<a href="http://purl.pt/24409">http://purl.pt/24409</a>	RES. 23 P.	16	1521
<a href="http://purl.pt/16506">http://purl.pt/16506</a>	RES. 882 A.	16	1523
<a href="http://purl.pt/22620">http://purl.pt/22620</a>	RES. 459 P.	16	1523
<a href="http://purl.pt/23037">http://purl.pt/23037</a>	RES. 1150 P.	16	1523
<a href="http://purl.pt/24411">http://purl.pt/24411</a>	RES. 6438 P.	16	1523
<a href="http://purl.pt/23728">http://purl.pt/23728</a>	RES. 223 P.	16	1524
<a href="http://purl.pt/22914">http://purl.pt/22914</a>	INC. 1486//2	16	1525
<a href="http://purl.pt/23115">http://purl.pt/23115</a>	RES. 1541 P.	16	1525
<a href="http://purl.pt/921">http://purl.pt/921</a>	RES. 454 V.	16	1526
<a href="http://purl.pt/11793">http://purl.pt/11793</a>	RES. 457 V.	16	1526
<a href="http://purl.pt/14402">http://purl.pt/14402</a>	RES. 26 A.	16	1526
<a href="http://purl.pt/24599">http://purl.pt/24599</a>	RES. 250 2 P.	16	1526
<a href="http://purl.pt/24271">http://purl.pt/24271</a>	RES. 1298 P.	16	1527
<a href="http://purl.pt/24656">http://purl.pt/24656</a>	RES. 253 P.	16	1528
<a href="http://purl.pt/15228">http://purl.pt/15228</a>	RES. 3959 V.	16	1529
<a href="http://purl.pt/25773">http://purl.pt/25773</a>	RES. 624 V.	16	1529
<a href="http://purl.pt/14549">http://purl.pt/14549</a>	RES. 3736 V.	16	1530
<a href="http://purl.pt/16678">http://purl.pt/16678</a>	RES. 166 A.	16	1531
<a href="http://purl.pt/24410">http://purl.pt/24410</a>	RES. 24 P.	16	1531
<a href="http://purl.pt/15298">http://purl.pt/15298</a>	RES. 4645 2 P.	16	1532
<a href="http://purl.pt/14913">http://purl.pt/14913</a>	RES. 3016 V.	16	1533
<a href="http://purl.pt/14928">http://purl.pt/14928</a>	RES. 129 A.	16	1534
<a href="http://purl.pt/21895">http://purl.pt/21895</a>	D.S. XVI-9	16	1534
<a href="http://purl.pt/23335">http://purl.pt/23335</a>	RES. 1485 P.	16	1534

<a href="http://purl.pt/24648">http://purl.pt/24648</a>	RES. 328 V.	16	1534
<a href="http://purl.pt/24598">http://purl.pt/24598</a>	RES. 250//1 P.	16	1535
<a href="http://purl.pt/120">http://purl.pt/120</a>	RES. 274 V.	16	1536
<a href="http://purl.pt/23036">http://purl.pt/23036</a>	RES. 1126 P.	16	1536
<a href="http://purl.pt/23117">http://purl.pt/23117</a>	RES. 1569 P.	16	1536
<a href="http://purl.pt/23151">http://purl.pt/23151</a>	RES. 1759 P.	16	1536
<a href="http://purl.pt/23165">http://purl.pt/23165</a>	RES. 1944 P.	16	1536
<a href="http://purl.pt/23270">http://purl.pt/23270</a>	RES. 4645//1 P.	16	1536
<a href="http://purl.pt/14445">http://purl.pt/14445</a>	RES. 410 V.	16	1537
<a href="http://purl.pt/14665">http://purl.pt/14665</a>	RES. 141 1 A.	16	1537
<a href="http://purl.pt/14701">http://purl.pt/14701</a>	RES. 124 A.	16	1538
<a href="http://purl.pt/14877">http://purl.pt/14877</a>	RES. 1633 A.	16	1538
<a href="http://purl.pt/14899">http://purl.pt/14899</a>	D.S. XVI-12	16	1538
<a href="http://purl.pt/22514">http://purl.pt/22514</a>	H.G. 14853//5 P.	16	1538
<a href="http://purl.pt/14280">http://purl.pt/14280</a>	ULFL RES. 247	16	1539
<a href="http://purl.pt/14849">http://purl.pt/14849</a>	RES. 1970 A.	16	1539
<a href="http://purl.pt/22313">http://purl.pt/22313</a>	RES. 3409 V.	16	1539
<a href="http://purl.pt/24071">http://purl.pt/24071</a>	RES. 74 A. e RES. 75 A.	16	1539
<a href="http://purl.pt/12148">http://purl.pt/12148</a>	RES. 5658//1 P.	16	1540
<a href="http://purl.pt/12149">http://purl.pt/12149</a>	RES. 5658//3 P.	16	1540
<a href="http://purl.pt/14634">http://purl.pt/14634</a>	RES. 3604 V.	16	1540
<a href="http://purl.pt/14829">http://purl.pt/14829</a>	RES. 2988//2 V.	16	1540
<a href="http://purl.pt/14830">http://purl.pt/14830</a>	RES. 2988//1 V.	16	1540
<a href="http://purl.pt/15169">http://purl.pt/15169</a>	RES. 156//2 V.	16	1540
<a href="http://purl.pt/15173">http://purl.pt/15173</a>	RES. 172 V.	16	1540
<a href="http://purl.pt/15191">http://purl.pt/15191</a>	RES. 264 V.	16	1540
<a href="http://purl.pt/15193">http://purl.pt/15193</a>	RES. 278 V.	16	1540
<a href="http://purl.pt/15320">http://purl.pt/15320</a>	RES. 412 V.	16	1540
<a href="http://purl.pt/11493">http://purl.pt/11493</a>	RES. 482 P.	16	1541
<a href="http://purl.pt/14687">http://purl.pt/14687</a>	RES. 145 A.	16	1541
<a href="http://purl.pt/14813">http://purl.pt/14813</a>	RES. 3985 V.	16	1541
<a href="http://purl.pt/15165">http://purl.pt/15165</a>	RES. 150 V.	16	1541

<a href="http://purl.pt/15311">http://purl.pt/15311</a>	RES. 71 V.	16	1541
<a href="http://purl.pt/23259">http://purl.pt/23259</a>	RES. 4448 P.	16	1541
<a href="http://purl.pt/24651">http://purl.pt/24651</a>	R. 7913 P.	16	1541
<a href="http://purl.pt/14446">http://purl.pt/14446</a>	RES. 397 P.	16	1542
<a href="http://purl.pt/14466">http://purl.pt/14466</a>	RES. 97 V.	16	1542
<a href="http://purl.pt/14568">http://purl.pt/14568</a>	RES. 58 A.	16	1542
<a href="http://purl.pt/15154">http://purl.pt/15154</a>	RES. 129 V.	16	1542
<a href="http://purl.pt/15155">http://purl.pt/15155</a>	RES. 120 V.	15	1542
<a href="http://purl.pt/15163">http://purl.pt/15163</a>	RES. 148 V.	16	1542
<a href="http://purl.pt/15281">http://purl.pt/15281</a>	RES. 6082 P.	16	1542
<a href="http://purl.pt/16536">http://purl.pt/16536</a>	RES. 476 A.	16	1542
<a href="http://purl.pt/23041">http://purl.pt/23041</a>	RES. 1200 P.	16	1542
<a href="http://purl.pt/23218">http://purl.pt/23218</a>	RES. 3132 P.	16	1542
<a href="http://purl.pt/23327">http://purl.pt/23327</a>	RES. 1333 P.	16	1542
<a href="http://purl.pt/24649">http://purl.pt/24649</a>	RES. 6437 P.	16	1542
<a href="http://purl.pt/14304">http://purl.pt/14304</a>	ULFL RES. 241 (ULFL)	16	1543
<a href="http://purl.pt/14770">http://purl.pt/14770</a>	RES. 851 A.	16	1543
<a href="http://purl.pt/14551">http://purl.pt/14551</a>	RES. 161 A.	16	1544
<a href="http://purl.pt/15293">http://purl.pt/15293</a>	RES. 405 V.	16	1544
<a href="http://purl.pt/15319">http://purl.pt/15319</a>	RES. 997 V.	16	1544
<a href="http://purl.pt/22945">http://purl.pt/22945</a>	RES. 94 P.	16	1544
<a href="http://purl.pt/14664">http://purl.pt/14664</a>	RES. 18 A.	16	1545
<a href="http://purl.pt/15168">http://purl.pt/15168</a>	RES. 156 1 V.	16	1545
<a href="http://purl.pt/16741">http://purl.pt/16741</a>	RES. 92 P.	16	1545
<a href="http://purl.pt/15201">http://purl.pt/15201</a>	RES. 292 V.	16	1546
<a href="http://purl.pt/23044">http://purl.pt/23044</a>	RES. 1280//1 P.	16	1546
<a href="http://purl.pt/23326">http://purl.pt/23326</a>	RES. 1280//2 P.	16	1546
<a href="http://purl.pt/24601">http://purl.pt/24601</a>	RES. 2984 P.	16	1546
<a href="http://purl.pt/14255">http://purl.pt/14255</a>	ULFL RES. 230	16	1547
<a href="http://purl.pt/23237">http://purl.pt/23237</a>	RES. 4185 P.	16	1547
<a href="http://purl.pt/24650">http://purl.pt/24650</a>	R. 7254 P.	16	1547
<a href="http://purl.pt/4066">http://purl.pt/4066</a>	RES. 126 1 A.	16	1548

<a href="http://purl.pt/6304">http://purl.pt/6304</a>	RES. 2480 P.	16	1548
<a href="http://purl.pt/6953">http://purl.pt/6953</a>	RES. 218//16 V.	16	1548
<a href="http://purl.pt/15157">http://purl.pt/15157</a>	RES. 133 V.	16	1548
<a href="http://purl.pt/15159">http://purl.pt/15159</a>	RES. 137 V.	16	1548
<a href="http://purl.pt/15167">http://purl.pt/15167</a>	RES. 153 V.	16	1548
<a href="http://purl.pt/15312">http://purl.pt/15312</a>	RES. 765 V.	16	1548
<a href="http://purl.pt/22944">http://purl.pt/22944</a>	RES. 85 P.	16	1548
<a href="http://purl.pt/23136">http://purl.pt/23136</a>	RES. 1683 P.	16	1548
<a href="http://purl.pt/23334">http://purl.pt/23334</a>	RES. 1481 P.	16	1548
<a href="http://purl.pt/14279">http://purl.pt/14279</a>	ULFL RES. 231	16	1549
<a href="http://purl.pt/14308">http://purl.pt/14308</a>	ULFL RES. 215	16	1549
<a href="http://purl.pt/14460">http://purl.pt/14460</a>	RES. 4706 P.	16	1549
<a href="http://purl.pt/23137">http://purl.pt/23137</a>	RES. 1692 P.	16	1549
<a href="http://purl.pt/23169">http://purl.pt/23169</a>	RES. 1966 P.	16	1549
<a href="http://purl.pt/23849">http://purl.pt/23849</a>	RES. 189//27 P.	16	1549
<a href="http://purl.pt/14546">http://purl.pt/14546</a>	RES. 4180 1 V.	16	1550
<a href="http://purl.pt/15300">http://purl.pt/15300</a>	RES. 4702 P.	16	1550
<a href="http://purl.pt/22946">http://purl.pt/22946</a>	RES. 97 P.	16	1550
<a href="http://purl.pt/23142">http://purl.pt/23142</a>	RES. 1706//1 P.	16	1550
<a href="http://purl.pt/23150">http://purl.pt/23150</a>	RES. 1755//2 P.	16	1550
<a href="http://purl.pt/23251">http://purl.pt/23251</a>	RES. 1663//1 P.	16	1550
<a href="http://purl.pt/23268">http://purl.pt/23268</a>	RES. 4621 P.	16	1550
<a href="http://purl.pt/14232">http://purl.pt/14232</a>	ULFL RES. 216	16	1551
<a href="http://purl.pt/14233">http://purl.pt/14233</a>	ULFL RES. 216	16	1551
<a href="http://purl.pt/23048">http://purl.pt/23048</a>	RES. 1404 P.	16	1551
<a href="http://purl.pt/23198">http://purl.pt/23198</a>	RES. 2778 1 P.	16	1551
<a href="http://purl.pt/23217">http://purl.pt/23217</a>	RES. 3131 P.	16	1551
<a href="http://purl.pt/23272">http://purl.pt/23272</a>	RES. 4681 P.	16	1551
<a href="http://purl.pt/23223">http://purl.pt/23223</a>	RES. 3296 P.	16	1551
<a href="http://purl.pt/23345">http://purl.pt/23345</a>	RES. 1673 P.	16	1551
<a href="http://purl.pt/23610">http://purl.pt/23610</a>	RES. 1079 2 V.	16	1551
<a href="http://purl.pt/14546">http://purl.pt/14546</a>	366214 RES. 4180//1 V.	16	1550

<a href="http://purl.pt/14856">http://purl.pt/14856</a>	RES. 4180 2 V.	16	1553
<a href="http://purl.pt/15151">http://purl.pt/15151</a>	ULFL RES. 245	16	1552
<a href="http://purl.pt/23225">http://purl.pt/23225</a>	RES. 3504//1 P.	16	1552
<a href="http://purl.pt/23226">http://purl.pt/23226</a>	RES. 3504//2 P.	16	1554
<a href="http://purl.pt/12671">http://purl.pt/12671</a>	RES. 842//3 P.	16	1552
<a href="http://purl.pt/13845">http://purl.pt/13845</a>	RES. 2094 A.	16	1552
<a href="http://purl.pt/14367">http://purl.pt/14367</a>	RES. 678//1 P.	16	1552
<a href="http://purl.pt/15294">http://purl.pt/15294</a>	RES. 423//1 V. e RES. 423//2 V. RES. 425//1 V. e RES. 425//2 V. RES. 427//1 V. a RES. 427//3 V.		
<a href="http://purl.pt/15371">http://purl.pt/15371</a>	RES. 454//2 P.	16	1552
<a href="http://purl.pt/15372">http://purl.pt/15372</a>	RES. 454//3 P.	16	1552
<a href="http://purl.pt/15373">http://purl.pt/15373</a>	RES. 454//4 P.	16	1552
<a href="http://purl.pt/22314">http://purl.pt/22314</a>	RES. 4122 P.	16	1552
<a href="http://purl.pt/22941">http://purl.pt/22941</a>	RES. 399//1 P.	16	1552
<a href="http://purl.pt/23177">http://purl.pt/23177</a>	RES. 2196 P.	16	1552
<a href="http://purl.pt/23225">http://purl.pt/23225</a>	RES. 3504//1 P.	16	1552
<a href="http://purl.pt/23246">http://purl.pt/23246</a>	RES. 4390 P.	16	1552
<a href="http://purl.pt/779">http://purl.pt/779</a>	BIB 288 P.	16	1553
<a href="http://purl.pt/15370">http://purl.pt/15370</a>	RES. 454//1 P.	16	1553
<a href="http://purl.pt/17370">http://purl.pt/17370</a>	RES. 198 A.	16	1553
<a href="http://purl.pt/22521">http://purl.pt/22521</a>	RES. 501 P.	16	1553
<a href="http://purl.pt/22619">http://purl.pt/22619</a>	RES. 412 A.	16	1553
<a href="http://purl.pt/23134">http://purl.pt/23134</a>	RES. 1672 P.	16	1553
<a href="http://purl.pt/23271">http://purl.pt/23271</a>	RES. 4650 P.	16	1553
<a href="http://purl.pt/23305">http://purl.pt/23305</a>	RES. 55 P.	16	1553
<a href="http://purl.pt/23349">http://purl.pt/23349</a>	RES. 1644//1 P e RES. 1644//2 P.	16	1553
<a href="http://purl.pt/24600">http://purl.pt/24600</a>	RES. 409 P.	16	1553
<a href="http://purl.pt/81">http://purl.pt/81</a>	RES. 3048 P.	16	1554
<a href="http://purl.pt/14281">http://purl.pt/14281</a>	ULFL RES. 248	16	1554
<a href="http://purl.pt/14403">http://purl.pt/14403</a>	RES. 28 A.	16	1554
<a href="http://purl.pt/14435">http://purl.pt/14435</a>	RES. 84 V.	16	1554

<a href="http://purl.pt/14796">http://purl.pt/14796</a>	RES. 119 A.	16	1554
<a href="http://purl.pt/15188">http://purl.pt/15188</a>	RES. 2613 V.	16	1554
<a href="http://purl.pt/15301">http://purl.pt/15301</a>	RES. 528 V.	16	1554
<a href="http://purl.pt/21896">http://purl.pt/21896</a>	DS XVI -22 a DS XVI -24	16	1554 1559
<a href="http://purl.pt/22948">http://purl.pt/22948</a>	RES. 99 P.	16	1554
<a href="http://purl.pt/23021">http://purl.pt/23021</a>	RES. 571 P.	16	1554
<a href="http://purl.pt/23040">http://purl.pt/23040</a>	RES. 1195 P.	16	1554
<a href="http://purl.pt/23161">http://purl.pt/23161</a>	RES. 1912//2 P.	16	1554
<a href="http://purl.pt/23226">http://purl.pt/23226</a>	RES. 3504//2 P.	16	1554
<a href="http://purl.pt/23232">http://purl.pt/23232</a>	RES. 4028 P.	16	1554
<a href="http://purl.pt/23351">http://purl.pt/23351</a>	RES. 4343 P.	16	1554
<a href="http://purl.pt/15041">http://purl.pt/15041</a>	RES. 147 A.	16	1555
<a href="http://purl.pt/22915">http://purl.pt/22915</a>	RES. 102 P.	16	1555
<a href="http://purl.pt/22935">http://purl.pt/22935</a>	RES. 275 V.	16	1555
<a href="http://purl.pt/23147">http://purl.pt/23147</a>	RES. 1742 P.	16	1555
<a href="http://purl.pt/23152">http://purl.pt/23152</a>	RES. 1787 P.	16	1555
<a href="http://purl.pt/23309">http://purl.pt/23309</a>	RES. 86 P.	16	1555
<a href="http://purl.pt/23315">http://purl.pt/23315</a>	RES. 228//1 P.	16	1555
<a href="http://purl.pt/14252">http://purl.pt/14252</a>	ULFL RES. 205	16	1556
<a href="http://purl.pt/14263">http://purl.pt/14263</a>	ULFL RES. 217	16	1556
<a href="http://purl.pt/14563">http://purl.pt/14563</a>	RES. 178//1 A.	16	1556
<a href="http://purl.pt/15007">http://purl.pt/15007</a>	RES. 3353 V.	16	1556
<a href="http://purl.pt/15042">http://purl.pt/15042</a>	RES. 148 A.	16	1556
<a href="http://purl.pt/17324">http://purl.pt/17324</a>	D.S. XVI -25 e D.S. XVI - 26	16	1556
<a href="http://purl.pt/17421">http://purl.pt/17421</a>	RES. 3138 1 V.	16	1556
<a href="http://purl.pt/23145">http://purl.pt/23145</a>	RES. 1712 P.	16	1556
<a href="http://purl.pt/14282">http://purl.pt/14282</a>	ULFL RES. 249	16	1557
<a href="http://purl.pt/14439">http://purl.pt/14439</a>	RES. 83 V.	16	1557
<a href="http://purl.pt/14929">http://purl.pt/14929</a>	RES. 2048 V.	16	1557
<a href="http://purl.pt/15295">http://purl.pt/15295</a>	RES. 429 V.	16	1557
<a href="http://purl.pt/16514">http://purl.pt/16514</a>	RES. 2120 P.	16	1557
<a href="http://purl.pt/16615">http://purl.pt/16615</a>	RES. 38 A. E RES. 39 A.	16	1557

<a href="http://purl.pt/17386">http://purl.pt/17386</a>	RES. 3138 2 V.	16	1557
<a href="http://purl.pt/17392">http://purl.pt/17392</a>	RES. 3138 3 V.	16	1557
<a href="http://purl.pt/17393">http://purl.pt/17393</a>	RES. 3138 4 V.	16	1557
<a href="http://purl.pt/22938">http://purl.pt/22938</a>	RES. 515 P.	16	1557
<a href="http://purl.pt/22942">http://purl.pt/22942</a>	RES. 131 P.	16	1557
<a href="http://purl.pt/23139">http://purl.pt/23139</a>	RES. 1697 P.	16	1557
<a href="http://purl.pt/23154">http://purl.pt/23154</a>	RES. 1852 P.	16	1557
<a href="http://purl.pt/23220">http://purl.pt/23220</a>	RES. 3243 P.	16	1557
<a href="http://purl.pt/23341">http://purl.pt/23341</a>	RES. 1539 2 P.	16	1557
<a href="http://purl.pt/26772">http://purl.pt/26772</a>	RES. 4703 P.	16	1557
<a href="http://purl.pt/14754">http://purl.pt/14754</a>	RES. 5944 P.	16	1558
<a href="http://purl.pt/17394">http://purl.pt/17394</a>	RES. 3138 5 V.	16	1558
<a href="http://purl.pt/21907">http://purl.pt/21907</a>	RES. 2095 A.	16	1558
<a href="http://purl.pt/22520">http://purl.pt/22520</a>	RES. 4338 P.	16	1558
<a href="http://purl.pt/23247">http://purl.pt/23247</a>	RES. 4391 P.	16	1558
<a href="http://purl.pt/23095">http://purl.pt/23095</a>	S.C. 4501 A.	17	16



### **XIII. Dimensões das obras na BNP**

#### **Série**

#### **Fundo Geral, Cartografia, Iconografia, Música e Reservados**

Dimensões máximas

- P. Série Preta 37x25,5 cm
- V. Série Vermelha 51x37 cm
- A. Série Azul 73,5x51 cm
- R. Série Roxa 101,5x73,5 cm

#### **Iconografia (cartazes)**

Dimensões máximas

- G. Pl. Série Gigante Plana 112x91 cm
- G.Cx. Série Gigante Caixa de 112x91 até 257x105 cm

#### **Periódicos**

Dimensões máximas

Cota P.P. ou R.E.

- B = P; Baixo = Preto 20 cm
- M = V; Médio = Vermelho 30 cm (entre 21 e 30 cm)
- A = A; Alto = Azul 31 cm (maior ou igual a 31 cm)

Cota J.

- B = P; Baixo = Preto 35 cm
- M = V; Médio = Vermelho 50 cm (entre 36 e 50 cm)
- A = A; Alto = Azul 51 cm (maior ou igual a 51 cm)

#### XIV. Lista das obras de Aldo Manuzio na Gallica

Thesaurus cornu copiae ; & Horti Adonidis ([Reprod.]) / [epistola  
dedicatoria Aldo Manucio ] - 1496

Omnia opera Angeli Politiani : et alia quaedam lectu digna, quorum nomina in sequenti  
indice videre licet ([Reprod.]) / ed. Aldus Manutius - 1498

Aristophanis Comoediae novem : Plutus, Nebulae, Ranae, Equites, Acharnes, Vespae,  
Aves, Pax, Contionantes / [Ed. Marcus Musurus] - 1498

Hypnerotomachia Poliphili - 1499

Epistole devotissime de sancta Catharina da Siena, adunate insemi per Bartholomeo da  
Alzano da Bergamo - 1500

Tragoediae septendecim , ex quibus quaedam habent commentaria ; et sunt hae :  
Hecuba, Orestes, Phoenissae, Medea, Hippolytus, Alcestis, Andromache,  
Supplices, Iphigenia in Aulide, Iphigenia in Tauris, Rhesus, Troades, Bacchae,  
Cyclops, Heraclidae, Helena, Ion - 1503

Gregorii, episcopi Nazanzeni, Carmina... e graeco in latinum ad verbum fere tralata  
[sic]... - "A la fin" : Venetiis, ex Aldi Academia, mense junio MDIIII - 1504

Habentur hoc volumine haec, videlicet : vita [auctore Planude] et fabellae Aesopi, cum  
interpretatio ne latina, ita tamen ut separari a graeco possit pro uniuscujusque  
arbitrio... / [Esopo de Phrygie et Babrius] ; éd. Manuce - 1505

Plutarchi Opuscula LXXXXII. Index Moralium omnium et eorum quae in ipsis tractantur habetur hoc quaternione... [Edidit Demetrius Ducas.] - "A la fin" : Venetiis, in aedibus Aldi et Andreae Asulani soceri, mense martio 1509 - 1509

Reliure de : Poemata : in quibus multa correctae sunt et institutiones suis locis positae commentariorum quodammodo vice funguntur. Undeviginti metrorum genera et quaedam sint et e quibus constant pedibus, et ante volumen simul habentur et intus in volumine suis locis. Adnotationes nonnullae in toto opere in quibus vel aliquid mutandum ostenditur vel cur mutatum sit ratio redditur / [curante Aldo Manutio ] - 1509

De conjuratione Catilinae ; ejusdem de bello Jugurthino ; ejusdem oratio contra M. T. Ciceronem. M. T. Ciceronis oratio contra C. Crispum Sallustium ; ejusdem orationes quatuor contra Lucium Catilinam. Porcii Latronis declamatio contra Lucium Catilinam. Orationes quaedam ex libris historiarum C. Crispi Sallustii - 1509

C. Plinii Secundi Novocomensis Epistolarum libri decem, in - 1511

Strozii poetae pater et filius ([Reprod.]) / [précédé d'une dédicace d'Alde Manuce] ; [et suivi d'un éloge prononcé par Celio Calcagnini aux obsèques d'Ercole Strozzi] - 1513

Habentur hoc volumine haec, Theodoro Gaza interprete : Aristotelis de Natura animalium lib. IX ; ejusdem de Partibus animalium lib. IIII ; ejusdem de Generatione animalium lib. V ; Theophrasti de Historia plantarum lib. IX et decimi principium duntaxat ; ejusdem de Causis plantarum lib. V ; Problematum

Aristotelis sectiones due de quadraginta ; Problematum Alexandri Aphrodisieii lib.

II - 1513

Catullus, Tibullus, Propertius ["sic"] - 1515

P. Ovidii Metamorphoseon libri XV : In fronte denique operis vita Ovidii ex ipsius operibus [ab Aldo M.] : index fabularum, et caeterorum, quae in hoc libro continentur, secundum ordinem alphabeti - 1515



## XV. Lista das obras de Aldo Manuziu no Internet Archive

Vol v. 1: Eis Organon Aristotelous. ; Anōnymon ..., 1495, Aristotle; Manuzio, Aldo, 1449 or 1450-1515; Forteguerra, Scipione, 1466-1515; Bondinus, Alexander, active 15th century; Galen; Pseudo-Galenus; Philo, of Alexandria; Theophrastus; Porphyry, approximately 234-approximately 305; Alexander, of Aphrodisias; Aristotle. Organon; Amboise, Jacques-Marius d', 1538-1611, former owner. DSI; Chiesa della missione (Mondovì, Italy), former owner. DSI; Burndy Library, donor. DSI

Petri Bembi De Aetna ad angelum Chabrielem liber., 1495, Manuzio, Aldo

Psalterion, 1496, Decadyus, Justinus, editor; Manuzio, Aldo, 1449 or 1450-1515, printer; Haleandius, Gregorius, former owner. MB; Bunnanus, Casparus, former owner. MB

Iamblichus de mysteriis Aegyptiorum, Chaldaeorum et Assyriorum. Proclus in Platonicum Alcibiadem de anima, atq[ue] daemone. Proclus de sacrificio & magia. Porphyrius de diuinis atq[ue] daemonibus. Synesius Platonicus de somniis. Psellus de daemonibus. Expositio Prisciani & Marsilii Theophrastu[m] de sensu,phantasia, & intellectu. Alcinoi Platonici philosophi liber de doctri[n]a Platonis. Speusippi Platonis discipuli liber de Platonis difinitionibus. Pythagorae philosophia aurea uerba. Symbola Pithagorae philosophi. Xenocratis philosophi platonici liber de morte. Marsili ficini liber de uoluptate, 1497, Jamblico; Manuzio, Aldo, 1450-1515

De morbo gallico, 1497, Nicolaus Leonicensus , Niccolò Leoniceno , Aldo Manuzio ,  
British Museum Dept . of Printed Books , Erasmus Press

Psalterium graecum, 1498, Manuzio, Aldo, 1450-1515

Institutiones graecae grammaticae, 1498, Bolzanius, Urbano, O.F.M., 1443-1524;  
Manuzio, Aldo, 1450-1515

Opera et alia quaedam lectu digna... (Omnia latine), 1498, Poliziano, Angelo, 1454-  
1494; Alejandro De Afrodiasias; Herodiano; Epicteto; Plutarco; Atanasio, Santo,  
Obispo de Alejandría, ca. 0297-0373; Manuzio, Aldo, 1450-1515

Hypnerotomachia Poliphili, 1499, Colonna, Francesco, 1433-1527; Manuzio, Aldo,  
1450-1515

Hypnerotomachia Poliphili : ubi humana omnia non nisi somnium esse docet atque  
obiter plurima scitu sane quam digna commemorat, 1499, Colonna, Francesco, -  
1527; Alberti, Leon Battista, 1404-1472; Colonna, Francesco, approximately  
1453-1517; Guidobaldo, da Montefeltro, Duke of Urbino, 1472-1508; Scythia,  
Johannes Baptista; Maro, Andreas, active 15th century-16th century; Manuzio,  
Aldo, 1449 or 1450-1515, printer; Crasso, Leonardo, publisher

Iulii Firmici Astronomicorum libri octo integri, & emendati, ex Scythicis ad nos nuper  
allati. Marci Manilii astronomicorum libri quinque. Arati phaenomena Germanico  
Caesare interprete cum commentariis & imaginibus. Arati eiusdem  
phaenomenon fragmentum Marco T.C. interprete. Arati eiusdem Phaenomena  
Ruffo Festo Auienio paraphraste. Arati eiusdem Phaenomena graece. Theonis  
commentaria copiosissima in Arati Phaenomena graece. Procli Diadochi Sphaera  
graece. Procli eiusdem Sphaera, Thoma Linacro Britanno interprete, 1499,

Manuzio, Aldo, 1450-1515; Arato de Soles, 0315-0240 a.J.C.; Proclus, ca. 410-485; Firmico Materno, Julio; Manilio, Marco; Germanico; Cicerón, Marco Tulio, 0106-0043 a.J.C.; Avieno, Rufo Festo; Linacre, Thomas, 1460-1524; Theon, Alexandrinus

Epistolae diversorum philosophorum, oratorum et rhetorum, 1499, Manuzio, Aldo, 1450-1515

Le cose volgari di messer Francesco Petrarca, 1501, Manuzio, Aldo, 1449 or 50-1515 & Torresanus, Andreas, de Asula

Iuuenalis. Persius., 1501, Manuzio, Aldo, 1449 or 50-1515 & Torresanus, Andreas, de Asula

Martialis., 1501, Manuzio, Aldo, 1449 or 50-1515 & Torresanus, Andreas, de Asula

[2]: Quae hoc libro continentur. Sedulii mirabiliam diuinorum libri quatuor carmine heroico. Eiusdem Elegia, in qua finis pentametri est similis principio hexametri. Eiusdem hymnus de Christo ab incarnatione, usque ad ascensionem. Iuueni de Euangelica historia libri quatuor. Aratoris cardinalis historiae apostolicae libri duo. ... Lactantii Firmiani de Resurrectione Elegia ... Cyprianus de ligno Crucis uersu heroico ... Vita S. Martini episcopi a Seuero Sulpitio prosa oratione ... Vita S. Nicolai e graeco in latinum a Leonardo Iustiniano patritio Veneto, 1501, Manuzio, Aldo, 1449 or 50-1515 & Torresanus, Andreas, de Asula

Vergilius, 1501, Manuzio, Aldo, 1449 or 50-1515 & Torresanus, Andreas, de Asula

Martialis., 1501, Manuzio, Aldo, 1449 or 50-1515 & Torresanus, Andreas, de Asula

Martialis., 1501, Manuzio, Aldo, 1449 or 50-1515 & Torresanus, Andreas, de Asula

Le cose volgari di messer Francesco Petrarca, 1501, Petrarca, Francesco, 1304-1374;

Bembo, Pietro, 1470-1547; Manuzio, Aldo, 1449 or 50-1515 printer

Le cose volgari di messer Francesco Petrarca., 1501, Manuzio, Aldo, 1449 or 50-1515

& Torresanus, Andreas, de Asula

Orthographia et flexus dictionum Graecarum omnium apud Statium cum accentib. et

generib. ex variis vtriusque linguae autorib, 1502, Manuzio, Aldo, 1449 or 50-

1515 & Torresanus, Andreas, de Asula

Lucanus., 1502, Manuzio, Aldo, 1449 or 50-1515 & Torresanus, Andreas, de Asula

Quae hoc volumine continentur. Ad Marinum Sannutum epistola qui apud graecos

scripserint metamorphoseos. Aldo priuilegium concessum ad reip. literariae

utilitatem. Orthographia dictionum graecarum per ordinem literarum. Vita Ouidij

ex ipsius operib. Index fabularum et caeterorum, quae insunt hoc libro secundum

ordinem alphabeti. Ouidii Metamorphoseon libri quindecim., 1502, Manuzio,

Aldo

Le terze rime di Dante., 1502, Manuzio, Aldo, 1449 or 50-1515 & Torresanus, Andreas,

de Asula

Le terze rime di Dante., 1502, Manuzio, Aldo, 1449 or 50-1515 & Torresanus, Andreas,

de Asula

Lucanus., 1502, Manuzio, Aldo, 1449 or 50-1515 & Torresanus, Andreas, de Asula

Le terze rime di Dante., 1502, Manuzio, Aldo, 1449 or 50-1515 & Torresanus, Andreas,

de Asula

Tade enestin en tÅ ide tÅ i bibliÅ i. Loukianou. Philostratou Eikones. Tou autou HÅ“rÅ ïka. Tou autou Bioi sophistÅ n. Philostratou neÅ terou Eikoneis. Kallistratou Ekphraseis. Que (!) hoc volumine continentur. Luciani Opera. Icones Philostrati. Eiusdem Heroica. Eiusdem uitæ Sophistarum. Icones Iunioris Philostrati. Descriptiones Callistrati .., 1503, Manuzio, Aldo, 1449 or 50-1515 & Torresanus, Andreas, de Asula

Publii Ouidii Nasonis, quae hoc in libello continentur. Fastorum. Libri 6. De tristibus. Libri 5. De Ponto. Libri 4, 1503, Manuzio, Aldo, 1449 or 50-1515 & Torresanus, Andreas, de Asula

Constantini Lascaris Byzantini De octo partibus orationis liber primus. Eiusdem de constructione liber secundus. Eiusdem de nomine & uerbo liber tertius. Eiusdem de pronomine secundum omnem linguam, & poeticum usum opusculum. Hæc omnia ... Cebetis tabula & graeca & latina, opus morale, & utile omnibus, & praecipue adolescentibus ... Oratio dominica & duplex salutatio ad beatiss. Virginem. Symbolum apostolorum. Euangelium Diui Ioannis Euangelistæ. Carmina aurea Pythagoræ. Phocylidis poema ad bene, beateque uiuendum. Omnia haec cum interpretatione latina. Introductio perbreuis ad hebraicam linguam, 1503, Manuzio, Aldo, 1449 or 50-1515 & Torresanus, Andreas, de Asula

Quae hoc in volumine tractantur Bessarionis Cardinalis ... in calumniatore[m] Platonis libri quatuor ; Eiusdem correctio librorum Platonis de legibus, 1503, Bessarion, Cardenal, 1403-1472; Georgius Trapezuntius , 1396?-1485?; Manuzio, Aldo, 1450-1515

Euripidou tragodiai heptakaideka on eniai met'exegeseon eisi de autai. Ekabe Orestes Phoinissai ... Euripidis tragoediae septendecim, ex quib. quaedam habent commentaria. & sunt hae. Hecuba Orestes Phoenissae Medea Hippolytus Alcestis Andromache Supplices Iphigenia in Aulide Iphigenia in Tauris Rhesus Troades Bacchae Cyclops Heraclidae Helena Ion. Vol. 1, 1503, Manuzio, Aldo

Euripidou tragodiai heptakaideka on eniai met'exegeseon eisi de autai. Ekabe Orestes Phoinissai ... Euripidis tragoediae septendecim, ex quib. quaedam habent commentaria. & sunt hae. Hecuba Orestes Phoenissae Medea Hippolytus Alcestis Andromache Supplices Iphigenia in Aulide Iphigenia in Tauris Rhesus Troades Bacchae Cyclops Heraclidae Helena Ion. Vol. 2, 1503, Manuzio, Aldo

Gregorii episcopi Nazanzeni Carmina ad bene, beatéque uiuendum utilissima nuper e græco in latinum ad uerbum ferè tralata..., 1504, Manuzio, Aldo

Dä“mosthenous logoi, dyo kai hexÄ“konta. Libaniou sophistou, hypotheseis eis tous autous logois. Bios dä“mosthenous, kat'auton libanion. Bios dä“mosthenous, kata ploutarchon. Demosthenis orationes duae & sexaginta. Libanii sophistae in eas ipsas orationes argumenta. Vita Demosthenis per Libanium. Eiusdem uita per Plutarchum, 1504, Manuzio, Aldo, 1449 or 50-1515 & Torresanus, Andreas, de Asula

emosthenis orationes duae & sexaginta. Libanii sophistae in eas ipsas orationes argumenta. Vita Demosthenis, 1504, Demóstenes, ca. 0384-0322 a.C.; Libanio; Manuzio, Aldo, Herederos de; Torresano, Andrea, de Asula, 1451-1529

Pontani Opera. Vrania, siue de stellis libri quinque. Meteororum liber unus. De hortis hesperidum libri duo. Lepidina siue postorales |! pompae septem. Item Meliseus,

Maeon Acon. Hendecasyllaborum libri duo. Tumulorum liber unus. Neniae duodecim. Epigrammata duodecim. Quae uero in toto opere habeantur in indice, qui in calce est, licet uidere., 1505, Manuzio, Aldo, 1449 or 50-1515 & Torresanus, Andreas, de Asula

Aurelius Augurellus., 1505, Manuzio, Aldo, 1449 or 50-1515 & Torresanus, Andreas, de Asula

Habentur hoc uolumine haec, uidelicet. Vita & fabellae Aesopi cum interpretatione Latina ... Gabriae Fabellae tres & quadraginta ex trimetris iambis, praeter ultimam ex scazonte, cum latina interpretatione ... Phurnutus seu, ut alii, Curnutus De natura deorum .., 1505, Manuzio, Aldo, 1449 or 50-1515 & Torresanus, Andreas, de Asula

Gli Asolani di messer Pietro Bembo., 1505, Manuzio, Aldo, 1449 or 50-1515 & Torresanus, Andreas, de Asula

Habentur hoc uolumine haec, uidelicet. Vita, & Fabellae Aesopi cum interpretatione latina, ita tamen ut separari a graeco possit pro uniuscuiusque arbitrio. quibus traducendis multum certe elaborauimus. nam quae ante tralata habebantur, infida admodu erant, quod facillimum erit conferenti cognoscere. Gabriae Fabellae tres & quadraginta ex trimetris iambis, praeter ultimam ex scazonte, cum latina interpretatioe. Quas idcirco bis curauimus in formadas, quia priores, ubi latinum a graeco seiungi potest, admodum quam incorrecte excusae fuerant exempli culpa. quare nacti emendatum exemplum, operae pretium uisum est iterum excudendas curare, ut ex secundis prima queant corrigi. Phurnutus seu, ut alii, Curnutus De natura deorum. Palaeophatus De non credendis historiis.

Heraclides Ponticus De allegoriis apud Homerum. Ori Apollinis Niliaci Hieroglyphica. Collectio prouerbiorum Tarrhæi, & Didymi, item eorum, quæ apud Sudam [sic], aliosq[ue] habentur per ordinem literarum. Ex Aphthonii Exercitamentis de fabula. Tum de formicis, & cicadis græce, & latine. De fabula ex Imaginibus Philostrati græce, & latine. Ex Hermogenis Exercitamentis de fabula Prisciano interprete. Apologus Aesopi de cassita apud Gellium, 1505, Aesop; Manuzio, Aldo, 1449 or 50-1515; Babrius

De constructione verborum, ... Aldus Manutius, 1510, Manuzio, Aldo

M[arci] T[ullii] Ciceronis epistolarum ad Atticum, ad Brutum, ad Quintum fratrem, libri XX, 1513, Cicero, Marcus Tullius; Manuzio, Aldo, 1449 or 1450-1515, printer; Torresanus, Andreas, de Asula, 1451-1529, printer; Atticus, Titus Pomponius. Epistolæ. 1513; Brutus, Marcus Junius, 85 B.C.-42 B.C. Epistolæ. 1513; Welch, Walter F., former owner. MMeT

Arcadia del Sannazaro., 1514, Manuzio, Aldo, 1449 or 50-1515 & Torresanus, Andreas, de Asula

Aldii Pii Manutii Institutionum grammaticarum libri quatuor. Quæ quoque libro continentur hanc uolenti chartam statim se offerunt., 1514, Manuzio, Aldo

Arcadia del Sannazaro., 1514, Manuzio, Aldo, 1449 or 50-1515 & Torresanus, Andreas, de Asula

Auli Gellii Noctium Atticarum libri vndeuginti, 1515, Manuzio, Aldo, 1449 or 50-1515 & Torresanus, Andreas, de Asula

Auli Gellii Noctium Atticarum libri vndeuginti, 1515, Manuzio, Aldo, 1449 or 50-1515 &

Torresanus, Andreas, de Asula

Lucanus., 1515, Manuzio, Aldo, 1449 or 50-1515 & Torresanus, Andreas, de Asula

Gli Asolani di messer Pietro Bembo., 1515, Manuzio, Aldo, 1449 or 50-1515 &

Torresanus, Andreas, de Asula

Catullus. Tibullus. Propertius., 1515, Manuzio, Aldo, 1449 or 50-1515 & Torresan



**UNIVERSIDADE DE ÉVORA**  
INSTITUTO DE INVESTIGAÇÃO  
E FORMAÇÃO AVANÇADA

---

**Contactos:**

Universidade de Évora

**Instituto de Investigação e Formação Avançada - IIFA**

Palácio do Vimioso | Largo Marquês de Marialva, Apart. 94

7002-554 Évora | Portugal

Tel: (+351) 266 706 581