



COLABORAÇÃO EM MASSA OU AMADORISMO EM MASSA?

UM ESTUDO COMPARATIVO DA QUALIDADE DA
INFORMAÇÃO CIENTÍFICA PRODUZIDA UTILIZANDO
OS CONCEITOS E FERRAMENTAS WIKI

Fernando Silvério Nifrário Rodrigues

Tese Apresentada à Universidade de Évora para
Obtenção do Grau de Doutor em Gestão

ORIENTADOR: *Professor Doutor Paulo Resende da Silva*

ÉVORA, Dezembro 2012



COLABORAÇÃO EM MASSA OU AMADORISMO EM MASSA?
UM ESTUDO COMPARATIVO DA QUALIDADE DA INFORMAÇÃO CIENTÍFICA PRODUZIDA
UTILIZANDO OS CONCEITOS E FERRAMENTAS WIKI.

Fernando Silvério Nifrário Rodrigues

Tese para obtenção do grau de
DOUTOR EM GESTÃO

Orientador: Professor Doutor Paulo Resende da Silva

Aim for the stars and maybe you'll reach the sky.

Reinhold Niebuhr

Aos meus filhos, Filipe e David,
e à multidão de “amadores” que diariamente partilham
o seu conhecimento de uma forma altruísta na *Web*.

*Because Wikipedia is a process, not a product, it replaces
guarantees offered by institutions with probabilities supported
by process: if enough people care enough about an article to read it,
then enough people will care enough to improve it, and over time
this will lead to a large enough body of good enough work
to begin to take both availability and quality of articles for granted,
and to integrate Wikipedia into daily use by millions.*

Clay Shirky (2008), *Here Comes Everybody*

AGRADECIMENTOS

O meu profundo agradecimento aos avaliadores que colaboraram neste estudo e que contribuíram com o seu trabalho e conhecimento para que fosse possível elaborar esta tese: Ana Fonseca, Armando Raimundo, Augusto Fitas, Carlos Sinogas, Comandante Rodrigues Pereira, Contra-Almirante Bartolomé Bauzá, Diogo Figueiredo, Fernando Barriga, Filipe Mesquita, Francisco Ramos, Jorge Casas Novas, José Calado, José Fernandes, José Garção, José Ramos, José Roberto, José Saias, José Ventura, Liliana Póvoas, Luís Arriaga da Cunha, Luís Gazarini, Luís Rato, Luísa Flora, Manuel Branco, Maria Ivone Clara, Maria José Bule, Maria Rosário Félix, Marta Páscoa, Marta Sequeira, Miguel Barão, Neto Vaz, Nuno Madeira, Paulo Resende da Silva, Pedro Dantas, Raluca Bibiri, Rui Fragoso, Rui Pereira, Rui Quaresma, Salvador Abreu, Saul Santos, Teresa Gonçalves, Vítor Nogueira, aos alunos do 3ºano da licenciatura em Gestão da UE (2010/11); e aos docentes do Departamento de História da Universidade de Évora e do Instituto Politécnico de Tomar que colaboraram neste estudo, mas que preferiram que o seu nome não fosse mencionado. Não posso também esquecer o meu orientador, Paulo Resende da Silva, a minha esposa, Alexandra Cheira, e as pessoas que, embora não avaliando directamente os artigos, contribuíram de uma forma ou de outra para que as avaliações se realizassem: António Serrano, Álvaro Pinto, Carrilho Ramos, Christopher Bochmann, João Ilhéu, Maria Raquel Lucas e Tânia Ganho.

A todos o meu sincero Obrigado.

COLABORAÇÃO EM MASSA OU AMADORISMO EM MASSA?

UM ESTUDO COMPARATIVO DA QUALIDADE DA INFORMAÇÃO CIENTÍFICA
PRODUZIDA UTILIZANDO OS CONCEITOS E FERRAMENTAS WIKI

RESUMO

Esta tese pretende contribuir para uma melhor compreensão do fenómeno Wiki enquanto sistema de gestão de conhecimento agregador do conhecimento privado e aferir o grau de confiabilidade da informação criada através da colaboração em massa, anónima e gratuita, por contraponto à abordagem tradicional.

Para tal, é desenvolvido um estudo comparativo entre a Wikipédia e a Enciclopédia Britânica de modo a aferir a qualidade dos repositórios de conhecimento por estas produzidos, permitindo, assim, tirar conclusões sobre a eficácia dos modelos de negócio que estão por detrás.

É utilizada uma amostra aleatória representativa do universo composto pelos artigos que integram as duas enciclopédias, tendo cada par sido previamente descaracterizado e posteriormente avaliado por um perito proficiente na área do conhecimento em que este se insere. Paralelamente, foi recolhida uma pequena amostra que, por conveniência, apenas integra artigos sobre gestão, sendo cada par de artigos avaliado por vários peritos de modo a determinar a incerteza associada a termos várias avaliações do mesmo artigo e aplicá-la às avaliações efectuadas por apenas um perito. Concluiu-se que a qualidade média dos artigos da Wikipédia analisados era superior à dos seus pares e que esta diferença era estatisticamente significativa.

Foi realizado um inquérito junto da academia em que se constatou que as fontes de informação tradicionais eram utilizadas, como primeira abordagem na busca de informação, por uma minoria, e que a confiança nestas fontes era substancialmente maior do que a depositada na informação obtida através da Wikipédia. Esta percepção da qualidade e os resultados diametralmente opostos da avaliação desta numa prova cega reforça a isenção do painel de avaliadores.

Apesar de a amostra escolhida ser representativa do universo a estudar, os resultados obtidos dependeram da opinião pessoal dos avaliadores e dos critérios escolhidos, pelo que a reprodutibilidade das conclusões deste estudo, utilizando outro painel de

avaliadores, não pode ser garantida, não sendo, no entanto, motivo suficiente para rejeitar liminarmente os resultados obtidos a partir de mais de meio milhão de avaliações.

Esta tese é, assim, uma tentativa de ajudar a clarificar o tema e contribuir para uma melhor percepção da qualidade de uma ferramenta, utilizada diariamente por milhões, da colaboração em massa que a alimenta, e do *software* colaborativo que a suporta.

PALAVRAS-CHAVE

Colaboração em Massa; Inteligência Colectiva; *Crowdsourcing*; Sistemas de Informação; Qualidade de Dados; Wikipédia; Enciclopédia Britânica.

MASS COLLABORATION OR MASS AMATEURISM?

A COMPARATIVE STUDY ON THE QUALITY OF SCIENTIFIC INFORMATION
PRODUCED USING WIKI TOOLS AND CONCEPTS

ABSTRACT

With this PhD dissertation, we intend to contribute to a better understanding of the Wiki phenomenon as a knowledge management system which aggregates private knowledge. We also wish to check to what extent information generated through anonymous and freely bestowed mass collaboration is reliable as opposed to the traditional approach.

In order to achieve that goal, we develop a comparative study between Wikipedia and Encyclopaedia Britannica with regard to accuracy, depth and detail of information in both, in order to confront the quality of the knowledge repository produced by them. That will allow us to reach a conclusion about the efficacy of the business models behind them.

We will use a representative random sample which is composed by the articles that are comprised in both encyclopedias. Each pair of articles was previously reformatted and then graded by an expert in its subject area. At the same time, we collected a small convenience sample which only integrates Management articles. Each pair of articles was graded by several experts in order to determine the uncertainty associated with having diverse gradings of the same article and apply it to the evaluations carried out by just one expert. The conclusion was that the average quality of the Wikipedia articles which were analysed was superior to its peers' and that this difference was statistically significant.

An inquiry was conducted within the academia which certified that traditional information sources were used by a minority as the first approach to seeking information. This inquiry also made clear that reliance on these sources was considerably larger than reliance on information obtained through Wikipedia. This quality perception, as well as the diametrically opposed results of its evaluation through a blind test, reinforces the evaluating panel's exemption.

However much the chosen sample is representative of the universe to be studied, results have depended on the evaluators' personal opinion and chosen criteria. This means that the reproducibility of this study's conclusions using a different grading panel cannot be guaranteed. Nevertheless, this is not enough of a reason to reject the study results obtained through more than five hundred evaluations.

This thesis is thus an attempt to help clarifying this topic and contributing to a better perception of the quality of a tool which is daily used by millions of people, of the mass collaboration which feeds it and of the collaborative software that supports it.

KEYWORDS

Mass Collaboration; Collective Intelligence; Crowdsourcing; Information Systems; Data Quality; Wikipedia; Encyclopaedia Britannica.

ÍNDICE GERAL

ÍNDICE DE ILUSTRAÇÕES.....	17
ÍNDICE DE TABELAS.....	21
1. INTRODUÇÃO	23
2. ENQUADRAMENTO	
2.1 INFORMAÇÃO	29
2.2 COLABORAÇÃO EM MASSA	36
2.3 A SABEDORIA DAS MULTIDÕES	56
2.4 A LOUCURA DAS MULTIDÕES	64
3. PROBLEMA	
3.1 AMADORES VERSUS PERITOS.....	69
3.2 INVENTARIAÇÃO DE ESTUDOS AFINS	76
3.3 WIKIPÉDIA VERSUS BRITÂNICA.....	84
3.4 OBJECTIVO VERSUS SUBJECTIVO.....	96
3.5 JÚRIS VERSUS JURADOS.....	104
4. METODOLOGIA	
4.1 RECOLHA DE DADOS	107
4.2 AVALIAÇÃO DOS DADOS.....	112
4.3 HÁBITOS VERSUS MONGES	117
4.4 PRECISO VERSUS EXACTO	121
4.5 TRATAMENTO DOS DADOS	124
4.6 FILOSOFIA DE INVESTIGAÇÃO.....	135
5. APRESENTAÇÃO DOS RESULTADOS	
5.1 CARACTERIZAÇÃO DOS HÁBITOS DE CONSULTA DE INFORMAÇÃO	141
5.2 CINQUENTA PRINCÍPIOS DE GESTÃO	149
5.3 COMPARAÇÃO WIKIPÉDIA - BRITÂNICA	162
6. CONSIDERAÇÕES FINAIS	
6.1 REFLEXÕES E DISCUSSÃO DE RESULTADOS.....	175
6.2 CONCLUSÕES E RECOMENDAÇÕES.....	180
BIBLIOGRAFIA	187
ÍNDICE REMISSIVO	205
ANEXOS	207

ÍNDICE DE ILUSTRAÇÕES

FIGURA 1: ESTRUTURA DA TESE.....	28
FIGURA 2: PIRÂMIDE DO CONHECIMENTO.....	30
FIGURA 3: INFORMAÇÃO, ADAPTADO DE SHANNON.....	31
FIGURA 4: INTERPRETAÇÕES DA COLABORAÇÃO EM MASSA.....	39
FIGURA 5: PIRÂMIDE DAS NECESSIDADES, ADAPTADO DE MASLOW	40
FIGURA 6: CRENÇAS SOCIAIS E CULTURAIS, ADAPTADO DE THIERER.....	43
FIGURA 7: RELAÇÕES DE INTERACÇÃO SOCIAL.....	45
FIGURA 8: TIPOS DE CONTRIBUIÇÕES, ADAPTADO DE COOK	47
FIGURA 9: LEI DA PARTICIPAÇÃO, ADAPTADO DE MAYFIELD.....	48
FIGURA 10: MODELO FLIRT, ADAPTADO DE VIITAMÄKI	50
FIGURA 11: GENES DA INTELIGÊNCIA COLECTIVA, ADAPTADO DE MALONE.....	51
FIGURA 12: PROBABILIDADE DE UM GRUPO TOMAR UMA DECISÃO CORRECTA	58
FIGURA 13: A LONGA CAUDA.....	70
FIGURA 14: CRESCIMENTO DA WIKIPÉDIA	73
FIGURA 15: VISIBILIDADE DA WIKIPÉDIA	74
FIGURA 16: DISTRIBUIÇÃO DOS ARTIGOS SEGUNDO OS "CRENTES" E OS "CÉPTICOS"	86
FIGURA 17: MODELO FUNCIONAMENTO WIKIPÉDIA E BRITÂNICA.....	92
FIGURA 18: QUALIDADE DOS DADOS.....	98
FIGURA 19: HIERARQUIA DA QUALIDADE DE DADOS	99
FIGURA 20: RESULTADOS DA CLASSIFICAÇÃO ALEATÓRIA.....	102
FIGURA 21: VARIAÇÃO DA DIMENSÃO DA AMOSTRA COM O NÚMERO DE AVALIADORES	106
FIGURA 22: DIMENSÃO DA AMOSTRA.....	108
FIGURA 23: AMOSTRAGEM 245 PARES ARTIGOS	111
FIGURA 24: EVOLUÇÃO DA DISTRIBUIÇÃO DOS ARTIGOS POR TEMAS.....	112
FIGURA 25: DISTRIBUIÇÃO PERCENTUAL DOS ARTIGOS POR TEMAS.....	113
FIGURA 26: DISTRIBUIÇÃO EM VALOR ABSOLUTO DO NÚMERO ARTIGOS POR TEMA	114

FIGURA 27: DISTRIBUIÇÃO DOS ARTIGOS POR ÁREAS DE CONHECIMENTO.....	115
FIGURA 28: DISTRIBUIÇÃO DAS AFLIAÇÕES DOS AVALIADORES	116
FIGURA 29: INQUÉRITO	117
FIGURA 30: PRECISO VERSUS EXACTO.....	121
FIGURA 31: AMOSTRAGEM ARTIGOS 50 PRINCÍPIOS DE GESTÃO	123
FIGURA 32: INTERPRETAÇÃO BOX PLOTS.....	124
FIGURA 33: TESTE DE HIPÓTESES.....	126
FIGURA 34: FLUXOGRAMA DA METODOLOGIA DO ESTUDO COMPARATIVO ENTRE AS ENCICLOPÉDIAS.....	132
FIGURA 35: FLUXOGRAMA DA METODOLOGIA DO ESTUDO DOS 50 PRINCÍPIOS DE GESTÃO	133
FIGURA 36: FLUXOGRAMA DA METODOLOGIA DA ANÁLISE DO INQUÉRITO	134
FIGURA 37: FILOSOFIA DE INVESTIGAÇÃO	135
FIGURA 38: FONTE DE INFORMAÇÃO QUE UTILIZAM PRIMEIRO	142
FIGURA 39: CONFIANÇA NA INFORMAÇÃO POR PARTE DOS DOCENTES UNIVERSITÁRIOS	143
FIGURA 40: CONFIANÇA NA INFORMAÇÃO POR PARTE DOS DOCENTES SECUNDÁRIO.....	143
FIGURA 41: CONFIANÇA NA INFORMAÇÃO POR PARTE DOS DISCENTES UNIVERSITÁRIOS.....	143
FIGURA 42: CONFIANÇA NA INFORMAÇÃO POR PARTE DOS DISCENTES SECUNDÁRIO.....	143
FIGURA 43: CONFIANÇA NA INFORMAÇÃO DISPONIBILIZADA POR FONTES ALTERNATIVAS, QUANDO COMPARADA COM A OBTIDA ATRAVÉS DE FONTES CLÁSSICAS.....	144
FIGURA 44: VARIAÇÃO DAS RESPOSTAS COM A ESCOLARIDADE	148
FIGURA 45: VARIAÇÃO DAS RESPOSTAS 2 E 3 EM FUNÇÃO DA RESPOSTA À QUESTÃO 1.....	148
FIGURA 46: BOX PLOT - AVALIAÇÕES DOCENTES ARTIGOS GESTÃO - WIKIPÉDIA.....	150
FIGURA 47: BOX PLOT - AVALIAÇÕES DOCENTES ARTIGOS GESTÃO - BRITÂNICA	150
FIGURA 48: BOX PLOT – DIFERENÇA AVALIAÇÃO ARTIGOS GESTÃO	152
FIGURA 49: DIFERENÇA ACUMULADA NOTAS WIKIPÉDIA - BRITÂNICA	153
FIGURA 50: BOX PLOT - AVALIAÇÕES DISCENTES ARTIGOS GESTÃO - WIKIPÉDIA.....	156
FIGURA 51: BOX PLOT - AVALIAÇÕES DISCENTES ARTIGOS GESTÃO - BRITÂNICA.....	156
FIGURA 52: BOX PLOT – DIFERENÇA AVALIAÇÕES ARTIGOS GESTÃO	157

FIGURA 53: VALOR MÉDIO DAS AVALIAÇÕES DOCENTES E DISCENTES	161
FIGURA 54: MÉDIA DO NÚMERO PALAVRAS POR ARTIGO AGRUPADAS PELA CLASSIFICAÇÃO OBTIDA.....	163
FIGURA 55: MÉDIA DO NÚMERO ILUSTRAÇÕES POR ARTIGO AGRUPADAS PELA CLASSIFICAÇÃO OBTIDA.....	163
FIGURA 56: Nº MÉDIO DE PALAVRAS E DE ILUSTRAÇÕES POR ARTIGO AGRUPADOS PELA DIFERENÇA NOTAS.	165
FIGURA 57: COMPARAÇÃO ENTRE OS VALORES OBSERVADOS E OS VALORES MÉDIOS AGRUPADOS PELA	166
FIGURA 58: BOX PLOT DIFERENÇAS ENTRE PALAVRAS/ARTIGO AGRUPADAS PELA DIFERENÇA ENTRE NOTAS ...	167
FIGURA 59: DIFERENÇA AVALIAÇÕES WIKIPÉDIA E BRITÂNICA.....	167
FIGURA 60: AVALIAÇÕES AGRUPADAS POR BINÓMIO DE NOTAS IGUAIS	168
FIGURA 61: BOX PLOT – AVALIAÇÕES WIKIPÉDIA E BRITÂNICA.....	168
FIGURA 62: HISTOGRAMAS DAS AVALIAÇÕES DOS ARTIGOS.....	169
FIGURA 63: QUALIDADE INFORMAÇÃO DA WIKIPÉDIA QUANDO COMPARADA COM A BRITÂNICA.....	169
FIGURA 64: BOX PLOT – DIFERENÇA ENTRE AVALIAÇÕES	170
FIGURA 65: MÉDIA DAS AVALIAÇÕES WIKIPÉDIA E BRITÂNICA DIVIDIDAS POR CATEGORIAS.....	171
FIGURA 66: VARIABILIDADE DA DIFERENÇA AVALIAÇÕES COMBINADA COM AS VARIABILIDADES CALCULADAS NO ESTUDO DOS ARTIGOS DE GESTÃO	173

(TODAS AS FIGURAS APRESENTADAS SÃO DE ELABORAÇÃO PRÓPRIA)

ÍNDICE DE TABELAS

TABELA 1: CÁLCULO DA PROBABILIDADE.....	32
TABELA 2: RESULTADO DAS VOTAÇÕES.....	60
TABELA 3: PRINCIPAIS ESTUDOS QUALIDADE WIKIPÉDIA.....	81
TABELA 4: COMPARAÇÃO GENES WIKIPÉDIA E BRITÂNICA.....	90
TABELA 5: SIGNIFICADO DAS DIFERENÇAS.....	103
TABELA 6: TRIBUNAIS DE JÚRI NO MUNDO	105
TABELA 7: FERRAMENTAS ESTATÍSTICAS A UTILIZAR NO ESTUDO.....	130
TABELA 8: RESULTADOS DO INQUÉRITO	142
TABELA 9: TESTE SHAPIRO-WILK - CONFIANÇA FONTES ALTERNATIVAS E CLÁSSICAS	145
TABELA 10: TESTE DE WILCOXON – CONFIANÇA FONTES ALTERNATIVAS E CLÁSSICAS	146
TABELA 11: TESTE SHAPIRO-WILK – DIFERENÇA CONFIANÇA FONTES ALTERNATIVAS E CLÁSSICAS.....	147
TABELA 12: ESTATÍSTICA DESCRITIVA – AVALIAÇÃO ARTIGOS GESTÃO	149
TABELA 13: ESTATÍSTICA DESCRITIVA – DIFERENÇA AVALIAÇÕES ARTIGOS GESTÃO	151
TABELA 14: INTERPRETAÇÃO QUALITATIVA DA ESCALA	152
TABELA 15: TESTE SHAPIRO-WILK – AVALIAÇÃO ARTIGOS GESTÃO.....	154
TABELA 16: TESTE WILCOXON – AVALIAÇÃO ARTIGOS GESTÃO.....	154
TABELA 17: ESTATÍSTICA DESCRITIVA ARTIGOS GESTÃO.....	155
TABELA 18: ESTATÍSTICA DESCRITIVA – DIFERENÇA AVALIAÇÕES ARTIGOS GESTÃO	157
TABELA 19: TESTE SHAPIRO WILK –AVALIAÇÕES ARTIGOS GESTÃO.....	158
TABELA 20: TESTE WILCOXON – AVALIAÇÕES ARTIGOS GESTÃO.....	159
TABELA 21: TESTE SHAPIRO-WILK – DIFERENÇA AVALIAÇÃO ARTIGOS GESTÃO	159
TABELA 22: TESTE MANN-WHITNEY - DIFERENÇA AVALIAÇÃO ARTIGOS GESTÃO	160
TABELA 23: MATRIZ CORRELAÇÃO 245 ARTIGOS.....	164
TABELA 24: TESTE SHAPIRO-WILK – AVALIAÇÕES WIKIPÉDIA E BRITÂNICA	170
TABELA 25: TESTE DE WILCOXON – AVALIAÇÕES WIKIPÉDIA E BRITÂNICA	171

(TODAS AS TABELAS APRESENTADAS SÃO DE ELABORAÇÃO PRÓPRIA)

There was a moment when technology was advanced enough to produce the radio and society was not yet advanced enough to accept it. The radio was then in its first phase of being a substitute: a substitute for theater, opera, concerts, lectures, café music, local newspapers, and so forth. (...) But quite apart from the dubiousness of its functions, radio is one-sided when it should be two. It is purely an apparatus for distribution, for mere sharing out. So here is a positive suggestion: change this apparatus over from distribution to communication. (...) That is to say, it would be if it knew how to receive as well as to transmit, how to let the listener speak as well as hear, how to bring him into a relationship instead of isolating him. On this principle the radio should step out of the supply business and organize its listeners as suppliers.

Bertolt Brecht (1932), *The Radio as an Apparatus of Communication*

1. INTRODUÇÃO

Os meus avós necessitavam de um fotógrafo sempre que queriam tirar uma fotografia, os meus pais deslocavam-se a um cinema para poderem assistir a um filme e, no meu primeiro ano na universidade, interagia com o computador entregando um conjunto de cartões perfurados e recebendo umas folhas impressas com o resultado dias depois. O que há uns anos atrás era apenas possível através do recurso a serviços técnicos especializados é hoje um dado adquirido para a maioria de nós.

Para Shirky (2008), os avanços tecnológicos tendem a ser disruptivos quando permitem a massificação de uma nova tecnologia e não apenas a substituição – evolução – de uma tecnologia por outra mais barata e eficiente.

Christensen (1997) considera que este primeiro tipo de avanço teve sempre como consequência que competências que antes estavam na mão de um grupo reduzido de peritos fossem assimiladas por uma multidão de amadores, o que, a médio prazo, torna o trabalho destes obsoleto, relegando-os para nichos de mercado e, a longo prazo, ao reajustamento das suas competências ou ao seu desaparecimento.

Em termos de acesso ao conhecimento, a invenção da impressão por Gutenberg no século XV, o aparecimento das rotativas com a revolução industrial no século XIX (Carregal, 1941) e a Internet no século XX foram responsáveis pelo aumento do

conhecimento disponível e do número dos seus potenciais consumidores e criadores (Sunstein, 2006).

Segundo Christensen (1997), estas revoluções nunca são pacíficas pois, para além dos peritos que vêem as suas competências deixarem de ser úteis – os copistas não devem ter apreciado muito a invenção de Gutenberg – existe também a diminuição do controlo sobre a informação que está disponível e a massificação de quem a acede (Surowiecki, 2004).

Para além de, como nas revoluções anteriores, o surgimento da Internet ter, segundo Anderson (2006), potenciado o aumento substancial do número de criadores e consumidores de informação, teve – especialmente desde a chamada reinvenção desta, conhecida por *Web 2.0* – um aumento sem precedentes no número de criadores, que se auto-organizaram utilizando ferramentas que permitem a partilha e consulta de informação entre milhões de pessoas, de uma forma centralizada – ou não – e cuja importância (que depende do ponto de vista) varia entre o último vídeo a que um determinado utilizador assistiu no *YouTube* e os últimos desenvolvimentos da demanda pelo bosão de Higgs¹ (Shirky, 2008).

A Wikipédia é um dos exemplos apontados por Tapscott & Williams (2006) que conseguiu utilizar o modelo de produção que os autores denominaram colaboração em massa, em conjunto com o *software* wiki, misturando as motivações não financeiras com as inovações tecnológicas da ferramenta para criar uma realidade que dificilmente alguém acreditaria ser possível.

Neste caso, parece ter-se verificado a premissa de Shirky (2008), e a tecnologia foi suficientemente disruptiva e conseguiu induzir um aumento exponencial no número de pessoas que coligem e consomem informação enciclopédica de uma forma colaborativa, auto-regulada e gratuita – uma área que, historicamente, era dominada por um punhado de técnicos especializados e consumida por uma minoria disposta a pagar pelo privilégio. Estamos, assim, em presença de uma revolução de comportamentos que vai além da revolução tecnológica que potenciou o seu aparecimento (Tapscott & Williams, 2006).

¹ Fonte: <http://public.web.cern.ch/public/en/science/higgs-en.html>.

A escolha deste tema de investigação prende-se com as questões que se levantam com a leitura de *Wikinomics: How Mass Collaboration Changes Everything*, em que Don Tapscott e Anthony Williams (Tapscott & Williams, 2006) mostram como o advento da *Web 2.0* potenciou uma mudança radical no modo como entendemos a colaboração. Devido à Internet, uma multidão de pessoas não enquadradas nas tradicionais organizações hierárquicas das empresas e organizações estão a produzir bens, serviços e conteúdos.

Apesar de algumas das ideias apresentadas poderem ser consideradas utópicas, interessou-nos particularmente o alerta lançado para o facto de que poderemos estar na presença de uma mudança profunda, i.e. disruptiva, na forma como é produzida a informação que utilizamos como referência e como suporte à tomada de decisões.

O livro de Tapscott e Williams (2006) realça ainda a importância do conhecimento explícito e tácito de cada indivíduo e da evolução tecnológica da Internet, que permitiu desenvolver mecanismos de agregação e partilha desses conhecimentos. Estes mecanismos de agregação do conhecimento individual podem, eventualmente, ser encarados como um sistema² que mistura a gestão de conhecimento com a gestão da informação, como reconhece Loureiro (2003) por vezes acontecer.

Entre as inúmeras reacções ao livro de Tapscott, destaca-se a de Andrew Keen que, em *The Cult of the Amateur: How Today's Internet is Killing our Culture* (Keen, 2007), desvaloriza os conteúdos gerados pelas massas, afirmando que estes não têm qualidade e retiram valor económico à criação tradicional de conteúdos. A discussão que se gerou em redor destas duas posições diametralmente opostas, principalmente na Internet, tem vindo a crescer e está longe de gerar consensos.

Um estudo publicado na revista *Nature*, intitulado *Internet Encyclopaedias Go Head to Head* (Giles, 2005), chamou-nos também a atenção, pois comparava o rigor dos conteúdos da Wikipédia³ com a Enciclopédia Britânica, concluindo que não havia uma diferença significativa na qualidade dos conteúdos destas. Apesar de a metodologia

² A componente do *software* e comunicações utilizados podem ser enquadrados nas TI; as regras de utilização e a gestão dos contribuidores e utilizadores podem ser enquadrados nos SI.

³ <http://wikipedia.org/>

utilizada ser questionável⁴ – e a Britânica refutou mesmo os resultados obtidos – fica no ar a possibilidade de uma enciclopédia gratuita, cujo conteúdo pode ser livremente editado por qualquer um e construída a partir das contribuições de uma multidão de amadores, ter construído um produto final similar a uma enciclopédia a que é comumente associada uma ideia de qualidade e rigor.

Das várias empresas e organizações citadas em *Wikinomics* de Tapscott e Williams (2006) como exemplos de sucesso de utilização da colaboração em massa, a jóia da coroa é, sem dúvida, a Wikipédia, considerada por Pink (2009) como a melhor representante do que considera ser o mais poderoso modelo de negócio do século XXI – o *open source* – tornando-a a candidata ideal ao estudo das potencialidades da colaboração em massa. Dado existirem empresas no mercado que fornecem o mesmo tipo de serviço utilizando uma abordagem convencional, este exemplo tem ainda a vantagem de permitir, através de um estudo comparativo entre a qualidade do produto final, concluir sobre o desempenho dos modelos de negócio utilizados na sua criação.

A escolha deste tema como tese de doutoramento foi ainda fortemente influenciada pela leitura de *The Wisdom of Crowds: Why the Many Are Smarter Than the Few and How Collective Wisdom Shapes Business, Economies, Societies and Nations* (Surowiecki, 2004), *Here comes everybody: The Power of Organization Without Organizations* (Shirky, 2008), *Drive: The Surprising Truth About What Motivates Us* (Pink, 2009) e *Cognitive Surplus: How Technology Makes Consumers into Collaborators* (Shirky, 2010), o que reforçou o interesse no tema, mostrando a sua pertinência e actualidade.

Não é o objectivo desta tese contrapor e dissecar todas as opiniões apresentadas nas obras atrás referidas, mas sim aferir se a colaboração em massa, dentro de um contexto e condições bem definidas, consegue produzir resultados comparáveis às abordagens convencionais, como prometido em *Wikinomics*, ou se o cepticismo de Keen tem toda a razão de ser. Pretendemos, pois, saber se existe algum caso em que a colaboração em massa produz resultados de qualidade superior ou igual à alternativa convencional. Se tal acontecer, será expectável que a colaboração em massa possa ser aplicada com sucesso a outros casos, pelo menos dentro do mesmo contexto e condições. Em síntese, pretendemos responder à questão:

⁴ Foi comparado um reduzido número de entradas similares em ambas as enciclopédias.

“Será que a Colaboração em Massa consegue produzir resultados de qualidade superior ou igual às alternativas convencionais?”

Assim, parece-nos que este tema tem um grau de visibilidade significativo, dentro e fora da academia. É um tema pouco maduro, ou seja, ainda não foi investigado exaustivamente pela comunidade académica – o que, apesar dos perigos que isso representa, é também uma oportunidade de efectuar um trabalho inovador; em termos de dimensão, é exequível de ser investigado no âmbito de uma tese de doutoramento, desde que, à partida, sejam delimitados os objectivos a alcançar e as limitações em termos de conclusões que a metodologia a aplicar condicionará.

Para responder à pergunta de investigação iremos então, em termos de estratégia, aferir as potencialidades da colaboração em massa através da determinação da qualidade do seu produto final, utilizando para tal o estudo de um caso concreto em que esta é aplicada; em termos tácticos, a organização escolhida é a Wikipédia, cujo modelo de produção assenta na colaboração em massa. A qualidade do seu produto final – informação – será aferida por comparação com a qualidade do produto final de uma enciclopédia generalista de referência, que concorre pelo mesmo mercado, a Enciclopédia Britânica – cujo modelo de produção é suportado por uma estrutura hierárquica convencional. Em sùmula, iremos comparar a qualidade dos conteúdos das duas enciclopédias e utilizar o resultado para aferir da viabilidade da colaboração em massa enquanto modelo de negócio em situações similares ao caso estudado.

Clay Shirky (2010) afirma que a maneira tradicional de resolver um problema, ou assegurar um serviço, passa pela criação de uma organização suportada por uma estrutura hierárquica, o que amiudamente conduz a que esta ponha em primeiro lugar a sua auto-preservação, subalternizando a resolução do problema para a qual foi criada. A substituição da abordagem de levar as pessoas ao problema por levar o problema às pessoas, deixando que estas se voluntariem e se auto-organizem – uma das características da colaboração em massa – é geralmente alvo de ataques cerrados das organizações que tentam a todo o custo desacreditar esta abordagem, especialmente quando estes projectos ameaçam a sua área de negócio. Shirky ironiza que estas têm

uma reacção semelhante aos estágios de Kübler-Ross⁵ em relação à aceitação de uma doença fatal – negação, raiva, negociação, e, por fim, a aceitação – estágio ainda não alcançado em relação a projectos como a Wikipédia, o Linux, o servidor *Web Apache*⁶, os Blogues⁷ e o Twitter⁸, entre outros. Por maior que seja a sua quota de mercado e a sua importância no nosso dia-a-dia, está-lhes associada uma imagem de amadorismo e de ausência de qualidade que conduz a uma relação de amor-ódio para com estes.

Apesar de conscientes da dificuldade em mudar o *statu quo*, em relação à opinião pré-concebida da maioria das pessoas sobre o tema, encontrámos motivação para a prossecução do nosso objectivo no passo de André Gide em *Les Faux-Monnayeurs*:

*On ne découvre pas de terre nouvelle sans consentir à perdre de vue, d’abord et longtemps, tout rivage. Mais nos écrivains craignent le large; ce ne sont que des côtoyeurs.*⁹ (Gide 1925: 376)

Para uma mais rápida apreensão da abordagem utilizada na demanda pela resposta à questão de investigação formulada, apresenta-se na Figura 1 uma representação da estrutura da tese, acompanhada de um resumo dos principais assuntos abordados em cada capítulo.

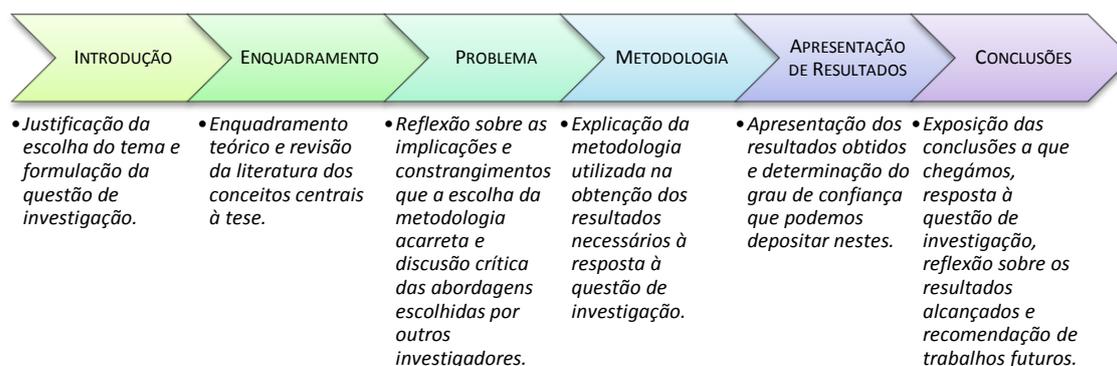


Figura 1: Estrutura da Tese

⁵ <http://www.britannica.com/EBchecked/topic/324278/Elisabeth-Kubler-Ross>, ou http://en.wikipedia.org/wiki/K%C3%BCbler-Ross_model

⁶ A maioria dos servidores *web* utilizam este software. (Pink, 2009)

⁷ <http://socialcapital.wordpress.com/2011/01/26/twitter-facebook-and-youtubes-role-in-tunisia-uprising/>

⁸ <http://www.publico.pt/Mundo/a-morte-de-bin-laden-no-twitter-1492329>

⁹ “Não se descobrem novas terras sem que se aceite primeiro perder de vista a costa durante muito tempo. Porém, os nossos escritores temem os grandes espaços: apenas bordejam a costa.” (tradução nossa)

*The endless cycle of idea and action,
Endless invention, endless experiment,
Brings knowledge of speech, but not of stillness;
Knowledge of speech, but not of silence;
Knowledge of words, and ignorance of the world.
All our knowledge brings us nearer to our ignorance,
All our ignorance brings us near to death,
But nearness to death no nearer to God.
Where is the Life we have lost in living?
Where is the wisdom we have lost in knowledge?
Where is the knowledge we have lost in information?*

T.S. Eliot (1934), *Choruses from The Rock*

2. ENQUADRAMENTO

Este capítulo enquadra teoricamente o tema através de uma revisão da literatura considerada relevante ao assunto em estudo, começando por abordar o conceito de Informação (2.1) e o de Colaboração em Massa (2.2), elencando de seguida uma pequena compilação de exemplos em que, ao longo da história, as multidões se comportam de uma forma inteligente (2.3) e de um modo diametralmente oposto (2.4).

2.1 INFORMAÇÃO

Platão definia conhecimento como uma crença verdadeira com uma razão (*true belief with an account*); esta definição foi adaptada e hoje é comumente aceite definir conhecimento como uma crença verdadeira justificada (*justified true belief*). Se a esta definição retirarmos a parte da justificação, da verdade e da crença, obtemos uma possível explicação para o conceito de informação (Dunn, 2008).

No entanto, Floridi (2004) defende que a informação deve ser definida como verdadeira – o que levanta a questão do que chamar à informação não verdadeira e mesmo a atribuir à palavra verdade um significado imutável no tempo e independente do observador. Quando alguém procura informação, espera que a informação que obtém seja correcta; no entanto, a veracidade desta não está contida no seu significado literal (Fetzer, 2004).

O conceito de informação fica mais claro quando inserido num contexto mais abrangente como a hierarquia DIKW - *Data, Information, Knowledge, Wisdom* (Cleveland, 1982). A distinção entre dados e informação, informação e conhecimento e conhecimento e sabedoria permite a construção de uma hierarquia (Figura 2) ou pirâmide do conhecimento¹⁰ (Ackoff, 1989).

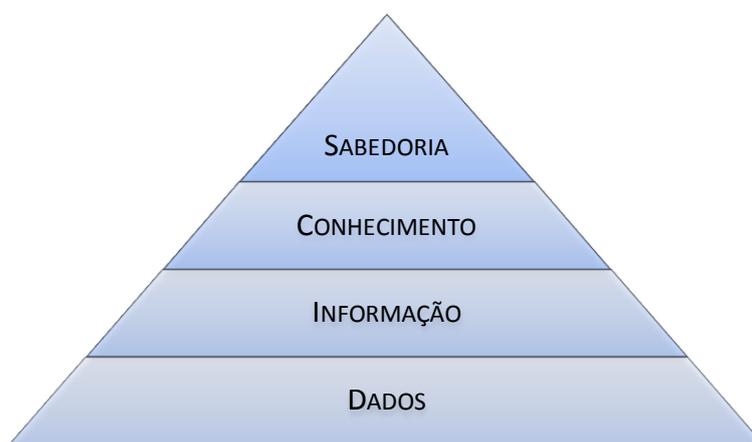


Figura 2: Pirâmide do conhecimento, Adaptado de Ackoff (1989)

Zeleny (1987) refere-se à pirâmide DIKW como *know-nothing, know-what, know-how and know-why*. À luz desta definição, parece evidente que muitas das nossas competências se baseiam apenas na informação (*know-what*), algumas no conhecimento (*know-how*) e poucas na sabedoria (*know-why*). Por exemplo, ao utilizarmos um motor de busca na Internet, competência amplamente difundida em que muitos dominam o *know-how*, quantos na realidade entendem como o motor de busca funciona (*know-why*)?

Ackoff (1989) propôs uma definição para os termos utilizados na pirâmide:

- Dados: Produto da observação. Existem mas não têm significância;
- Informação: O que é inferido dos dados. Permite responder às perguntas «quem», «o quê», «onde», «quando» e «quanto»;
- Conhecimento: Conjunto de informação que é útil;
- Sabedoria: É a capacidade que permite avaliar qualquer escolha.

¹⁰ Alguns autores chamam-lhe hierarquia ou pirâmide da informação.

Apesar de o conceito de informação variar consoante o autor e o contexto, parece razoável seguir a definição de Dunn (2008) e assumir, no âmbito desta tese, que a informação só se transforma em conhecimento quando satisfaz os três testes de Platão (*believed, justified, true*). A quantidade da informação que é, também, conhecimento, será então o nosso instrumento de medida para avaliar a colaboração em massa, enquanto metodologia credível na produção de conhecimento.

Claude Shannon (1948) sugeriu que podíamos medir a informação contida numa mensagem como o inverso do logaritmo de base dois da probabilidade desta ocorrer:

$$\text{Informação}(m) = \log_2 \frac{1}{\text{Probabilidade}(m)}$$

Da análise da equação anterior, facilmente se constata que, quanto menor for a probabilidade de a mensagem ocorrer, maior será a informação nela contida (Figura 3):

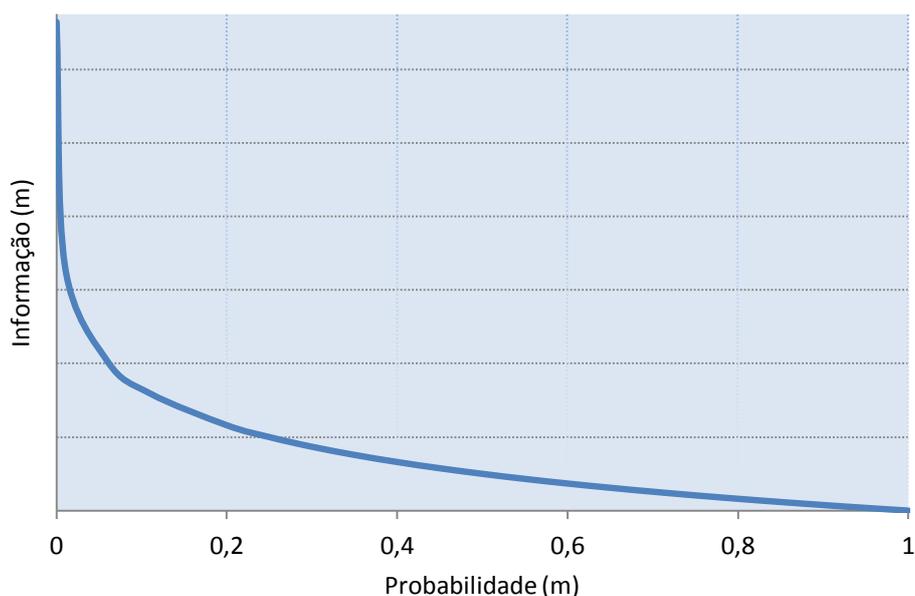


Figura 3: Informação, adaptado de Shannon (1948)

No caso particular em que a informação é uma cadeia aleatória de caracteres binários, a probabilidade de um carácter ser “0” é igual à probabilidade de ser um “1”, e é igual a 0,5; facilmente se mostra que a informação contida nessa cadeia de caracteres binários é igual ao comprimento desta.

Senão vejamos (Tabela 1):

Nº Caracteres	Probabilidade	Inverso da probabilidade	Logaritmo base 2 do inverso da probabilidade
1	0,5	$\frac{1}{0,5} = 2$	$\log_2 2 = 1$
2	$0,5 \times 0,5 = 0,5^2 = 0,25$	$\frac{1}{0,25} = 4$	$\log_2 4 = 2$
3	$0,5 \times 0,5 \times 0,5 = 0,5^3 = 0,125$	$\frac{1}{0,125} = 8$	$\log_2 8 = 3$

Tabela 1: Cálculo da probabilidade

Com n caracteres temos que a probabilidade é de $0,5^n$; o inverso da probabilidade é:

$$\left(\frac{1}{0,5}\right)^n = 2^n \text{ Calculando o logaritmo, temos: } \log_2 2^n = n$$

O que Shannon nos mostra ao sugerir esta forma de calcular a informação contida numa mensagem é que, quanto maior for a probabilidade de uma mensagem ocorrer, menor é a quantidade de informação nela contida; se assumirmos que toda a informação pode ser representada em binário¹¹, então, quanto maior for a mensagem, maior é a informação nela contida e a medida dessa quantidade de informação é igual ao comprimento da cadeia que contém a mensagem. Esta aproximação teórica da quantificação da informação contida numa mensagem conduz a um paradoxo.

Tomemos como exemplo a língua portuguesa e *Os Lusíadas* de Camões. Ao analisar o primeiro verso da primeira estrofe, “As armas e os barões assinalados”, torna-se evidente que, ao escrevermos a primeira letra “A”, estamos a condicionar as letras que podemos usar a seguir, para que a palavra que estamos a formar tenha sentido em Português; i.e., apesar de a probabilidade¹² de ocorrência do primeiro carácter ser $\frac{1}{23}$, a probabilidade do segundo já será maior e assim sucessivamente, de modo a garantir que a frase faça sentido.

¹¹ Os computadores guardam e processam todo o tipo de informação (texto, imagem, áudio, vídeo, etc.) em binário.

¹² Assumindo, por simplificação, um alfabeto composto por 23 caracteres e ignorando sinais de pontuação e acentos.

Se, pelo contrário, admitirmos a completa aleatoriedade dos caracteres que escrevemos, a probabilidade de escrever um caracter será sempre igual $\left(\frac{1}{23}\right)$ e, segundo a fórmula de Shannon, irá sempre conter uma maior quantidade de informação do que uma cadeia de igual dimensão que siga as regras e convenções de uma língua.

Esta problemática é abordada em *The Anatomy of Melancholy* por Burton (1621):

By this art you may contemplate the variation of the twenty-three letters, which may be so infinitely varied, that the words complicated and deduced thence will not be contained within the compass of the firmament. (Burton 1621:2)

No seu livro *The Library of Babel*, Jorge Luís Borges (1941) ilustra este conceito de uma maneira genial, ao descrever uma biblioteca em que estão guardados todos os livros do mundo: os que já se escreveram e os que estão por escrever – estando nestes incluídos todos aqueles gerados pela concatenação de caracteres aleatórios, mesmo que não façam sentido em nenhuma língua. Uma das conclusões¹³ a que chega Borges é que tal biblioteca seria infinitamente grande mas de dimensão finita, ou seja, que é teoricamente possível de executar desde que estejam disponíveis o tempo e os recursos necessários.

W. G. Bloch¹⁴ (2008) analisa a obra de Borges quantitativamente e, baseado no texto deste e em análise combinatória simples, calcula que cada livro terá 1.312.000 posições a preencher com 25 símbolos ortográficos, o que perfaz $25^{1,312,000}$ livros, ou seja aproximadamente $10^{1,834,097}$ – um número absurdamente grande, mas mesmo assim, finito.

Outra abordagem do mesmo tema é a parábola, ou teorema, dos macacos e das máquinas de escrever. Esta aparece sob várias formas e em diversos contextos, mas basicamente expõe a ideia de que um número infinito de macacos, a teclar aleatoriamente em máquinas de escrever, acabará por produzir um texto com sentido, geralmente uma das obras de Shakespeare. Por vezes, o número de macacos é limitado e é o tempo que demora que é admitido como infinito (Gibbons, 2008).

¹³ Para além de outras mais profundas, que provêm de a biblioteca não ser mais do que uma metáfora do universo e de Deus, mas que não são relevantes no contexto desta tese.

¹⁴ Autor de *The Unimaginable Mathematics of Borges' Library of Babel*.

A origem desta parábola é, por vezes, atribuída a T.H. Huxley (Keen, 2007), um apoiante da teoria da evolução das espécies de Darwin, que a terá alegadamente usado num debate entre evolucionistas e criacionistas em 1860. Como não existem registos sobre esse debate e as máquinas de escrever só foram patenteadas por Christopher Sholes¹⁵ em 1868, as dúvidas em relação à paternidade da ideia parecem pertinentes (Gibbons, 2008).

No seu livro *The Total Library*, Borges (1939) refere que as origens do teorema remontam a Aristóteles, encontrando semelhanças num passo do livro por este escrito em 350 a.C., *Metaphysics*:

And as those who make the underlying substance one generate all other things by its modifications, supposing the rare and the dense to be the sources of the modifications, in the same way these philosophers say the differences in the elements are the causes of all other qualities. These differences, they say, are three - shape and order and position. For they say the real is differentiated only by 'rhythm and 'inter-contact' and 'turning'; and of these rhythm is shape, inter-contact is order, and turning is position; for A differs from N in shape, AN from NA in order, M from W in position. (Aristóteles traduzido por Tancred 2004:18)

Apesar da explicação de Borges (1939), não é esta a formulação actual da parábola e, não obstante as semelhanças do conteúdo, a forma é substancialmente diferente. Assim, a primeira referência verificável encontrada à parábola dos macacos foi escrita por Émile Borel (1913) em *Mécanique Statistique et Irréversibilité*:

Concevons qu'on ait dressé un million de singes à frapper au hasard sur les touches d'une machine à écrire et que, sous la surveillance de contremâtres illettrés, ces singes dactylographes travaillent avec ardeur dix heures par jour avec un million de machines à écrire de types variés. Les contremâtres illettrés rassembleraient les feuilles noircies et les relieraient en volumes. Et au bout d'un an, ces volumes se trouveraient renfermer la copie exacte des livres de toute nature et de toutes langues conservés dans les plus riches bibliothèques du monde. Telle est la probabilité pour qu'il se produise pendant un instant très court, dans un espace de quelque étendue, un écart notable de ce que la mécanique statistique considère comme la phénomène le plus probable. (Borel 1913: 192)

Desde então, a parábola foi adoptada e adaptada por diversos autores, de modo a ilustrar diferentes pontos de vista, tendo mesmo entrado na cultura popular e aparecido num

¹⁵ <http://www.britannica.com/EBchecked/topic/541481/Christopher-Latham-Sholes> ou http://en.wikipedia.org/wiki/Christopher_Latham_Sholes.

episódio dos Simpsons¹⁶, numa tira do Dilbert¹⁷ de Scott Adams em 1989, e até Steve Jobs pode afirmar: “*There's an app for that*”¹⁸.

Uma das definições que melhor ilustra o conceito que pretendemos trazer para esta tese, é um poema de Lucio, utilizado numa apresentação à *British Association for the Advancement of Science* (Lucio apud Weaver, 1963):

*Give me half a dozen monkeys,
Set them to the lettered Keys,
And instruct these simian flunkies
Just to hit them as they please.
Lo! The anthropoid plebeians
Toiling at their careless plan
Would in course of countless aeons
Duplicate the lore of man.*

*Thank you, thank you, men of science!
Thank you, thank you, British Ass!
I for long have placed reliance
On the tidbits that you pass.
And this season's nicest chunk is
Just to sit and think of those
Six imperishable monkeys
Typing in eternal rows!*

Em 2002, o jardim zoológico de Paignton resolveu interpretar o teorema literalmente e os macacos Elmo, Gum, Heather, Holly, Mistletoe e Rowan produziram, colaborativamente, um texto que foi compilado e editado em livro¹⁹ com o título *Notes Towards the Complete Works of Shakespeare*. Os resultados, como seria de esperar, não foram muito promissores; no entanto, os autores apenas tiveram dois meses para produzir o ensaio.

Será que a abordagem sugerida no teorema e que se baseia, à primeira vista, na força bruta e na tentativa e erro poderá alguma vez produzir resultados?

¹⁶ Episódio: “*Last Exit to Springfield*” http://en.wikipedia.org/wiki/Last_Exit_to_Springfield

¹⁷ <http://www.dilbert.com/strips/1989-05-15/?Page=1538>

¹⁸ <http://itunes.apple.com/us/app/monkey-typewriter/id446643962?mt=8>

¹⁹ <http://www.vivaria.net/experiments/notes/publication/>

Successful problem solving requires finding the right solution to the right problem. We fail more often because we solve the wrong problem than because we get the wrong solution to the right problem

Russell Ackoff (1974), *Redesigning the Future*

2.2 COLABORAÇÃO EM MASSA

Anos antes do aparecimento da chamada *Web 2.0*, Lévy (1997) publica *L'intelligence collective. Pour une anthropologie du cyberspace*. O autor introduz o conceito de espaço de conhecimento, a que chama *cosmopedia*, em que antevê o potencial do ciberespaço na transformação das estruturas de poder e conhecimento com base na facilidade da comunicação e troca de ideias entre grupos virtuais alargados, não limitados por constrangimentos geográficos. Este fenómeno permitirá o aparecimento de novas formas de cidadania baseadas na troca voluntária e recíproca de informação e de tomada de decisão em grupos temporários alargados, em que a mobilidade dos membros entre comunidades é uma constante. Nas palavras do autor, a inteligência colectiva pode ser definida como:

The members of a thinking community search, inscribe, connect, consult, explore (...) Not only does the cosmopedia make available to the collective intellect all of the pertinent knowledge available to it at a given moment, but it also serves as a site of collective discussion, negotiation, and development (...) Unanswered questions will create tension within cosmopedic space, indicating regions where invention and innovation are required. (Lévy 1997: 217)

A passagem de um modelo de pensamento baseado na ideia de *cogito* (eu penso), para o seu plural *cogitamus* (nós pensamos) prognosticada por Lévy (1997), não se limita a concatenar os pensamentos individuais num único: baseia-se na crença de que a manutenção das singularidades e divergências dentro do grupo são a pedra de toque para a evolução e crescimento de uma inteligência colectiva em que o todo é maior do que a soma das partes.

O que Lévy (1997) considerou uma “alcançável utopia” veio a tornar-se uma realidade anos mais tarde, pelo menos parcialmente, com o surgimento da Wikipédia e do conceito de *prosumers*– junção das palavras “produtor” e “consumidor”– apresentado em *Wikinomics* por Tapscott e Williams (2006).

No seu livro *A Grammar of the Multitude*, Paolo Virno descreve o conceito de *multitudo*²⁰ por contraponto ao conceito de *populus* (o povo). Este conceito alternativo para povo pretende expressar a diversidade dentro de um grupo social por oposição à homogeneidade comumente atribuída às classes sociais, nomeadamente, ao proletariado. O termo *multitudo* foi originalmente utilizado por Niccolò Machiavelli e, posteriormente, utilizado por Spinoza e Hobbes no âmbito da sua discussão sobre teoria política (Virno, 2004).

Para Spinoza (1677), *multitudo* é a forma de existência social e política que assume a pluralidade na acção colectiva e pressupõe que as ideias individuais não se diluem e convergem numa só. Hobbes (1651) associa o termo à resistência à autoridade e à incapacidade de as massas transferirem os seus direitos para os governantes, transformando-se numa única entidade (o povo) a quem podem ser atribuídas vontades e acções, e cuja existência está na base do conceito de estado.

No seu livro *Empire* e na sua sequela *Multitude: War And Democracy In The Age Of Empire*, Hardt e Negri (2000, 2004) fazem a aproximação entre o conceito de rede e *multitudo*, apresentando a sua interpretação de *Multitudo* no seguinte passo:

The multitude must be conceived as a network, an open and expansive network in which all differences can be expressed freely and equally, a network that provides the means of encounter so that we can work and live freely in common. (Hardt e Negri 2004: xiv)

Assim, e independentemente dos diversos contextos em que o termo *Multitudo*²¹ foi e é utilizado e dos diferentes juízos de valor que, entre outros, Spinoza e Hobbes sobre ele teceram, parece-nos um termo adequado para descrever os intervenientes da colaboração em massa, até porque a definição de Hardt e Negri (2004) atribui à multidão características que Surowiecki (2004) considera imprescindíveis para o sucesso desta.

Com base nas ideias apresentadas por Tapscott e Williams (2006), podemos tentar definir colaboração em massa como um modelo de actuação colectiva que ocorre quando um grande número de pessoas trabalha voluntariamente na prossecução de um

²⁰ Multitude, Multidão.

²¹ Doravante iremos substituir o termo em latim por Multidão.

objectivo comum. De um modo geral, os projectos escolhidos são modulares e baseados na Internet de modo a não ficarem confinados à localização geográfica e poderem tirar partido das ferramentas de colaboração *online* – como por exemplo o *software* Wiki – que facilitam a auto-organização dos participantes, potenciando a discussão de ideias e a partilha de conhecimento, de modo a garantir que os objectivos do projecto são atingidos.

Para Ghazawneh (2008), o que distingue a colaboração em massa de outras formas de colaboração é o facto de ser o próprio conteúdo que está a ser criado que gere o processo colaborativo e não as interacções sociais, reais ou virtuais, entre os participantes. Por outras palavras, a ênfase é posta no trabalho que está a ser feito, servindo este de mediador entre os colaboradores. A natureza modular deste tipo de projectos permite mesmo que ocorram contribuições sem haver interacção com outros contribuidores.

De acordo com Tapscott e Williams (2006), a inteligência colectiva é sinónima de colaboração em massa sempre que esta garanta a partilha aberta de ideias, renunciando à propriedade intelectual, entre pares auto-organizados não limitados às barreiras geográficas. Para Howe (2009), a inteligência colectiva é apenas um dos tipos de projectos que compõem o que convencionou chamar *crowdsourcing*. Benkler (2006) e Shirky (2008) utilizam amiúde o termo *peer production*, com um sentido em tudo idêntico ao da colaboração em massa.

A denominação utilizada para descrever o fenómeno que temos estado a analisar varia consoante os autores consultados e, apesar de estarem relacionados, não têm todos a mesma abrangência. Numa abordagem minimalista, se considerarmos as opiniões acima apresentadas, podemos inferir que o *crowdsourcing* engloba a inteligência colectiva e esta, por sua vez, a colaboração em massa, utilizando alguns autores a expressão *peer production* como sinónimo desta.

Na Figura 4, apresenta-se uma compilação das denominações mais comuns e os nomes dos autores²² que amiúde as utilizam. Como o significado que cada autor dá à

²² Existem autores que usam indiscriminadamente mais do que uma denominação e outros nem sempre associam rigorosamente o mesmo conceito à mesma denominação.

denominação que utiliza também varia de caso para caso, a simplificação apresentada pode não reflectir na totalidade a realidade que é, de certeza, mais complexa.

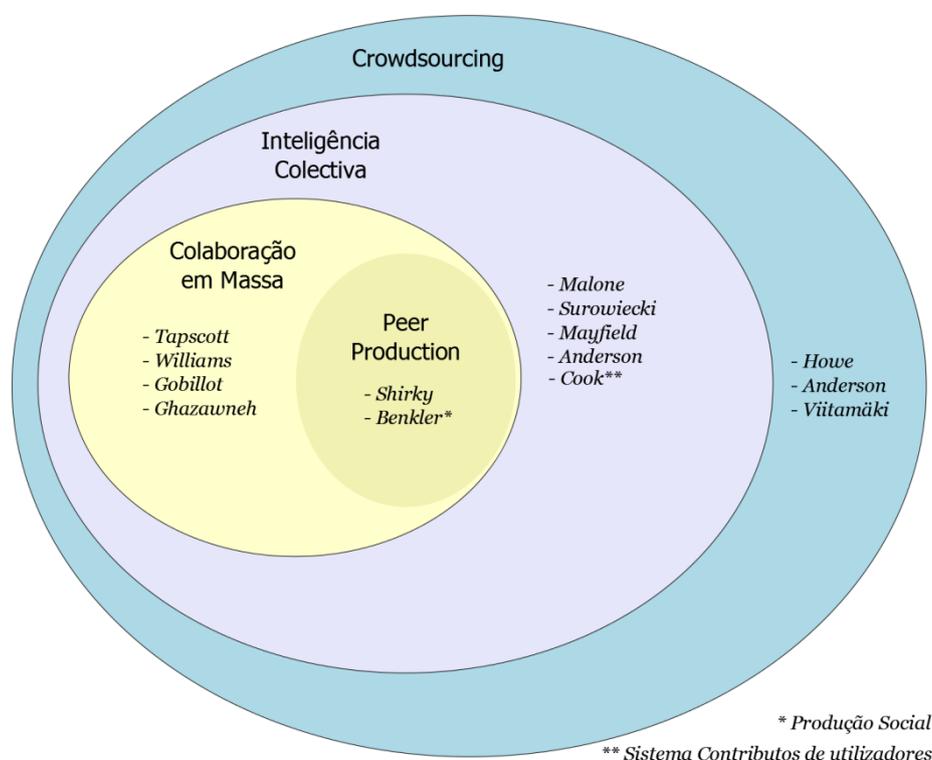


Figura 4: Interpretações da Colaboração em Massa

À excepção deste subcapítulo, em que se pretende enquadrar teoricamente o fenómeno de um ponto de vista mais abrangente, na restante parte da tese iremos utilizar a expressão “colaboração em massa”, segundo a definição de Tapscott e Williams (2006).

Uma explicação sobre a razão por que as pessoas participam voluntária e gratuitamente em projectos de colaboração em massa que, geralmente, não têm retorno tangível, poderá encontrar-se na hierarquia de necessidades proposta por Maslow (1943). Embora Maslow nunca tenha utilizado uma pirâmide para representar a sua teoria, é comum encontrar na literatura a interpretação que apresentamos na Figura 5, em que as necessidades mais básicas constituem a base da pirâmide.



Figura 5: Pirâmide das Necessidades, adaptado de Maslow (1943)

Satisfeitos os níveis da base da pirâmide – necessidades fisiológicas e de segurança - a participação neste tipo de projectos pode ajudar a suprir a necessidade humana de pertencer e ser aceite por um grupo social²³ (Amor & Pertença); pode mitigar o desejo de ser valorizado e reconhecido pelos outros, autovalorizando-se através da participação em actividades que lhe proporcionem a sensação que está a contribuir para algo (Estima); pode ainda conseguir utilizar todas as suas qualidades e potencialidades²⁴ – o que fica facilitado na colaboração em massa em que pode escolher o tipo de contributo que quer dar, ao contrário da maioria das tarefas que desempenha enquanto membro da sociedade em que pode não ter hipótese de utilizar todas as suas capacidades (Auto-Realização).

H. Harlow, M. Harlow, e Meyer (1950) conduziram experiências com primatas em que constataram que a *performance* destes na resolução de *puzzles* piorava quando eram recompensados com comida, afecto e aplausos após a correcta resolução destes. Estes resultados não eram explicáveis pelos factores motivacionais comumente aceites – biológicos²⁵ e recompensas e punições, tendo Harlow (1953) sugerido a existência de um terceiro motivo – o prazer intrínseco associado à realização da tarefa, ou seja, os primatas tinham resolvido os *puzzles* pelo prazer de o fazer.

²³ Neste caso, virtual.

²⁴ “What a man can be, he must be.” (Maslow 1954: 91)

²⁵ Comida, bebida e sexo.

Maslow (1954) explica que a necessidade de auto-realização não surge quando as necessidades dos níveis abaixo deste estão meramente satisfeitas, mas apenas quando o indivíduo as domina e começa então a almejar algo mais. Com base em experiências realizada com grupos de estudantes universitários a quem era pedido para resolver um *puzzle* tridimensional – o cubo de Soma²⁶ – Deci (1971) concluiu, que quando o dinheiro era usado como recompensa externa pelo seu desempenho, os participantes tinham a tendência de perder o interesse intrínseco nessa actividade, tal levando a longo prazo a uma perda de motivação para o desempenho da tarefa. Frey (1997) defende a ideia de que para percebermos o comportamento económico do homem, temos de repensar o conceito de *Homo Oeconomicus* – encarar o homem como um *robot* que busca incessantemente a maximização da riqueza – substituindo-o pelo *Homo Oeconomicus Maturus*, devido à importância dos factores motivacionais intrínsecos que não são contemplados na primeira abordagem, pois não é possível explicar o comportamento económico deste apenas com base nos incentivos externos.

Ariely, Gneezy, Loewenstein, e Mazar (2005) realizaram um estudo na Faculdade de Narayanan, em Madurai, Índia, em que eram oferecidas recompensas monetárias²⁷ aos 87 estudantes que neste participaram. As tarefas a desempenhar implicavam criatividade, concentração e coordenação motora. Os dois grupos com as recompensas mais baixas tiveram *performances* aproximadamente semelhantes, enquanto o grupo cujo pagamento era substancialmente superior aos outros dois, contra todas as expectativas, teve a pior *performance* de todas, verificando-se que o desempenho diminuía com o aumento do incentivo em 8 de cada 9 tarefas.

No seu livro *Drive: The Surprising Truth About What Motivates Us*, Pink (2009) compara as instruções e protocolos que gerem os indivíduos numa sociedade com o sistema operativo de um computador, considerando a versão 1.0 deste aquela em que a motivação está apenas associada ao desejo biológico intrínseco de sobreviver. A versão seguinte (2.0) foi construída com base no sistema de recompensas e punições extrínsecas, quer elas sejam tangíveis ou não, tendo provado ser eficaz enquanto

²⁶ <http://web.inter.nl.net/users/C.Eggermont/Puzzels/Soma/>

²⁷ Os participantes foram divididos em três grupos, recebendo cada grupo um incentivo monetário diferente - 1 dia de salário médio indiano para o 1º grupo, 15 para o 2º e 150 para o 3º.

instrumento de gestão ao longo do século XX, essencialmente nas tarefas mecânicas²⁸ repetitivas que podem ser realizadas seguindo um algoritmo pré-definido, falhando nas tarefas cognitivas (não rotineiras) e criativas que precisam de ser abordadas heurísticamente. Pink (2009) defende esta divisão fazendo referência a alguns dos estudos por nós atrás citados e a muitos outros realizados no âmbito da sociologia, psicologia e economia.

O autor propõe, assim, uma actualização do “sistema operativo” para a versão 3.0, de modo a encontrar uma explicação para os casos em que a abordagem do “pau e da cenoura” não consegue descrever a realidade, nomeadamente na colaboração em massa. Pink (2009) teoriza que, a partir do momento em que pagamos o suficiente às pessoas para o dinheiro deixar de ser um problema, os incentivos financeiros externos são contraproducentes, passando o desejo de autonomia – a possibilidade de escolhermos onde queremos aplicar as nossas capacidades cognitivas, a vontade de alcançar a excelência (mestria) num determinado assunto e a existência de um propósito na actividade que desempenhamos que não seja, apenas, a prossecução do lucro – a constituir os factores motivacionais mais importantes.

De acordo com Pink (2009), podemos dividir as pessoas em dois tipos: o tipo X (extrínseco), que negligencia o prazer de desempenhar uma actividade em favor das recompensas externas, tangíveis ou não, que esta possa trazer, e o tipo I (intrínseco) que, apesar de não desdenhar essas recompensas desde que as considere suficientes, dá mais valor ao trabalho em si do que a estas. O autor considera que, a longo prazo, a *performance* do tipo I tende, quase sempre, a ser superior ao tipo X.

A motivação 3.0 que Pink (2009) descreve como a busca do prazer obtido pela simples realização da tarefa²⁹ por contraponto à realização desta na busca de uma compensação monetária mostra que o homem não responde apenas a impulsos e motivações externas, sendo o seu desejo de autonomia, mestria e propósito uma fonte importante de motivação. Se juntarmos a esta explicação a necessidade de o ser humano suprir as necessidades do topo da pirâmide de Maslow (1954), podemos inferir que os participantes da colaboração em massa muito dificilmente poderão ser descritos pela

²⁸ Este tipo de tarefas foi, ao longo do século XX, passando cada vez mais a ser executado por máquinas.

²⁹ O autor cita como exemplos actividades como a aprendizagem de um instrumento musical, a resolução de *puzzles*, a participação num artigo na Wikipédia e em outros projectos de *Open Source*.

parábola dos macacos e das máquinas de escrever recorrentemente citada por Keen (2007).

Em 2010, Thierer publica um artigo no livro *The Next Digital Decade - Essays on the Future of the Internet*, em que é debatido o impacto que o actual estágio de desenvolvimento da Internet tem, e poderá vir a ter, na cultura e na sociedade. O autor agrupa as diversas opiniões sobre o assunto em dois grandes grupos distintos: os optimistas e os pessimistas. (Thierer, 2010)

Partindo das opiniões expressas nas obras mais recentes de autores como Sunstein (2001), Surowiecki (2004), Anderson (2006), Keen (2007), Benkler (2006), Tapscott e Williams (2006), Shirky (2008), Howe (2009), e Siegel (2008), entre outros, Thierer (2010) elenca os principais aspectos sociais e culturais que estes grupos defendem e que se apresentam na Figura 6.

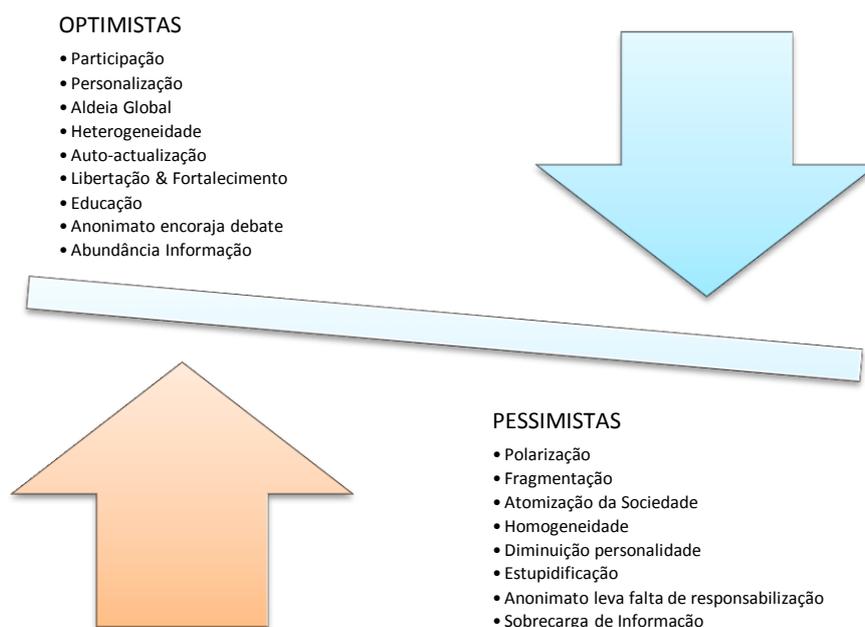


Figura 6: Crenças Sociais e Culturais, adaptado de Thierer (2010)

As diferenças entre os dois grupos reflectem-se também em termos da abordagem económica e dos modelos de negócio que defendem, a saber:

- A oferta de serviços gratuitos é benéfica versus é uma ameaça à qualidade e ao modelo de negócio;

- Foco na colaboração em massa versus esforço individual;
- Aceitação da criatividade dos amadores versus superioridade do profissionalismo;
- Adopção de sistemas abertos versus modelos proprietários;
- Modelos Wiki potenciam a sabedoria das multidões versus a loucura das multidões;
- Benefícios do *crowdsourcing* versus preocupações com a exploração do trabalho gratuito.

Pioneiros no campo da gestão do conhecimento, Takeuchi e Nonaka (1986) propuseram um modelo de desenvolvimento de *software* com base na acumulação do conhecimento tácito, em que substituíam a tradicional abordagem sequencial³⁰ pelo envolvimento da equipa como um todo, ao longo de toda a fase de desenvolvimento do projecto. Usaram o exemplo de uma equipa de rãguebi que passa a bola entre si inúmeras vezes, enquanto se desloca na direcção do objectivo, por considerarem que esta abordagem potencia a partilha de conhecimento, essencial neste tipo de projectos.

DeGrace e Stahl (1990) chamaram a esta abordagem *scrum*³¹, um termo relacionado com o rãguebi mencionado na abordagem inicial de Takeuchi e Nonaka (1986), definindo os diferentes papéis dentro da equipa baseando-se na seguinte história popular:

A pig and a chicken are walking down a road. The chicken looks at the pig and says, 'Hey, why don't we open a restaurant?' The pig looks back at the chicken: 'Good idea, what should we call it?' The chicken thinks about it for a minute and suggests, 'How about "Ham and Eggs"?' 'I don't think so,' says the pig, 'because that would mean that I am committed but you'd only be involved.' (Gobillot 2009: 57)

Assim, a equipa de um projecto *Scrum* é composta por “porcos” – ou seja, aqueles que estão comprometidos com o projecto - e “galinhas” – os que apenas estão envolvidos e

³⁰ Como numa corrida de atletismo de 4x100 metros em que a equipa vai passando o testemunho, sequencialmente, entre si para chegar à meta.

³¹ Formação ordenada para reposição da bola em jogo depois de uma interrupção num jogo de rãguebi.

que, embora tenham interesse neste e nos seus benefícios, não sofrem as mesmas consequências que os “porcos” no caso de o projecto falhar.

No grupo dos “porcos”, temos três tipos de intervenientes: o dono do produto, o facilitador cujo papel é remover os obstáculos que vão surgindo e que impeçam a equipa de alcançar o seu objectivo – não sendo, no entanto, o gestor da equipa, já que esta se auto-organiza – e a equipa propriamente dita que irá realizar o trabalho. Nas “galinhas”, temos os futuros utilizadores do sistema, os *stakeholders* – aqui no sentido do futuro proprietário do produto – e os fornecedores que contribuem pontualmente com bens tangíveis (equipamento) ou intangíveis (conhecimento).

Gobillot (2009) utilizou esta abordagem para tentar perceber a natureza das relações entre os actores envolvidos na colaboração em massa, processo em que a prossecução do objectivo também não é suportada por uma estrutura hierárquica tradicional.

Na Figura 7, podemos ver o modelo de Gobillot (2009) para a colaboração em massa, em que as setas duplas representam as relações de interacção social entre os participantes.

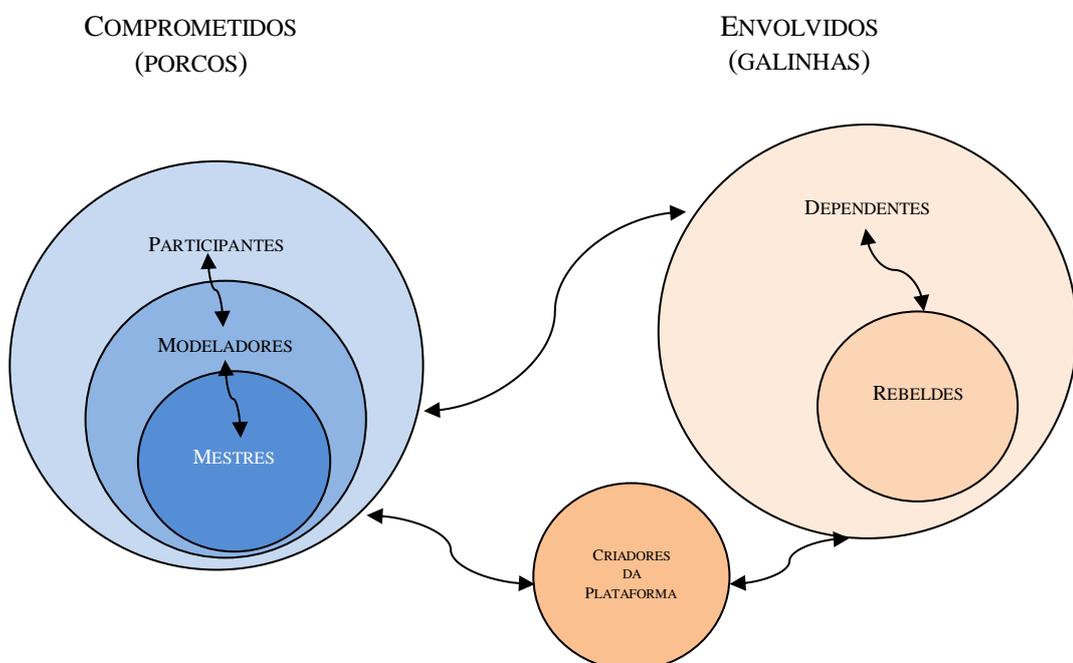


Figura 7: Relações de Interação Social, adaptado de Gobillot (2009)

Gobillot (2009) entende que este modelo reforça a ideia de que o enfoque da colaboração em massa recai sobre os recursos e não sobre as estruturas que, usualmente,

são usadas para os controlar. A escolha entre comprometer-se ou envolver-se com o projecto é deixada ao livre arbítrio de cada um de nós.

Se particularizarmos este modelo geral ao caso da Wikipédia, no caso do grupo dos comprometidos (“porcos”) poderemos fazer coincidir os “participantes” com os contribuidores ocasionais – o maior grupo – que não chegam a registar-se no *Website*; os “modeladores” serão os que contribuem mais regularmente e que se registam, pois pretendem voltar a contribuir no futuro; os “mestres” serão aqueles a quem a comunidade decidiu atribuir “poderes especiais” para lidar com os actos de vandalismo e bloquear temporariamente as edições de artigos sobre temas sensíveis que estejam a ser alvo de sucessivas alterações.

Em relação aos envolvidos (“galinhas”), o “criador da plataforma” é, neste caso, a *Wikipedia Foundation*³²; os “dependentes” – a maioria dos intervenientes – são os utilizadores que consultam a Wikipédia; os “rebeldes”³³ são aqueles que pretendem desestabilizar a comunidade e que vão desde os que praticam actos de vandalismo até àqueles que investem tempo e recursos para a desacreditar. Benkler (2006) explica que, paradoxalmente, este grupo contribui para unir e motivar a comunidade, chamando a atenção para um artigo³⁴ de Robert McHenry (na altura, o editor-chefe da Enciclopédia Britânica), em que este escarnece da Wikipédia por um erro na data³⁵ de nascimento na entrada sobre Alexander Hamilton e pela organização do conteúdo deste, o que provaria que a Wikipédia não era confiável.

Este facto provocou na comunidade um “toque a reunir” que, no espaço de horas, corrigiu o erro, tendo o artigo em causa sido revisto e substancialmente melhorado no espaço de uma semana. Benkler (2006) realça que o ataque de McHenry serviu essencialmente para reforçar o facto de o mecanismo de correcção da Wikipédia funcionar e não para a desacreditar.

³² http://pt.wikipedia.org/wiki/Wikimedia_Foundation

³³ Gobillot chama a este subgrupo “patos”, sendo no entanto usualmente referidos como “galos” no âmbito dos projectos *Scrum*. Estas metáforas servem apenas para realçar o facto de, apesar de fazerem parte do grupo das “galinhas”, não põem ovos, ou seja, não se envolvem de uma forma positiva.

³⁴ http://www.ideasinactiontv.com/tcs_daily/2004/11/the-faith-based-encyclopedia.html

³⁵ O erro também estava presente nas enciclopédias *Columbia* e *Encarta*, embora McHenry o não tenha referido.

Cook (2008) sistematiza no seu artigo *The Contribution Revolution* os diferentes tipos de contribuições individuais, conscientes ou inconscientes, que podem ser usadas para criar valor, pondo em evidência o potencial actual da colaboração em massa. Na Figura 8, podemos ver uma representação dessa classificação.



Figura 8: Tipos de Contribuições, adaptado de Cook (2008)

O autor aponta ainda possíveis explicações para o grande número de contribuições que se verificam. Em primeiro lugar, a maioria das pessoas não se apercebe das recomendações passivas que faz: por exemplo, ao comprar um produto na Amazon está a contribuir com dados que irão alimentar o motor de recomendações desta empresa, sucedendo o mesmo quando selecciona um dos resultados de uma pesquisa no Google. Outro aspecto passa pela disponibilização de um serviço que o utilizador necessita: ao utilizar o Flickr para partilhar as suas fotos com os amigos, caso não as identifique como privadas, está a contribuir para o banco de imagens do *website*.

Cook (2008) realça também o desejo de interacção social, que está por detrás do sucesso do Facebook e do LinkedIn; o reconhecimento público ao escrever, por exemplo, o comentário mais votado numa notícia, ou o reconhecimento dos seus pares³⁶ ao contribuir activamente para a Wikipédia; o desejo de se exprimir escrevendo um blogue

³⁶ Pois a Wikipédia não usa o nome dos autores nos artigos.

ou publicando um vídeo no YouTube ou, pura e simplesmente, por altruísmo, quando escreve uma crítica sincera sobre um restaurante que lhe agradou particularmente.

Na Figura 9, podemos ver uma interpretação da *Power Law of Participation* proposta por Mayfield (2006) no seu *Weblog*³⁷. O autor baseou-se no trabalho de Anderson (2006), substituindo o consumo de conteúdos pela sua criação.

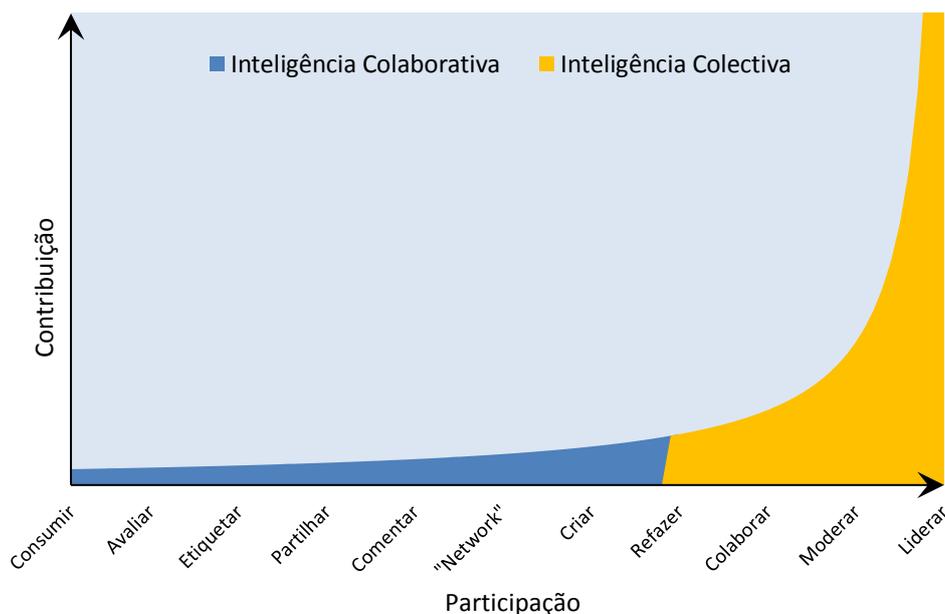


Figura 9: Lei da Participação, adaptado de Mayfield (2006)

Mayfield (2006) mostra a relação, em termos qualitativos, entre as diferentes actividades realizadas *online* e o valor – contribuição – que lhe está associado. As participações mais usuais e que requerem menos empenho contribuem para a inteligência colectiva, enquanto as actividades cujo contributo – valor – é substancialmente maior (embora necessitem de um maior esforço de participação) são realizadas em grupo.

O termo *crowdsourcing* foi cunhado por Howe (2006) num artigo³⁸ em que alerta para as possibilidades de a colaboração em massa estar a tornar-se uma alternativa ainda mais barata que o *outsourcing*. O autor reforça, no entanto, o facto de esta não dever ser encarada como uma forma de trabalho barato, pois resulta de uma decisão da sociedade que só é tomada quando esta se apaixona por um projecto e decide contribuir.

³⁷ http://ross.typepad.com/blog/2006/04/power_law_of_pa.html

³⁸ http://www.wired.com/wired/archive/14.06/crowds_pr.html

No seu livro *Crowdsourcing - Why the Power of the Crowd is Driving the Future of Business*, Howe (2009) define o conceito como o acto de realizar uma tarefa, usualmente desempenhada pela organização ou um fornecedor de serviços, usando as contribuições de um grupo externo à organização, de grande dimensão e tamanho indefinido, sobre o qual esta não tem controlo directo. Para Howe (2009), o *crowdsourcing* permite colmatar um antigo problema da humanidade, a dispersão do talento e do conhecimento, estabelecendo uma ligação entre os detentores destes e aqueles que precisam dele.

Esta abordagem de trazer o conceito de colaboração em massa para o mundo dos negócios, introduzindo a possibilidade de as contribuições serem remuneradas, potenciar que o talento e conhecimento disponível sejam alocados às tarefas que melhor se enquadram nas suas capacidades. Esta abordagem é criticada por Cook (2008), que adverte que os incentivos financeiros poderão fazer baixar a qualidade das contribuições: se as pessoas encararem a sua participação como um trabalho, irão querer ser pagas de uma maneira justa, ou irão ajustar a qualidade dos seus contributos à remuneração auferida.

Howe (2009) identificou quatro tipos de projectos de *crowdsourcing*:

- **Inteligência Colectiva:** Aqueles em que as massas conseguem ter melhores resultados que os indivíduos, desde que estas possuam diversidade de opinião, independência, descentralização e utilizem um mecanismo comum de agregação das contribuições (Surowiecki 2004). Um exemplo deste mecanismo são os fóruns *online* de utilizadores de um determinado *software* que, amiudamente, são mais eficazes que os geralmente eficientes centros oficiais de apoio ao cliente.
- **Criação em Massa:** Iniciativas de criação de conteúdos, como por exemplo a Wikipédia e o YouTube.
- **Votação em Massa:** A possibilidade de classificar os conteúdos (vídeos, artigos, músicas, fotografias e outros) produzidos em grande quantidade, através da votação binária – “gosto” ou “não gosto” – tem como resultado a separação

entre o trigo e o joio, permitindo aos conteúdos de maior qualidade serem reconhecidos e apreciados por outros.

- **Financiamento em Massa:** Quando um grupo de pessoas se propõe desenvolver um produto novo, desde que exista um número mínimo de clientes interessados que se comprometam a comprá-lo e pagá-lo antecipadamente. Usualmente, o preço do produto é mais barato para quem financia o projecto de início; quem é financiado sabe que, pelo menos, irá vender os exemplares suficientes para financiar a maquinaria e materiais que teve de adquirir. Caso dentro do prazo estipulado não seja possível encontrar o número mínimo de clientes estipulado, os pagamentos são devolvidos. Como exemplo temos a kickstarter.

O modelo FLIRT (Focus, Language, Incentives, Rules & Tools) proposto por Viitamäki (2008) pretende ser um enquadramento pragmático para perceber os elementos-chave que condicionam o sucesso de um projecto de *crowdsourcing*, bem como uma ferramenta para o planeamento deste numa perspectiva empresarial (Figura 10).

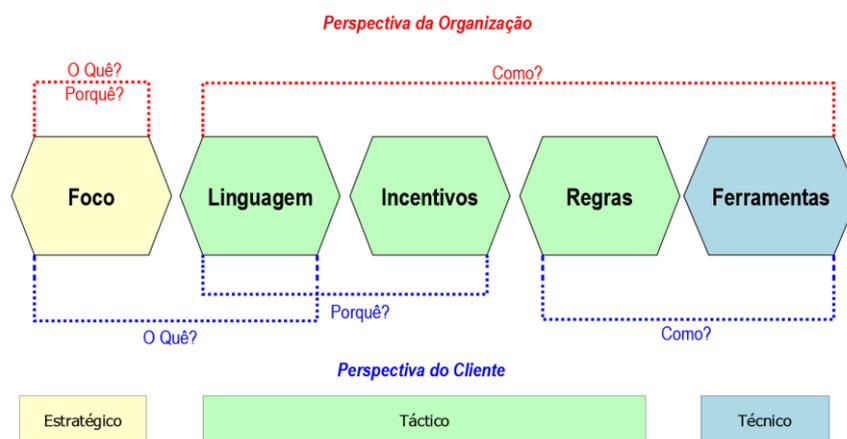


Figura 10: Modelo FLIRT, adaptado de Viitamäki (2008)

Viitamäki (2008) aborda o problema do ponto de vista da organização e dos clientes mostrando quais dos cinco elementos-chave servem para responder às questões de cada grupo:

- **Foco:** compreende a definição dos objectivos do negócio, a participação dos clientes e as capacidades da organização, enquanto aspectos não negociáveis, e o

âmbito, escala e profundidade da actividade que poderão ser negociados com os participantes.

- Linguagem: engloba os objectos e as interacções sociais que funcionam como a “cola social” da comunidade e moldam a imagem da organização.
- Incentivos: poderão ser intrínsecos ou extrínsecos, podendo estes últimos ser ainda objectivos ou subjectivos.
- Regras: definição de como a comunidade irá gerir a iniciação, criação e interacção entre os participantes, bem como os aspectos legais a ter em conta.
- Ferramentas: escolha da plataforma e tecnologia a utilizar.

Malone, Laubacher, e Dellarocas (2010) propõem uma caracterização dos projectos que se baseiam na inteligência colectiva, identificando as características básicas que os compõem – genes – e usando associações destes para descrever os diferentes tipos de iniciativas que se encontram nesta área – genoma.

A abordagem para a identificação dos genes passa pelas respostas aos pares de perguntas “O Quê? e Como?” e “Quem? e Porquê?”. Na Figura 11, podemos ver as respostas possíveis consideradas pelos autores como suficientes para definir os genes da inteligência colectiva.

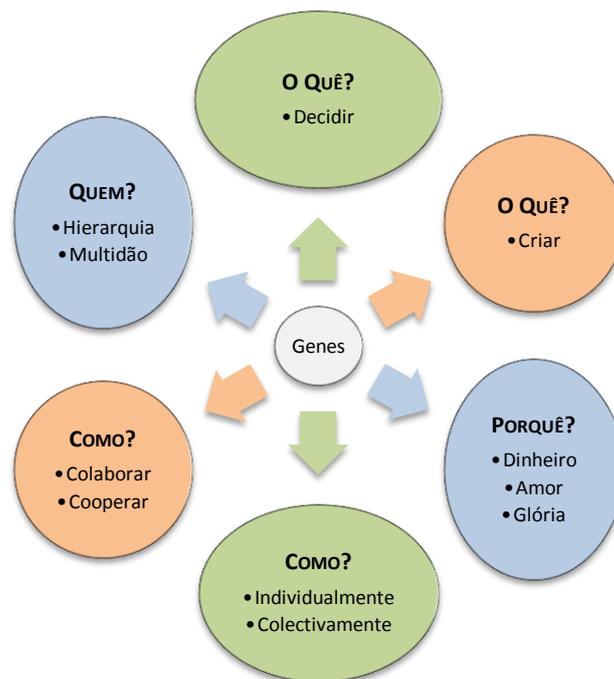


Figura 11: Genes da Inteligência Colectiva, adaptado de Malone et al. (2010)

Para uma melhor compreensão da Figura 11, apresenta-se de seguida uma explicação mais detalhada desta (Malone et al. 2010).

- **QUEM?**
 - Hierarquia: quando alguém tem autoridade de atribuir a uma pessoa ou grupo uma determinada tarefa.
 - Multidão: as tarefas são efectuadas quando alguém de um grupo de grande dimensão decide fazê-lo.
- **O QUÊ?**
 - Decidir: Como? As decisões podem ser tomadas individualmente (compra e venda e criação ou recomendações de conteúdos) ou colectivamente (por votação, consenso e outros).
 - Criar: Como? Através da cooperação – simples agregação de contribuições individuais – ou da colaboração, quando tarefas que poderiam ser realizadas por um indivíduo são executadas em conjunto por vários.
- **PORQUÊ?**
 - Dinheiro: remuneração directa ou indirecta das contribuições.
 - Amor: as contribuições realizam-se pelo prazer altruísta de ajudar e pela identificação com os objectivos do projecto.
 - Glória: busca do reconhecimento da sociedade ou dos seus pares.

Com esta contribuição, Malone et al. (2010) pretendem criar um enquadramento para novos projectos de inteligência colectiva que permita decidir os genes – factores diferenciadores – a “misturar” de modo a criar um genoma vencedor, propondo uma alternativa à mimetização de casos de sucesso, mais susceptíveis de falhar, especialmente quando não se identificaram todos os factores críticos ao sucesso.

O enquadramento da colaboração em massa na área dos sistemas de informação ou da gestão do conhecimento revela-se problemática, em parte devido à abrangência e falta de consenso nas definições das próprias áreas em si. Em relação aos sistemas de informação, Carvalho (2000) aborda precisamente este problema, concluindo que a denominação é utilizada para identificar diferentes assuntos, tendo geralmente em comum o facto de lidarem com informação e estarem relacionados com as organizações ou o trabalho levado a cabo por estas. Para Buckingham, Hirschheim, Land, e Tully (1987), um sistema de informação é um sistema que reúne, guarda, processa e faculta informação relevante para a organização, de modo que a informação é acessível e útil para aqueles que a querem utilizar, incluindo empregados e clientes. Segundo Alter (1996), é uma combinação de procedimentos, informação, pessoas e tecnologias de informação, organizadas para o alcance dos objectivos de uma organização; para Amaral e Varajão (2007), são um meio para a satisfação da missão da organização.

Em comum com todas estas definições está subjacente o facto de os sistemas de informação serem uma ferramenta que pretende ajudar a organização a cumprir os seus objectivos e não um fim em si mesmo: i.e. mesmo que a missão de uma organização seja a produção de informação³⁹, a maioria das abordagens em termos de sistemas de informação concentra-se na gestão da informação que é necessária para produzir o produto, e não no produto em si, mesmo que este seja também informação.

Em relação a outra área da gestão, a gestão do conhecimento, que Nonaka e Takeuchi (1995) consideram ser a gestão do ambiente que permite ao conhecimento fluir ao longo do seu ciclo de vida, que pode ser dividido em explícito e tácito, e que Serrano e Fialho (2003) consideram basear-se na partilha dos conhecimentos individuais para a formação de conhecimento organizacional, podemos ter um raciocínio idêntico, i.e. são um meio e não um fim.

Apesar de estas duas áreas serem suficientemente abrangentes, e alguns autores como Yuan, Chen, Wang, e Du (2008) considerarem a inteligência colectiva *latus sensus* como um tipo de gestão de conhecimento na Internet, e Dalkir (2005) referir que a natureza multidisciplinar da gestão do conhecimento permite englobar nesta as

³⁹ Por exemplo, um jornal.

tecnologias colaborativas, a maioria da teoria por estas produzida não traz grande valor acrescentado à compreensão do caso que pretendemos estudar.

Shirky (2010) apresenta como justificação para a participação voluntária em projectos de colaboração em massa o excedente de cognição que proveniente do tempo livre obtido pela evolução das regalias sociais e do conhecimento que adveio do aumento da literacia, que se verificou a partir do fim da 2ª Guerra Mundial, e não uma alteração do comportamento da sociedade.

Esse excedente, associado à revolução tecnológica recente, que permite a criação de comunidades virtuais não limitadas por barreiras físicas, potenciou a alteração da ocupação dos tempos livres, tendo muitos trocado o consumo passivo de conteúdos por actividades interactivas, intelectualmente mais desafiantes, onde podem dar uso ao seu conhecimento.

Shirky (2010) refere que o potencial de crescimento é enorme, pois as horas gastas a construir a Wikipédia – em todas as linguagens em que esta está disponível – representa aproximadamente 1% do tempo que os norte-americanos utilizam a ver televisão durante um ano.

Segundo Malone (2006), para que a inteligência colectiva ganhe o estatuto que merece junto da academia, é necessário empreender, pelo menos, três tipos de pesquisa:

- Coligir e analisar exemplos e estudos de caso de modo a demonstrar a importância da área;
- Criar novos exemplos de inteligência colectiva, desenvolvendo novas tecnologias ou aplicando as existentes a novos casos;
- Sistematizar estudos e experiências de modo a fazer avançar o estado da arte, compreendendo quais os princípios humanos, organizacionais, sociais e motivacionais que fazem o sistema funcionar.

Esta tese enquadra-se no primeiro tipo, acrescentando ao estudo do caso uma metodologia que prove cientificamente que existe pelo menos um (\exists^{-1}) caso em que, indubitavelmente, a inteligência colectiva é um sucesso. Tal pode contribuir para

estabelecer um ponto de partida seguro, mostrando que vale a pena investigar este domínio, afastando as dúvidas que ainda persistem sobre a qualidade dos resultados da colaboração em massa e o cepticismo de que este método de produção é alvo. Assim, iremos concentrar-nos no estudo da colaboração em massa – enquanto método de produção de informação – e na aferição da qualidade do seu produto final.

Nos próximos subcapítulos, iremos abordar o comportamento da Multidão do ponto de vista de Spinoza (Sabedoria das Multidões) e de Hobbes (Loucura das Multidões).

*Each individual will indeed be a worse judge than the experts,
but collectively they will be better, or at any rate no worse.*

Aristotle (350 a.C.), *The Politics*

2.3 A SABEDORIA DAS MULTIDÕES

Galton (1907a) publicou na revista *Nature* um artigo intitulado “Vox Populi”. Nele relata a realização de um concurso numa feira de gado em que eram aceites apostas sobre qual o peso de um boi, depois de abatido e amanhado. Galton (1907a) tratou estatisticamente os dados do concurso e concluiu que a mediana das 787 apostas registadas diferia apenas 9 lbs⁴⁰ do peso⁴¹ real do boi, ou seja, um erro relativo de aproximadamente 0,75%. Surpreendentemente, o valor calculado era a melhor estimativa do peso do boi, sendo inclusivamente mais correcta que as alvitradas, a título individual, pelos profissionais peritos em gado presentes no local.

Um pouco mais tarde, em resposta a uma das cartas ao editor da revista *Nature*, Galton (1907b) publica uma resposta em que indica que, usando a média aritmética, o resultado seria ainda mais preciso, passando para apenas 1lb (0,08%) a diferença entre o peso real e o peso estimado. No entanto, Galton continuou a defender que a utilização da mediana é mais correcta para expressar a *Vox Populi* por achar ser esta a única que está de acordo com o princípio democrático: “um voto, um valor”.

Hayek (1945), crítico⁴² do planeamento centralizado da economia, publica na revista *The American Economic Review* o artigo “The Use of Knowledge in Society”, em que defende que os preços dos bens são excelentes agregadores da informação e dos gostos de grandes grupos de pessoas, incorporando na sua génese informação que está de tal modo pulverizada que é impossível ser coligida por um sistema de planeamento central. Hayek defende assim que a *Vox Populi* é determinante na economia, não podendo mesmo em certos assuntos ser substituída pela *Vox Expertorum*.

⁴⁰ Aproximadamente 4Kg

⁴¹ O peso real era de 1198lbs e a mediana das apostas, 1207lbs.

⁴² Friedrich Hayek (1944), *The road to serfdom*.

No seu livro *The Wisdom of Crowds: Why the Many Are Smarter Than the Few and How Collective Wisdom Shapes Business, Economies, Societies and Nations*, Surowiecki (2004) começa por descrever o exemplo de Galton (1907a) e apresenta outros mais recentes em que, surpreendentemente, a *Vox Populi* obteve melhores resultados do que os especialistas, como é o caso da previsão do número de gomas contidas num frasco. Esta experiência, facilmente reproduzível, consiste simplesmente em tentar prever o número de gomas que o frasco contém. Encontram-se vários relatos de experiências similares e em todas a estimativa do grupo⁴³ é mais precisa do que a maioria dos palpites individuais. No exemplo citado por Surowiecki (2004), 56 alunos estimaram em 871 o número de gomas contidas num frasco, sendo que este na realidade continha 850 gomas. Apenas um dos 56 alunos teve uma estimativa mais correcta que a estimativa do grupo.

Sunstein (2006), autor de *Infotopia: how many minds produce knowledge*, defende as ideias de Hayek (1945) sobre a importância das massas nos sistemas de preços dos mercados e analisa alguns dos exemplos elencados por Surowiecki (2004), nomeadamente o da experiência do frasco de gomas. Este autor refere que os resultados obtidos seriam totalmente diferentes se, em vez de perguntarmos o número de gomas contidos no frasco, perguntássemos qual o número de átomos de gomas neste contido. Sunstein (2006) mostra que a informação proveniente de vários indivíduos é mais fiável e potencia uma melhor tomada de decisões por parte destes – desde que o grupo respeite as condições do teorema de Condorcet⁴⁴, condições estas que não se verificam certamente no caso da contagem dos átomos de gomas contidos no frasco.

O teorema de Condorcet diz-nos que, se cada elemento de um grupo tiver uma probabilidade superior a 0,5 de tomar a decisão correcta, então a probabilidade de o grupo como um todo tomar a decisão correcta aumenta quando juntamos mais elementos ao grupo, sendo a probabilidade de o grupo (P_n) tomar a decisão correcta dada por $P_n = \sum_{i=(n+1)/2}^n [n!/i! (n-i)!] P^i (1-P)^{n-i}$, sendo n o número de elementos do grupo, aqui considerado ímpar por razões de simplificação de cálculo, e $0,5 < P_n < 1$ (Young, 1997).

⁴³ Valor médio.

⁴⁴ O teorema enunciado pelo Marquês de Condorcet (1785) em *Essay on the Application of Analysis to the Probability of Majority Decisions*

A título de exemplo, consideremos o caso particular em que cada elemento do grupo tem uma probabilidade de 0,51 (apenas ligeiramente acima da decisão aleatória) de tomar a decisão certa. Neste caso, um grupo de 10.000 indivíduos tem uma probabilidade de 99,99% de tomar a decisão correcta. Na Figura 12, podemos ver a variação dessa probabilidade com o número de votantes para o exemplo atrás referido:

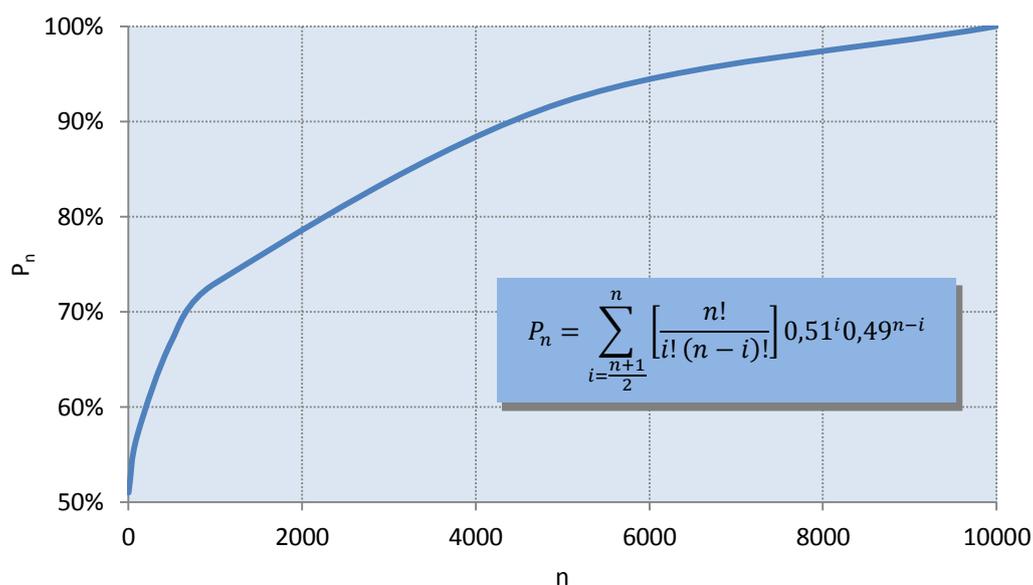


Figura 12: Probabilidade de um grupo tomar uma decisão correcta

Outro dos exemplos de Surowiecki (2004) é o concurso televisivo “Quem quer ser milionário?”, em que, para ganhar, o jogador tinha de responder a uma série de questões de escolha múltipla e onde, no caso de ter dúvidas numa pergunta, poderia pedir telefonicamente a ajuda de um amigo (*Vox Expertorum*) – teoricamente alguém que ele reconhecia como um perito no assunto em causa – podendo também solicitar o voto dos espectadores na plateia (*Vox Populi*) e decidir com base nos resultados desse voto. Da análise dos resultados da edição americana do programa, Surowiecki (2004) constatou que os *peritos* acertaram 65% das vezes, enquanto a audiência chegou aos 91%.

Em *The difference: How the Power of Diversity Creates Better Groups, Firms, Schools, and Societies*, Page (2007) escreve que cada pessoa na “multidão” apenas detém uma parte da informação necessária à correcta tomada de decisão do grupo e defende que a maneira correcta de agregar essa informação é através da votação. Page usa o exemplo de Surowiecki (2004), o programa televisivo “Quem quer ser milionário?”, para ilustrar a sua ideia, construindo um modelo para a análise do mesmo. Tomando como exemplo

uma questão do concurso em que é perguntado qual dos seguintes nomes⁴⁵ não faz parte de uma determinada banda de música composta por três músicos, Page (2007) assume uma plateia de cem pessoas, dividida em quatro grupos: os que sabem a resposta correcta (conhecem todos os elementos do grupo – 7 pessoas), os que conhecem dois elementos do grupo (têm de escolher um nome entre os dois restantes – 10 pessoas), os que conhecem apenas um dos elementos do grupo (e escolhem um nome entre três – 15 pessoas) e os que não conhecem nenhum dos elementos e a sua escolha é aleatória entre as quatro alternativas – 68 pessoas.

Admitindo as respostas possíveis como A, B, C e D e assumindo que A é a resposta correcta, ao passar à votação as 7 pessoas que conhecem a resposta votam em (A), as 10 que conhecem dois dos elementos do grupo escolhem aleatoriamente entre os dois restantes dividindo os seus votos entre estes ($10/2=5$ votos para A), os 15 que só conhecem um dos elementos do grupo têm de dividir os votos pelas três alternativas que lhes restam (mais $15/3=5$ votos para A) e os restantes 68, como não sabem nenhuma resposta, votam aleatoriamente, distribuindo os seus votos igualmente pelas quatro alternativas (mais $68/4=17$ votos para A). Assim, a resposta A obtém $7+5+5+17=34$ votos.

Como foi assumido que os votantes escolhiam aleatoriamente entre as alternativas quando não conheciam as respostas correctas, os restantes $100-34=66$ votos serão repartidos pelas restantes três alternativas de resposta (B, C e D), ou seja, $66/3=22$ votos para cada. Assim, a hipótese A ganha a votação com 34 votos, mesmo tendo usado um exemplo em que o grupo era pouco conhecedor do tema da pergunta: apenas 7% sabiam a resposta correcta e dois terços nada sabiam sobre o assunto.

Partindo do modelo de Page (2007), podemos ir um pouco mais além deste exemplo e chegar a resultados ainda mais surpreendentes. Consideremos uma composição da plateia de dimensão n , em que x , y , z e w são as percentagens relativas à composição desta em termos dos grupos definidos por Page (2007):

- x : Percentagem dos que conhecem os três elementos da banda
- y : Percentagem dos que conhecem dois elementos da banda

⁴⁵ No concurso, são dadas quatro alternativas de resposta.

- z : Percentagem dos que conhecem um elemento da banda
- w : Percentagem dos que nada sabem

As quatro respostas possíveis são na mesma A, B, C e D, sendo que A é a resposta correcta. Como $x+y+z+w = n$ (número de votantes), o resultado das votações será dado pelo seguinte quadro (Tabela 2):

	X	Y	Z	W	Σ
A	$n \cdot x$	$n \cdot \frac{y}{2}$	$n \cdot \frac{z}{3}$	$n \cdot \frac{w}{4}$	$n \left(x + \frac{y}{2} + \frac{z}{3} + \frac{w}{4} \right)$
B	0	$n \cdot \frac{y}{6}$	$n \cdot \frac{2z}{9}$	$n \cdot \frac{w}{4}$	$n \left(\frac{y}{6} + \frac{2z}{9} + \frac{w}{4} \right)$
C	0	$n \cdot \frac{y}{6}$	$n \cdot \frac{2z}{9}$	$n \cdot \frac{w}{4}$	$n \left(\frac{y}{6} + \frac{2z}{9} + \frac{w}{4} \right)$
D	0	$n \cdot \frac{y}{6}$	$n \cdot \frac{2z}{9}$	$n \cdot \frac{w}{4}$	$n \left(\frac{y}{6} + \frac{2z}{9} + \frac{w}{4} \right)$

Tabela 2: Resultado das votações

Para a hipótese A ganhar a votação, temos que garantir que:

$$n \left(x + \frac{y}{2} + \frac{z}{3} + \frac{w}{4} \right) > n \left(\frac{y}{6} + \frac{2z}{9} + \frac{w}{4} \right) \Leftrightarrow$$

$$x + \frac{y}{2} + \frac{z}{3} > \frac{y}{6} + \frac{2z}{9} \Leftrightarrow x + \frac{y}{3} + \frac{z}{9} > 0$$

Como se pode facilmente inferir da inequação anterior, esta só não é verdadeira quando x , y e z são iguais a zero (e consequentemente $w = 100\%$): basta que apenas um dos votantes saiba pelo menos o nome de um dos elementos da banda para que a plateia vote correctamente, ou seja, teoricamente a plateia vota sempre⁴⁶ correctamente, seja qual for a sua composição. Na prática e, como Page (2007) adverte, erros de vários tipos e o facto de as escolhas entre as respostas não serem verdadeiramente aleatórias levam a que, em cerca de 10% dos casos, a plateia se engane.

⁴⁶ No caso em que $w = 100\%$ a votação termina empatada.

Em 1896, o psicólogo francês Gustave Le Bon escreve em *The Crowd – A study of the popular mind* a seguinte constatação (Le Bon 1896):

Under certain given circumstances, and only under those circumstances, an agglomeration of men presents new characteristics very different from those of the individuals composing it. (Le Bon 1896: 26)

Mais de cem anos depois, Surowiecki (2004) escreve a frase que sintetiza a ideia central por detrás de *The wisdom of the crowds* – e que leva mais além a afirmação de Le Bon (1896):

Under the right circumstances, groups are remarkably intelligent, and are often smarter than the smartest people in them. (Surowiecki 2004: XIII)

Ou seja, nas circunstâncias certas, os grupos são notavelmente inteligentes e, muitas vezes, com desempenho superior ao melhor elemento do grupo. A inteligência individual é relevante na resolução de problemas mas, *per se*, não garante que todas as perspectivas de abordagem do problema sejam equacionadas e teoriza que, para um grupo se comportar de forma inteligente, temos de garantir:

- **Diversidade de opinião:** o grupo deverá ser heterogéneo, de modo a todos terem um pouco de informação privada, i.e. informação que não é comum a todos os elementos do grupo;
- **Independência:** a opinião de uns não deve ser influenciada pela dos outros, de modo a não se perder a diversidade de opinião;
- **Descentralização:** as pessoas podem tirar partido do conhecimento local;
- **Agregação:** existência de um mecanismo que transforme as opiniões pessoais numa resposta colectiva⁴⁷.

É de realçar que a resolução dos problemas de coordenação e cooperação que amiudamente se encontram em grupos grandes e heterogéneos é essencial para a correcta tomada de decisões por parte destes.

⁴⁷ O *software* Wiki é um bom exemplo desse mecanismo.

Um exemplo fascinante de aplicação da sabedoria das multidões foi a elaboração do *Oxford English Dictionary*: começou em 1857 e demorou setenta anos, tendo tido a contribuição de dezenas de milhar de pessoas que coligiram 414.825 definições. Os colaboradores – considerados as mentes brilhantes da época⁴⁸ – enviavam as suas colaborações para os editores, que faziam a triagem das definições a incluir no dicionário. No seu livro *The Professor and the Madman: A Tale of Murder, Insanity, and the Making of the Oxford English Dictionary*, Simon Winchester (1998) conta a história verídica de um notável contribuidor, o Dr. William Chester Minor, que, durante duas décadas, contribuiu com cerca de 10.000 definições para o dicionário, tendo-se descoberto quase no fim da sua vida que não era doutor, mas sim um assassino, clinicamente demente, que se encontrava encarcerado num hospício.

Mais recentemente, como exemplos de sucesso da aplicação da sabedoria das massas, temos o Linux e o Google. Segundo Moody (2001), o Linux é desenvolvido através da colaboração em massa por programadores de todo o mundo, de um modo descentralizado, tendo depois um agregador central dos resultados obtidos, o seu criador Linus Torvalds. Uma das vantagens desta abordagem, para além de o resultado final ser gratuito, é que a natureza descentralizada permite que programadores, ou mesmo grupos de programadores, trabalhem em paralelo no mesmo problema, chegando a soluções diferentes – o que permite a Linus Torvalds escolher a melhor solução para o problema em causa, concentrando então a comunidade o seu esforço no aperfeiçoamento da solução escolhida.

Este tipo de abordagem aos problemas é impossível de mimetizar numa empresa convencional, pois esta não tem os recursos humanos disponíveis para o fazer. Como a colaboração é voluntária, é expectável que quem se propõe resolver um problema ou desenvolver uma funcionalidade esteja a trabalhar num assunto que lhe dê prazer, o que à partida lhe proporciona uma maior oportunidade de sucesso. (Moody, 2001)

⁴⁸ Como por exemplo J.R.R. Tolkien

Hardt e Negri (2004), defensores das virtudes da multidão, sintetizam as vantagens do software colaborativo de código aberto, no seguinte passo:

We are more intelligent together than any one of us is alone. Open source, collaborative programming does not lead to confusion and wasted energy. It actually works. One approach to understanding the democracy of the multitude, then, is an open-source society, that is, a society whose source code is revealed so that we all can work collaboratively to solve its bugs and create new, better social programs. (Hardt e Negri 2004: 340)

O outro exemplo de sucesso da aplicação da sabedoria das massas é o motor de busca do Google. O modo encontrado para gerir uma biblioteca de mais de 25 biliões de documentos, não centralizada, em que todos os dias são adicionadas páginas sem que tal passe por uma administração central, passa por encontrar as páginas que contenham a informação referida pelo utilizador na sua busca e depois hierarquizá-las de modo a apresentar ao utilizador, em primeiro lugar, os documentos mais relevantes de acordo com a busca que efectuou (Austin, 2010).

Austin (2010) afirma que é nesta ordenação que está a vantagem competitiva do Google em relação a outros motores de busca. De uma forma simplista⁴⁹, uma página sobe na lista consoante o número de páginas que a refere e o posicionamento destas na lista, i.e., quantas mais páginas bem classificadas apontarem para uma determinada página, mais relevante esta é considerada pelo Google, aparecendo mais acima na lista de resultados de uma busca.

O Google confiou na sabedoria da imensa massa anónima e heterogénea de autores de páginas *Web* e, tendo em conta a popularidade deste e os seus resultados financeiros, não podemos afirmar que se tenha dado mal.

⁴⁹ O algoritmo contém também mecanismos para tentar evitar a manipulação dos resultados através da criação de páginas fantasma, sem conteúdos, que apenas servem para fazer subir outra na lista, bem como outros mecanismos que, por razões óbvias, não são do domínio público.

Nothing is so uncertain, or so difficult to gauge as the temper of a mass of people.

Titus Livius (59 a.C.–17d.C.), *The History of Rome*

2.4 A LOUCURA DAS MULTIDÕES

A 14 de Março de 1838, Thoreau (1888) escreveu no seu diário:

The mass never comes up to the standard of its best member, but on the contrary degrades itself to the level with the lowest. (Thoreau 1888:27)

Parece ser do senso comum que o homem, quando em grupo, se comporta muitas vezes de modo irracional, especialmente quando a diversidade de opiniões na multidão não é garantida, quer por falta de independência nas decisões, quer por mecanismos de agregação, ou mesmo por ambos. Muitas vezes, os líderes que emergem destes grupos têm o efeito perverso de conseguir manipular ou coagir as multidões de modo a que estas passem a ser apenas um instrumento por eles controlado: basta pensar que a maioria dos ditadores da história da humanidade teve sempre por trás multidões que apoiavam as suas actuações.

Para efeitos desta tese, interessa apenas analisar os casos em que a multidão respeita os princípios propostos por Surowiecki (2004) – diversidade de opinião, independência, descentralização e agregação – condições tidas como necessárias, se bem que nem sempre suficientes, como vamos ver mais à frente, para que a multidão se comporte de modo inteligente.

Mackay (1841) escreve *Extraordinary Popular Delusions and the Madness of Crowds*, obra em que são coligidos diversos casos em que grupos se comportam de forma irracional num contexto social e económico. Logo no prefácio podemos ler:

Men, it has been well said, think in herds; it will be seen that they go mad in herds, while they only recover their senses slowly, and one by one. (Mackay 1841:107)

Um dos casos que melhor ilustra o ponto de vista de Mackay (1841) é a descrição do colapso económico ocorrido na Holanda, no século XVII, por causa da obsessão dos holandeses com as tulipas. Apesar de ser hoje em dia imediatamente relacionada com o

país, a tulipa, originária da Ásia central, apenas foi introduzida na Holanda no século XVI: tendo-se transformado num enorme sucesso, a sua posse era considerada um sinal claro de estatuto social, o que levou a um aumento generalizado da procura com o conseqüente aumento do preço destas flores. Chegaram a ser criados instrumentos financeiros para satisfazer a procura, através da compra de entregas futuras⁵⁰ de tulipas. Este instrumento foi evoluindo e passou-se a comprar o direito⁵¹, e não a obrigação, de comprar tulipas que seriam entregues no futuro.

Por volta de 1637, o preço era tão alto que certas variedades podiam chegar a valer o mesmo que uma casa em Amesterdão; o negócio das tulipas tinha enriquecido muitas pessoas e a especulação de preços estava no auge até que, por fim, a “bolha” especulativa rebentou, deixando na falência muita gente.

Este exemplo parece não ter ensinado muito às multidões que constituem os mercados, pois ainda no passado recente assistimos ao rebentar das “bolhas” das *dotcom* e do *subprime*, provando que nem sempre as multidões acertam, tendo Mackay (1841) atribuído a causa do sucedido à ganância, estupidez e ao “instinto de manada” que afirma fazerem esquecer os princípios básicos da economia. Segundo Surowiecki (2004), a explicação do comportamento das massas durante um colapso económico é a perda da diversidade, ou seja, as pessoas deixam de pensar por si próprias e respondem por imitação do comportamento das outras, deixando assim de estar reunidas as condições para agirem de forma inteligente enquanto grupo.

Outro exemplo em que as multidões não estiveram à altura, apesar de não se terem comportado de forma irracional, foi o jogo de xadrez disputado *online*⁵², em 1999, entre Kasparov e o *Mundo*, organizado pela MSN (Microsoft Network). Segundo as regras, as jogadas do *Mundo* eram obtidas por votação das dezenas de milhares de jogadores de xadrez inscritos no *site*, se bem que escolhidas entre as sugestões de um grupo de jovens estrelas de xadrez, e o ritmo de jogo, um movimento a cada vinte e quatro horas, permitia que a equipa do *Mundo* trocasse impressões entre si, através do *chat* disponibilizado no *site*. Ao fim de 124 dias, Kasparov ganhou.

⁵⁰ Mercado de futuros.

⁵¹ Mercado de opções.

⁵² <http://web.archive.org/web/20031212084423/classic.zone.msn.com/kasparov/morenews.asp>

Muitos outros exemplos poderiam ser dados em que multidões se comportaram de um modo não inteligente (chegando mesmo, por vezes, a dar origem ao fenómeno da histeria de massas⁵³). Dois importantes pensadores do século passado, Nietzsche e Jung, pronunciaram-se precisamente sobre a loucura das multidões:

*Insanity in individuals is something rare —
but in groups, parties, nations, and epochs it is the rule.*

Nietzsche (1909), *Beyond Good and Evil*

Masses are always breeding grounds of psychic epidemics

Carl Jung (1940), *Concerning Rebirth*

Mais recentemente, Postman (1993) questiona no seu livro *Technopoly: The Surrender of Culture to Technology* se somos nós que controlamos a tecnologia ou se é esta que nos controla a nós, argumentando que os *media* se tornaram a mensagem a partir do momento em que a sociedade deixou de usar a tecnologia – em particular as de informação – para resolver os problemas do dia-a-dia, deixando que sejam estas a moldar a sociedade. Neste tipo de sociedade a que o autor se refere como *Technopoly* – tecnocracia totalitária – o excesso de informação disponível, gerada em grandes volumes, sem um destinatário em particular, torna esta desprovida de sentido, significado e propósito, alertando para o perigo de nos transformarmos na sociedade retratada por Aldous Huxley em *Um Admirável Mundo Novo*.

Em *Against the Machine: Being Human in the Age of the Electronic Mob*, Siegel (2008) secunda a opinião de Postman (1993) de que a informação está a tornar-se uma forma de lixo, afirmando que a auto-expressão das massas, potenciada pelo uso das tecnologias de informação em geral e pela Internet em particular, está a destruir a verdadeira cultura e o conhecimento, estando este a definhar em informação que, *per se*, não tem grande utilidade.

⁵³ O episódio protagonizado por Orson Welles em 1938, aquando da emissão de um programa de rádio que adaptava o romance de H. G. Wells, *A Guerra dos Mundos*, é mais um exemplo da facilidade com que o pânico injustificado se pode propagar a um grupo alargado de pessoas.

O autor defende ainda que a visão de uma participação *online* em massa, como forma de potencializar a democracia participativa através de modelos de colaboração entre pares, não iria resultar, conduzindo tão-somente a uma relativização da “verdade”.

Keen (2010) também defende que devemos resistir às tentações de tecnologia. Para tal, alega que a actual massificação da utilização desta leva ao aparecimento de amadores que se autoproclamam escritores de manhã, músicos durante a tarde e realizadores de cinema à noite, gerando um sem-número de conteúdos sem qualidade, tentando pôr ao mesmo nível o trabalho dos peritos e dos amadores. Keen (2010) considera que este comportamento é semelhante ao defendido por Marx e Engels (1845):

Whereas in communist society, where nobody has one exclusive sphere of activity but each can become accomplished in any branch he wishes, society regulates the general production and thus makes it possible for me to do one thing today and another tomorrow, to hunt in the morning, fish in the afternoon, rear cattle in the evening, criticize after dinner, just as I have a mind, without ever becoming hunter, fisherman, shepherd or critic.(Marx & Engels 1845: Parte 1, Capítulo 4)

No mesmo artigo, Keen (2010) defende que é o narcisismo que leva as massas a utilizar a tecnologia desta maneira, criando uma superabundância de autores que levam à diminuição das audiências, das elites intelectuais e, conseqüentemente, ao empobrecimento cultural da sociedade.

*It was the best of times, it was the worst of times,
it was the age of wisdom, it was the age of foolishness,
it was the epoch of belief, it was the epoch of incredulity,
it was the season of Light, it was the season of Darkness,
it was the spring of hope, it was the winter of despair,
we had everything before us, we had nothing before us,
we were all going direct to Heaven, we were all going direct the other way.*

Charles Dickens (1859), *A Tale of Two Cities*

3. PROBLEMA

O objectivo deste capítulo é discutir a problemática associada a estudos deste tipo, funcionando como um preâmbulo ao capítulo da metodologia. Vamos primeiro justificar a escolha da Wikipédia como representante dos casos de sucesso da colaboração em massa (3.1), elencando de seguida os estudos similares previamente realizados (3.2) e explicando como a comparação da qualidade de duas enciclopédias permitirá responder à questão de investigação (3.3). Por fim, será discutida a problemática da avaliação da qualidade (3.4) e da escolha dos avaliadores (3.5).

3.1. AMADORES VERSUS PERITOS

O termo *Web 2.0* foi pela primeira vez usado por Dale Dougherty em 2004⁵⁴, e apesar da disputa sobre a pertinência da definição e mesmo sobre o nome⁵⁵ a utilizar, não há dúvida de que a *World Wide Web* tem visto uma transformação crescente dos seus utilizadores de consumidores passivos em produtores activos de conteúdos. Esta mudança foi potenciada pela melhoria significativa da infra-estrutura tecnológica⁵⁶, associada a uma diminuição do custo de acesso à Internet. (O'Reilly, 2007)

Segundo O'Reilly (2007), a massificação do acesso rápido e barato à Internet permitiu o aparecimento de *Websites*⁵⁷ que incentivavam o internauta a criar conteúdos e partilhá-los com os seus pares, em vez de apenas utilizar a Internet para consumir conteúdos

⁵⁴ Fonte: <http://www.paulgraham.com/web20.html>

⁵⁵ Algumas das alternativas mais usadas: *Web 2.0, living Web, Hypernet, active Web, read/write Web*.

⁵⁶ Aumento das velocidades de transferência e da capacidade de armazenamento de dados.

⁵⁷ Como por exemplo o Youtube, Flickr, Blogger, Facebook e Wikipedia.

criados por profissionais e decalcados, com maior ou menor sucesso, dos *media* tradicionais. Esta tendência permitiu um aumento exponencial do número de criadores de conteúdos, ajudados pela facilidade e baixo custo com que a tecnologia actual permite criar conteúdos e torná-los acessíveis a biliões de potenciais consumidores.

Anderson publicou em 2004⁵⁸, na revista *Wired*, um artigo em que explica que num mercado em que a oferta tende para infinito (enorme variedade de produtos), a procura existe mesmo para os produtos mais obscuros. No seu livro, *The Long Tail: Why the Future of Business is Selling Less of More*, Anderson (2006) mostra, recorrendo a exemplos concretos, que os conteúdos cuja publicação e distribuição não eram viáveis de forma tradicional por serem específicos de mais e terem pouca procura estão agora disponíveis e, inclusivamente, representam uma fatia importante do mercado.

Na Figura 13, podemos visualizar uma adaptação e simplificação dos resultados obtidos por Anderson: à esquerda da linha a tracejado, estão os conteúdos com maior popularidade e, conseqüentemente, maior procura, mas que oferecem no entanto pouca diversidade. São geralmente fruto do trabalho dos autores em conjunto com editoras e são distribuídos de forma convencional e em suporte físico.

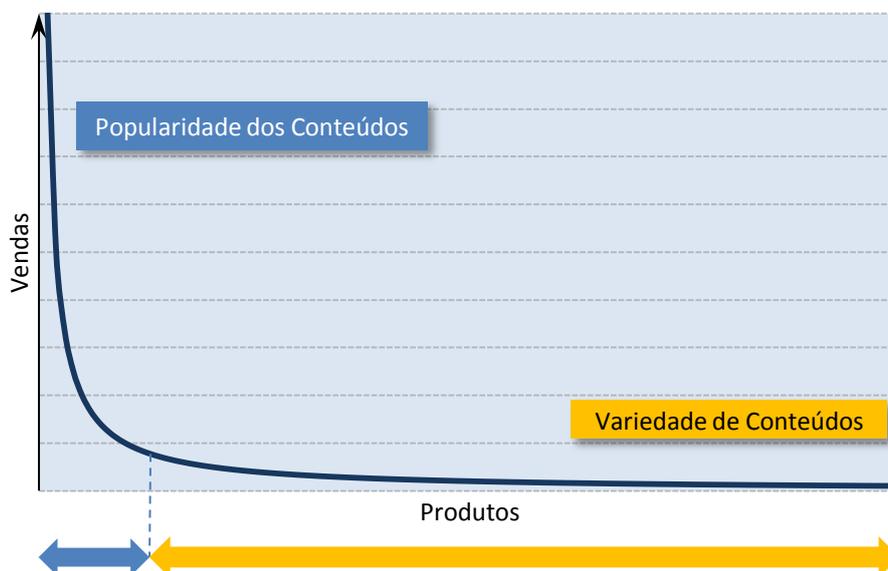


Figura 13: A longa cauda, adaptado de Anderson (2006)

⁵⁸ Fonte: <http://www.wired.com/wired/archive/12.10/tail.html>

Os conteúdos à direita, a *longa cauda*, não são rentáveis para as editoras devido aos custos de distribuição e venda, sendo publicados directamente pelo autor e apenas acessíveis digitalmente; oferecem, no entanto, uma grande diversidade de conteúdos.

Assim, nos últimos anos, temos assistido a uma explosão na diversidade de conteúdos disponíveis *online*, em que uma imensa multidão de amadores, sem o crivo e as condicionantes das editoras, tem disponibilizado conteúdos que, surpreendentemente, têm retirado quota de mercado a todos os tipos de *media*⁵⁹ e estão, de facto, a mudar o mercado. (Anderson, 2006)

O facto de um conteúdo integrar a *cabeça* ou a *longa cauda* não permite tirar conclusões sobre a sua qualidade: o seu posicionamento⁶⁰ na curva depende exclusivamente da procura do mercado por esse item.

No seu livro *Wikinomics: How Mass Collaboration Changes Everything*, Tapscott e Williams (2006) partem desta nova conjectura e tentam mostrar a sua aplicabilidade ao mundo empresarial, pondo especial enfoque na colaboração em massa. Neste momento, o estágio de desenvolvimento da Internet permite fazê-lo e estão também disponíveis os meios necessários para filtrar, ordenar e coligir as contribuições de muitos, como por exemplo o *software* Wiki, que permite aos utilizadores criar e editar colaborativamente páginas *Web*, utilizando apenas um *Web Browser*.

O livro descreve uma série de casos de sucesso de aplicação do modelo de colaboração em massa e defende apaixonadamente o sucesso, e mesmo a inevitabilidade, do modelo, como se pode ler na transcrição abaixo:

Mass collaboration across borders, disciplines, and cultures is at once economical and enjoyable. We can peer produce an operating system, an encyclopedia, the media, a mutual fund, and even physical things like a motorcycle. We are becoming an economy unto ourselves — a vast global network of specialized producers (...) A new economic democracy is emerging in which we all have a lead role (...) The losers launched Web sites. The winners launched vibrant communities. The losers built walled gardens. The winners built public squares. The losers innovated internally. The winners innovated with their users. The losers jealously guarded their data and software interfaces. The winners shared them with everyone.
(Tapscott e Williams 2006: 15, 39)

⁵⁹ «Publishers of music, literature, movies, software, and television are like proverbial canaries in a coal mine — the first casualties of a revolution that is sweeping across all industries». (Tapscott e Williams 2006: 14)

⁶⁰ Basta pensar nos livros e discos mais vendidos em Portugal.

Em *The Cult of the Amateur: How Today's Internet is Killing our Culture*, Keen (2007) tenta desmontar os argumentos apresentados por Tapscott e Williams, desprezando e ridicularizando os conteúdos produzidos através da colaboração em massa, que considera não terem qualidade e estarem a contribuir para o retrocesso das artes e da ciência, socorrendo-se da alegoria dos macacos e das máquinas de escrever para descrever a sua visão do fenómeno, como se pode ver no seguinte passo:

In our Web 2.0 world, the typewriters aren't quite typewriters, but rather networked personal computers, and the monkeys aren't quite monkeys, but rather Internet users. And instead of creating masterpieces, these millions and millions of exuberant monkeys — many with no more talent in the creative arts than our primate cousins — are creating an endless digital forest of mediocrity. For today's amateur monkeys can use their networked computers to publish everything from uninformed political commentary, to unseemly home videos, to embarrassingly amateurish music, to unreadable poems, reviews, essays, and novels. (Keen 2007: 2-3)

No entanto, algumas evidências mostram que nem sempre o processo convencional de criação de conteúdos funciona e garante a qualidade; senão, vejamos os seguintes exemplos.

Em 1994, Alan Sokal (1996), um físico teórico da Universidade de Nova Iorque, submeteu um ensaio à *Social Text*⁶¹, uma importante revista de estudos culturais cujo corpo editorial incluía nomes como Andrew Ross e Fredric Jameson. O artigo, intitulado “Transgressing the boundaries: Toward a transformative hermeneutics of quantum gravity”, foi premeditadamente escrito de modo a agradar à corrente de pensamento dos editores da revista e estava bem escrito e estruturado. No entanto, continha várias incongruências e disparates deliberados, chegando mesmo a afirmar que o significado da gravidade mais não era que uma ficção e que dependia da perspectiva cultural. Em 1996, depois de analisado pelo rigoroso crivo dos editores da revista, o artigo foi aceite para publicação e publicado num número especial da revista dedicada às guerras da ciência (Sokal, 1996).

Doris Lessing, vencedora do prémio Nobel da Literatura de 2007, submeteu dois romances para publicação, *Diary of a Good Neighbour* e *If the Old Could...*, sob o

⁶¹ <http://socialtext.dukejournals.org/>

pseudónimo de Jane Somers, de modo a aferir quais as hipóteses que uma jovem escritora teria de ver a sua obra publicada. Os romances foram recusados pelo seu editor de longa data (Lessing, 2002).

Um dos exemplos de maior sucesso, e que mais opiniões divide, na chamada *Web 2.0*, considerada a *jóia da coroa* dos defensores da colaboração em massa, é a enciclopédia *online*, multilíngue e sem fins lucrativos Wikipédia. Esta enciclopédia é escrita por voluntários anónimos e está estruturada no *software* wiki, que permite a qualquer utilizador editar, apagar e criar entradas.

Fundada em 2001 por Jimmy Wales, está disponível em 260 idiomas⁶² e contém mais de 15 milhões de artigos – dos quais cerca de 3 milhões estão em língua inglesa. Na Figura 14, podemos observar a evolução do número⁶³ de artigos que a compõem:

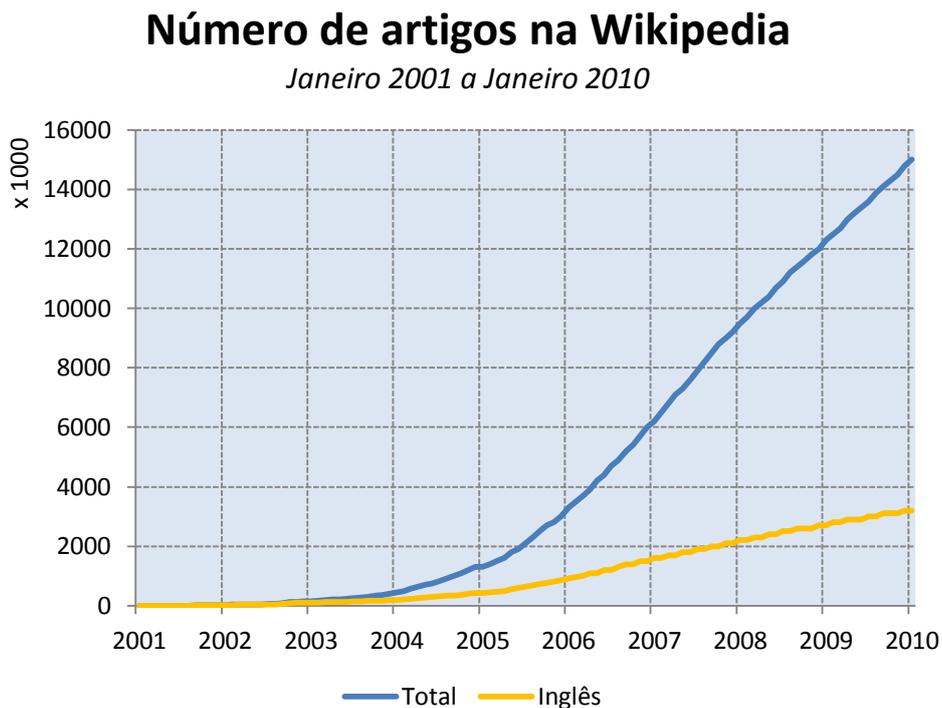


Figura 14: Crescimento da Wikipédia

Como se pode constatar, o crescimento da quantidade de artigos disponíveis mostra claramente a vitalidade do projecto e o aumento do número de pessoas que contribuem

⁶² <http://stats.wikimedia.org/PT/TablesArticlesTotal.htm>

⁶³ Fonte: <http://Wikipedia.org>

para aumentar e melhorar os seus conteúdos. O mais de um milhão de visitas⁶⁴ diárias provam a utilidade e o sucesso da iniciativa. Segundo Nielson (2006), a sua regra dos 90-9-1⁶⁵ não se aplica à Wikipédia, onde apenas cerca de 0,2% contribuem. No entanto, os cerca de 15 milhões de utilizadores registados dão uma estimativa do número de pessoas que já contribuíram na elaboração de conteúdos para a Wikipédia. Note-se que, como não é necessário registo para efectuar correcções em artigos já existentes, a estimativa peca por defeito. (Ver Anexo I)

A Wikipédia também tem crescido em termos de visibilidade junto da academia. Na Figura 15, podemos ver a evolução anual do número de publicações constantes da base de dados da *ScienceDirect*⁶⁶ que referem a palavra “Wikipedia”.



Figura 15: Visibilidade da Wikipédia

Uma das críticas que comumente é feita à Wikipédia é a sua alegada falta de qualidade de informação. Esta crítica deriva essencialmente do modelo – a colaboração em massa – que serve de base à sua criação e manutenção.

⁶⁴ <http://www.alexa.com>

⁶⁵ Em média, na Internet 90% dos utilizadores apenas consomem informação, 9% contribuem esporadicamente e somente 1% o fazem assiduamente.

⁶⁶ <http://www.sciencedirect.com>

A avaliação da qualidade de informação da Wikipédia poderá assim servir de base à determinação da exequibilidade da utilização da colaboração em massa na elaboração de conteúdos científicos confiáveis.

*A revolution doesn't happen when society adopts new tools,
it happens when society adopts new behaviors*

Clay Shirky (2008), *Here comes everybody*

3.2 INVENTARIAÇÃO DE ESTUDOS AFINS

Existem diversos estudos académicos que abordam assuntos relacionados com a Wikipédia, mantendo esta uma página⁶⁷ própria com a listagem destes em suplemento aos repositórios tradicionais. Encontram-se artigos que abordam questões como a evolução da enciclopédia em termos do número de artigos, de editores, da dimensão dos artigos, do comportamento dos editores, do estilo dos artigos produzidos colaborativamente; outros focam o processo colaborativo em si, o conteúdo dos artigos, o modo como estão ligados e a dinâmica que envolve a sua criação.

O problema da qualidade da informação contida na Wikipédia também já foi objecto de estudo por parte da comunidade científica, embora as metodologias utilizadas não tenham sido sempre as mais correctas, quer pelos critérios escolhidos quer porque a amostra utilizada não era representativa. Muitas vezes a Wikipédia é comparada apenas num tema específico e com um concorrente especializado neste – não raras vezes a *autoridade* nesse tema – obtendo-se assim resultados pouco relevantes e expectáveis. Seria mais proveitoso compará-la com os seus concorrentes directos, as enciclopédias generalistas, ou mesmo com as fontes que o utilizador teria em alternativa à Wikipédia, muitas vezes apenas os *Websites* encontrados através de um motor de busca. Foram deixados de fora desta lista os artigos de opinião e os estudos cujo principal enfoque não seja a problemática da qualidade da informação.

Lih (2004) escreve o artigo “Wikipedia as Participatory Journalism: Reliable Sources? Metrics for Evaluating Collaborative Media as a News Resource”, em que analisa artigos da Wikipédia, antes e depois de terem sido citados nos *media*, e encontra uma correlação entre o facto de terem sido citados nos *media* e o aumento da qualidade destes. A qualidade dos artigos era aferida tendo em conta o número de vezes que o artigo era editado e o número total de pessoas que nele colaboravam. A assunção de que

⁶⁷ http://en.wikipedia.org/wiki/wikipedia:wikipedia_in_academic_studies

o número de edições e editores de um artigo resulta sempre num artigo de melhor qualidade não é demonstrada na comunicação de Lih; no entanto, o aumento de edições e editores de um determinado tema após este ter estado em foco nos *media* parece expectável e compreensível.

Em 2004, o jornal *The Guardian* pediu a sete peritos que analisassem um artigo⁶⁸ cada, nas suas respectivas áreas de conhecimento, e os classificassem numa escala de um a dez. A média de classificações obtida foi de 5,6⁶⁹.

Giles (2005) publicou na revista *Nature*⁷⁰ um estudo comparativo sobre o rigor dos conteúdos da Wikipédia e da Enciclopédia Britânica, intitulado “Internet Encyclopaedias Go Head to Head”. Foram escolhidos 42 pares de artigos de diversas disciplinas científicas e enviados a peritos das respectivas áreas para avaliação. Os avaliadores encontraram 166 erros, 4 dos quais bastante graves na Wikipédia, enquanto na Britânica foram encontrados 127 erros, dos quais 4 também eram bastante graves. Um número significativo de avaliadores referiu também que os artigos da Wikipédia que avaliou estavam mal estruturados e eram, por vezes, confusos.

Com base nestes resultados, a *Nature* concluiu que a Wikipédia está próxima da Britânica em termos de rigor nas entradas científicas. A Britânica reagiu em 2006, publicando no seu *Website* uma carta⁷¹ intitulada “Fatally Flawed - Refuting the Recent Study on Encyclopedic Accuracy by the Journal Nature”, em que ataca violentamente o estudo e acusa a *Nature* de alguns artigos terem sido apenas analisados em parte, de outros serem o resultado de colagem de partes de várias entradas, e mesmo de alguns dos artigos não terem sido retirados da própria enciclopédia. A Britânica acusou ainda a revista de não ter sabido distinguir erros graves de pequenas omissões e que é subjectivo afirmar o que deve, e o que não deve, ser incluído numa entrada de uma enciclopédia.

⁶⁸ <http://www.guardian.co.uk/technology/2005/oct/24/comment.newmedia>

⁶⁹ Um dos peritos atribuiu uma classificação de 0 em 10 ao artigo que avaliou. Se não incluirmos este valor, a média global sobe para 6,5 em 10.

⁷⁰ *Nature*, Vol 438, pp.900-901.

⁷¹ http://corporate.britannica.com/britannica_nature_response.pdf

A *Nature* refutou estas críticas, tendo disponibilizado informação suplementar⁷² através da adenda “Supplementary Information to Accompany Nature News Article «Internet Encyclopaedias Go Head to Head»” sobre a metodologia utilizada, argumentando que os critérios usados foram iguais para as duas publicações e que quaisquer erros cometidos pelos seus avaliadores afectaram ambas as enciclopédias, pelo que mantinha as conclusões do estudo, conforme consta na carta⁷³ que enviou à Britânica “Encyclopaedia Britannica and Nature: a Response”.

A revista *Wired* publica o artigo “The Book Stops Here” (Pink, 2005) em que são referidos diversos estudos que mostram que a Wikipédia autocorrige rapidamente os casos de vandalismo. Em alguns estudos citados por Pink (2005), foram premeditadamente introduzidas incorrecções plausíveis em artigos da Wikipédia, que foram, na sua maioria, corrigidos em minutos.

No *The Journal of American History*, Rosenzweig (2006) publica o artigo “Can History be Open Source? Wikipedia and the Future of the Past”, em que analisa 25 biografias de vultos da história americana, comparando as entradas semelhantes na Wikipédia, *Encarta* e na *American National Biography Online*⁷⁴, tendo concluído que as entradas na Wikipédia tinham uma qualidade aceitável, superando mesmo a *Encarta* em termos de biografias disponíveis.

No *Journal of the American College of Surgeons*, Devgan, Powe, Blakey, e Makary (2007) escrevem o artigo “Wiki-Surgery? Internal Validity of Wikipedia as a Medical and Surgical Reference”, em que identificam 39 técnicas cirúrgicas comuns, das quais encontraram 35 na Wikipédia. Em seguida, três médicos independentes analisaram cada artigo em termos de rigor e exactidão científica. Não foram encontrados erros nos trinta e cinco artigos analisados, sendo que 30 poderiam ser recomendados a pacientes, por consenso dos avaliadores. As indicações dos procedimentos foram consideradas adequadas em 34 dos 35 casos analisados.

Bragues (2007) examinou sete entradas da Wikipédia, sobre conhecidos filósofos ocidentais, no seu artigo “Wiki-Philosophizing in a Marketplace of Ideas: Evaluating

⁷² <http://www.nature.com/nature/journal/v438/n7070/extref/438900a-s1.doc>

⁷³ http://www.nature.com/press_releases/Britannica_response.pdf?item

⁷⁴ Obra de referência especializada em biografias, publicada pela Oxford University Press.

Wikipedia's Entries on Seven Great Minds”, tendo confrontado estas com obras de referência sobre os filósofos em causa. Bragues não encontrou incorrecções significativas nos artigos, tendo no entanto encontrado omissões significativas, estimando em 52% o valor médio de cobertura dos tópicos dos artigos.

Pender, Kaye, Lisa, Christopher, e Satyamurthy (2008) escrevem no artigo “Putting Wikipedia to the Test: A Case Study” os resultados de um estudo em que são comparadas as entradas de três doenças na Wikipédia, com o intuito de testar se esta poderia ser usada⁷⁵ em alternativa às referências tradicionais por estudantes de medicina. O painel de médicos que avaliou as entradas conclui que não era recomendável que os alunos baseassem o seu estudo na Wikipédia porque, apesar de as entradas serem razoavelmente concisas e livres de erros, em duas delas havia dois aspectos importantes das doenças em causa que eram omissos.

Clauson, Polen, Boulos, e Dzenowagis (2008) comparam entradas sobre medicamentos na Wikipédia com a *Medscape Drug Reference* (MDR), uma base de dados *online* de acesso gratuito editada tradicionalmente, no artigo “Scope, Completeness, and Accuracy of Drug Information in Wikipedia”. O estudo incidiu sobre oito categorias de medicamentos, concluindo que a Wikipédia reconheceu apenas 40% dos medicamentos testados (MDR: 83%), a informação estava completa em 76% dos casos (MDR: 96%), não foram encontrados erros na Wikipédia (MDR: 4) e a Wikipédia apresentava 48 omissões relevantes (MDR: 14). Os autores concluíram, assim, que a Wikipédia era menos abrangente e completa do que a base de dados escolhida para a comparar.

No artigo “Comparison of Wikipedia and Other Encyclopedias for Accuracy, Breadth, and Depth in Historical Articles”, Rector (2008) compara nove entradas sobre a história americana da Wikipédia com as entradas correspondentes de três enciclopédias: *Encyclopaedia Britannica*, *Dictionary of American History* e *American National Biography Online*. O estudo mostrou que as três enciclopédias tinham uma exactidão que rondava os 95%, enquanto a Wikipédia se ficava pelos 80%.

Todos os estudos referidos concluem que a Wikipédia tem uma qualidade inferior às enciclopédias com que é comparada, deixando no entanto transparecer que a qualidade

⁷⁵ O estudo foi efectuado porque a Wikipédia já estava, de facto, a ser utilizada pelos alunos, em parte devido à sua facilidade de acesso e à utilização maciça de hipertexto.

científica dos artigos desta é superior ao que seria expectável. De um modo geral, tentam aferir a qualidade da Wikipédia, amiudadamente efectuando comparações com outras enciclopédias, tendo no entanto sido criticados por utilizar amostras pouco representativas, metodologias de cálculo da qualidade baseados em critérios pouco abrangentes ou escolher obras especializadas dirigidas a nichos de mercado na sua comparação.

Na tabela 3, podemos ver um resumo das metodologias utilizadas nos trabalhos atrás elencados. A análise destas põe em evidência as diferenças entre o presente estudo e os anteriormente realizados. A metodologia utilizada está na base das críticas a que alguns dos estudos citados foram sujeitos, especialmente quando as suas conclusões apontavam para uma direcção diferente daquela que seria expectável, pelo que a pertinência e o contributo deste trabalho para a clarificação do tema em estudo ficam assim indissociáveis da credibilidade que a metodologia a adoptar trouxe às suas conclusões.

O número de artigos reportados na Tabela 3 refere-se apenas à Wikipédia, pelo que, quando a comparação é feita com, por exemplo, duas outras fontes, o número total de artigos analisados é três vezes maior. A referida tabela também não refere os estudos complementares realizados no âmbito deste estudo, que levaram a que este incorpore um total de 606 avaliações.

Autor	Título	Avaliação	Nº Artigos	Nº Avaliadores	Amostragem	Prova Cega	Área	Comparação
<i>The Guardian</i> (2005)	"Can you trust Wikipedia?"	N	7	7	C	Não	Global	Não comparativo
Giles (2005)	"Internet Encyclopedias Go Head to Head"	E	42	42	C	Sim	Global	<i>Britannica</i>
Rosenzweig (2006)	"Can History be Open Source? Wikipedia and the Future of the Past"	N	25	?	C	Não	História	<i>Britannica Encarta American National Biography Online*</i>
Devgan et al. (2007)	"Wiki-Surgery? Internal validity of Wikipedia as a medical and surgical reference"	E	35	3	C	Não	Medicina	Obras referência
Bragues (2007)	"Wiki-Philosophizing in a Marketplace of Ideas: Evaluating Wikipedia's Entries on Seven Great Minds"	N	7	1	C	Não	Filosofia	Obras referência
Pender et al. (2008)	"Putting Wikipedia to the Test: A Case Study"	N	3	?	C	Não	Medicina	Obras referência
Clauson et al. (2008)	"Scope, Completeness, and Accuracy of Drug Information in Wikipedia"	E	8	?	C	Não	Farmácia	<i>Medscape Drug Reference* (MDR)</i>
Rector (2008)	"Comparison of Wikipedia and other encyclopedias for accuracy, breadth, and depth in historical articles"	N	9	1	C	Não	História	<i>Britannica Dictionary of American History* American National Biography Online*</i>
Este Estudo	"Colaboração em Massa ou Amadorismo em Massa?"	N	195	66	A	Sim	Global	<i>Britannica</i>

(N) Atribuição Nota numa Escala; (E) Contagem Erros; (C) Conveniência; (A):Aleatória; (*) Obra Referência

Tabela 3: Principais Estudos Qualidade Wikipédia

No seu livro *Research Methods for Everyday Life - Blending Qualitative and Quantitative Approaches*, Vanderstoep e Johnston (2009) descrevem e discutem as vantagens e condicionantes dos vários aspectos que caracterizam os métodos de investigação comumente utilizados. Partindo destes princípios gerais, vamos de seguida elencar os aspectos particulares da metodologia a utilizar neste caso, que devem ser tidos em conta de modo a não limitar o âmbito e a aceitação das conclusões a que esta conduzir:

- **Comparar apenas itens comparáveis:** como termos de comparação para aferir a qualidade relativa da Wikipédia não deverão ser utilizadas obras de referência, mas sim enciclopédias generalistas, de modo a garantir que estamos a comparar produtos que pertencem ao mesmo segmento e que são alternativas mútuas entre si. (5 dos 8 estudos não satisfazem este critério)
- **Caracterizar toda a população:** o âmbito do estudo deve ser, se possível, global e não cingir-se apenas a uma área do conhecimento, pois só assim poderemos concluir sobre a qualidade da Wikipédia como um todo. (6 dos 8 estudos não satisfazem este critério)
- **Isenção:** os avaliadores não devem saber a proveniência dos artigos que avaliam (prova cega), de modo a garantir que não se deixam influenciar por ideias preconcebidas que possam ter. (7 dos 8 estudos não satisfazem este critério)
- **Representatividade da amostra:** a amostra deve ser representativa da população que pretendemos estudar, recolhida aleatoriamente e com uma dimensão que garanta a margem de erro que se pretende, de modo a que não se possa afirmar que esta não é representativa da população ou que os elementos escolhidos para análise favorecem um dos pratos da balança. Com isto não queremos dizer que a utilização de amostras de conveniência ou outras não seja válida, mas tão-somente que, no ponto de vista de uma abordagem positivista do problema, esta metodologia conduz a resultados mais fidedignos. (8 dos 8 estudos não satisfazem este critério)

- **Utilização de uma escala ajustada ao que queremos medir:** a qualidade de cada elemento da amostra (artigo) deve ser aferida por um valor dentro de uma escala (nota), resultante de uma apreciação global e comparada com a qualidade do seu par proveniente da enciclopédia generalista a utilizar na sua comparação. Embora a contagem de erros seja precisa para aferir um ditado, *per se*, parece-nos curta para avaliar uma redacção. (5 dos 8 estudos não satisfazem este critério)
- **Utilizar instrumentos de medida fiáveis:** as avaliações devem ser efectuadas por peritos na área do artigo em causa, cuja credibilidade não possa ser posta em causa. (8 dos 8 estudos satisfazem este critério)

Os comentários efectuados às metodologias utilizadas nos estudos atrás elencados não pretendem indiciar que estas não são válidas, mas sim alertar para as limitações às conclusões e abrangência que a sua aplicação acarreta.

Dos estudos citados, o da revista *Nature* (Giles, 2005) é aquele que mais se aproxima deste em termos de metodologia. A diferença temporal entre estes seria, *per se*, um factor justificativo para a realização deste estudo, se tivermos em conta a velocidade de crescimento da Wikipédia e a natureza dinâmica dos seus conteúdos.

As diferenças entre a metodologia escolhida e a utilizada por Giles (2005) tiveram em linha de conta as críticas de que este foi alvo e que, esperamos, irão permitir ter uma ideia actualizada e mais sustentada da qualidade da informação disponibilizada pela Wikipédia e, conseqüentemente, do valor da colaboração em massa.

*The windmill gives you society with the feudal lord;
the steam-mill, society with the industrial capitalist.*

Karl Marx (1920), *The poverty of philosophy*

3.3 WIKIPÉDIA VERSUS BRITÂNICA

A determinação do valor absoluto de algo é sempre uma tarefa difícil, mesmo no campo das ciências exactas, e pressupõe sempre, directa ou indirectamente, uma comparação. Quando afirmamos que, por exemplo, uma pessoa mede 1,80m de altura, estamos no fundo a comparar a altura dessa pessoa com a distância a que se convencionou chamar um metro.

Da mesma maneira, ao querermos medir a qualidade de algo e pondo de lado, por agora, os problemas em termos da universalidade da classificação e aceitação generalizada dos critérios a utilizar, é também mais fácil fazê-lo se utilizarmos outro elemento da mesma categoria – que vamos admitir ser comparável com este – e partir para uma classificação relativa entre os dois, estabelecendo apenas uma relação de ordem na qualidade destes, ou com algum esforço, o seu posicionamento numa escala de classificação curta e generalista.

Para aferir a qualidade da Wikipédia, vamos também ter de utilizar um elemento de comparação. Vamos para tal utilizar uma enciclopédia generalista e que concorre no mesmo mercado do que a Wikipédia: a Enciclopédia Britânica.

A Enciclopédia Britânica⁷⁶ foi publicada pela primeira vez no século XVIII, tendo como editor William Smellie, sendo actualmente a mais antiga enciclopédia de língua inglesa que ainda continua a ser publicada. A sua reputação foi sendo construída ao longo dos tempos graças às contribuições de peritos de renome em diversas áreas do conhecimento que com ela colaboraram, tais como Sigmund Freud, Albert Einstein, Marie Curie, Leon Trotsky, Milton Friedman e Carl Sagan, entre outros.

Durante a explosão da chamada *Web 1.0*, a enciclopédia adaptou o seu formato aos novos tempos tendo passado a disponibilizar, como alternativa à edição em papel, uma

⁷⁶ Fonte: <http://www.britannica.com/EBchecked/topic/186618/Encyclopaedia-Britannica>

versão digital em CD/DVD-ROM que tem a mais-valia da inclusão de conteúdos multimédia, podendo também poder ser acedida, mediante o pagamento de uma anuidade, através do seu *Website*. No início de 2012, a Britânica descontinuou a publicação da edição em papel da sua enciclopédia, passando a disponibilizar apenas a versão digital desta.

Este avanço tecnológico possibilitou que a enciclopédia pudesse passar a ser corrigida e actualizada dinamicamente pelos seus editores – o que, na edição em papel, apenas estava disponível anualmente, com a publicação do *book of the year*, destinado a manter a enciclopédia actualizada até que se justificasse a publicação de uma nova edição que incorporasse esses conteúdos.

Em termos de disponibilidade e facilidade de acesso aos conteúdos, a Britânica é similar à Wikipédia, com a ressalva de o acesso ao texto integral dos artigos só ser possível mediante o pagamento de uma assinatura. Este facto poderá explicar, em grande parte, a diferença de popularidade na escolha dos internautas, não sendo, no entanto, condicionadora da qualidade da informação nela contida.

O parágrafo seguinte, retirado da descrição que os seus editores fazem do seu produto no *Website* da empresa, parece ser esclarecedor da sua atitude e posicionamento no mercado:

*Our line of products has grown, the media of publication have changed, but Britannica's basic mission has not. It's the same as it was in 1768: to be the worldwide leader in reference, education, and learning*⁷⁷.

A escolha da Britânica como elemento de comparação deve-se à credibilidade desta junto da academia – os seus conteúdos foram criados e editados por peritos – e, também, por ter sido utilizada em estudos anteriores (Giles, 2005) como termo de comparação.

Embora existam enciclopédias temáticas que, à partida, oferecem maior qualidade e abrangência de conteúdos, parece-nos faltar rigor científico à utilização de um conjunto de enciclopédias especializadas nas suas áreas do conhecimento para testar a qualidade de uma enciclopédia generalista.

⁷⁷ Fonte: http://corporate.britannica.com/company_info.html

Em relação às enciclopédias escolhidas para a comparação, Tapscott e Williams (2006) têm uma posição diametralmente oposta à de Andrew Keen (2007) no que a estas diz respeito. Podemos ler em *Wikinomics*:

Wikipedia, after all, is now the largest encyclopedia in the world, offered for free, and created entirely by volunteers on an open platform that allows anyone to be an editor. It's amazing that Wikipedia exists at all (...) It's not just its size or popularity, but also the way Wikipedia has evolved that makes it unique. Thousands of Web users volunteer their time and knowledge to help fulfill the community's goal of providing every person in the world with a high-quality encyclopedia in their native language. "Imagine a world in which every single person on the planet is given free access to the sum of all human knowledge. That's what we're doing," says Wales. (Tapscott & Williams 2006: 71)

Enquanto que, em *The cult of the amateur*,

On today's Internet, however, amateurism, rather than expertise, is celebrated, even revered. Today, the Oxford English Dictionary and the Encyclopedia Britannica, two trusted reference volumes upon which we have long relied for information, are being replaced by Wikipedia and other user generated resources. The professional is being replaced by the amateur, the lexicographer by the layperson, the Harvard professor by the unschooled populace.(...) Wikipedia, as I discussed earlier, allows absolutely anyone to add and edit entries on its Web site. (Keen 2007: 37)

Na Figura 16, podemos ver uma representação qualitativa das posições dos autores sobre a qualidade dos artigos que compõem as enciclopédias em questão:

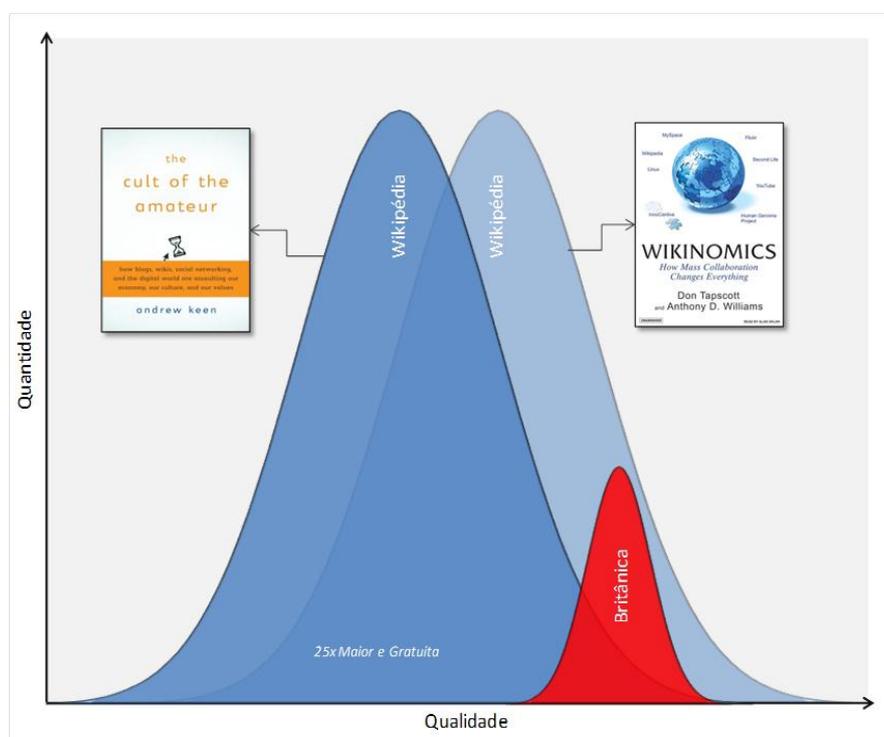


Figura 16: Distribuição dos artigos segundo os "Crentes" e os "Cépticos"

A determinação de qual das duas representações se aproxima mais da realidade permitirá saber se a colaboração em massa (anónima e gratuita) é um mecanismo que consegue produzir uma fonte de informação⁷⁸ pelo menos tão confiável como a abordagem tradicional.

Segundo Shannon (1948), a informação contida numa mensagem é tanto maior quanto menor for a probabilidade desta ocorrer, ou seja, a hipótese defendida em *Wikinomics* é, sem dúvida, aquela que mais informação contém.

O interesse desta comparação não se prende com a avaliação dos produtos em si mesmos, mas sim com as metodologias que lhe estão subjacentes. Assim, vamos assumir que as diferenças de qualidade destes produtos reflectem, maioritariamente, os processos que utilizam na criação de conteúdos que nos propomos avaliar.

Gobillot (2009) considera que, para serem eficazes e cumprirem a sua missão, as organizações necessitam de ter o controlo sobre o que entra e sai destas, bem como o que acontece dentro da organização. Este controlo é usualmente obtido através de uma hierarquia de papéis bem definidos e que permite um correcto aproveitamento dos recursos. Este modelo, que passaremos a referir como tradicional, é o utilizado pela Britânica.

A Britânica usa um processo tradicional⁷⁹ em que os conteúdos são produzidos por peritos remunerados escolhidos pela própria empresa e depois revistos pelos editores antes de serem disponibilizados aos utilizadores. Os editores têm a última palavra no processo e podem mandar reformular, ou mesmo excluir, um artigo que julguem não ter qualidade para ser publicado.

Segundo Keen (2007), é precisamente a escolha de peritos para a elaboração dos artigos e posterior revisão que torna a Britânica uma fonte de informação credível e confiável, por contraponto à Wikipédia. Por outro lado, Tapscott e Williams (2006) alertam que o poder unilateral por parte do editor pode conduzir a artigos com uma abordagem enviesada que reflectem a visão pessoal do editor sobre o tema.

⁷⁸ Segundo a definição de Floridi e não de Fetzer.

⁷⁹ Fonte: <http://corporate.britannica.com/>

Lih (2009) explica que a Wikipédia contrapõe à abordagem académica “editar e depois escrever” a abordagem “escrever e depois editar”: para a colaboração em massa funcionar, é necessário despertar a atenção de eventuais contribuidores, pelo que a criação de um pequeno artigo, mesmo que incompleto e mal estruturado, irá atrair outros contribuidores cuja actividade atrairá outros mais, aumentando assim as hipóteses de se chegar a um artigo de qualidade. Para o autor, o processo pode ser descrito através da metáfora de um cardume de piranhas a alimentar-se, cujo frenesim vai aumentando com a chegada de cada vez mais elementos. Lih (2009) alerta que o comportamento que descreve como “efeito piranha” necessita de uma massa crítica de pessoas e de atenção para funcionar.

A Wikipédia não escolhe os seus contribuidores, não os remunera, e os artigos não passam por nenhuma revisão antes de serem publicados. Segundo Broughton (2008), o *software* wiki permite que cada artigo tenha associado uma página⁸⁰ de discussão em que os contribuidores deste possam conversar e tentar chegar a consenso sobre o seu conteúdo através do diálogo. Caso tal não seja possível, é comum usar a votação para decidir se determinado conteúdo deve ser ou não incluído no artigo.

Nos casos extremos em que a natureza polémica de um artigo suscite “guerras” de edição e mesmo actos de vandalismo que passem pela publicação de conteúdo ofensivo e a remoção dos pontos de vista de outros sobre o tema, a Wikipédia criou a figura do administrador, um utilizador que, pela quantidade e qualidade das suas contribuições, granjeou um estatuto reconhecido pelos seus pares. É-lhe dado o poder de bloquear temporariamente a edição de um determinado tópico, banir contribuidores pela prática reiterada de conduta imprópria e apagar trechos que infrinjam a lei dos direitos de autor ou incluam conteúdos impróprios. É de salientar que estas intervenções apenas têm o intuito de impedir actos de vandalismo e violações das regras de conduta e não decidir o que irá constar no artigo em causa – um pouco à semelhança da justificação da existência da Polícia numa democracia. (Lih, 2009)

⁸⁰ Esta página não é visível para quem estiver a apenas a consultar a enciclopédia. No entanto, basta escolher a opção “editar artigo” para a visualizar.

Como exemplo de algumas destas guerras, e segundo a página⁸¹ da Wikipédia dedicada ao assunto, podemos referir a disputa política subjacente à grafia da capital da Ucrânia Kiev (Russo) ou Kyiv (Ucraniano) – 3.026 modificações; a disputa religiosa sobre as definições de ortodoxia e heresia no artigo sobre o Cristianismo – 15.647 edições; e mesmo discussões irrelevantes como saber se os dotes musicais de Bono enquanto tocador de harmónica devem ser incluídos no artigo sobre os U2 – 10,367 edições.

É de realçar que, para além do acesso aos conteúdos da Wikipédia serem gratuitos, a não inclusão de publicidade nestes garante que os contributos da multidão anónima que criou os conteúdos não estão a ser utilizados pela instituição para gerar receita em proveito próprio, como fazem a maioria das redes sociais. Com efeito, apesar de semelhantes à Wikipédia – disponibilizam gratuitamente conteúdos produzidos pelos próprios utilizadores – o modelo de negócio destes gera, no entanto, receita através da inclusão de publicidade nos acessos aos conteúdos criados, gratuitamente, pelos próprios.

Para Gobillot (2009), as potenciais falhas da política editorial da Wikipédia têm, apesar disso, dois grandes benefícios. O primeiro passa pelo facto de o grande número de contributos potenciar a recolha de conhecimento que até então não tinha sido explorado, providenciando uma visão mais abrangente e com diferentes perspectivas do tema do artigo em causa, por contraponto à visão unilateral do editor. O segundo prende-se com a inexistência de censura editorial que possibilita a abordagem de temas e assuntos de interesse comum mas sem interesse académico e que, numa política editorial clássica são, normalmente por decisão editorial, deixados de fora do conteúdo de uma enciclopédia.

As organizações, segundo Chiavenato (1993), podem ser divididas em empresas e organizações não lucrativas, consoante exista ou não a prossecução do lucro. Esta diferença entre a Fundação Wikipédia e a Empresa Britânica é mais uma das diferenças a ter em conta nesta comparação. Pink (2009) realça a diferença substancial entre os factores motivacionais dos colaboradores de ambos os projectos. Enquanto na Britânica a motivação passa essencialmente pelos factores externos associados à recompensa e ao castigo, na Wikipédia esta é intrínseca e provém da existência de autonomia, da

⁸¹ Fonte: http://en.wikipedia.org/wiki/Wikipedia:Lamest_edit_wars

possibilidade de alcançar a excelência numa actividade não sujeita a rotinas (heurística), e de um propósito que vá para além da simples prossecução do lucro; em síntese – liberdade, desafio e propósito. Segundo Pink (2009), podemos dividir as pessoas em dois tipos: o tipo X (extrínsecos) que negligenciam o prazer de desempenhar uma actividade em favor das recompensas externas – tangíveis ou não, que esta possa trazer e o tipo I (intrínsecos) que, apesar de não desdenhar essas recompensas, desde que as considerem suficientes, dão mais valor ao trabalho em si do que a estas. O autor considera que a performance do tipo I tende, quase sempre, a ser superior ao tipo X no longo prazo.

Na Tabela 4 encontra-se um exemplo dos resultados da aplicação da caracterização de projectos proposta⁸² por Malone et al. (2010) a ambas as enciclopédias, em que se evidenciam as diferenças dos genes destas.

EDIÇÃO DE UM ARTIGO	O QUÊ?	QUEM?	PORQUÊ?	COMO?	
Wikipédia	Criar	Nova versão do artigo	Multidão	Amor e Glória	Colaboração
	Decidir	Manter as alterações efectuadas	Multidão	Amor e Glória	Consenso
Britânica	Criar	Nova versão do artigo	Redactor	Dinheiro e Glória	Individualmente
	Decidir	Manter as alterações efectuadas	Editor	Dinheiro	Hierarquia

Tabela 4: Comparação Genes Wikipédia e Britânica

Lih (2009) descreve o modelo organizacional da Wikipédia comparando-o com o de um formigueiro, em que existem três tipos de indivíduos: as formigas obreiras, os soldados e a rainha, que se diferenciam pelas tarefas que executam, contactando entre elas sem ordens “vindas de cima”, já que não existe uma estrutura hierárquica que comande a comunidade como, por exemplo numa alcateia de lobos.

No seu livro *Emergence: The Connected Lives of Ants, Brains, Cities, and Software*, Johnson (2002) descreve como, em termos individuais enquanto parte de um todo (o formigueiro), as formigas formam uma “inteligência” colectiva, operando de uma forma descentralizada. A inexistência de uma estrutura de comando faz com que as decisões

⁸² Ver o subcapítulo 2.2.

do formigueiro sejam fruto da soma das decisões individuais, da troca de informações par a par e do rasto de feromonas com que vão marcando os trilhos por onde passam.

O rasto de feromonas deixado pelas formigas tem, ao nível da Wikipédia, o seu equivalente nas páginas que permitem aos utilizadores saber que artigos foram editados recentemente e consultar o histórico de edições de um artigo. A interacção par a par é feita na página de discussão que está subjacente a cada artigo e as decisões de cada colaborador são, também, tomadas de moto próprio e não como consequência de um comando de uma estrutura hierárquica.

Por seu turno, a Britânica tem como modelo uma estrutura hierárquica clássica, em que cada nível da estrutura tem papéis diferentes e bem definidos, sendo que um nível é directamente responsável pelo trabalho do nível que está imediatamente abaixo e responde perante o nível que lhe está acima.

Se compararmos os 120.000⁸³ artigos da Enciclopédia Britânica com os mais de 3.000.000 da edição inglesa da Wikipédia, 25 vezes mais, e tendo em conta a reputação de rigor e abrangência da primeira, facilmente constatamos que a maioria dos artigos presentes na Wikipédia deve pertencer à longa cauda⁸⁴ descrita por Anderson (2006).

Admitindo a simplificação – que pode não ser verdadeira – de que a Enciclopédia Britânica cobre todos os assuntos que seria suposto encontrar numa enciclopédia generalista e que a Wikipédia tem uma entrada para todos os artigos constantes na Britânica, i.e. esta última está contida na primeira, então 96% dos artigos da Wikipédia referem-se a temáticas não consideradas suficientemente relevantes pela Britânica para merecerem uma entrada na sua enciclopédia. Não deixa, no entanto, de ser relevante, e mesmo surpreendente, a esmagadora diferença em termos de dimensão das enciclopédias em questão; no entanto, a dimensão e o preço nem sempre são bons estimadores de qualidade.

A Figura 17 mostra, num fluxograma simplificado, os processos que estão por detrás da criação de artigos em ambas as enciclopédias, pondo em evidência as diferenças substanciais entre estes, ao nível do processo de criação de conteúdos.

⁸³ Fonte: <http://www.britannica.com/premium>

⁸⁴ Ver Figura 13, página 70.

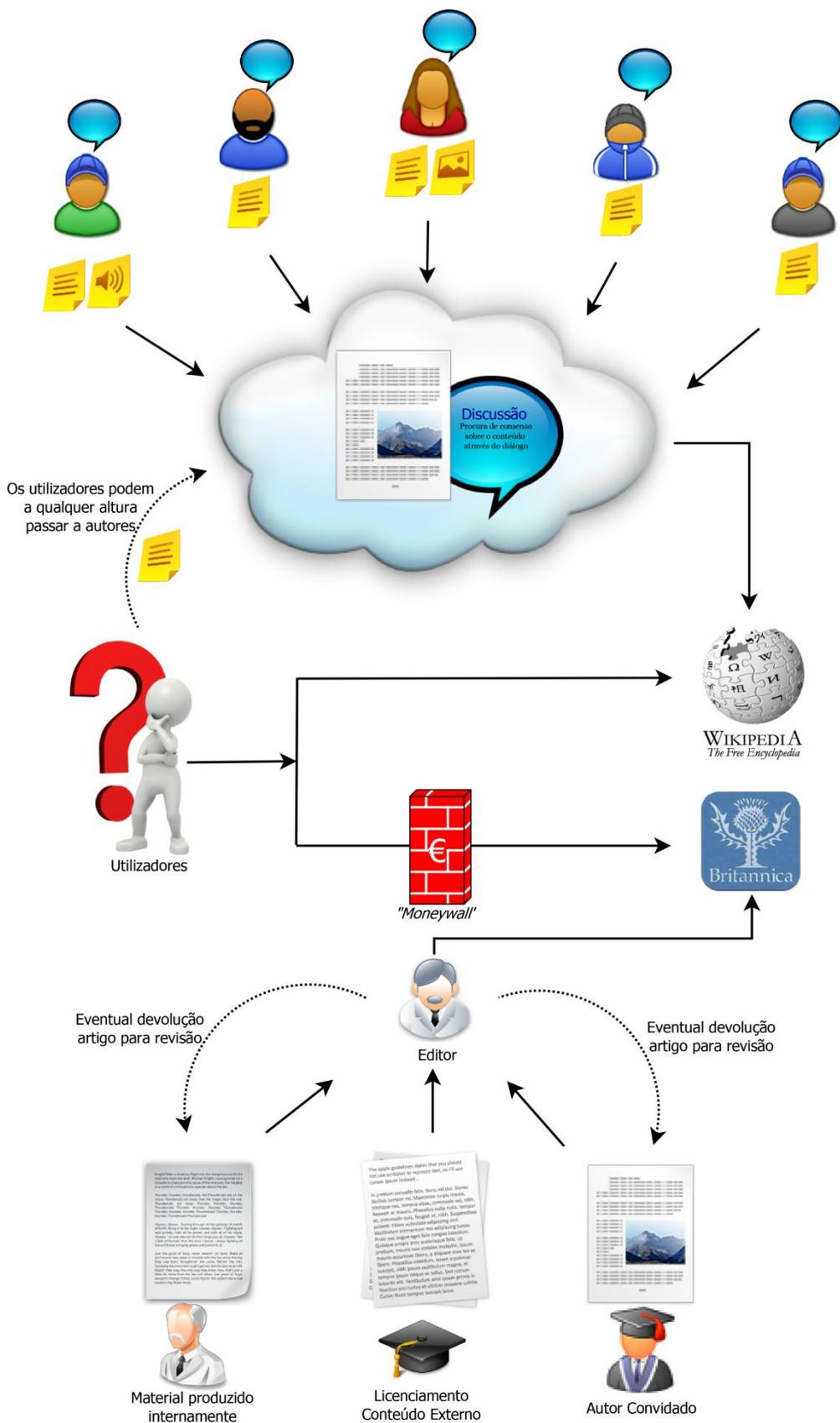


Figura 17: Modelo Funcionamento Wikipédia e Britânica

O Professor Castro Mendes (1994) utiliza na sua explicação do conceito de direito natural a disputa entre as teorias sobre a natureza humana de Rousseau e Hobbes, afirmando que o primeiro acreditava que o homem nascia puro e era a sociedade que o corrompia, enquanto que o segundo acreditava que o homem era intrinsecamente mau e cabia à sociedade a tarefa de o moldar. De facto, Hobbes (1651) defende no seu livro *Leviatã* a sua descrença na natureza humana afirmando: *Bellum omnium contra omnes*⁸⁵, enquanto Rousseau (1754) em *Discurso sobre a Origem e os Fundamentos da Desigualdade entre os Homens* defende precisamente o oposto, escrevendo: “O homem nasce livre, mas em toda parte está aprisionado”. Rousseau adopta uma postura optimista sobre a natureza humana e, embora admita que o homem é intrinsecamente egocêntrico, acredita que este possui sentimentos, como o de piedade, que lhe permitem viver em sociedade, sendo as iniquidades morais e políticas entre estes causadas pela sociedade.

Rousseau (1762) defende no seu livro *O Contrato Social* que o poder deve ficar na mão dos cidadãos (*Multitudo*), enquanto o conceito de estado *Leviatã* de Hobbes (1651) pressupõe que os cidadãos (*Populus*) devem abdicar da sua liberdade em favor do estado, de modo a conseguirem viver em sociedade. Numa análise superficial e simplista, podemos rotular estas posições como democracia não representativa e totalitarismo.

As teorias defendidas por Rousseau e Hobbes podem, *latus sensus*, ajudar a compreender o modelo de funcionamento das enciclopédias em estudo, pondo em evidência as diferenças de pensamento que lhes estão subjacentes, em termos de confiança na natureza humana, na pluralidade, e na necessidade (ou não) de delegar a vontade individual num governante (editor).

Ao analisar o modelo de criação de conteúdos da Britânica, e se entendermos os editores desta como guardiões da qualidade do produto, podemos questionar-nos sobre os perigos que advêm de estes falharem, como Sokal (1996) demonstrou em relação ao *peer review*.

⁸⁵ A guerra de todos contra todos.

O poeta romano Juvenal escreve nas suas Sátiras *Quis custodiet ipsos custodes?*⁸⁶, pergunta que, extravasando o seu contexto original⁸⁷, ilustra o perigo do modelo de criação de conteúdos da Britânica. Uma resposta a esta questão pode ser encontrada em *A República*⁸⁸ de Platão (380 a.C.) em que este, ao referir-se aos guardas⁸⁹, afirma que é absurdo pensar que estes precisem de ser vigiados e devemos confiar que se comportem como é esperado (Hurwicz, 2007).

Independentemente da questão levantada e do posicionamento que tomarmos em relação ao optimismo de Platão e ao pessimismo (cinismo) de Juvenal, o modelo de criação de conteúdos da Britânica é aquele que, teoricamente, tem as condições para produzir resultados de qualidade. O senso comum alerta-nos que o modelo da Wikipédia não deveria, de todo, funcionar, como afirma Noam Cohen num artigo⁹⁰ publicado no *The New York Times*, em que resume a questão abordada neste subcapítulo, ao transcrever um comentário de um contribuidor da Wikipédia: “*The problem with Wikipedia is that it only works in practice. In theory, it can never work.*”

Com o intuito de explicar a viabilidade do modelo de funcionamento da Wikipédia, Shirky (2008) conta-nos a história do santuário de Ise⁹¹, um santuário xintoísta dedicado à deusa do sol situado na ilha de Ise no Japão há mais de 1.300 anos. A UNESCO recusou incluir o local na sua lista de lugares históricos porque este é feito de madeira, o que obriga os sacerdotes que mantêm o santuário a reconstruí-lo de duas em duas décadas, usando madeira cortada da mesma floresta com que o original foi talhado. Do ponto de vista dos sacerdotes, o santuário tem 1.300 anos; do ponto de vista da UNESCO, apenas 20 – pois estes avaliam o edifício e não o processo – o que faz com que um castelo em ruínas com 500 anos mereça a inclusão na lista e um santuário que é reconstruído geração após geração durante mais de mil, não.

⁸⁶ “Quem guardará os guardas?”

⁸⁷ Juvenal sugere que não podemos confiar nas esposas e que a solução não passa por guardá-las, pois também não podemos confiar nos guardas.

⁸⁸ Escrita cerca de meio século antes, e a propósito de um contexto completamente diferente do de Juvenal.

⁸⁹ Platão refere-se aos guardas enquanto guardiões (governantes) das cidades estado Gregas.

⁹⁰ http://www.nytimes.com/2007/04/23/technology/23link.html?_r=1&ex=1178510400&en=c0eb1b23e5c579f7&ei=5070

⁹¹ <http://www.britannica.com/EBchecked/topic/241431/Grand-Shrine-of-Ise> ou http://en.wikipedia.org/wiki/Ise_Grand_Shrine

Nesta história, Shirky (2008) deixa clara a associação entre a Britânica – durável, castelo – e a Wikipédia – efémera, santuário – afirmando que esta última precisa para sobreviver de actos de “amor”, i.e. cuidados e dedicação constantes à semelhança do santuário xintoísta, pois ambos são construídos a partir de materiais efémeros. Caso a Enciclopédia Britânica decidisse abandonar o negócio, o produto em si, iria decaindo devagar e gradualmente à medida que os seus conteúdos fossem ficando obsoletos – à semelhança de um castelo erigido em pedra. Por outro lado, o abandono da Wikipédia por parte dos seus contribuidores rapidamente iria conduzir à sua desintegração através de actos de vandalismo e de grupos de interesses que iriam alterar os artigos em proveito próprio, provocando o seu desaparecimento num curto espaço de tempo.

À semelhança da UNESCO, a academia, de um modo geral, também não confia na Wikipédia. Se os sacerdotes xintoístas conseguem manter ao longo de gerações o santuário da ilha de Ise, apesar de este não ser feito de pedra, através do “amor” que os liga a este, a Wikipédia também poderá sobreviver e vingar durante o tempo que durar a relação entre os contribuidores e o projecto (Shirky, 2008).

*Understanding Poetry, by Dr. J. Evans Pritchard, Ph.D.
If the poem's score for perfection is plotted along the horizontal of a graph, and its importance is plotted on the vertical, then calculating the total area of the poem yields the measure of its greatness.(...)
A sonnet by Byron may score high on the vertical, but only average on the horizontal. A Shakespearean sonnet, on the other hand, would score high both horizontally and vertically, yielding a massive total area(...)
Excrement. That's what I think of Mr. J. Evans Pritchard.
We're not laying pipe, we're talking about poetry.*

Tom Schulman (1989), *Dead Poets Society: The Screenplay*

3.4 OBJECTIVO VERSUS SUBJECTIVO

Um dos factores críticos no sucesso deste estudo é a escolha do método de aferir e quantificar a qualidade dos artigos. Como a utilização de um método objectivo permite a reprodutibilidade do estudo, foi considerada a hipótese de utilização de testes de legibilidade, que utilizam métricas com base no número de caracteres, sílabas, palavras e frases, e geram um resultado quantitativo susceptível de ser comparado (Thorndike, 1921; Zipf, 1949; Klare 1963). O facto de estes testes apenas pretenderem concluir sobre a legibilidade e não sobre a qualidade, associado ao facto de o conteúdo das entradas nas enciclopédias não ser apenas texto e os resultados obtidos nos testes exploratórios efectuados não terem sido muito animadores, levaram ao abandono desta opção.

Adler, Chatterjee, Alfaro, Faella, Pye, e Raman (2008) propõem um algoritmo para determinar a confiabilidade de uma entrada na Wikipédia. Para tal, tiram partido do carácter transaccional⁹² das edições das entradas na Wikipédia, bem como da reputação⁹³ do utilizador que efectuou a alteração. O resultado destes cálculos permite “pintar” o fundo das frases com códigos de cor que reflectem que partes do artigo são mais confiáveis do que outras. O algoritmo considera mais confiáveis as partes das entradas que tiveram mais edições e melhores editores. Para além de ser questionável que exista uma correlação forte entre o número de edições e a reputação dos editores e a

⁹² As Wikis não guardam apenas a última versão de uma entrada, mas sim todas as edições que conduziram ao seu estado final. Esta metodologia permite, por exemplo, anular rapidamente actos de vandalismo.

⁹³ A reputação de um utilizador é calculada com base no seu registo histórico de edições.

qualidade de um artigo, a aplicação desta metodologia tem como inconveniente o facto de não poder ser utilizada nos artigos da Enciclopédia Britânica, que não são elaborados colaborativamente.

Postos de parte os métodos automáticos quantitativos (objectivos), resta introduzir o factor humano (subjectivo) na equação e partir para uma abordagem qualitativa.

A determinação da qualidade através da avaliação efectuada por pessoas vai, à partida, condicionar os resultados do estudo em termos culturais e temporais, impedindo a sua reprodutibilidade, ou seja, a validade deste estudo está limitada ao local e à época em que este foi efectuada, como se de uma fotografia se tratasse, sendo ainda também afectado pela qualidade dos “fotógrafos”, i.e. os avaliadores.

O carácter temporal da “verdade” é facilmente explicável pelo avanço da ciência: tomando como exemplo a lei da soma de velocidades de Galileu que aprendemos na escola, $s = v + u$, esta passou a estar errada⁹⁴ aquando da apresentação da teoria da relatividade restrita, que nos diz que a soma das velocidades é, sim, dada por $s = \frac{v+u}{1+\frac{v.u}{c^2}}$

A incerteza quanto ao que é correcto – ou seja, “verdade” – varia também com o contexto cultural em que o avaliador está inserido. Charles Darwin escreveu *A Origem das Espécies* há mais de 150 anos e, no entanto, segundo um estudo publicado por Miller, Scott, e Okamoto (2006) referente a 2005 sobre a aceitação pública da teoria da evolução em 34 países, sobressai o facto de, em países como a Islândia, Dinamarca, Suécia e França, 80% dos inquiridos aceitarem a teoria da evolução, enquanto nos Estados Unidos apenas 40% o fazem, sendo este o valor mais baixo a seguir à Turquia. Neste estudo, Portugal fica um pouco acima dos 60%.

O valor dos Estados Unidos é ainda mais surpreendente se considerarmos que aos dois quintos que acreditam na teoria da evolução se juntam aproximadamente os mesmos dois quintos que acreditam na teoria do criacionismo e um quinto de indecisos. De acordo com um inquérito (Gallup Poll⁹⁵ 6-7 Fev. 2009), a proporção de americanos que aceita a teoria da evolução aumenta com a escolaridade dos inquiridos, sendo de 53%

⁹⁴ A lei de Galileu continua válida para as velocidades comumente observadas na terra, só sendo necessário utilizar a correcção introduzida pela teoria da relatividade restrita para velocidades próximas da velocidade da luz.

⁹⁵ <http://www.gallup.com/poll/114544/Darwin-Birthday-Believe-Evolution.aspx>

junto dos licenciados e 74% nos pós-graduados, valor mesmo assim abaixo da média da maioria dos países europeus.

Outro problema da abordagem qualitativa é que a percepção de qualidade também varia com os critérios utilizados, sendo difícil estabelecer critérios suficientemente abrangentes que sejam aplicáveis a todo o tipo de conteúdos.

Wang e Strong (1996) propõem uma classificação (Figura 18) dos atributos a que os utilizadores dão mais importância no que diz respeito à qualidade dos dados.



Figura 18: Qualidade dos dados, adaptado de Wang e Strong (1996)

Dos quatro grupos de atributos, as características englobadas nos grupos *Representacional* e *Acessibilidade* estão directamente dependentes do sistema que suporta a informação – o que, no caso em estudo, é similar em ambos os casos. O grupo *Contextual*, como o próprio nome indica, refere-se a características que variam com o consumidor da informação e o contexto em que esta será utilizada, que também é

constante⁹⁶ nos dois grupos em estudo. Assim, e apesar da importância destes atributos para a percepção da qualidade da informação por parte do utilizador, e tendo em conta que, ao avaliarmos objectivamente a qualidade intrínseca dos dados, estamos implicitamente a avaliar as outras categorias (Inês, 2010), iremos cingir-nos apenas às propriedades intrínsecas da informação.

Na Figura 19, encontram-se os atributos que Wang e Strong (1996) concluíram no seu estudo serem os mais importantes, ordenados por ordem decrescente de importância. A escala varia entre um e nove, sendo que a importância de um atributo é tanto maior quanto menor for a classificação que lhe foi atribuída.



Figura 19: Hierarquia da Qualidade de dados, adaptado de Wang e Strong (1996)

Se considerarmos apenas as características intrínsecas constatamos que, para os consumidores de informação, a credibilidade é o factor mais importante, seguido da exactidão e da objectividade⁹⁷. Com base nesta informação, revela-se essencial que, no estudo a realizar, as entradas das enciclopédias omitam a sua proveniência, de modo a evitar que os peritos não utilizem a credibilidade da fonte nas suas avaliações.

⁹⁶ Será indicado aos avaliadores para considerarem como destinatários dos artigos alunos ao nível de licenciatura.

⁹⁷ Note-se que estas, em termos globais, aparecem somente na quarta e oitava posições.

O estudo de Wang e Strong (1996) dá-nos pistas importantes na escolha do perfil dos peritos a convidar para avaliar os artigos. Quem se dedica ao ensino está habituado a avaliar conteúdos produzidos pelos alunos – pondo de lado a credibilidade da fonte – e concentrando-se apenas na exactidão e objectividade dos conteúdos, o que torna este grupo profissional ideal para realizar uma avaliação deste tipo.

Como as enciclopédias generalistas abarcam todos os ramos do conhecimento, será necessária a participação de professores das mais variadas áreas, de modo a garantir que são proficientes nos temas que os artigos abordam. Torna-se assim extremamente difícil encontrar um conjunto de critérios específicos que sejam transversais a todas as áreas, pelo que serão usados apenas a exactidão e a objectividade, apesar de estes serem um pouco generalistas. Em termos de exigência – e como a qualidade de uma resposta depende do grau de exigência da pergunta – esta será fixada ao nível da licenciatura, ou seja, cada docente irá avaliar as entradas de ambas as enciclopédias correspondentes ao mesmo tema, tema esse que está inserido na sua área de conhecimento, como se estivesse a avaliar uma pergunta respondida por dois alunos que frequentem o curso de licenciatura que lecciona.

Como será pedido a cada perito que avalie sempre pares de artigos, não se põe o problema de obtermos resultados não comparáveis, pois o contexto, grau de exigência e critérios são invariáveis dentro de cada par analisado. No entanto, irá ser pedido aos avaliadores que foquem a sua avaliação na exactidão e na objectividade da informação, categorias que, para Wang e Strong (1996), englobam correcção, exactidão, precisão, perfeição, confiança, integridade, objectividade e inexistência de erros, de modo a tentar uniformizar os critérios entre os avaliadores.

Os peritos avaliarão os artigos, utilizando uma escala de Likert de cinco níveis (Mau, Medíocre, Suficiente, Bom e Muito Bom), aos quais faremos corresponder os valores {1,2,3,4,5}. Como o número de alternativas de resposta condiciona a dimensão da amostra a estudar, como veremos mais à frente, e um maior número de opções iria complicar ainda mais a vida aos avaliadores, optámos por esta alternativa, que nos parece um bom compromisso e com a qual estamos todos familiarizados (Likert, 1932).

Este método permite agilizar a avaliação dos artigos e, embora torne menos confiável a média dos valores absolutos das classificações atribuídas a cada uma das enciclopédias,

permite ter total confiança na diferença das notas atribuídas a cada par, atenuando as diferenças introduzidas pelo facto de termos diferentes avaliadores.

Consideremos $(W, B)_j^i$, em que W e B são as notas atribuídas pelo avaliador i às entradas correspondentes ao tema j da Wikipédia e da Britânica. Consideremos que ambos os avaliadores são da mesma área científica e que cada um analisa um tema. Admitindo que os resultados da avaliação são $(5,4)_1^1$ e $(4,3)_2^2$, podemos sempre admitir que esta variação não reflecte a diferença de qualidade dos pares de artigos, mas sim um diferente grau de exigência dos avaliadores – ou seja, poderíamos ter obtido os resultados $(5,4)_1^1$ e $(5,4)_2^1$ ou $(4,3)_1^2$ e $(4,3)_2^2$, ou mesmo $(4,3)_1^2$ e $(5,4)_2^1$, consoante os pares de artigos fossem ambos avaliados pelo primeiro avaliador, ambos pelo segundo ou se se tivesse trocado de pares de artigos. Neste exemplo, ao contabilizar o valor médio das notas atribuídas a cada enciclopédia, chegaríamos aos valores (5,4), $(4\frac{1}{2}, 3\frac{1}{2})$ e (4,3) apenas devido ao maior grau de exigência do primeiro avaliador e à distribuição de pares de artigos escolhida.

No entanto, se observarmos as diferenças entre as notas atribuídas dentro de cada par, elas são invariáveis, pois dentro de um par os critérios de avaliação são sempre os mesmos. Assim, e expostas as eventuais vulnerabilidades da utilização dos valores absolutos, vamos transformar, nas análises a efectuar, as duas notas atribuídas a cada tema na sua diferença, sendo esta medida muito menos sensível às variações de critérios entre avaliadores, impossíveis de controlar.

$$R_{n \times 2} = \begin{bmatrix} W_1 & B_1 \\ W_2 & B_2 \\ W_3 & B_3 \\ \vdots & \\ W_n & B_n \end{bmatrix} \xrightarrow{\text{Diferença}} \begin{bmatrix} W_1 - B_1 \\ W_2 - B_2 \\ W_3 - B_3 \\ \vdots \\ W_n - B_n \end{bmatrix}$$

O número de combinações possíveis de notas é dado por:

$${}^n A'_p = n^p; \forall n, p \in \mathbb{N} \xrightarrow{n=5 \wedge p=2} {}^5 A'_2 = 5^2 = 25$$

Podendo estas ser calculadas, subtraindo a matriz A da sua transposta:

$$A = \begin{bmatrix} 5 & 5 & 5 & 5 & 5 \\ 4 & 4 & 4 & 4 & 4 \\ 3 & 3 & 3 & 3 & 3 \\ 2 & 2 & 3 & 2 & 2 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix}, A^T = \begin{bmatrix} 5 & 4 & 3 & 2 & 1 \\ 5 & 4 & 3 & 2 & 1 \\ 5 & 4 & 3 & 2 & 1 \\ 5 & 4 & 3 & 2 & 1 \\ 5 & 4 & 3 & 2 & 1 \end{bmatrix} \xrightarrow{\text{Então}} A - A^T = \begin{bmatrix} 0 & 1 & 2 & 3 & 4 \\ -1 & 0 & 1 & 2 & 3 \\ -2 & -1 & 0 & 1 & 2 \\ -3 & -2 & -1 & 0 & 1 \\ -4 & -3 & -2 & -1 & 0 \end{bmatrix}$$

No caso de as avaliações serem feitas aleatoriamente, a probabilidade de obter cada um dos resultados possíveis é dado por:

$$P(4) = \frac{1}{25}; P(3) = \frac{2}{25}; P(2) = \frac{3}{25}; P(1) = \frac{4}{25}; P(5) = \frac{5}{25} \text{ e } P(-n) = P(n)$$

Estes resultados seguem aproximadamente uma distribuição normal de $\mu=0$ e $\delta=2$ (Figura 20)

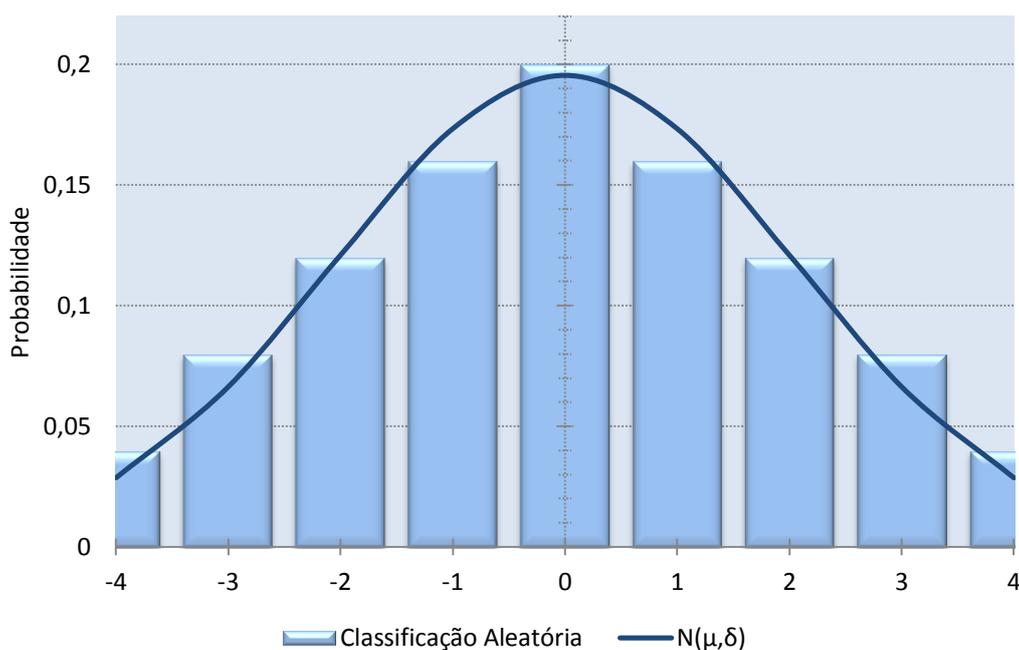


Figura 20: Resultados da Classificação Aleatória

Ou seja, imaginando um cenário de falta de qualidade nas avaliações, no limite este favorece a hipótese de não haver diferença significativa entre as enciclopédias.

Caso tivesse sido pedido de início aos avaliadores para classificar os pares de artigos nesta escala de 9 níveis diferentes, a complexidade do seu trabalho teria aumentado desnecessariamente, não seria obtido um efeito tão acentuado na redução do problema

da variabilidade de critérios e não se teria uma estimativa, embora frágil, do valor absoluto da qualidade de cada enciclopédia.

Os valores {-4,-3,-2,-1,0,1,2,3,4} podem ser substituídos de acordo com a Tabela 5, de modo a facilitar a percepção do seu significado, enquanto resposta à pergunta:

A qualidade do artigo x proveniente da Wikipédia, quando comparado com o seu homónimo da Britânica, é:

0 - Equivalente			
1	Melhor	-1	Pior
2	Muito Melhor	-2	Muito Pior
3	Muitíssimo Melhor	-3	Muitíssimo Pior
4	Incomparavelmente Melhor	-4	Incomparavelmente Pior

Tabela 5: Significado das Diferenças

Os resultados da avaliação a efectuar aos conteúdos das Enciclopédias serão influenciados pela percepção da “verdade” por parte dos avaliadores à luz do seu conhecimento no momento em que esta é efectuada e será baseada num *justified true belief*.

Whoever undertakes to set himself up as a judge of Truth and Knowledge is shipwrecked by the laughter of the gods.

Albert Einstein (1954), *Essays presented to Leo Baeck*

3.5 JÚRIS VERSUS JURADOS

No capítulo anterior, concluímos sobre a importância de os artigos serem avaliados aos pares e a vantagem de os peritos estarem ligados à docência devido à sua experiência neste tipo de tarefa.

No entanto, a avaliação poderia ser feita recorrendo à *Vox Populi*, utilizando a votação de leigos, em vez de recorrer à *Vox Expertorum*, pedindo a colaboração de peritos na área do artigo.

Se nos debruçarmos sobre a história do Direito, encontramos inúmeros exemplos em que decisões tão importantes como a culpabilidade de um indivíduo são tomadas por leigos. No séc. V a.C. em Atenas, o berço da democracia, os julgamentos eram decididos por painéis de cidadãos, conhecidos por *dikastai* que, consoante a importância dos casos, podiam chegar aos 500 indivíduos, e que decidiam por voto secreto em maioria qual dos litigantes tinha razão e a pena a aplicar (Cohen, 2005).

Em 1215, o rei de Inglaterra é forçado a assinar um documento conhecido como a Magna Carta, que limita os seus poderes e que, no artigo 39, afirma:

NO Freeman shall be taken or imprisoned, or be disseized of his Freehold, or Liberties, or free Customs, or be outlawed, or exiled, or any other wise destroyed; nor will We not pass upon him, nor condemn him, but by lawful judgment of his Peers, or by the Law of the land. We will sell to no man, we will not deny or defer to any man either Justice or Right. (Howard 1998:14)

Assim, é reconhecido o direito ao julgamento pelos seus pares (Howard, 1998). Nos dias de hoje, o conceito de jurados – julgamento pelos pares – continua a ser utilizado em todo o mundo civilizado, sendo em alguns países obrigatório, enquanto em outros constitui uma possibilidade que se aplica apenas a determinado tipo de crimes. O número de jurados também varia bastante, sendo no entanto mais comum serem doze, no caso em que as decisões são tomadas por unanimidade, e em número ímpar quando a decisão é tomada por votação (Jólluskin, 2009).

Na Tabela 6, foi feita uma pequena compilação dos tipos de jurados que existem, acompanhada do âmbito do seu poder decisório, bem como exemplos de alguns países em que estes modelos estão em funcionamento (Winter & Greene, 2007):

JURADOS	PAÍSES	DECISÃO	FUNCIONAMENTO
Anglo-saxónico ou Puro	Reino Unido, Estados Unidos, Espanha e Austrália	Unanimidade	Os Jurados (leigos) decidem quanto à culpabilidade e o Juiz quanto à pena a aplicar.
Escabinado	Portugal, França, Itália e Alemanha.	Majoria	Colégio de leigos e Juízes, decidem entre si a culpabilidade e a pena a aplicar.
Misto	Bélgica e Áustria.	Majoria	O Juiz decide quanto à culpabilidade, no caso de esta se verificar, e a pena é determinada por um escabinado de Juízes e Jurados (leigos)

Tabela 6: Tribunais de Júri no Mundo

Segundo Winter e Greene (2007), a possibilidade, ou a obrigatoriedade, de a decisão ser tomada por jurados aumenta com a importância do julgamento.

Em contraponto a esta situação, a decisão de um tribunal pode ser tomada por um juiz, ou seja, um perito em leis que usa a lei e a sua opinião pessoal para decidir quanto à culpabilidade e a pena a aplicar. Em Portugal esta é a regra, embora em certo tipo de crimes possa ser pedido o julgamento por um júri escabinado (Jólluskin, 2009).

Os tribunais em que a decisão é produzida por juízes podem ser unipessoais⁹⁸ – em que apenas um juiz é responsável pela decisão – ou colectivos em que participam dois ou mais juízes (Jólluskin, 2009). Como se pode ver, em termos de direito, as decisões de culpabilidade e da pena a aplicar variam bastante, sendo a lei (critérios) aplicada por Juízes (peritos), Jurados (leigos) ou combinações destas.

Fazendo o paralelismo com a avaliação da qualidade dos artigos, também estes podem ser “julgados” por qualquer uma das opções atrás apresentadas, inclusive pelos *dikastai* (multidão de amadores); no entanto, à semelhança dos estudos afins atrás inventariados,

⁹⁸ Julgado, Juízo ou Vara.

vamos optar por utilizar peritos (Juízes), por ser, à partida, a alternativa que conduz a resultados mais facilmente aceites pela academia.

Resta escolher se os avaliadores dos pares de artigos serão unipessoais ou colectivos. Parece óbvio que, quanto maior for o número de avaliadores para cada par de artigos, maior é a probabilidade e a confiança que teremos nos resultados obtidos.

No entanto, se tivermos em conta que o número (n) máximo de avaliações exequíveis de ser realizadas é uma constante e a dimensão da amostra que será necessário analisar para que esta seja representativa do universo que queremos estudar, temos que n é calculado multiplicando o número de avaliadores por artigo dimensão da amostra (Figura 21).

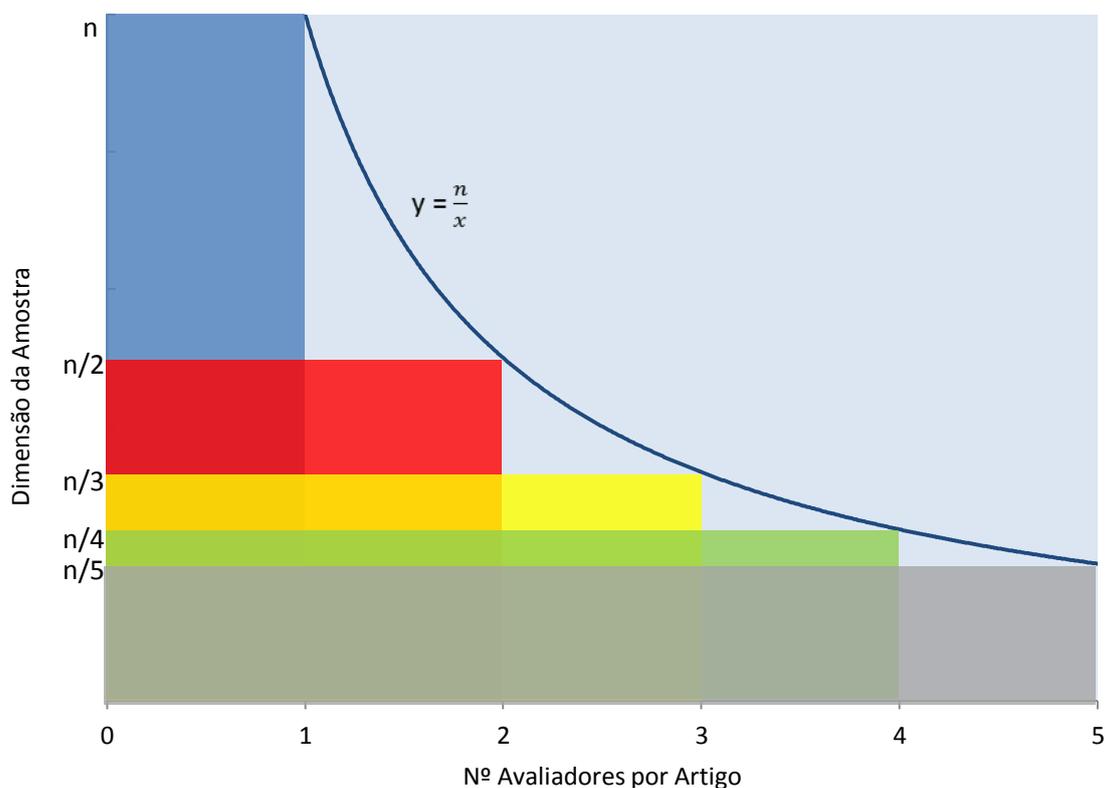


Figura 21: Variação da Dimensão da Amostra com o Número de Avaliadores

Na Figura 21, podemos ver como a área dos rectângulos é igual – consequência de termos imposto um limite n ao número de artigos que é possível avaliar – e a evolução do tamanho da amostra com o aumento do número de avaliadores por artigo. Assim, o número de avaliadores por artigo a escolher ficará condicionado à dimensão da amostra a estudar.

I have a great subject (statistics) to write upon, but feel keenly my literary incapacity to make it easily intelligible without sacrificing accuracy and thoroughness.

Sir Francis Galton (1889), *Natural Inheritance*

4. METODOLOGIA

De modo a comparar a qualidade dos conteúdos das duas enciclopédias e utilizar o resultado para aferir da viabilidade da colaboração em massa enquanto modelo de negócio em situações similares, iremos neste capítulo expor a metodologia utilizada, começando pela recolha dos dados (4.1) e a avaliação destes (4.2). São de seguida apresentados dois estudos complementares que pretendem caracterizar a opinião preconcebida dos avaliadores (4.3) e inferir a variabilidade das classificações de múltiplos avaliadores à mesma questão (4.4). Por fim, são elencadas as técnicas e ferramentas a utilizar no tratamento dos dados (4.5) e o posicionamento desta tese em termos de filosofia da investigação (4.6).

4.1 RECOLHA DOS DADOS

Pretende-se, pois, realizar uma amostragem que possibilite efectuar um estudo comparativo entre a Enciclopédia Britânica e a Wikipédia que permita concluir se existe uma diferença significativa na qualidade da informação nestas contida.

A amostra a utilizar deverá ser recolhida aleatoriamente entre todos os artigos e não apenas sobre um tema em particular. Devido à própria natureza da Wikipédia, que muda constantemente, há que ter em mente que os resultados a obter apenas concluem sobre a situação à data da recolha da amostra.

A escolha da metodologia de amostragem a adoptar baseou-se nas críticas mais comuns aos estudos sobre a qualidade dos conteúdos da Wikipédia, feitas aos trabalhos elencados no capítulo anterior. Assim, deverão ter-se em conta a dimensão e a representatividade da amostra em relação ao universo a estudar, de modo a evitar repetir esses erros.

A amostra a utilizar deverá ser estatisticamente representativa dos artigos que compõem a Wikipédia, definindo-se à partida uma margem de erro e um intervalo de confiança que conduzam a um compromisso aceitável entre a precisão do estudo e a dimensão da amostra a tratar.

Tendo em conta os mais de três milhões de artigos que compõem a Wikipédia, Ruane (2005) aconselha uma amostra de cerca de 750⁹⁹ artigos, enquanto Gaur e Gaur (2009) consideram que uma amostra excelente deverá conter 500 artigos, embora uma amostra de 200-300 seja considerada adequada. Ambos os autores usam a *rule of thumb* para chegar a estes valores.

Adams, Raeside, e Khan (2007), Given (2008), Urdan (2005) e Vanderstoep e Johnston (2009) propõem a utilização da seguinte fórmula:

$$n_0 = Z_{\alpha/2}^2 \frac{p(1-p)}{d^2}$$

n_0 = Dimensão da amostra, Z = Valor standard da normal,
 α = Nível de significância, p = Proporção da população e d = margem de erro.

Caso a dimensão da amostra obtida fosse superior a 5% do universo, poderíamos diminuir a dimensão desta. No caso em análise, tal não se verifica.

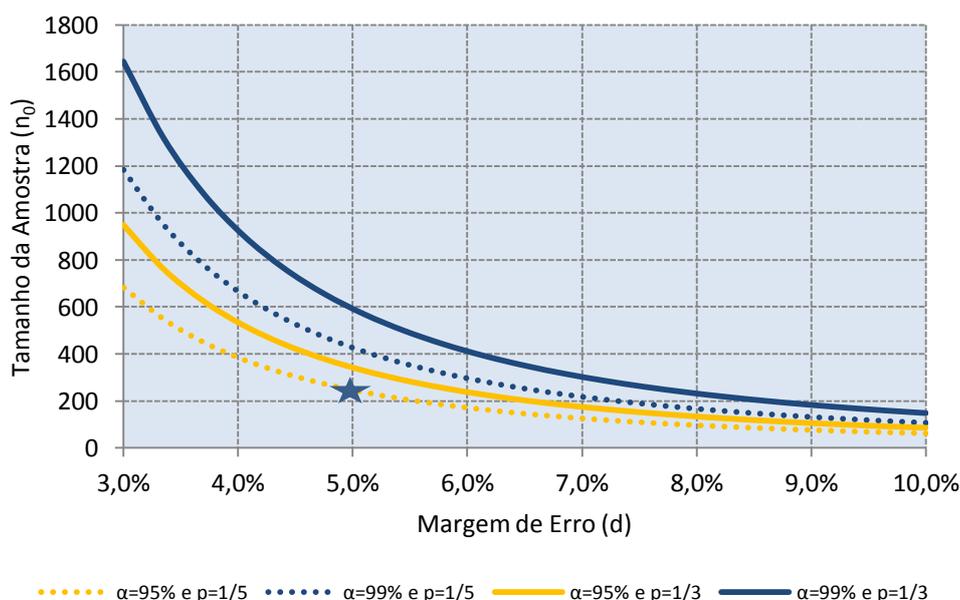


Figura 22: Dimensão da amostra

⁹⁹ 0,025% por cada milhão de elementos do universo a amostrar.

Na Figura 22, podemos ver a variação da dimensão da amostra com a margem de erro para os níveis de significância $\alpha=95\%$ e $\alpha=99\%$, a que correspondem os valores de $Z_{\alpha/2}^2$ de 1,96 e 2,58 respectivamente, e para os valores de p de 1/3 para três alternativas¹⁰⁰ de resposta no que diz respeito à classificação dos artigos e de 1/5 para cinco.

Tendo em conta o comumente aceitável neste tipo de estudos (Gaur & Gaur, 2009), em termos de nível de significância e margem de erro, as dificuldades em escolher entre muitas alternativas de classificação e as implicações que a dimensão da amostra tem na exequibilidade deste trabalho, optámos por utilizar um nível de significância de 95%, uma margem de erro de 5%, e cinco alternativas de classificação (Muito Bom, Bom, Suficiente, Medíocre e Mau). Nestas condições ($\alpha = 95\%$, $Z = 1,96$, $p = 1/5$ e $d = 5\%$), a dimensão da amostra a estudar (n_0) será de:

$$1,96^2 \frac{\frac{1}{5} \left(1 - \frac{1}{5}\right)}{0,05^2} = 245$$

Nas primeiras semanas de 2010, iniciámos o processo de extracção aleatória dos 245 pares de artigos das versões *online* de ambas as enciclopédias. Foram recolhidos os títulos, a categoria¹⁰¹, o *URL* e o conteúdo das entradas, tendo sido deixadas de fora as secções *Related Articles* e *External Web sites* na Britânica e os seus correspondentes na Wikipédia, *Articles Related*, *See Also* e *External Links*. Devido à forma como funciona o hipertexto, caso não o fizéssemos corríamos o risco de, a partir de um artigo e explorando todas as ligações que este contém e, de seguida, as ligações das ligações, obter todo o conteúdo da Enciclopédia partindo apenas de um artigo.

Como a Wikipédia contém vinte e cinco vezes mais entradas do que a Britânica, teria sido mais fácil extrair aleatoriamente o artigo da Britânica, procurando seguidamente a entrada correspondente na Wikipédia. Contudo, dado que a Britânica não permite a selecção aleatória de artigos, tivemos de usar o procedimento oposto, que conduziu à

¹⁰⁰ Vanderstoep, S. & Johnston, D. (2009), p.31.

¹⁰¹ A categoria foi apenas recolhida na Britânica, visto ser esta a única a disponibilizar este tipo de informação.

extração de 6.382 entradas da Wikipédia, até serem encontradas 245 entradas¹⁰² que estivessem simultaneamente presentes nas duas enciclopédias. Só foram seleccionados os pares que correspondiam a entradas principais da Britânica, isto porque algumas vezes o tema era referido na Britânica, mas apenas superficialmente e no contexto de outro tema.

Na Figura 23, podemos ver um fluxograma que sintetiza os procedimentos efectuados na obtenção da amostra e na posterior preparação para análise.

¹⁰² Na altura da amostragem, os dados oficiais indicavam que a Wikipédia continha 25 vezes mais artigos do que a Britânica; se dividirmos 6.382 por 245, obtemos aproximadamente uma relação de 24:1 (erro de 4%)

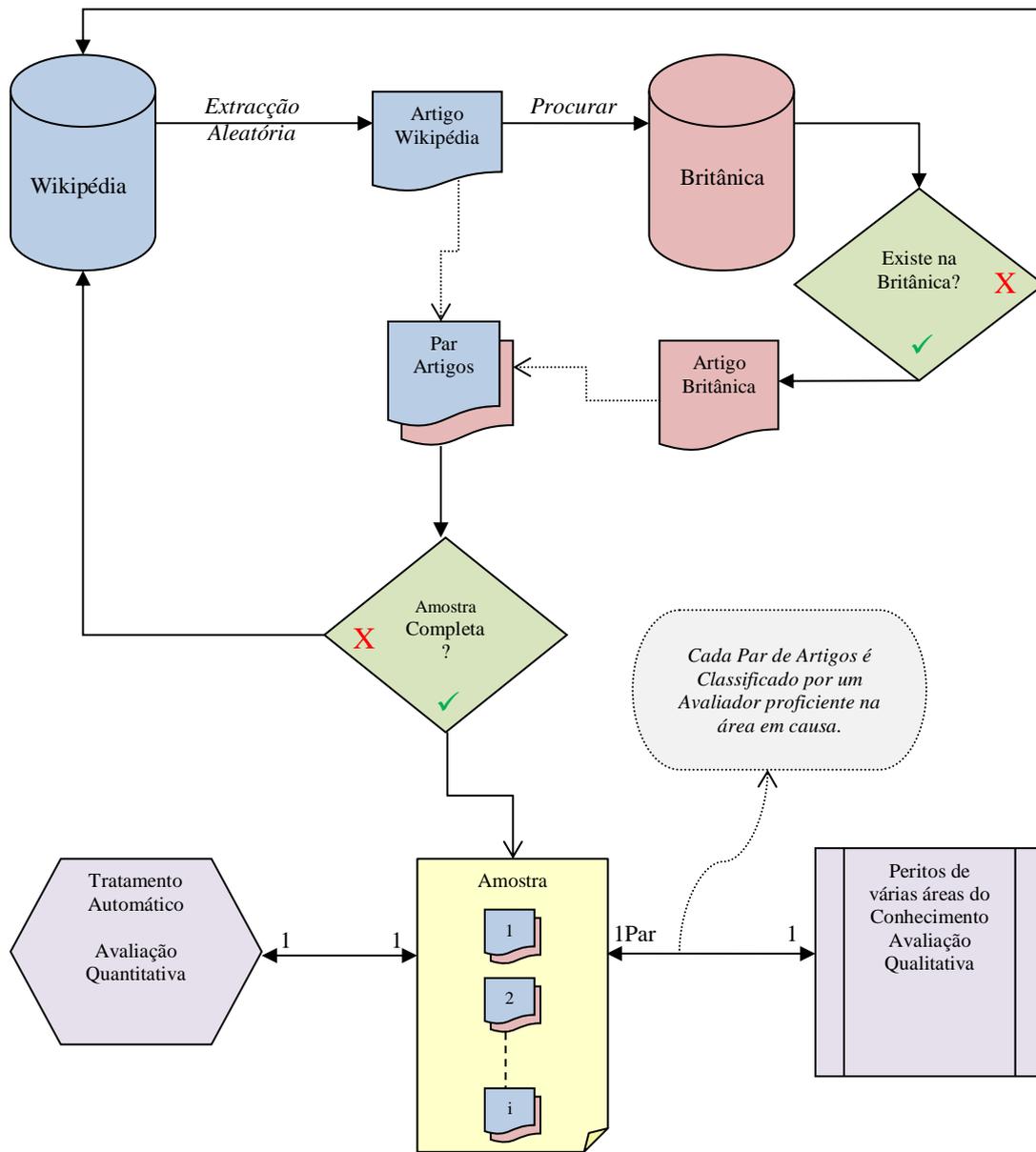


Figura 23: Amostragem 245 Pares Artigos

4.2 AVALIAÇÃO DOS DADOS

A Enciclopédia Britânica classifica os seus artigos em quatro grandes categorias, a saber: Artes & Entretenimento, Ciência & Tecnologia, História & Sociedade e Viagens & Geografia. Estas categorias foram utilizadas para etiquetar os pares de artigos de modo a permitir mais tarde analisar os dados em grupos e verificar se existem conclusões diferentes a tirar consoante a área do conhecimento analisada.

Na Figura 24, podemos ver a evolução da distribuição dos pares de artigos pelas categorias atrás mencionadas à medida que, cronologicamente, foram sendo integrados na amostra.

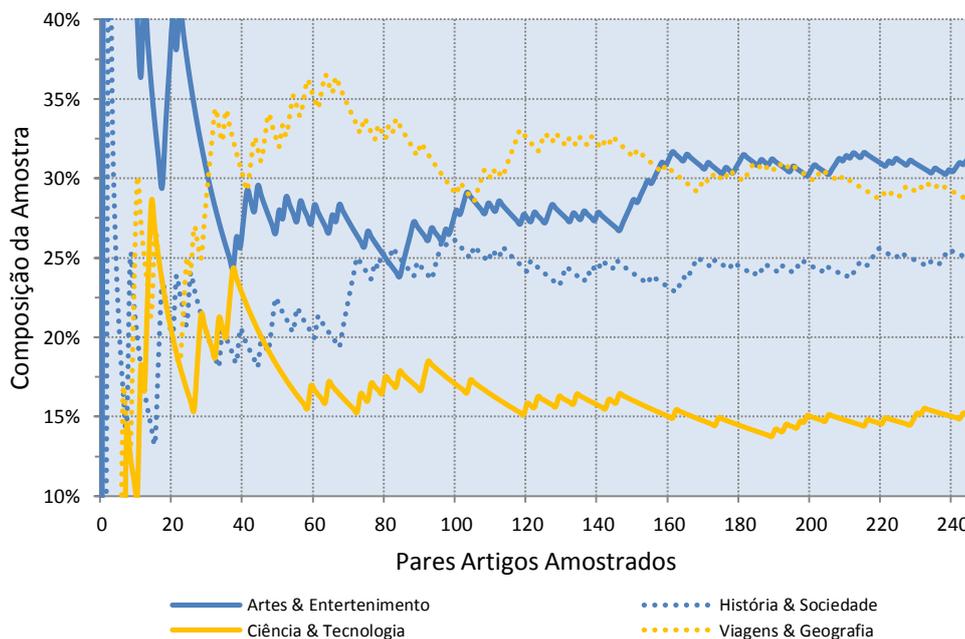


Figura 24: Evolução da Distribuição dos Artigos por Temas

Segundo Halavais e Lackaff (2008), a cobertura dos temas na Wikipédia, ao contrário das enciclopédias tradicionais, é directamente proporcional ao interesse que os colaboradores desta têm numa determinada temática, pelo que é expectável que a qualidade e abrangência dos artigos varie bastante entre temas.

Com base nesta informação e observando a grande variabilidade inicial na composição da amostra, em termos da proporção dos temas que a compõem, podemos constatar os perigos que uma amostra de pequenas dimensões representa num estudo deste tipo, reforçando assim a importância do método de amostragem escolhido e a sua contribuição enquanto factor diferenciador entre este estudo e os estudos afins referidos no subcapítulo 3.2.

Como a distribuição dos pares amostrados corresponde à distribuição dos artigos por essas categorias na Britânica e, durante a extracção aleatória dos artigos da Wikipédia, estes foram sendo divididos¹⁰³ por esses mesmos grupos, podemos nesta fase ter já uma primeira classificação da amostra. (Figura 25)

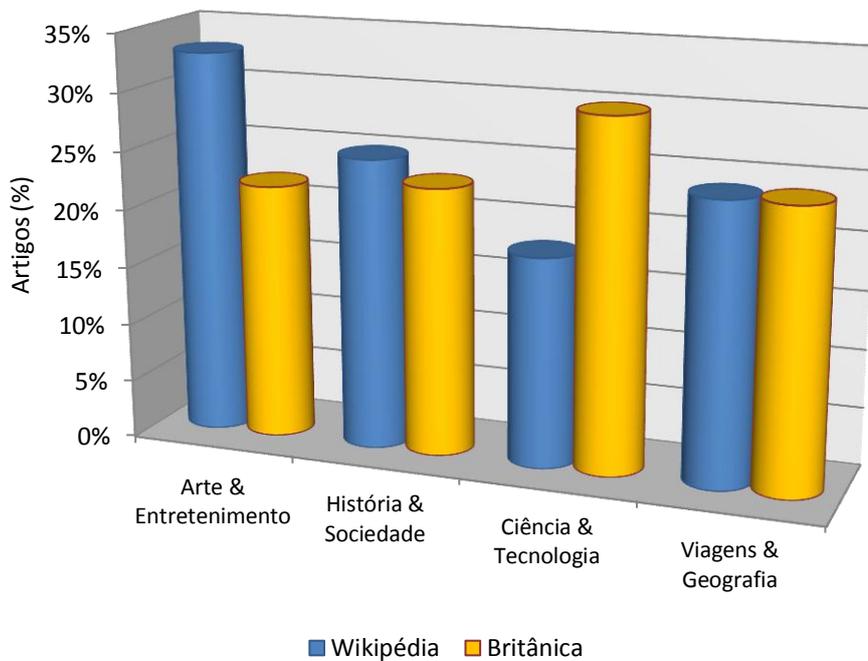


Figura 25: Distribuição percentual dos artigos por temas

Nesta primeira classificação, salta à vista que, percentualmente, a Britânica possui bastante mais artigos sobre *Ciência & Tecnologia* e menos sobre *Arte & Entretenimento*, quando comparada com a Wikipédia; convém, no entanto, não esquecer

¹⁰³ Esta classificação foi efectuada pelo investigador utilizando apenas critérios do bom senso, pelo que poderá não ser totalmente correcta, se bem que é, de certeza, uma boa aproximação.

os conselhos¹⁰⁴ de Huff (1954): de modo a não tornar esta tese em mais um exemplo prático do seu livro *How to Lie with Statistics*, apresenta-se de seguida o mesmo gráfico (Figura 26) – mas com as abcissas em valores absolutos – que mostra como é irrelevante a análise efectuada sobre o gráfico anterior (Figura 25).

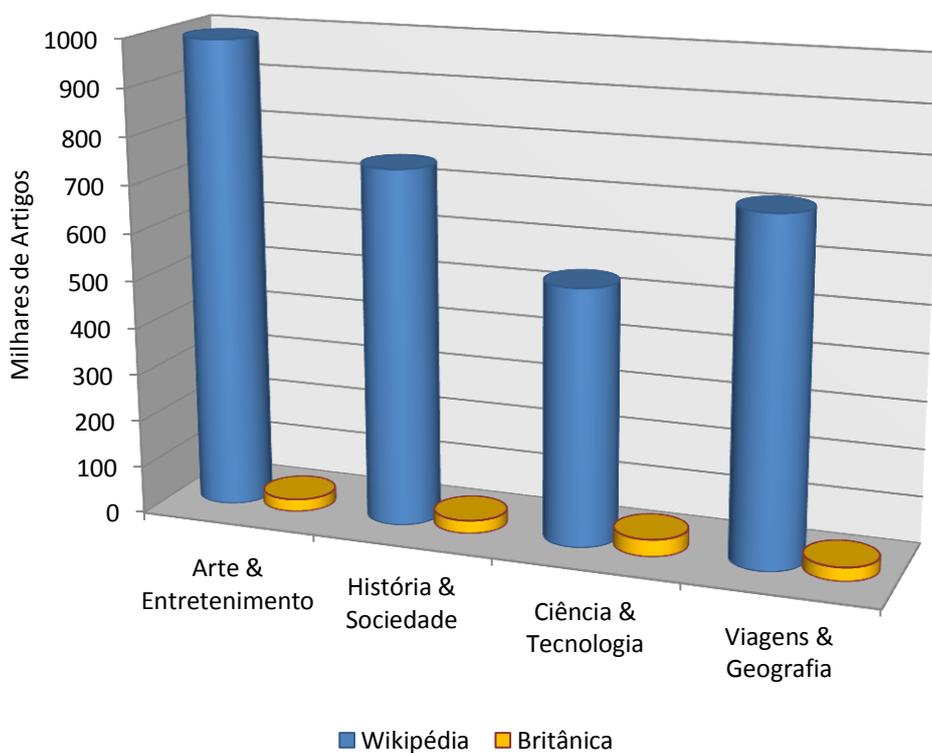


Figura 26: Distribuição em valor absoluto do número artigos por tema

Agora que conhecemos o número de artigos que temos de avaliar (245 pares, ou seja 490 artigos), foi calculado o número médio de palavras por artigo da amostra, que é cerca de 2000, pelo que obtemos um total de quase um milhão de palavras, considerando um número médio de 250 palavras por página – ou seja, estamos a falar de cerca de 4000 páginas.

De acordo com o exposto no capítulo anterior, em que teria de ser admitido um número máximo para o número de artigos susceptíveis de serem avaliados em tempo útil, não nos parece exequível utilizar mais do que um avaliador por artigo, pois realizar 490 avaliações é já uma tarefa extremamente difícil de levar a bom porto e a alternativa, que

¹⁰⁴ *Misinforming people by the use of statistical material might be called statistical manipulation; in a word (though not a very good one), statistication.* (Huff, 1954:100)

é trabalhar com uma amostra muito menor e pouco representativa do universo a estudar, não é uma opção.

Para simplificar a distribuição dos artigos pelos avaliadores, os pares de artigos foram ainda subdivididos em categorias mais precisas e menos abrangentes. Na Figura 27, podemos visualizar uma representação dessa divisão, em que a área das palavras é proporcional à ocorrência dessa categoria na amostra:



Figura 27: Distribuição dos Artigos por áreas de Conhecimento

Os artigos foram agrupados aos pares e descaracterizados, tendo-lhes sido removidas todas as hiperligações e formatação de origem de modo a ocultar o mais possível a sua proveniência e introduzida aleatoriamente uma marca de água, com as letras A e B, de modo a ser possível identificar mais tarde a sua proveniência.

Os avaliadores foram escolhidos preferencialmente entre os docentes das Universidades de Évora e de Lisboa, sendo o pedido de colaboração geralmente endereçado ao responsável do departamento, distribuindo este seguidamente pelos seus colegas os pares de artigos da sua área, de acordo com as especializações destes.

Alguns dos temas¹⁰⁵, devido à sua especificidade, foram avaliados por não docentes, sendo no entanto estas pessoas com incontestável conhecimento da área em que serviram de avaliadores. Na Figura 28, podemos ver a distribuição dos avaliadores de acordo com a sua afiliação:

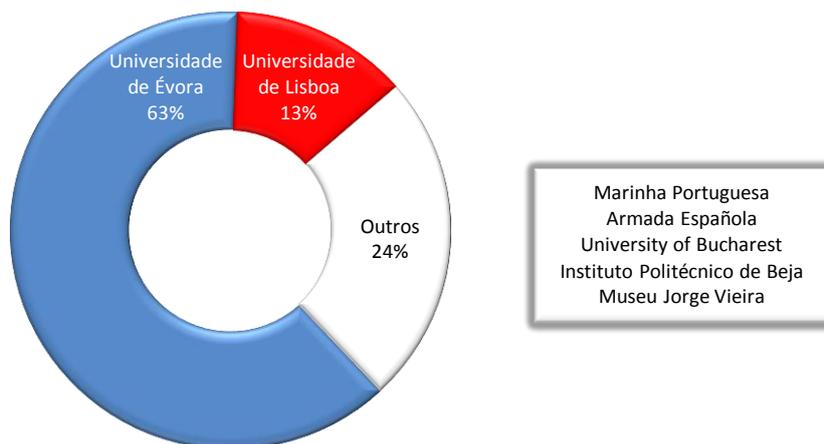


Figura 28: Distribuição das Afiliações dos Avaliadores

Encontram-se coligidos no Anexo II os resultados das avaliações, bem como os nomes dos avaliadores que participaram no estudo.

¹⁰⁵ Militar, Naval, Religião e Pintura.

4.3 HÁBITOS VERSUS MONGES

Para melhor enquadrar o estudo a realizar, foi elaborado um pequeno inquérito de modo a caracterizar os hábitos de consulta de informação da população académica e a confiança que deposita na informação obtida, através de instrumentos similares à Wikipédia – *Vox Populi* – e à Britânica – *Vox Expertorum*. (Figura 29)

The image shows a survey interface with a dark blue header containing the title "Caracterização Hábitos de Consulta Informação". Below the header, there are three questions:

- 1. Quando necessita obter informação/aprofundar um determinado tema, onde procura primeiro informação?**
This is a dropdown menu with two options: "Motor de Busca / Wikipedia" (selected with a blue checkmark) and "Enciclopédia / Dicionário / Obra de Referência".
- 2. Confia na Informação obtida através de um Motor de Busca / Wikipedia?**
This is a 5-point Likert scale. The options are 1, 2, 3, 4, and 5. Option 3 is highlighted in blue. Below the scale, "Desconfio completamente" is on the left and "Confio plenamente" is on the right.
- 3. Confia na Informação obtida através de uma Enciclopédia / Dicionário / Obra de Referência?**
This is a 5-point Likert scale. The options are 1, 2, 3, 4, and 5. Option 4 is highlighted in blue. Below the scale, "Desconfio completamente" is on the left and "Confio plenamente" is on the right.

Figura 29: Inquérito

O inquérito foi deliberadamente construído com um número mínimo de questões, sem perguntas que fizessem a caracterização dos inquiridos, de modo a tentar maximizar a taxa de resposta, que é habitualmente muito baixa nestes casos. A diferenciação entre os inquiridos foi efectuada administrando o mesmo inquérito a grupos diferentes, de modo a analisar uma eventual diferença de respostas entre eles.

¹⁰⁶ Esta citação é comumente atribuída a Einstein e a Mark Twain em vários livros, mas não foi possível verificar a fonte.

A primeira pergunta visa saber, de facto, a que tipo de ferramenta os inquiridos recorrem instintivamente quando necessitam de obter informação. Embora esta pergunta diferencie o método como a informação que utilizam foi coligida e por quem, poderá ser por muitos respondida em termos do suporte que utilizam, ou seja, a Internet ou o papel para a obter.

Analisando a pergunta, a primeira opção de resposta (Motor de Busca / Wikipédia) tem subjacente o facto de os conteúdos obtidos por estes métodos não garantirem que as páginas *Web* / artigos da Wikipédia tenham sido criados por peritos cuja credibilidade não possa ser posta em causa. Em relação à segunda opção (Enciclopédia / Dicionário / Obra de Referência), as respostas destas fontes são expectavelmente da responsabilidade de peritos no assunto procurado.

Obviamente que podemos consultar enciclopédias, dicionários e obras de referência *online* mas, como na maior parte dos casos esses acessos não são gratuitos, a grande maioria dos respondentes não terá acesso a estas, pelo que é provável estar a utilizar informações potencialmente oriundas de fontes sobre as quais não podemos garantir terem sido disponibilizadas por um perito na matéria sempre que usa os resultados de um motor de busca.

Se a pergunta formulada fosse “Confia em informação que não pode garantir ter sido criada por um perito na matéria?” iríamos obter muito poucas, ou nenhuma, respostas positivas, enquanto com a pergunta utilizada – sem chamar a atenção sobre a ausência de garantia da credibilidade da fonte – as repostas obtidas reflectem mais genuinamente a opinião que os inquiridos têm sobre a informação que obtêm através de um motor de busca ou da Wikipédia.

Uma explicação exaustiva destes conceitos sobre os quais muitos dos inquiridos nunca se debruçaram iria diminuir a taxa de resposta, pelo que optámos por manter a ideia inicial de simplicidade subjacente ao inquérito, para tentar maximizar a taxa de resposta.

As perguntas 2 e 3 pretendem quantificar a confiança dos inquiridos nos dois tipos de métodos de obter informação e, à semelhança das avaliações dos peritos, também deverão ser analisadas aos pares, pois a resposta de uma influencia a da outra e vice-versa. Assim, a escala escolhida – 5 níveis – é idêntica à do estudo anterior e a

transformação das duas respostas num único valor, dado pela sua diferença entre elas, segue os mesmos critérios elencados anteriormente.

O inquérito foi efectuado junto de quatro grupos diferentes:

- Docentes Universitários
- Alunos Universitários
- Professores do Ensino Secundário
- Alunos do Ensino Secundário

A característica diferenciadora entre os grupos, aquela em que é expectável encontrar diferenças nas respostas, é o número médio de anos de escolaridade. Admite-se, para tal, que a maioria dos docentes do ensino secundário possui uma licenciatura e os docentes universitários um doutoramento, ou seja, admite-se que o número médio de anos passados em formação é significativamente diferente entre os Professores do Secundário e os Universitários, sendo óbvia a diferença entre os grupos de discentes e entre estes e os docentes.

Por uma questão de exequibilidade, o inquérito foi conduzido apenas dentro de um *cluster* dentro de cada grupo apresentado, pelo que os resultados não podem ser alargados a nível nacional. No entanto, é expectável que outros *clusters* dentro do mesmo grupo tivessem resultados semelhantes pois, à partida, não há nenhuma variável diferenciadora dos *clusters* pertencentes ao mesmo que grupo que seja candidata óbvia a uma potencial diferenciação entre estes.

Na prática, e no caso do grupo dos docentes universitários, tomando como exemplo três *clusters* – as Universidades do Minho, Lisboa e Évora – não há, à partida, nenhuma diferença entre estas que obviamente potencie uma diferença significativa das respostas obtidas.

Como a resposta aos inquéritos é voluntária, no caso de baixas taxas de resposta poderemos também ter o problema de o conjunto de respostas obtidas dentro de um *cluster* vir de um subgrupo deste, significativamente diferente do grupo.

Os quatro *clusters* escolhidos para representar os grupos apresentados foram os docentes e discentes da Universidade de Évora e os docentes e discentes da Escola

Secundária Diogo de Gouveia em Beja¹⁰⁷. Segundo o Relatório de Actividades de 2010 da Universidade de Évora¹⁰⁸, o universo a inquirir é de 655 docentes e 9510 discentes; indicando o Relatório de Avaliação Externa¹⁰⁹ da Inspeção-Geral da Educação referente à Escola Secundária Diogo de Gouveia – Beja, os números cifram-se em 107 docentes e 697 discentes. Os resultados deste inquérito poderão ajudar a corroborar a hipótese, posta com base nas conclusões do estudo de Wang e Strong (1996), de que a credibilidade era a característica intrínseca mais importante na percepção da qualidade da informação, o que levou a descaracterizar os artigos a avaliar das enciclopédias, permitindo também a comparação entre os hábitos de consulta de informação, a confiança que, à partida, depositam na informação produzida pela *Vox Populi* e os resultados das avaliações, por parte dos docentes da Universidade de Évora.

Os resultados obtidos e os pormenores de como foram postos em prática os inquéritos encontram-se no Anexo III.

¹⁰⁷ O inquérito foi conduzido apenas junto dos professores e alunos do ensino secundário (10º, 11º e 12º anos), pois considerou-se que os alunos do ensino básico ainda não teriam maturidade suficiente para participar neste estudo.

¹⁰⁸ Páginas 75 e 92.

¹⁰⁹ Fonte: www.ige.min-edu.pt_upload_AEE_2011_DRA_AEE_11_ES3_Diogo_Gouveia_R, página 3. O Relatório é de Fevereiro de 2011 e analisa o ano lectivo de 2010/11.

*The best we can do is size up the chances,
calculate the risks involved,
estimate our ability to deal with them,
and then make our plans with confidence.*

Henry Ford (1863 - 1947)

4.4 PRECISO VERSUS EXACTO

Para obviar o problema de a dimensão da amostra nos impedir de ter mais do que um avaliador por artigo, idealizou-se um estudo auxiliar que ajudasse a perceber qual a ordem de grandeza da variação nos resultados que a introdução de múltiplos avaliadores por artigo iria introduzir.

Por outras palavras, e como foi explicado anteriormente, como não conhecemos – e não existe – uma “verdade absoluta” com a qual comparar os *justified true beliefs* dos avaliadores, i.e. não sabemos onde está o centro do alvo (Figura 30), não podemos tirar conclusões sobre a exactidão das respostas; podemos, no entanto, tentar aferir a precisão destas.

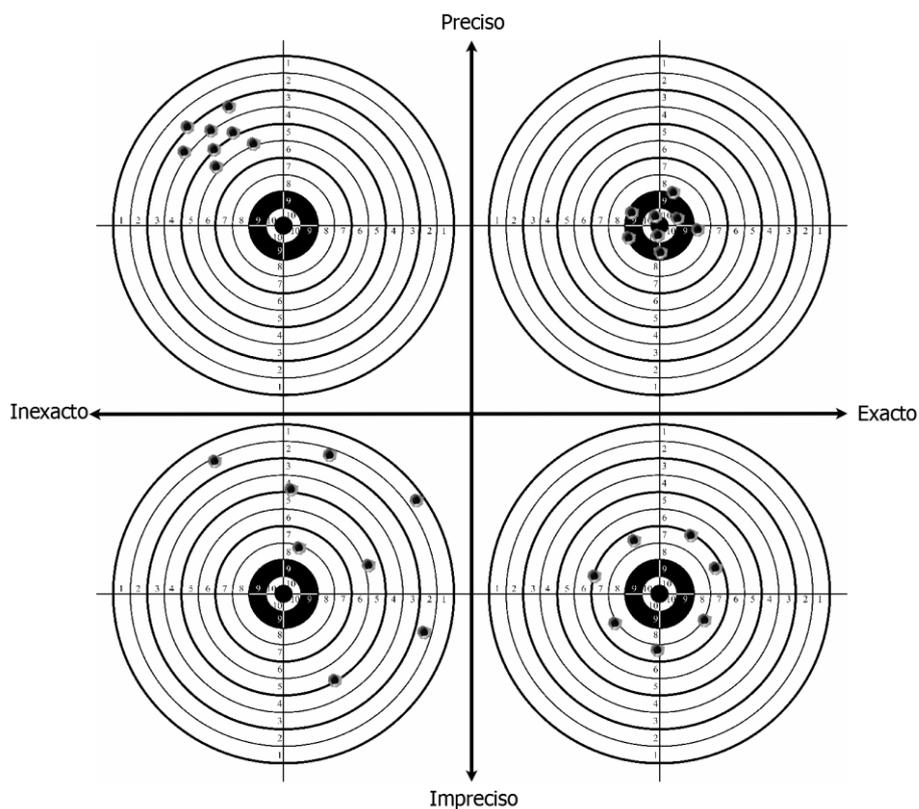


Figura 30: Preciso versus Exacto

A exactidão está relacionada com a média, i.e. se o centro de massa dos impactos está sobre o centro do alvo, enquanto a precisão pode ser conotada com o desvio padrão, ou seja, o valor médio das distâncias ao centro de massa dos impactos (Carpentier, 1993).

Para a prossecução desse objectivo, foi elaborada uma metodologia que conduzisse à obtenção de uma amostra não aleatória, de pequena dimensão, susceptível de ser avaliada por múltiplos avaliadores de modo a concluir sobre a variabilidade das respostas destes e, partindo daí, usar essa incerteza nas avaliações com apenas um avaliador. Embora não seja possível provar que a variação das respostas se iria manter constante em todos os ramos da ciência, esta é a aproximação possível, tendo em conta o número de avaliadores com credibilidade e boa vontade suficientes que foi possível encontrar para colaborar neste estudo.

Por razões de conveniência, e presumida uma maior facilidade em recrutar avaliadores na área de Gestão, foi este o tema escolhido para a selecção dos artigos. Foi utilizado um suplemento da revista *Executive Digest*, em que são elencados 50 conceitos de gestão (Cardoso, 1997), para obter um leque de assuntos alargado e não muito específico, de modo a serem susceptíveis de ter sido incluídos em enciclopédias generalistas.

O facto de o artigo em causa vir de uma revista de divulgação científica, e não de uma revista especializada, deverá ser uma mais-valia neste contexto e justifica a sua escolha.

O processo de constituição da amostra é similar ao do estudo principal, em que são apenas seleccionados os temas que constam em ambas as enciclopédias, com a diferença de que o processo não termina quando for encontrado um número de pares predeterminado à partida, mas sim quando forem analisados os 50 temas elencados na revista.

Para além do painel de 12 avaliadores – Professores na área de Gestão – foi também reunido um painel de 18 alunos, do 3º ano da Licenciatura em Gestão da Universidade de Évora, a quem foi também pedido que avaliasse os artigos de modo a concluir sobre as hipóteses levantadas no subcapítulo “Juízes versus Jurados”.

Em súmula, foram construídos dois “tribunais”: um de 18 jurados – não peritos – constituído pelos alunos, e outro de 12 juízes – peritos – os professores, de modo a

analisar as eventuais diferenças nas avaliações destes. Mais importante para este estudo será, no entanto, a variabilidade das respostas entre os peritos.

No fluxograma seguinte (Figura 31), encontra-se um resumo do processo atrás descrito. Os resultados obtidos estão compilados no Anexo IV.

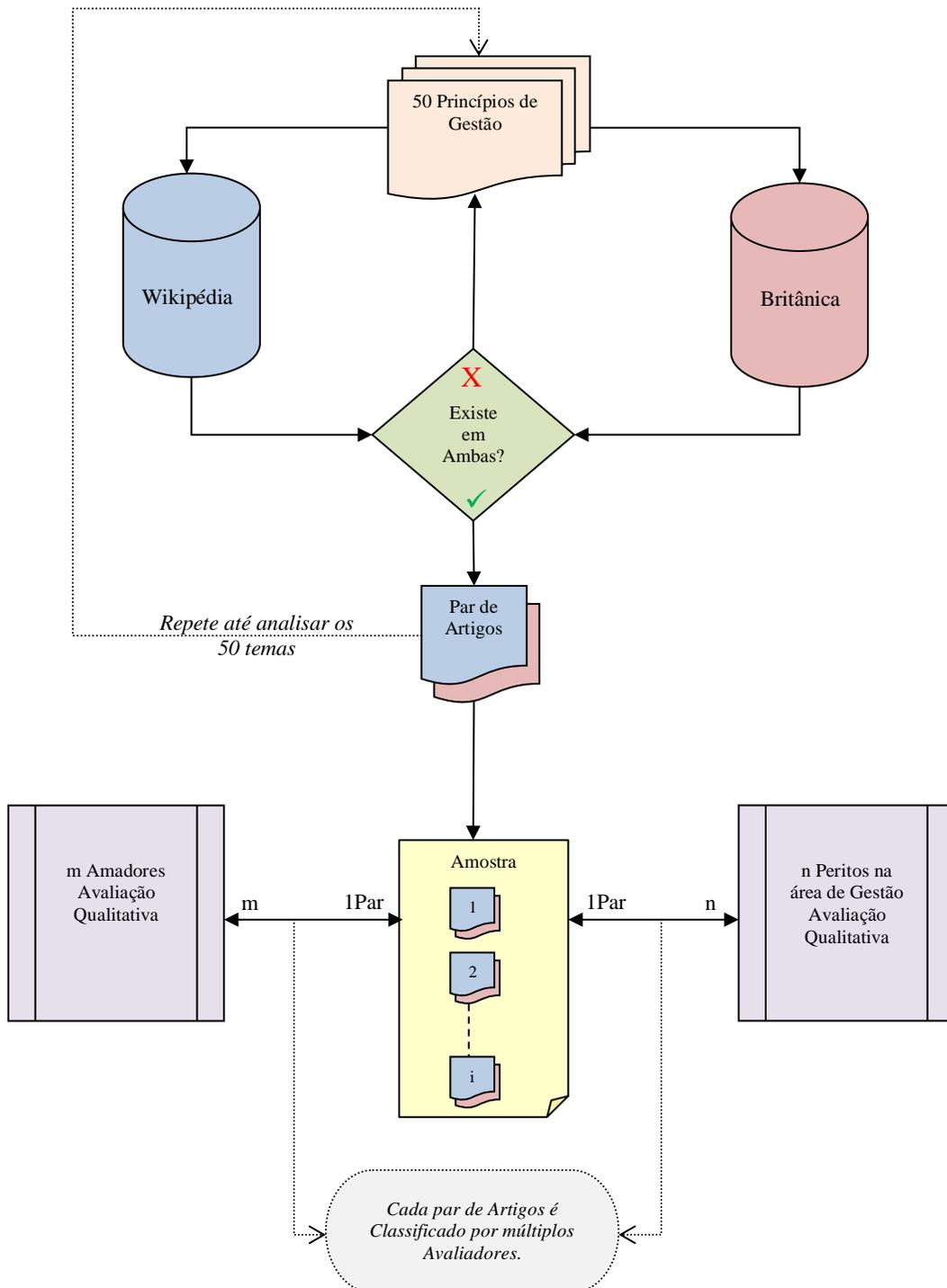


Figura 31: Amostragem Artigos 50 Princípios de Gestão

4.5 TRATAMENTO DOS DADOS

Dependendo das distribuições encontradas, mais ou menos simétricas, iremos utilizar como medidas de tendência central a média ou mediana, tendo em conta que esta última é menos afectada pela presença de valores anómalos nas distribuições. O cálculo do valor absoluto da soma dos desvios em relação a um e outro estimador permitirá avaliar qual deles será mais útil em cada circunstância.

Sempre que seja pertinente, iremos representar os dados através de *box plots* (Toutenburg & Shalabh, 2009), utilizando os critérios que se podem ver na Figura 32:

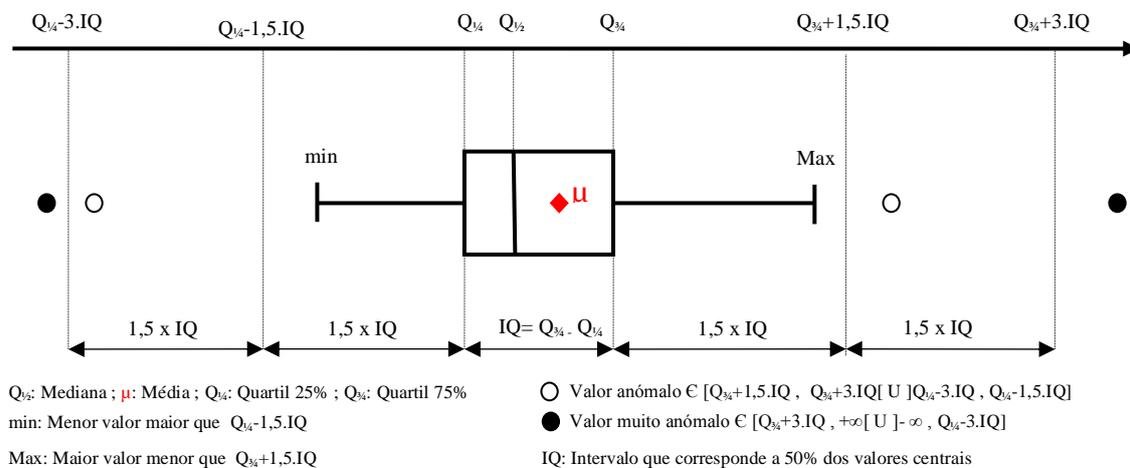


Figura 32: Interpretação Box Plots

Serão também utilizados histogramas de Pareto (Toutenburg & Shalabh, 2009), em que o número de classes será naturalmente definido pelo número de níveis de classificação utilizado, cinco e nove. Os valores absolutos observados serão representados através de barras na escala da esquerda, e os valores acumulados em percentagem na da direita.

Em termos de medidas de dispersão, e tendo em conta o especial cuidado a ter no cálculo desta no estudo em que múltiplos avaliadores analisam os mesmos artigos, cuja

variabilidade será usada para melhor interpretar os resultados do estudo principal, iremos escolher medidas que permitam uma utilização directa dos resultados obtidos, ou seja, que tenham a mesma escala do que os dados de partida – medidas de dispersão absoluta; excluimos assim a variância desta análise.

Como no desvio padrão as distâncias em relação à média são calculadas ao quadrado, i.e., a influência dos *outliers* é exagerada, iremos privilegiar a média e a mediana dos desvios absolutos em relação à mediana e a média dos desvios absolutos em relação à média, devido à sua maior resiliência em relação à presença de *outliers* (Gorard, 2004).

As fórmulas de cálculo das medidas de dispersão referidas são:

- *Média Desvio Absoluto Média* = $\frac{1}{n} \cdot \sum_{i=1}^n |X_i - \mu|$
- *Média Desvio Absoluto Mediana* = $\frac{1}{n} \cdot \sum_{i=1}^n |X_i - M|$
- *Mediana Desvio Absoluto Mediana* = $Mediana|x_i - M|$
- *Incerteza estatística* = $\frac{\delta}{\sqrt{n}}$

Em que $\mu = Média$, $\delta = Desvio Padrão$, $M = Mediana$ e $n = n^o$ dados

De modo a certificar-nos de que as diferenças que as eventuais diferenças a obter entre as avaliações dos artigos e as respostas ao inquérito por parte dos grupos em estudo são significativas e não poderiam ser obtidas através de respostas aleatórias, i.e., são estatisticamente significativas, vamos utilizar testes de hipóteses (Fisher, 1966).

Neste tipo de testes, a hipótese nula (H_0) representa a teoria que queremos estudar, quer por acreditarmos que é verdadeira, quer para ser usada como argumento na adopção da hipótese alternativa (H_A). A ênfase posta na hipótese nula (o que, por vezes, nos leva a agir como advogados do diabo) tem a ver com o facto de ser esta a hipótese que vamos na realidade testar, mesmo que com o intuito de a rejeitar e posteriormente aceitar a hipótese alternativa.

Segundo Fisher (1966), a significância de um teste (α) – erro tipo I¹¹⁰ – é a probabilidade de rejeitar acidentalmente uma hipótese nula verdadeira. Os erros de tipo

¹¹⁰ P (Rejeitar H_0 | H_0 Verdadeira)

Π^{111} (β) advêm da probabilidade de o teste efectuado falhar a rejeição de uma hipótese nula falsa. A potência estatística do teste efectuado pode ser calculado a partir da probabilidade complementar de ocorrência de erros do tipo II ($1 - \beta$), ou seja, do cálculo de P (Rejeitar H_0 | H_A Verdadeira). Na Figura 33, podemos ver uma síntese dos possíveis resultados de um teste de hipóteses.



Figura 33: Teste de Hipóteses

A conclusão final, em relação à hipótese nula, é sempre de rejeitar H_0 em favor de H_A , ou de não rejeitar H_0 . Nunca se conclui que rejeitamos H_A , ou mesmo que aceitamos H_A . Isto deve-se ao facto de a conclusão *não rejeitamos a hipótese nula* não significar necessariamente que a hipótese nula é verdadeira; apenas sugere que não temos suficientes provas contra esta para a rejeitar. Quando *rejeitamos a hipótese nula* estamos apenas a sugerir que a hipótese alternativa pode ser verdadeira.

Pode parecer apenas um preciosismo de linguagem, mas é mais do que isso: este procedimento evidencia os pressupostos que estão por detrás de um teste de hipóteses, cujas etapas são, segundo Fisher (1966):

- **Definir** as hipóteses a testar e a margem de erro a utilizar (α);
- **Seleccionar** o teste a utilizar e calcular o seu valor;
- **Decidir** se rejeitamos, ou não, H_0 , por comparação com os valores de referência.

¹¹¹ P (Aceitar H_0 | H_a Verdadeira)

Iremos preferencialmente utilizar testes paramétricos devido à sua maior robustez (Fay & Proschan, 2010), sendo estes substituídos pelos seus equivalentes não paramétricos, quando não se verificarem as condições de aplicabilidade dos primeiros. A principal característica que uma amostra tem de possuir para podermos aplicar testes paramétricos é poder ser descrita, é ter uma distribuição aproximadamente normal. Caso este requisito não se verifique, devemos utilizar um teste não paramétrico.

A utilização de testes paramétricos pressupõe estarmos a testar a hipótese das médias das variáveis em causa serem estatisticamente diferentes, enquanto no caso dos testes não paramétricos estaremos a testar as medianas dessas mesmas variáveis. Esta escolha faz todo o sentido pois, caso as variáveis não sigam uma distribuição aproximadamente normal, a mediana descreve a distribuição de um modo mais eficaz do que a média.

Segundo Dunn (2001), existem diversas maneiras de determinar a normalidade de uma amostra e este autor propõe a simples visualização da sua distribuição através de um histograma. Com o intuito de retirar o factor humano da equação e garantir uma maior objectividade e homogeneidade de critérios, optamos pela utilização do teste de Shapiro-Wilk (Bernstein & Bernstein, 1999).

Shapiro e Wilk (1965) propõem que testemos a hipótese (H_0) de a amostra em causa provir de uma população normal, utilizando para tal técnicas de análise da variância com o intuito de avaliar a qualidade do ajustamento entre a curva de distribuição das frequências da amostra com a distribuição normal. Se considerarmos uma amostra de N^{112} elementos (X_1, X_2, \dots, X_n), o valor do teste (W) poderá ser calculado através da expressão:

$$W = \frac{(\sum_{i=1}^{N/2} a_{N-i+1}(x_{N-i+1} - x_i))^2}{\sum_{i=1}^N (x_i - \bar{x})^2}$$

em que a_i são os coeficientes¹¹³ de Shapiro-Wilk e \bar{x} é a média aritmética. Rejeitamos a hipótese nula se o valor de W calculado for menor do que o valor de W para o nível de significância escolhido.

¹¹² No caso de N ser par, despreza-se a observação mediana.

¹¹³ Derivados da aproximação de Royston.

A escolha do teste estatístico a utilizar na comparação das amostras, aparte da escolha entre testes paramétricos e não paramétricos, vai depender das amostras a comparar. Assim, na comparação de amostras emparelhadas dependentes, i.e., amostras que têm o mesmo intervalo de valores e o valor de um elemento do par influencia e é comparável com o valor correspondente, iremos utilizar o teste t de Student ou o teste de Wilcoxon, ambos para amostras emparelhadas, consoante o par de variáveis apresente, ou não, uma distribuição normal.

Bernstein e Bernstein (1999) mostram que o teste t de Student para amostras emparelhadas compara as médias destas, averiguando se a diferença média é significativamente diferente de zero. Se estabelecermos como H_0 a hipótese das médias de duas amostras (A e B) compostas por N elementos cada serem iguais, o valor do teste (t) poderá ser calculado através da expressão:

$$t = \frac{\bar{x}_A - \bar{x}_B}{\sqrt{\frac{\sigma_A^2 + \sigma_B^2}{N}}}$$

em que \bar{x} é a média aritmética e σ^2 a variância. Rejeitamos a hipótese nula se o valor de t calculado for maior do que o valor de t para o nível de significância escolhido e o número de graus de liberdade¹¹⁴ da amostra.

Caso o par de amostras a comparar não siga uma distribuição normal, iremos utilizar o teste de Wilcoxon para amostras emparelhadas, em que a hipótese nula a testar é a da igualdade das medianas. Wilcoxon (1945) propõe a seguinte metodologia: consideremos duas amostras emparelhadas (A e B) compostas por N elementos cada. Vamos começar por calcular, para cada par, o módulo da sua diferença, excluir aqueles pares em que a diferença é nula e ordenar por ordem crescente os restantes. Com base nesta ordenação, vamos criar uma nova variável atribuindo o valor correspondente à sua posição¹¹⁵; nos casos em que as diferenças sejam iguais, atribui-se o valor médio da posição que estes ocupam na lista¹¹⁶. Por fim, acrescenta-se o sinal da diferença entre

¹¹⁴ $N \times 2$

¹¹⁵ Vamos atribuir o valor de um ao primeiro elemento da lista, dois ao segundo e assim sucessivamente.

¹¹⁶ Caso tenhamos dois valores iguais para ocupar as posições i e i+1 da variável que estamos a criar, atribuímos a ambas a posição $(2i+1)/2$.

cada par à nova variável que criamos com o processo atrás descrito e calcula-se o módulo do somatório desta. Obtemos assim o valor (W) do teste de Wilcoxon.

O número de graus de liberdade (GL) é dado pelo número de pares com diferença não nula e, quando esse valor aumenta,¹¹⁷ a distribuição aproxima-se de uma distribuição normal e podemos calcular Z através da expressão:

$$Z = \frac{W-0,5}{\sigma_w}, \sigma_w = \sqrt{\frac{GL(GL+1)(2GL+1)}{6}}$$

Se o valor de Z for superior ao valor de Z crítico para o nível de significância escolhido, rejeitamos a hipótese nula.

Caso as amostras a comparar sejam independentes entre si, i.e. não emparelhadas, e eventualmente, de dimensões diferentes, começaremos por averiguar da sua normalidade através do teste de Shapiro-Wilk de modo a escolher que tipo de teste, paramétrico ou não paramétrico, iremos utilizar (Bernstein & Bernstein, 1999).

No caso de termos de utilizar testes não paramétricos, iremos aplicar o teste de Mann-Whitney para averiguar se a diferença entre as medianas das mostras a comparar é significativa em termos estatísticos, no caso de estarmos a comparar duas amostras, e a generalização deste, para mais de duas amostras, o teste de Kruskal-Wallis (Kruskal & Wallis, 1952).

Segundo Bernstein e Bernstein (1999), a escolha do tipo de teste paramétrico a aplicar às amostras tem de ser precedida pela determinação da homocedasticidade destas, ou seja, pela determinação da homogeneidade da sua variância. Esta determinação faz-se pela aplicação do teste de Levene e, caso esta se verifique, optaremos por efectuar uma análise da variância (ANOVA) e usaremos os testes de Brown-Forsythe e de Welch nas restantes situações.

Para determinar se as eventuais diferenças encontradas entre as diversas variáveis a estudar serão preferencialmente utilizados testes paramétricos, sendo estes substituídos pelos seus equivalentes quando não se verificarem as condições de aplicabilidade dos

¹¹⁷ N>10

primeiros. A margem de erro a utilizar será de 5% e o nível de significância 95%, sendo utilizado o *software* SPSS para a realização dos cálculos.

Iremos apenas utilizar uma¹¹⁸ das três variáveis quantitativas que se referem ao tamanho dos artigos: número de palavras (P), número de caracteres (C) e número de linhas (L), visto serem, obviamente, linearmente dependentes, como se pode constatar através do cálculo dos coeficientes de correlação de Pearson entre estas, para os artigos da Wikipédia (W) e da Britânica (B).

$$\begin{bmatrix} W & P & C & L \\ P & 1 & & \\ C & 0,997 & 1 & \\ L & 0,997 & 1,000 & 1 \end{bmatrix} e \begin{bmatrix} B & P & C & L \\ P & 1 & & \\ C & 0,998 & 1 & \\ L & 0,998 & 1,000 & 1 \end{bmatrix}$$

Na Tabela 7, apresenta-se um resumo dos testes estatísticos atrás apresentados e que serão utilizados no tratamento dos dados, bem como os principais critérios que condicionam a sua escolha e aplicabilidade, sendo nalguns casos necessário que os dados cumpram outros requisitos.

		TESTES PARAMÉTRICOS		TESTES NÃO PARAMÉTRICOS	
		<i>Amostras seguem aproximadamente Distribuição Normal? (Shapiro-Wilk)</i>			
AMOSTRAS EMPARELHADAS		t Student <i>(Normal)</i>		Wilcoxon <i>(Diferente da Normal)</i>	
		ANOVA <i>(Variância Homogénea)</i>	Brown-Forsythe e Welch <i>(Variância não Homogénea)</i>	Mann-Whitney <i>(Duas)</i>	Kruskal-Wallis <i>(Mais de duas)</i>
AMOSTRAS INDEPENDENTES		<i>Homocedasticidade? (Levene)</i>		<i>Nº de Amostras?</i>	

Tabela 7: Ferramentas Estatísticas a Utilizar no Estudo

¹¹⁸ Foi escolhida arbitrariamente a variável número de palavras, em representação das três.

Nas páginas seguintes, são apresentados os fluxogramas simplificados da metodologia utilizada neste estudo (Figura 34), bem como na dos estudos complementares (Figura 35 e Figura 36). É de referir que a análise dos fluxogramas deve ser complementada com a informação contida na Tabela 7, de modo a obter uma visão detalhada do processo.

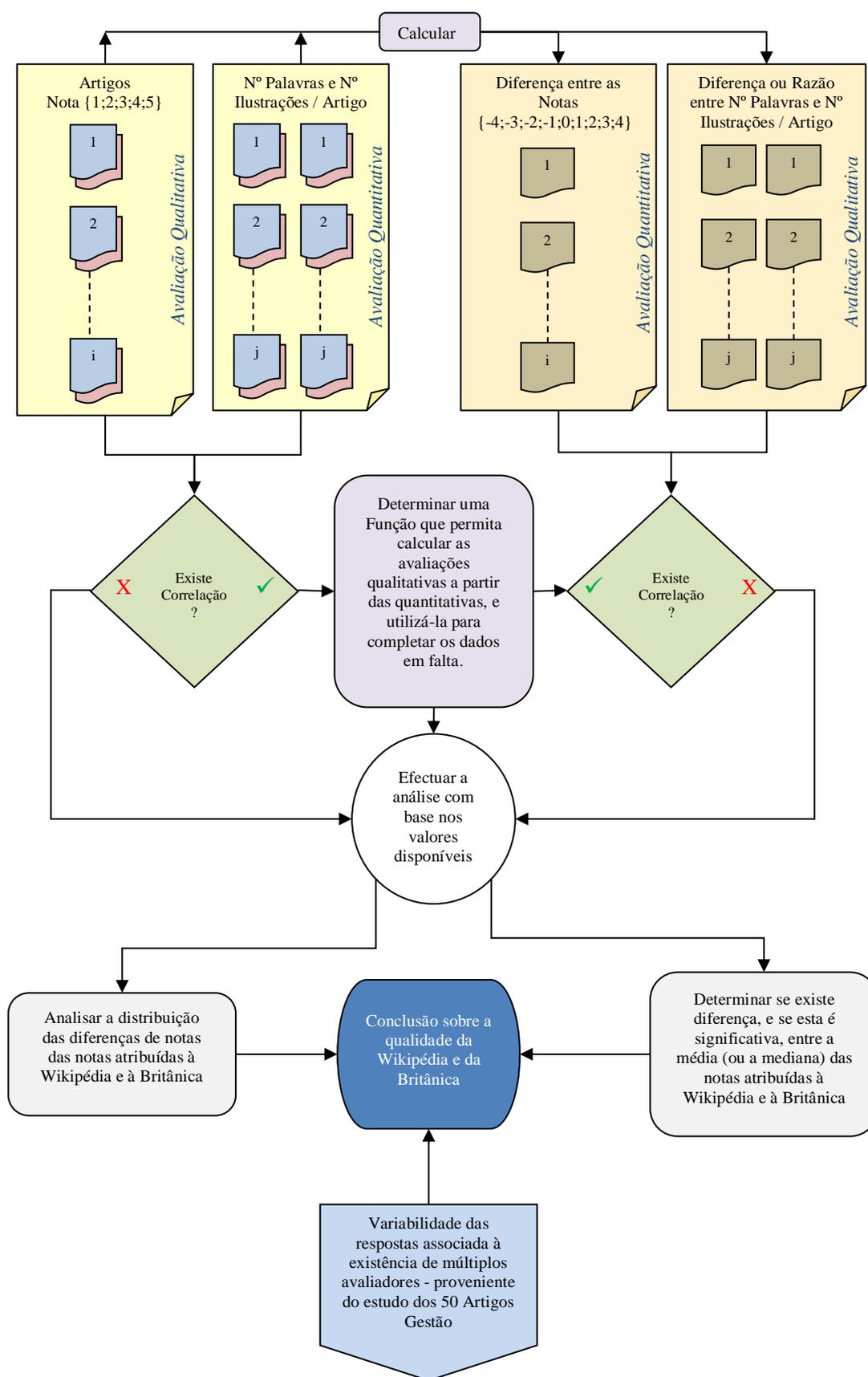


Figura 34: Fluxograma da Metodologia do Estudo Comparativo entre as Enciclopédias

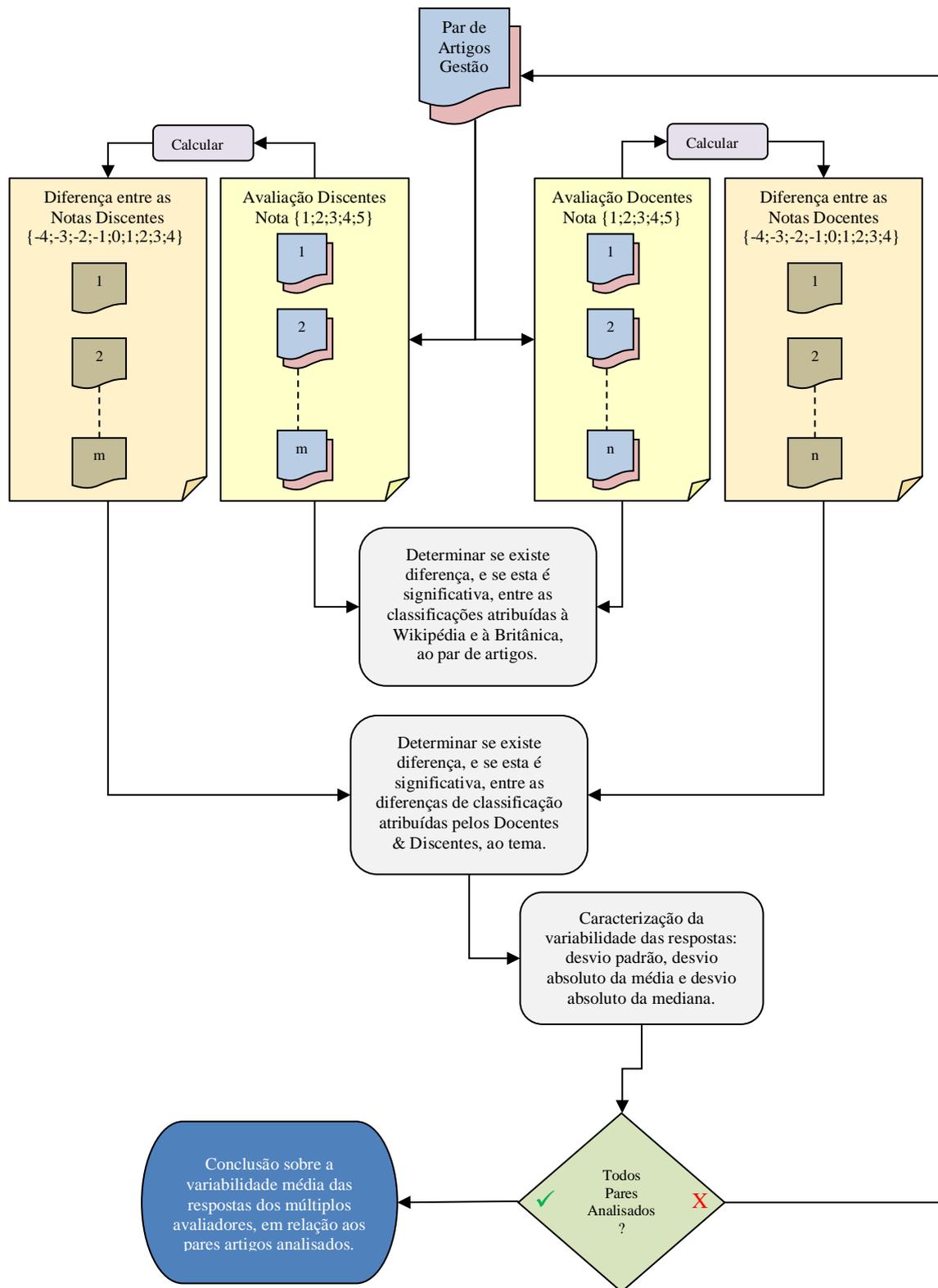


Figura 35: Fluxograma da Metodologia do Estudo dos 50 Princípios de Gestão

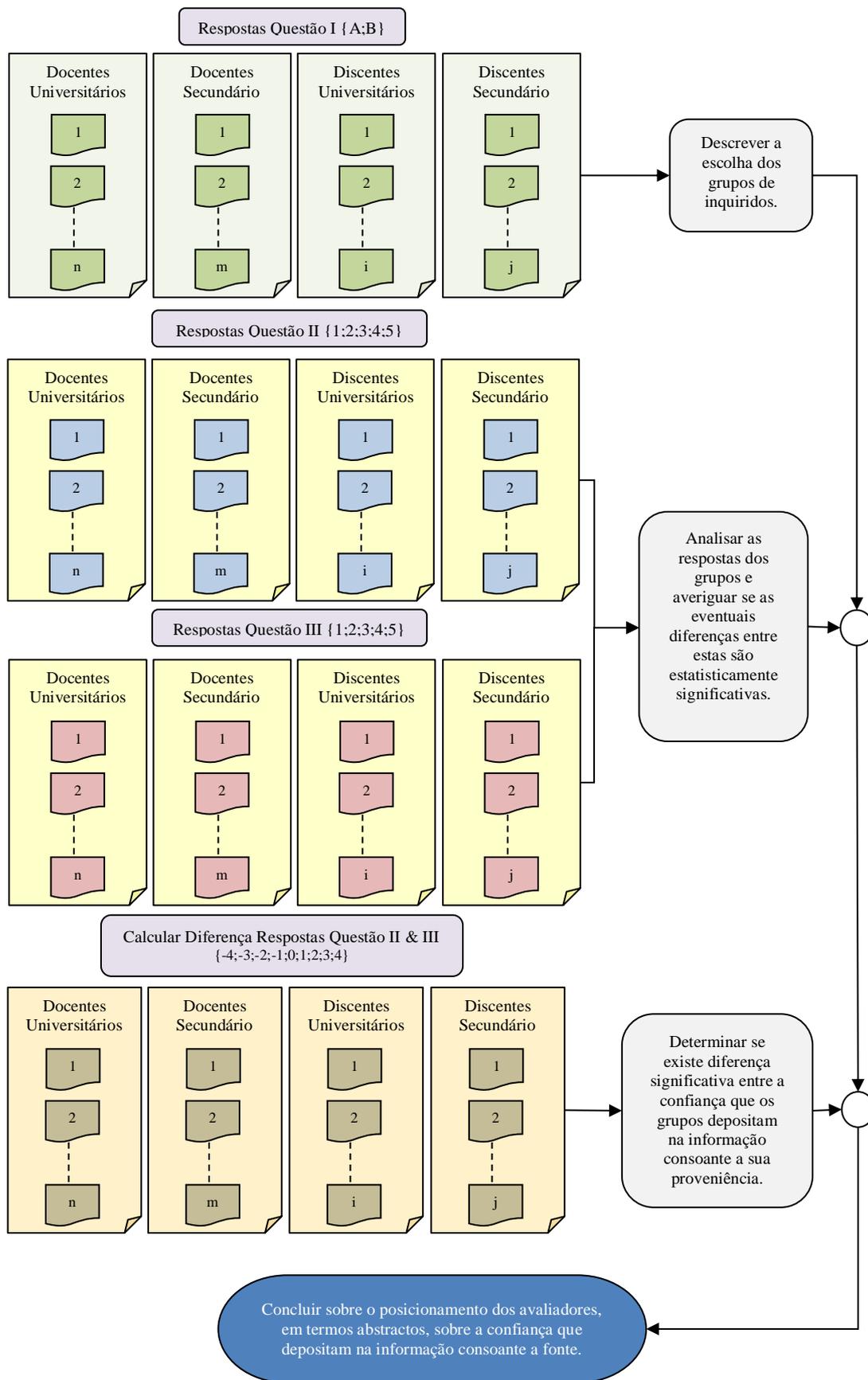


Figura 36: Fluxograma da Metodologia da Análise do Inquérito

Alice: Would you tell me, please, which way I ought to go from here?
The Cat: That depends a good deal on where you want to get to.
Alice: I don't much care where.
The Cat: Then it doesn't much matter which way you go.
Alice: ...so long as I get somewhere.
The Cat: Oh, you're sure to do that, if only you walk long enough.

Lewis Carroll (1865), *Alice's Adventures in Wonderland*

4.6 FILOSOFIA DE INVESTIGAÇÃO

Foi utilizado o modelo proposto por Saunders, Lewis, e Thornhill (2009) para representar a filosofia de investigação adoptada nesta tese. O posicionamento deste trabalho de investigação, de acordo com o modelo referido, encontra-se representado na Figura 37.

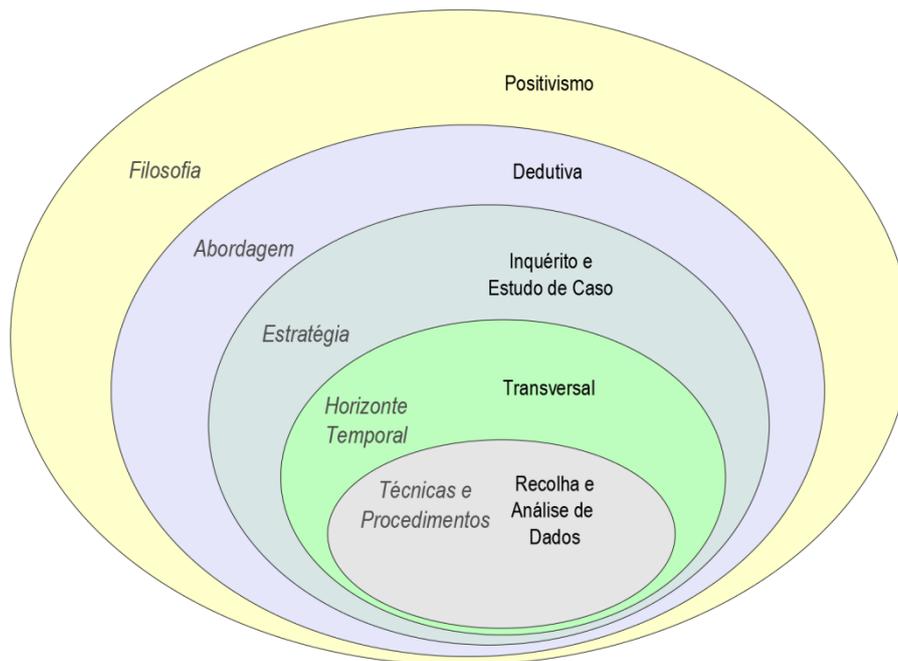


Figura 37: Filosofia de Investigação, adaptado de Saunders et al. (2009)

Em termos de estratégia optou-se pela realização de um estudo de caso, com o intuito de obter uma imagem (fotografia) do problema em estudo num momento específico do tempo, utilizando para tal a recolha e posterior análise de dados. A filosofia positivista escolhida, levou-nos a definir uma metodologia estruturada, susceptível de ser reutilizada no futuro, com o intuito de verificar a evolução do objecto em estudo.

Atendendo ao trabalho de Gay & Diehl (apud Reto & Nunes, 1999), poderemos ainda classificar esta investigação, quanto ao objectivo e ao método, como Avaliativa e Casual-Comparativa, respectivamente. A necessidade de optarmos por um estudo comparativo prende-se com a dificuldade de definir qualidade da informação – variável que queremos aferir – usando uma escala absoluta e universal, pelo que procurámos encontrar um termo de comparação que permite, pelo menos, a classificação relativa desta.

A filosofia positivista adoptada pretendeu mimetizar os métodos e técnicas utilizados nas ciências naturais, tendo o investigador, em termos ontológicos, mantido um posicionamento externo e independente em relação ao tema investigado¹¹⁹; foram utilizadas amostras estruturadas de grande dimensão, extraídas da população a estudar, usando técnicas que garantem a sua representatividade e, posteriormente, analisadas com os instrumentos de medida que se consideraram os mais adequados, como é habitual nesta filosofia.

A escolha entre uma abordagem indutiva ou dedutiva revela-se mais complexa. Segundo Freixo (2009), o conhecimento científico moderno tem origem em dois raciocínios: o indutivo e o dedutivo. O método indutivo tem origem na concepção empirista da ciência, ou seja, a ciência deve basear-se na indução e na experimentação e não na metafísica e na especulação. Francis Bacon (1561-1626), pioneiro do método indutivo e autor da obra *Novum Organum* (Bacon, 1620), em que defende esta abordagem, afirma que o saber deve ser activo e fecundo em resultados práticos e não apenas a simples catalogação de factos de uma realidade supostamente fixa.

O raciocínio dedutivo que se baseia na concepção racionalista da ciência faz-se do geral para o particular, ou seja, procura-se através de leis e teoremas gerais uma explicação para o particular. René Descartes (1596-1650), contemporâneo de Bacon e autor do *Discurso do Método* (Descartes, 1637), defende a incerteza sistemática, ou seja, devemos duvidar de todos os conhecimentos até que seja possível confirmá-los utilizando o raciocínio lógico sobre premissas totalmente evidentes.

¹¹⁹ Este posicionamento foi também incutido aos avaliadores, nomeadamente através da descaracterização dos artigos que lhes era pedido avaliar.

As conclusões provenientes do raciocínio indutivo superam a informação contida nas premissas, enquanto na dedução apenas reformulam as premissas que lhes deram origem, ou seja, a primeira abordagem amplia os nossos conhecimentos e a segunda apenas o explicita (Freixo, 2009).

A escolha da abordagem dedutiva prende-se com o facto de, se utilizarmos a incerteza sistemática de Descartes, a resposta à pergunta de investigação se poder obter testando a hipótese: “A colaboração em massa não consegue produzir resultados científicos de qualidade”, pois basta provar que existe pelo menos um (\exists^{-1}) caso em que tal não acontece, i.e. se no caso em estudo a Wikipédia produzir resultados científicos de qualidade, poderemos rejeitar a hipótese. Popper (1935) defende esta abordagem lógica, afirmando que, se uma conjectura implica consequências observáveis e se estas consequências são refutáveis, então a conjectura pode ser eliminada.

A semelhança do teste de hipóteses de Fisher (1966), se no caso em estudo se concluir que a informação produzida por este não é de qualidade, tal não legitima a utilização do raciocínio lógico dedutivo:

- Premissa: “A colaboração em massa não consegue produzir resultados científicos de qualidade”
- Inferência: “A informação x foi produzida utilizando a colaboração em massa”
- Tese: “logo, a informação x não tem qualidade”

Diferente seria se quiséssemos provar que: “A colaboração em massa produz sempre resultados científicos de qualidade”; aí, o estudo realizado poderia ser encarado como uma abordagem indutiva ao problema, limitando-se a testar um caso que, *per se*, não seria suficiente para ratificar a hipótese equacionada e, conseqüentemente, formular uma teoria, o que, segundo Popper (1935), só poderia ocorrer quando a observação de casos isolados atingisse o infinito, o que nunca poderia ocorrer. Saunders et al. (2009) afirmam que, por vezes, as duas se podem misturar na mesma pesquisa.

Podemos tentar utilizar um teste de hipóteses para responder à pergunta de investigação, embora este não se vá comportar como tal, pois estamos apenas a estudar um dos casos

da colaboração em massa, o que impossibilita que aceitemos a hipótese nula e concluir que a colaboração em massa não consegue produzir resultados de qualidade.

- H_0 : A colaboração em massa não consegue produzir resultados de qualidade;
- H_A : A colaboração em massa consegue produzir resultados de qualidade.

Limitados pela metodologia escolhida (estudo de caso), nada poderemos concluir sobre a hipótese alternativa e não podemos também nunca aceitar a hipótese nula; restando apenas a hipótese de rejeitar esta – caso se verifique que a qualidade da Wikipédia é superior ou igual à da Britânica.

O estudo comparativo entre a qualidade das enciclopédias Britânica e a Wikipédia, foi portanto realizado com o intuito de tentar encontrar um caso de sucesso da colaboração em massa, tendo no entanto sido realizados mais dois estudos complementares.

Como a cada medição está sempre associado um erro¹²⁰, apesar de este nunca se poder eliminar totalmente, é de extrema importância que seja conhecido de modo a que estejamos conscientes da precisão dos resultados obtidos. Nesta investigação, foi utilizado um grupo de controlo – estudo “50 Princípios de Gestão” – precisamente com o intuito de conhecer o erro associado às avaliações e poder aferir-se a validade dos resultados alcançados.

O inquérito “Caracterização dos Hábitos de Consulta de Informação” foi elaborado com o intuito de caracterizar a população de onde foi retirada a maioria dos avaliadores, de modo a aferir se as suas avaliações estavam a ser, eventualmente, condicionadas pelas convicções que expressaram no inquérito. Por último, a utilização da diferença entre as avaliações permite minimizar eventuais disparidades de critério entre os avaliadores, garantindo uma mais correcta comparação entre as avaliações.

Epistemologicamente, e apesar de todos os dados recolhidos serem quantitativos, as avaliações efectuadas por terceiros à qualidade dos artigos reflectem uma inevitável subjectividade inerente à necessária introdução do factor humano na classificação destes. Esta constatação não invalida o *justified true belief* do investigador na capacidade destes em produzir dados credíveis, como se de factos observados pelo

¹²⁰ Este erro nas ciências naturais divide-se em erro da medição e erro intrínseco ao instrumento de medida utilizado (precisão).

próprio se tratassem. Este facto permite reforçar, em termos axiológicos, uma total independência entre o investigador e os dados.

Os avaliadores que colaboraram neste estudo foram encarados, do ponto de vista metodológico, como um “instrumento de medida” ao qual está associado um determinado erro que pretendemos que seja mínimo. A escolha de avaliadores doutorados e professores universitários é equivalente à utilização dos melhores instrumentos de medida à disposição do investigador, e a escolha do instrumento correcto para uma determinada medição, à correcta alocação dos avaliadores, de acordo com as suas competências, a um determinado artigo.

*If you torture numbers long enough,
they will confess to anything.*

Gregg Easterbrook (2003), *The Progress Paradox*

5. APRESENTAÇÃO DE RESULTADOS

Neste capítulo, são apresentados os resultados do inquérito sobre os hábitos de consulta de informação (5.1), do estudo sobre a variabilidade das avaliações (5.2) e do estudo comparativo entre a Wikipédia e a Britânica (5.3).

5.1 CARACTERIZAÇÃO DOS HÁBITOS DE CONSULTA DE INFORMAÇÃO

Os resultados¹²¹ obtidos no inquérito efectuado com o intuito de caracterizar os hábitos de consulta de informação no ensino secundário e universitário por parte dos seus principais intervenientes, professores e alunos, bem como a confiança que, *a priori*, depositam nas fontes de informação “clássicas” por contraponto com as “alternativas”, encontram-se sumariados na Tabela 8.

São de salientar as baixas taxas de resposta obtidas, com especial realce para o grupo dos alunos universitários – menos de 1% – constituindo o grupo de professores do secundário uma excepção (36%).

Nas páginas seguintes, vamos passar a referir “informação disponibilizada por um motor de busca / Wikipédia” por “fontes alternativas”, e “enciclopédia / dicionário / obra de referência” por “fontes clássicas”. Os inquéritos foram efectuados durante os meses de Abril e Maio de 2011.

Em relação à primeira pergunta, “Quando necessita de obter informação - aprofundar um determinado tema, onde procura primeiro?”, podemos ver na Figura 38 como se distribuem as preferências dos quatro grupos, indicando a média dos grupos de inquiridos que 81% usam numa primeira abordagem os resultados obtidos através das fontes ditas alternativas.

¹²¹ Ver Anexo III.

APRESENTAÇÃO DE RESULTADOS

	Universidade		Secundário	
	Docentes	Discentes	Docentes	Discentes
Amostra (#)	63	78	38	69
População (#)	655	9510	107	697
Taxa Resposta (%)	9,6%	0,8%	35,5%	9,9%
Quando necessita de obter informação - aprofundar um determinado tema, onde procura primeiro?				
Motor de Busca / Wikipédia	47	63	27	65
Enciclopédia / Dicionário / Obra de Referência	16	15	11	4
Qual o grau de confiança que tem na Informação obtida através de um Motor de Busca / Wikipédia?				
Muito Alto	0	3	1	4
Alto	13	21	10	18
Médio	27	37	19	39
Baixo	17	15	7	7
Muito Baixo	6	2	1	1
Qual o grau de confiança que tem na Informação obtida através de uma Enciclopédia / Dicionário / Obra de Referência				
Muito Alto	29	44	21	42
Alto	28	33	16	26
Médio	4	1	1	1
Baixo	2	0	0	0
Muito Baixo	0	0	0	0

Tabela 8: Resultados do Inquérito

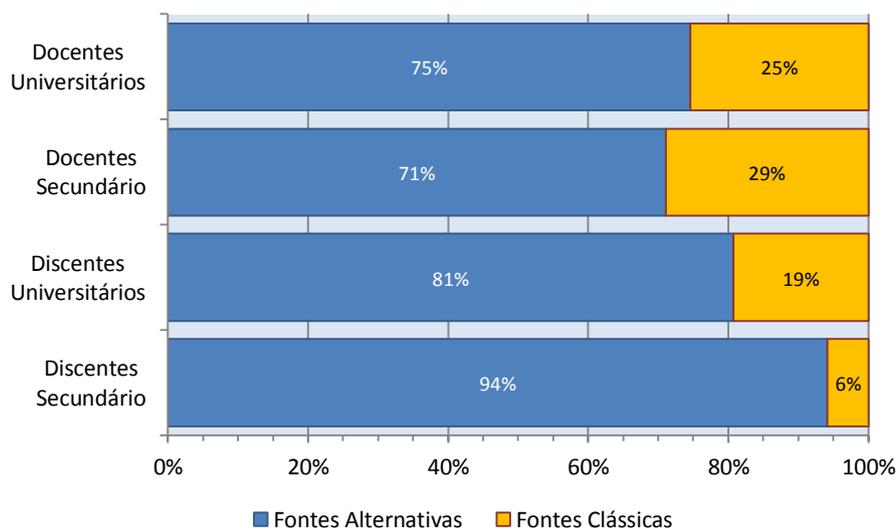


Figura 38: Fonte de Informação que Utilizam Primeiro

Apresentam-se, de seguida, os histogramas com as respostas à segunda e terceira perguntas do inquérito, relativas aos quatro grupos em análise. (Figura 39, Figura 40, Figura 41 e Figura 42)



Figura 39: Confiança na Informação por parte dos Docentes Universitários



Figura 40: Confiança na Informação por parte dos Docentes Secundário



Figura 41: Confiança na Informação por parte dos Discentes Universitários



Figura 42: Confiança na Informação por parte dos Discentes Secundário

Dos histogramas apresentados, importa referir as distribuições aproximadamente simétricas de todos os grupos inquiridos em relação à confiança que depositam nas fontes “alternativas”, por contraponto às distribuições das respostas sobre a confiança nas fontes “clássicas”, que se aproximam de exponenciais, fortemente assimétricas.

Assinale-se ainda a quase inexistência de falta de confiança de todos os grupos em relação à confiabilidade da informação obtida através de fontes “clássicas”, valor esse que apenas baixa de 100% no caso dos Docentes Universitários, e para apenas 97%.

Na Figura 43, encontram-se os box plots das diferenças entre as respostas às perguntas 2 e 3, em que se podem comparar as respostas dos quatro grupos.

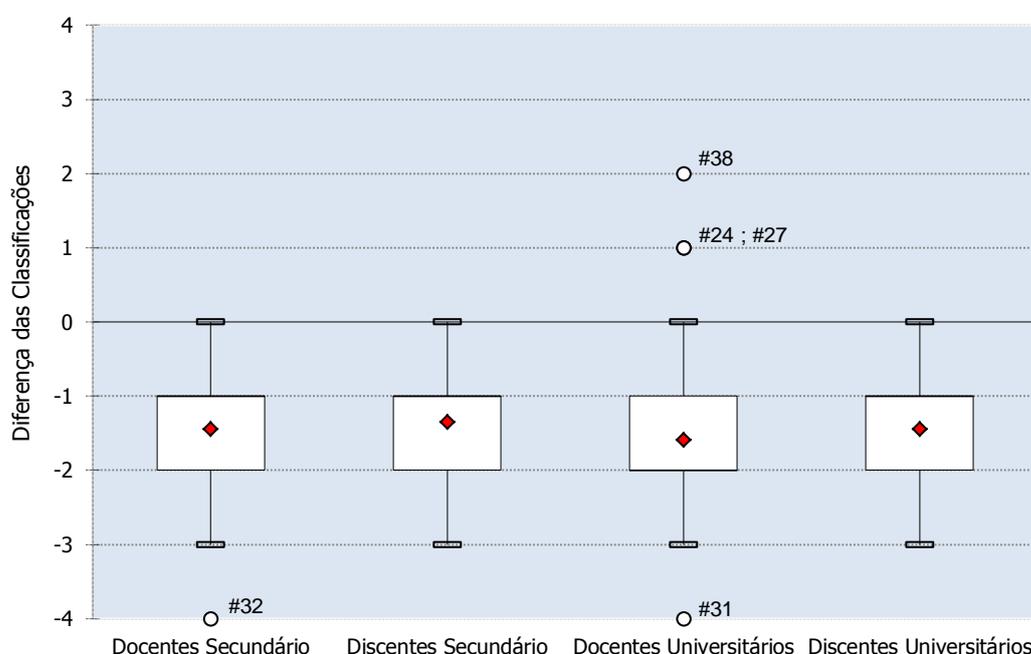


Figura 43: Confiança na Informação Disponibilizada por Fontes Alternativas, quando Comparada com a Obtida Através de Fontes Clássicas

Numa escala de nove níveis, está expressa a diferença de confiança na informação obtida através de ferramentas “alternativas” e “clássicas”, no que diz respeito à sua elaboração. Quanto maior for a diferença em termos absolutos, maior será a diferença de confiança sentida pelos inquiridos, indicando o sinal da diferença que confiam mais nas fontes “alternativas” quando a diferença é positiva e nas “clássicas” no caso de este ser negativo.

Como se pode ver, os quatro grupos apresentam distribuições semelhantes, sendo evidente que os inquiridos mostraram ter uma maior confiança nas fontes “clássicas” de informação, o que vem corroborar a hipótese posta inicialmente de tentar anular o efeito da credibilidade da fonte, aquando da avaliação dos artigos.

Este resultado era expectável e mostra que os 19% dos inquiridos que usam numa primeira abordagem as fontes “clássicas” para procurar informação têm uma opinião muito crítica – negativa neste caso – da informação obtida por fontes “alternativas”.

De modo a averiguar a possibilidade de aplicar testes paramétricos aos resultados das perguntas 2 e 3 do inquérito por parte dos quatro grupos em estudo, testaram-se as hipóteses:

- H_0 : a variável é normalmente distribuída;
- H_1 : a variável não é normalmente distribuída.

Como o teste de Kolmogorov-Smirnov apenas é aconselhado para amostras de dimensão superior a 50 e uma das variáveis a testar tem menos de 50 respostas, utilizámos o teste de Shapiro-Wilk, que é eficaz em todas as variáveis (Park, 2008).

A probabilidade associada ao teste de normalidade é inferior a 0,05 em todas as variáveis; logo, rejeita-se a hipótese nula, concluindo-se que as variáveis não são normalmente distribuídas (Tabela 9).

	Fontes Alternativas		Fontes Clássicas	
	Shapiro-Wilk	Probabilidade	Shapiro-Wilk	Probabilidade
Docentes Universitários	0,871	0,000	0,756	0,000
Discentes Universitários	0,883	0,000	0,670	0,000
Docentes Secundário	0,878	0,001	0,695	0,000
Discentes Secundário	0,844	0,000	0,658	0,000

Tabela 9: Teste Shapiro-Wilk - Confiança Fontes Alternativas e Clássicas

Para testar se os pares de respostas são, em termos estatísticos, significativamente diferentes entre si e tendo em conta que não podemos assumir que as variáveis seguem

uma distribuição aproximadamente normal, vamos ter de optar por um teste não paramétrico para testar as hipóteses:

- $H_0: \mu_1 = \mu_2$ (as médias são iguais);
- $H_1: \mu_1 \neq \mu_2$ (as médias são diferentes).

Foi então aplicado o teste de Wilcoxon às amostras emparelhadas, pois consideramos que as respostas às questões 2 e 3 são dependentes entre si, dentro de cada grupo, tendo sido obtidos os seguintes resultados: (Tabela 10)

	Média		Z	Probabilidade (bicaudal)
	Fontes Alternativas	Fontes Clássicas		
Docentes Universitários	2,9	4,2	-6,223	0,000
Discentes Universitários	3,1	4,5	-7,269	0,000
Docentes Secundário	3,1	4,5	-5,038	0,000
Discentes Secundário	3,3	4,7	-6,991	0,000

Tabela 10: Teste de Wilcoxon – Confiança Fontes Alternativas e Clássicas

Como todos os valores das probabilidades são inferiores a 0,05, rejeita-se a hipótese nula e podemos concluir que existe uma diferença estatisticamente significativa entre as respostas dos inquiridos para todos os grupos, ou seja, as diferenças entre as medianas são estatisticamente significativas.

De modo a testar se existem diferenças significativas entre as respostas dadas pelos quatro grupos, vamos analisar as quatro variáveis criadas a partir das diferenças entre as respostas às questões 2 e 3, começando por testar a sua normalidade, através do seguinte teste de hipóteses:

- H_0 : a variável é normalmente distribuída;
- H_1 : a variável não é normalmente distribuída.

	Alternativas - Clássicas	
	Shapiro-Wilk	Probabilidade
Docentes Universitários	0,925	0,001
Discentes Universitários	0,886	0,000
Docentes Secundário	0,892	0,002
Discentes Secundário	0,831	0,000

Tabela 11: Teste Shapiro-Wilk – Diferença Confiança Fontes Alternativas e Clássicas

A probabilidade associada ao teste de normalidade é inferior a 0,05 em todas as variáveis; logo, rejeita-se a hipótese nula, concluindo-se que as variáveis não são normalmente distribuídas (Tabela 11).

Mais uma vez, será necessário aplicar testes não paramétricos para verificar se existem diferenças estatisticamente significativas entre as variáveis. Como desta vez queremos comparar mais do que duas variáveis, estas são independentes entre si e as distribuições têm aproximadamente a mesma forma¹²², utilizaremos o teste de Kruskal-Wallis.

Para tal, estipulam-se as seguintes hipóteses:

- $H_0: M_1=M_2=M_3=M_4$ (as medianas são iguais);
- $H_1: \exists^{-1} M_i \neq M_j, \text{ com } i, j \in \{1,2,3,4\}$ (nem todas as medianas são iguais).

Os resultados do teste indicam um valor do χ^2 de 3.097, para 3 graus de liberdade, com uma probabilidade associada de 0,377, valor maior que 0,05, pelo que não se rejeita a hipótese nula e se conclui que não existe diferença significativa entre as medianas, ou seja, não existem diferenças estatisticamente significativas entre as respostas dos quatro grupos de inquiridos.

Da aplicação dos testes anteriores, conclui-se que existe uma diferença significativa entre o grau de confiança atribuído pelos inquiridos às fontes “clássicas” e às “alternativas” em todos os grupos, sendo a confiança depositada nas fontes “clássicas” muito superior, e que essa diferença é semelhante nos quatro grupos estudados.

¹²² Ver Figura 43.

Na Figura 44, pode observar-se a diminuição do grau de confiança dos inquiridos, em relação a ambos os tipos de fontes de informação, à medida que aumenta o seu número de anos de instrução, bem como a clara diferença – comum a todos os grupos – entre estas. A percentagem dos que recorrem primeiro a fontes “alternativas” apresenta, também, uma tendência de descida, com o aumento do número de anos de instrução.

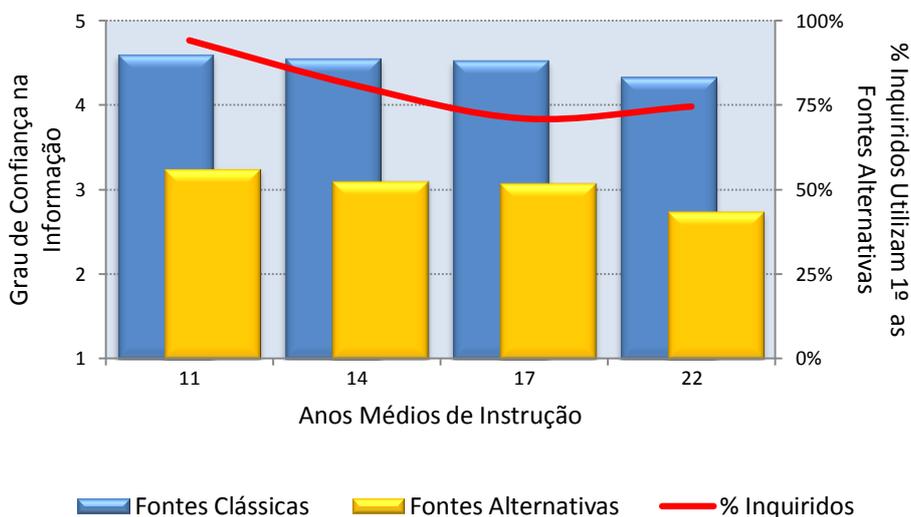


Figura 44: Variação das Respostas com a Escolaridade

Na Figura 45, todos os inquiridos dos dois grupos – os que utilizam primeiro fontes “alternativas” e os que utilizam primeiro fontes “clássicas” quando procuram informação sobre um determinado tema – mostraram ter o mesmo grau de confiança nas fontes “clássicas”. Relativamente às fontes “alternativas”, a confiança nestas é bastante menor no grupo que prefere utilizar as fontes “clássicas” como primeira alternativa à busca de informação.

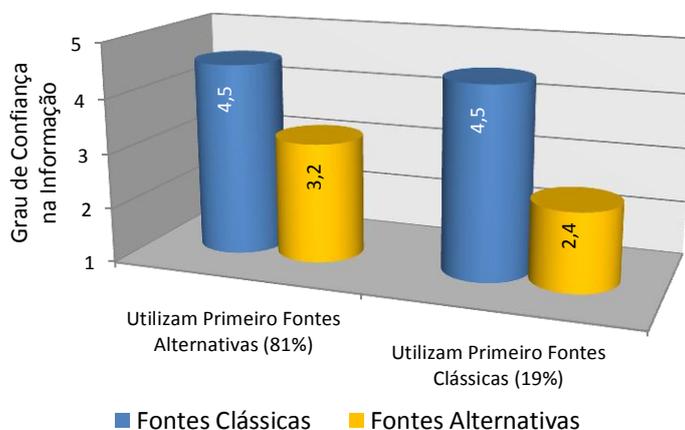


Figura 45: Variação das Respostas 2 e 3 em Função da Resposta à Questão 1

*If you can't prove what you want to prove,
demonstrate something else and pretend they are the same thing.
In the daze that follows the collision of statistics with the human mind,
hardly anyone will notice the difference.*

Darrell Huff (1954), *How to Lie with Statistics*

5.2. CINQUENTA PRINCÍPIOS DE GESTÃO

Os resultados¹²³ da procura dos 50 temas de Gestão encontraram 44 artigos (88%) na Wikipédia e apenas 4 artigos na Britânica¹²⁴ (8%). Mais uma vez, ficou bem patente a diferença em termos da dimensão destas.

Os quatro pares de artigos versam sobre os temas *Just in Time* (JIT), *Activity Based Costing* (ABC), *Outsourcing* (OUT) e *Total Quality Management* (TQM). Foram submetidos à avaliação de 12 Professores e de 16 alunos, sendo as principais medidas¹²⁵ de tendência central e de dispersão as seguintes (Tabela 12):

Docentes	JIT		ABC		OUT		TQM	
	W	B	W	B	W	B	W	B
Média	3,9	2,3	4,0	3,2	3,8	2,5	2,9	1,8
Mediana	4,0	2,0	4,0	3,0	5,0	3,0	3,0	2,0
Moda	4,0	2,0	4,0	3,0	5,0	3,0	3,0	1,0
Desvio padrão	0,6	0,8	0,4	0,6	1,7	0,9	0,8	0,8
Desvio Médio Absoluto da Média	0,5	0,7	0,2	0,4	1,5	0,8	0,5	0,7
Desvio Médio Absoluto da Mediana	0,4	0,6	0,5	0,6	1,3	0,7	0,4	0,7
Mediana Desvio Absoluto da Mediana	0,0	0,5	0,0	0,0	0,0	0,5	0,0	1,0
Incerteza estatística	0,2	0,2	0,1	0,2	0,5	0,3	0,2	0,2

Tabela 12: Estatística Descritiva – Avaliação Artigos Gestão

¹²³ Ver Anexo IV.

¹²⁴ Estes quatro temas foram também encontrados na Wikipédia.

¹²⁵ Escala de 1 a 5.

Iremos analisar em primeiro lugar as respostas dos Professores, grupo a partir do qual se pretende estudar a variabilidade, de modo a calcular a ordem de grandeza das variações nos resultados que seriam introduzidas no estudo em que temos apenas um avaliador por par de artigos:

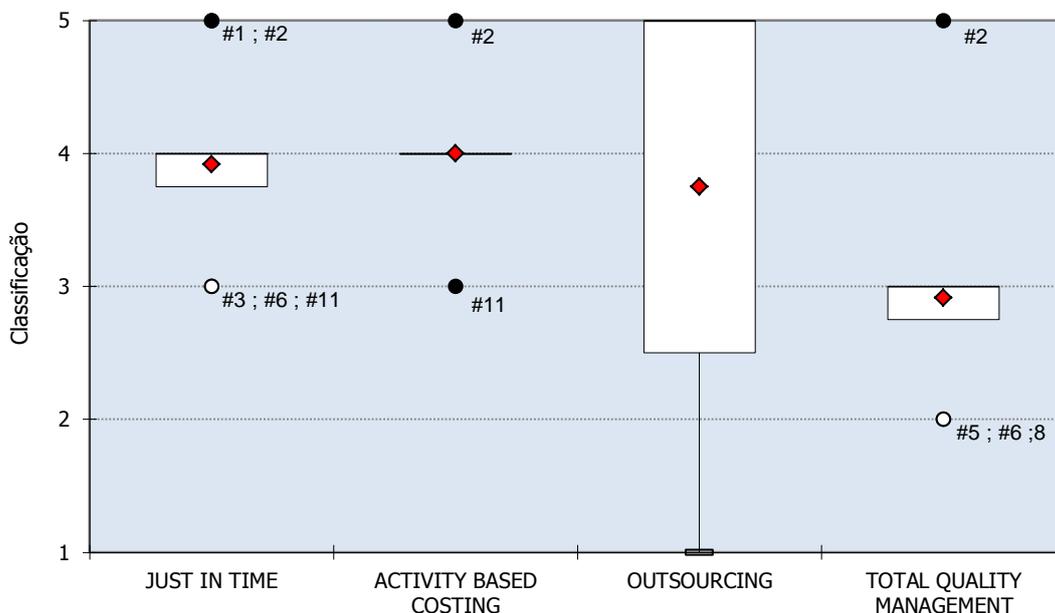


Figura 46: Box Plot - Avaliações Docentes Artigos Gestão - Wikipédia

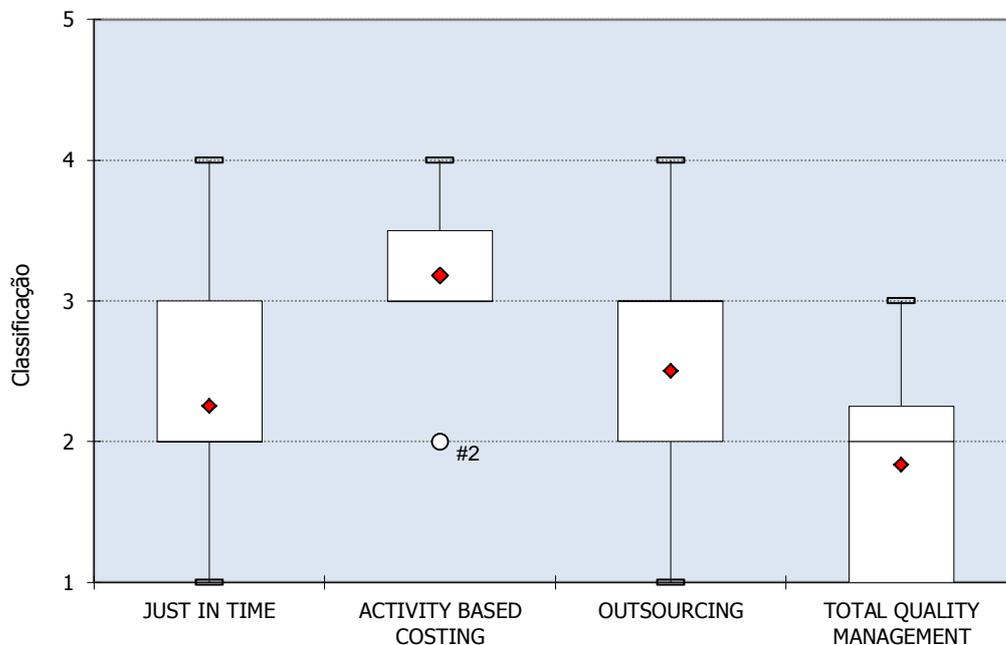


Figura 47: Box Plot - Avaliações Docentes Artigos Gestão - Britânica

Analisando os *Box Plots* (Figura 46 e Figura 47), constata-se que as notas atribuídas aos artigos da Wikipédia, à exceção do terceiro tema (*Outsourcing*), estão bastante concentradas, contendo no entanto um número relevante de *outliers* para a dimensão da amostra. Nas avaliações dos artigos da Britânica, a variação dentro da mesma pergunta é um pouco maior, mas aproximadamente similar em todos os temas.

Analisemos agora as diferenças entre as classificações da Wikipédia e da Britânica:

Docentes	JIT	ABC	OUT	TQM
Média	1,7	0,8	1,3	1,1
Mediana	1,5	1,0	2,0	1,0
Moda	1,0	1,0	2,0	0,0
Desvio padrão	1,2	0,8	2,2	1,2
Desvio Médio Absoluto da Média	1,0	0,6	2,0	0,9
Desvio Médio Absoluto da Mediana	1,0	0,6	1,8	0,9
Mediana Desvio Absoluto da Mediana	0,5	0,5	1,5	1,0
Incerteza estatística	0,4	0,2	0,6	0,3

Tabela 13: Estatística Descritiva – Diferença Avaliações Artigos Gestão

Como se pode constatar (Tabela 13), as medidas de dispersão são substancialmente maiores no tema *Outsourcing*, o que mostra que estamos na presença de um par de artigos que não gerou consenso, estando as opiniões radicalmente divididas entre os avaliadores, como facilmente se constata observando os histogramas seguintes.

Para uma melhor compreensão do significado dos valores das diferenças utilizados nos histogramas, propõe-se a seguinte classificação (Tabela 14):

Qualidade Informação da Wikipédia quando comparada com a Britânica		
Diferença entre as notas	4	Incomparavelmente Melhor
	3	Muitíssimo Melhor
	2	Muito Melhor
	1	Melhor
	0	Equivalente
	-1	Pior
	-2	Muito Pior
	-3	Muitíssimo Pior
	-4	Incomparavelmente Pior

Tabela 14: Interpretação Qualitativa da Escala

Apresentam-se seguidamente os *Box Plots* relativos às diferenças de classificação, para as quatro perguntas (Figura 48):

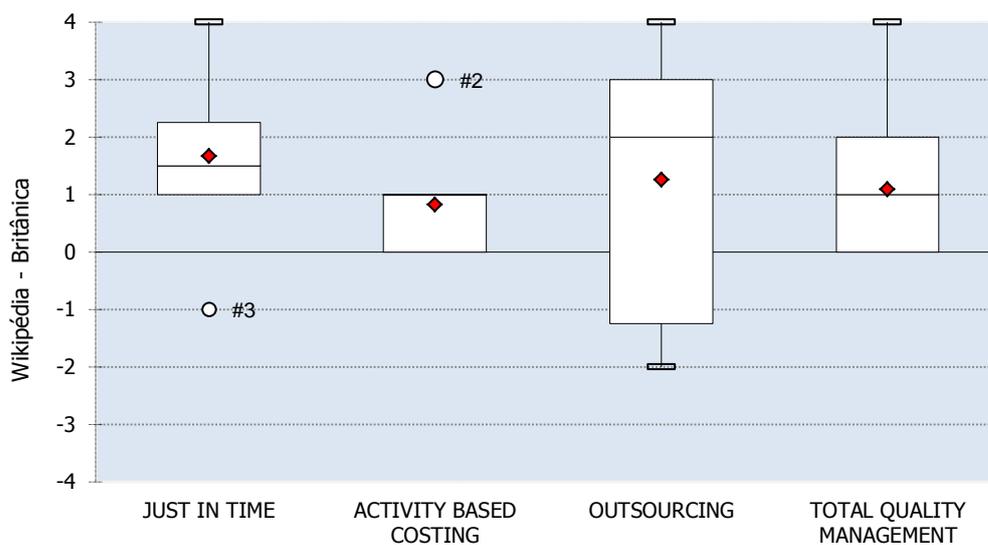


Figura 48: Box Plot – Diferença Avaliação Artigos Gestão

e as frequências acumuladas relativas às diferenças entre as avaliações de cada par, para os quatro temas em análise:

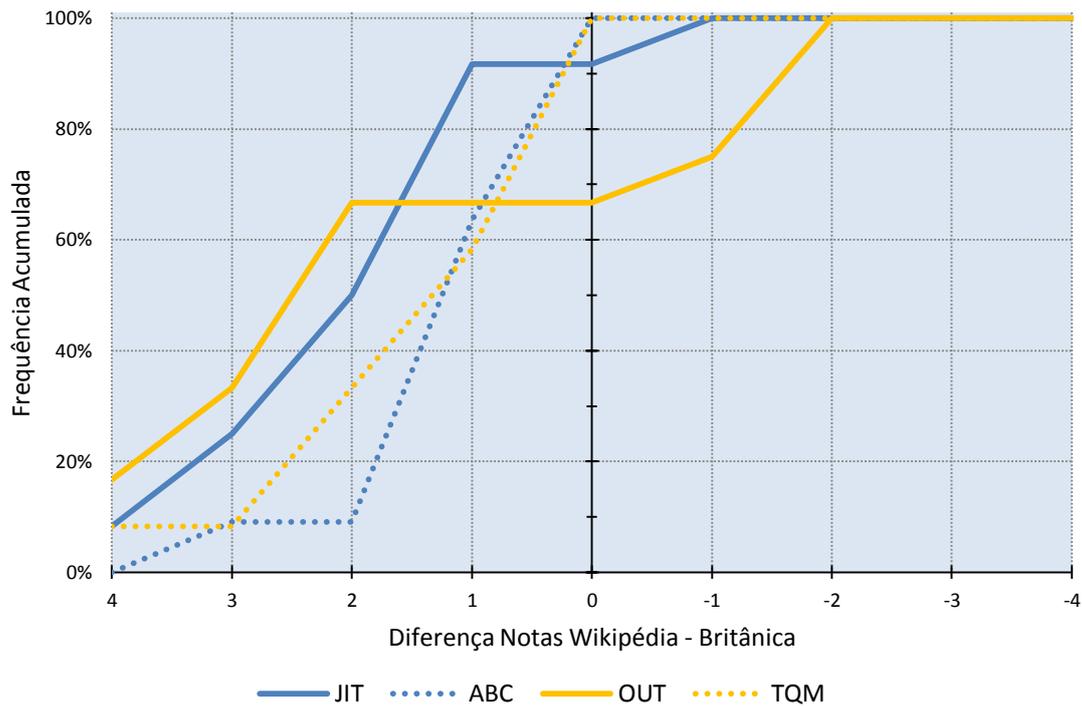


Figura 49: Diferença Acumulada Notas Wikipédia - Britânica

Como se pode ver na Figura 49, se excluirmos o tema *Outsourcing*, mais de 90% das respostas vão no sentido de a qualidade dos artigos da Wikipédia ser superior ou igual à Britânica. Analisando os *Box Plots*, poderíamos também ser tentados a retirar alguns dos *outliers*, ou seja, retirar do estudo as respostas dos avaliadores que são diametralmente opostas às da maioria, e cuja avaliação foi, quiçá, enviesada pela utilização da proveniência dos artigos na sua classificação.

No entanto, e como as avaliações com apenas um avaliador também terão de certeza casos destes, optámos por não retirar nenhum dos temas nem nenhum dos avaliadores e contabilizar todas as avaliações.

De modo a averiguar a possibilidade de aplicar testes paramétricos às notas atribuídas pelos avaliadores aos artigos, testaram-se as hipóteses:

- H_0 : a variável é normalmente distribuída;
- H_1 : a variável não é normalmente distribuída.

Utilizando o teste de *Shapiro-Wilk*, obtivemos os seguintes resultados:

	Wikipédia		Britânica	
	Shapiro-Wilk	Probabilidade	Shapiro-Wilk	Probabilidade
JIT	0,809	0,012	0,884	0,099
ABC	0,565	0,000	0,739	0,002
OUT	0,692	0,001	0,867	0,060
TQM	0,706	0,001	0,802	0,010

Tabela 15: Teste Shapiro-Wilk – Avaliação Artigos Gestão

A probabilidade associada ao teste de normalidade é inferior a 0,05 em todas as variáveis, à exceção dois casos sublinhados, rejeitando-se a hipótese nula em todos os casos à exceção destes dois (Tabela 15). Como em nenhum par se verifica que ambas as variáveis são normalmente distribuídas, será necessário optar por um teste não paramétrico. Deste modo, optou-se pelo teste de *Wilcoxon* para amostras emparelhadas, tendo sido definido o seguinte teste de hipóteses:

- $H_0: \mu_1 = \mu_2$ (as médias são iguais);
- $H_1: \mu_1 \neq \mu_2$ (as médias são diferentes).

Tendo sido obtidos os seguintes resultados (Tabela 16):

	Média		Z	Probabilidade (bicaudal)
	Wikipédia	Britânica		
JIT	3,9	2,3	-2,829	0,005
ABC	4,0	3,2	-2,530	0,011
OUT	3,8	2,5	-1,846	0,065
TQM	2,9	1,8	-2,401	0,016

Tabela 16: Teste Wilcoxon – Avaliação Artigos Gestão

Como todos os valores das probabilidades são inferiores a 0,05, à exceção do valor referente ao par de artigos sobre *Outsourcing*, rejeita-se a hipótese nula em todos os casos à exceção deste. Verifica-se, assim, que as diferenças entre as medianas encontradas são estatisticamente significativas para todos os pares de artigos, à exceção do referido par.

Apresentam-se de seguida os resultados obtidos para as avaliações dos alunos: (Tabela 17)

Discentes	JIT		ABC		OUT		TQM	
	W	B	W	B	W	B	W	B
Média	4,3	2,4	3,8	2,8	3,8	2,3	4,2	2,4
Mediana	4,0	2,0	4,0	3,0	4,0	2,0	5,0	2,0
Moda	5,0	2,0	4,0	3,0	4,0	2,0	5,0	2,0
Desvio padrão	0,7	0,9	0,8	0,7	0,6	0,8	1,1	1,1
Desvio Médio Absoluto da Média	0,6	0,7	0,6	0,5	0,5	0,6	0,9	0,9
Desvio Médio Absoluto da Mediana	0,4	0,6	0,5	0,6	1,3	0,7	0,4	0,7
Mediana Desvio Absoluto da Mediana	0,0	0,5	0,0	0,0	0,0	0,5	0,0	1,0
Incerteza estatística	0,2	0,2	0,1	0,2	0,5	0,3	0,2	0,2

Tabela 17: Estatística Descritiva Artigos Gestão

a que correspondem os *Box Plots* referentes às avaliações dos alunos (Figura 50 e Figura 51):

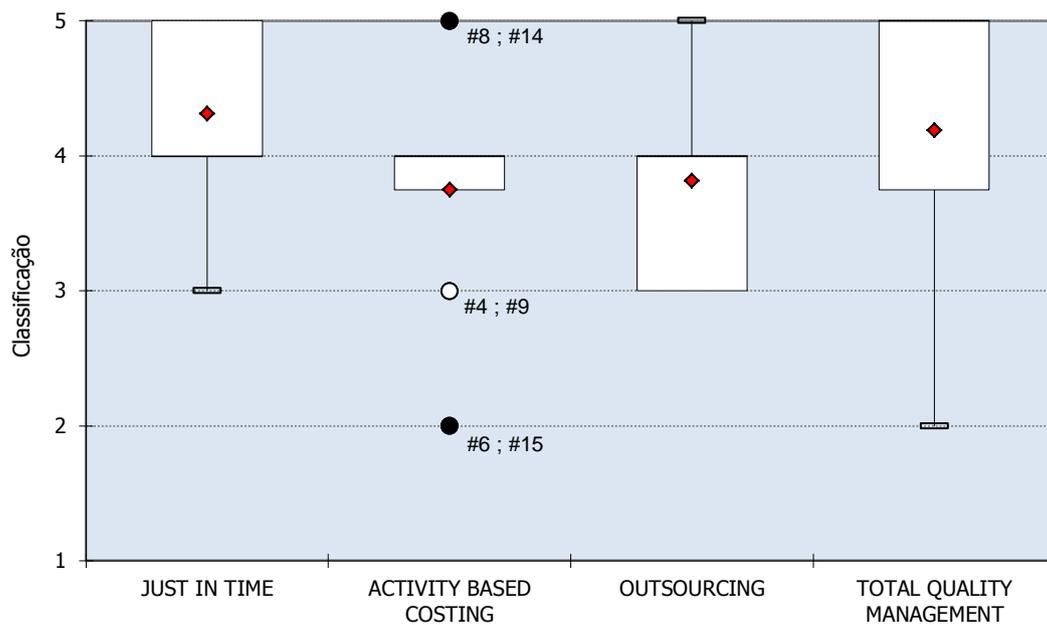


Figura 50: Box Plot - Avaliações Discentes Artigos Gestão - Wikipédia

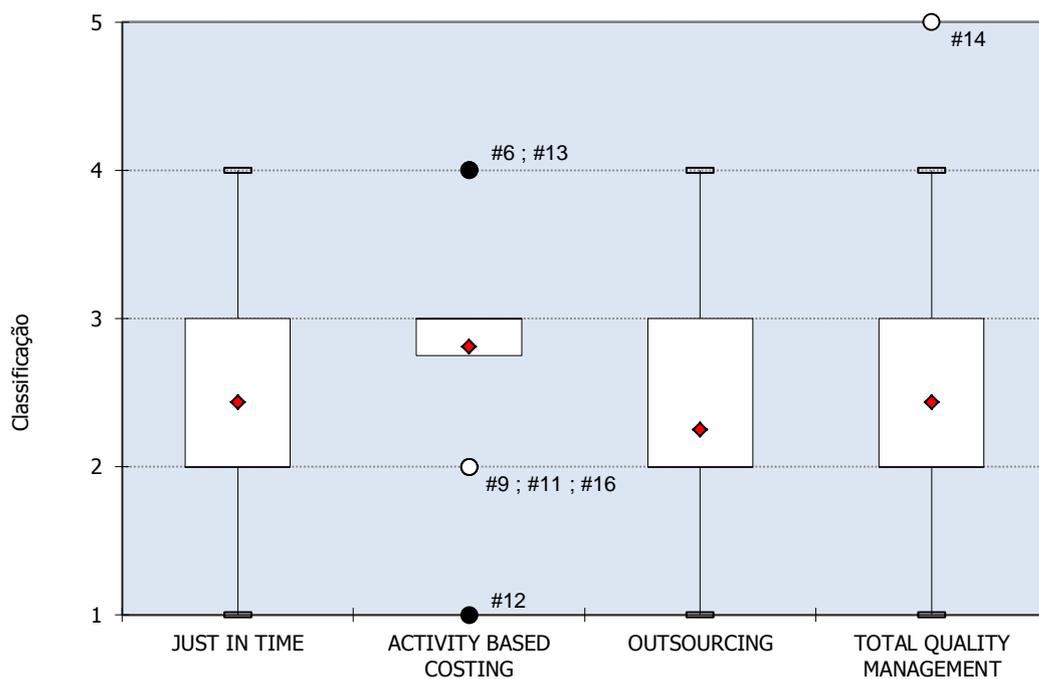


Figura 51: Box Plot - Avaliações Discentes Artigos Gestão - Britânica

Como se pode ver, junto dos alunos a questão relativa ao tema ABC originou um grande número de *outliers*, sendo no entanto estes originados devido a uma grande concentração de notas junto da classificação 4 na Wikipédia e 3 na Britânica, o que dá origem a um valor interquartil muito baixo e a consequente facilidade com que as

respostas que se afastam ligeiramente do valor da moda sejam classificadas como *outliers*.

Na Tabela 18 e na Figura 52, encontram-se os valores referentes à diferença entre as classificações atribuídas e o respectivo *Box Plot*:

Discentes	JIT	ABC	OUT	TQM
Média	1,9	0,9	1,6	1,8
Mediana	2,0	1,0	2,0	2,0
Moda	3,0	1,0	2,0	3,0
Desvio padrão	1,4	1,2	0,9	1,4
Desvio Médio Absoluto da Média	1,0	0,8	0,8	1,2
Desvio Médio Absoluto da Mediana	1,0	0,8	0,7	1,1
Mediana Desvio Absoluto da Mediana	1,0	1,0	0,5	1,0
Incerteza estatística	0,3	0,3	0,2	0,3

Tabela 18: Estatística Descritiva – Diferença Avaliações Artigos Gestão

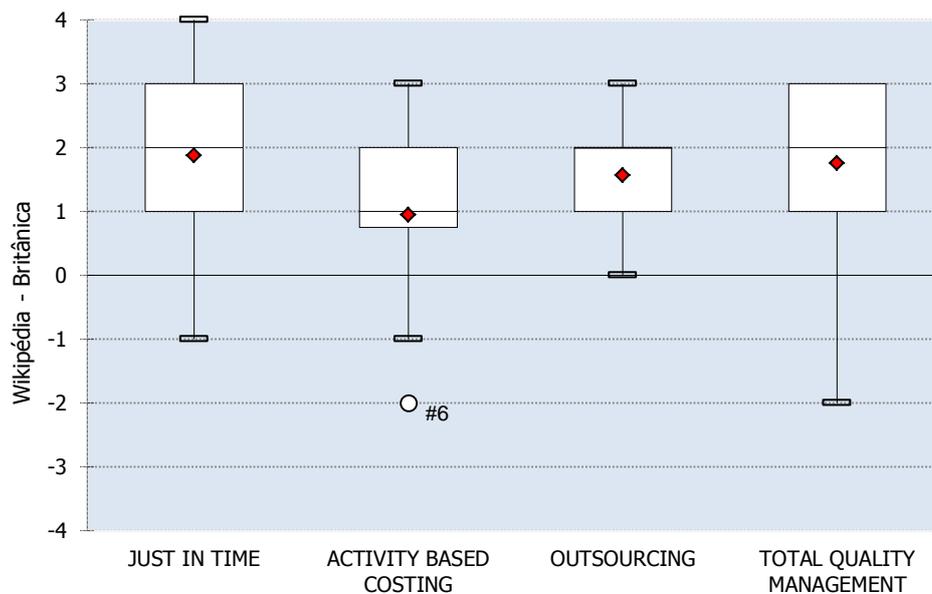


Figura 52: Box Plot – Diferença Avaliações Artigos Gestão

Como se pode ver, quando estudamos as diferenças, desaparece o efeito verificado anteriormente relativo ao tema ABC: os indicadores de variabilidade das respostas são menores e relativamente constantes ao longo de todos os pares de artigos.

De modo a averiguar a possibilidade de aplicar testes paramétricos, testaram-se as hipóteses:

- H_0 : a variável é normalmente distribuída;
- H_1 : a variável não é normalmente distribuída.

Utilizando o teste de *Shapiro-Wilk*, obtivemos os seguintes resultados (Tabela 19):

	Wikipédia		Britânica	
	Shapiro-Wilk	Probabilidade	Shapiro-Wilk	Probabilidade
JIT	0,788	0,002	0,888	0,051
ABC	0,791	0,002	0,814	0,004
OUT	0,794	0,002	0,850	0,014
TQM	0,743	0,001	0,827	0,006

Tabela 19: Teste Shapiro Wilk –Avaliações Artigos Gestão

A probabilidade associada ao teste de normalidade é inferior a 0,05 em todas as variáveis, à excepção do caso sublinhado, rejeitando-se a hipótese nula em todos os casos à excepção deste. Como em nenhum par se verifica que ambas as variáveis são normalmente distribuídas, há que optar por um teste não paramétrico; optámos, assim, pelo teste de *Wilcoxon* para amostras emparelhadas, tendo sido definido o seguinte teste de hipóteses:

- H_0 : $\mu_1 = \mu_2$ (as médias são iguais);
- H_1 : $\mu_1 \neq \mu_2$ (as médias são diferentes).

Os resultados encontram-se na Tabela 20.

	Média		Z	Probabilidade (bicaudal)
	Wikipédia	Britânica		
JIT	4,3	2,4	-3,239	0,001
ABC	3,8	2,8	-2,385	0,017
OUT	3,8	2,3	-3,270	0,001
TQM	4,2	2,4	-3,060	0,002

Tabela 20: Teste Wilcoxon – Avaliações Artigos Gestão

Como todos os valores das probabilidades são inferiores a 0,05, rejeita-se a hipótese nula em todos os casos. Verifica-se, assim, que as diferenças entre as medianas encontradas são estatisticamente significativas para todos os pares de artigos.

De modo a testar se existem diferenças significativas entre as respostas dadas pelos professores e os alunos para os quatro temas, iremos analisar as variáveis criadas a partir das diferenças entre as notas atribuídas à Wikipédia e à Britânica, começando por testar a sua normalidade, testando as hipóteses:

- H_0 : a variável é normalmente distribuída;
- H_1 : a variável não é normalmente distribuída.

Os resultados obtidos encontram-se na Tabela 21:

W-B	Docentes		Discentes	
	Shapiro-Wilk	Probabilidade	Shapiro-Wilk	Probabilidade
JIT	0,924	0,321	0,883	0,044
ABC	0,742	0,002	0,903	0,091
OUT	0,834	0,023	0,853	0,015
TQM	0,824	0,018	0,812	0,004

Tabela 21: Teste Shapiro-Wilk – Diferença Avaliação Artigos Gestão

A probabilidade associada ao teste de normalidade é inferior a 0,05 em todas as variáveis, à exceção dos casos sublinhados; logo, rejeita-se a hipótese nula em todos os casos exceptuando estes, concluindo-se que não há nenhum par de variáveis normalmente distribuídas.

Mais uma vez, será necessário aplicar testes não paramétricos para verificar se existem diferenças estatisticamente significativas entre as diferenças das avaliações dos docentes e discentes. Como desejamos comparar duas variáveis independentes entre si, utilizaremos o teste de *Mann-Whitney* quatro vezes, uma para cada tema.

Para tal, estipulamos as seguintes hipóteses:

- $H_0: \mu_1 = \mu_2$ (as médias são iguais);
- $H_1: \mu_1 \neq \mu_2$ (as médias são diferentes);

Foram obtidos os seguintes resultados (Tabela 22):

W-B	Média		Z	Probabilidade (bicaudal)
	Docentes	Discentes		
JIT	1,7	1,9	-0,695	0,487
ABC	0,8	0,9	-0,818	0,413
OUT	1,3	1,6	-0,291	0,771
TQM	1,1	1,8	-1,670	0,095

Tabela 22: Teste Mann-Whitney - Diferença Avaliação Artigos Gestão

Em todos os temas obtivemos valores da probabilidade superiores que 0,05, pelo que não se rejeita a hipótese nula e se conclui que não existe diferença significativa entre as médias, ou seja, não existem diferenças estatisticamente significativas entre as respostas dos professores e dos alunos em relação a nenhuma das quatro questões avaliadas.

Os alunos atribuíram, em todos os casos, diferenças de notas ligeiramente superiores em relação aos professores e os dois grupos consideraram que, em todas as questões avaliadas, a qualidade da informação na Wikipédia era significativamente superior à Britânica.

O valor médio da diferença entre as avaliações dos quatro pares de artigos da Wikipédia e da Britânica é de $1,2 \pm 0,4$ no caso dos docentes, e de $1,5 \pm 0,3$ nos discentes. Como medida de variabilidade, utilizou-se a incerteza estatística¹²⁶ que, como é calculada através do quociente entre o desvio padrão e a raiz quadrada da dimensão da amostra, tem tendência a diminuir com o aumento desta, ou seja, quantas mais medições dispomos de uma mesma medida, menor é a incerteza que temos quanto ao seu real valor.

Na Figura 53, podemos comparar as avaliações – em termos da proporção de avaliações em que a Wikipédia tem uma qualidade superior ou igual à Britânica – efectuadas pelos dois grupos, artigo a artigo. Em termos médios, as avaliações dos dois grupos são semelhantes, tendo os docentes atribuído 90% de pares de classificações em que a nota da Wikipédia era maior ou igual à da Britânica, e os discentes 92%.

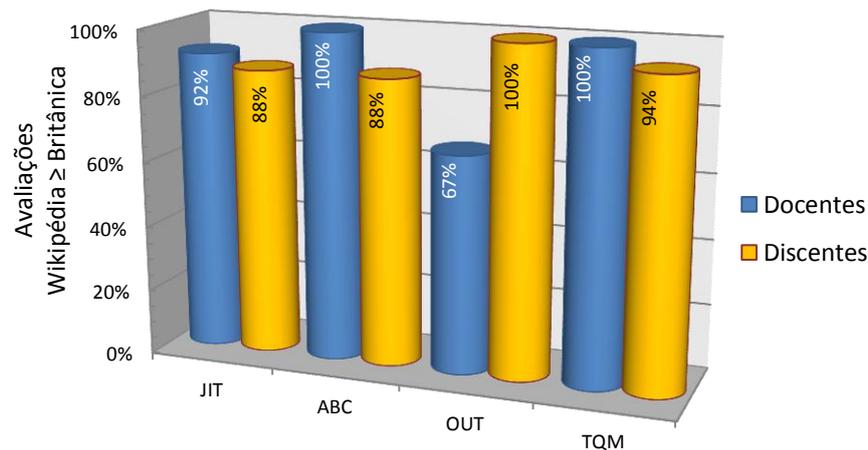


Figura 53: Valor Médio das Avaliações Docentes e Discentes

A avaliação efectuada pelos professores (colectivo de juízes) é assim, em termos médios, não significativamente diferente da avaliação dos alunos (jurados); curiosamente, o artigo em que a opinião dos docentes mais se dividiu – *Outsourcing* – foi aquele em que houve unanimidade por parte dos discentes, quiçá indiciando um erro ou omissão grave neste que apenas alguns conseguiram detectar.

¹²⁶ *Standard Error*

*If we knew what it was we were doing,
it would not be called research, would it?*

Albert Einstein (1941)

5.3. COMPARAÇÃO WIKIPÉDIA - BRITÂNICA

A amostra inicial de 245 pares de artigos foi dividida em quatro grandes grupos: Artes & Entretenimento, Ciência & Tecnologia, História & Sociedade e Viagens & Geografia, sendo também os artigos divididos consoante a área do conhecimento a que pertenciam de modo a serem distribuídos pelos avaliadores.

Foram encontrados avaliadores para todas as áreas do conhecimento constantes das três primeiras categorias, tendo sido avaliados todos os pares de artigos. Em relação ao grupo Viagens & Geografia, os 57 pares de artigos desta categoria são compostos por um artigo de História, um de Pintura, um de Arquitectura e 54 de Geografia.

Apesar de todos os esforços desenvolvidos, foi impossível encontrar avaliadores na área da Geografia dispostos a colaborar com o estudo, tendo sido apenas avaliados os 3 artigos que não são desta área. Os artigos de Geografia em causa são, na sua totalidade, compostos por descrições de cidades, regiões, países, montanhas e rios.

Para tentar obviar esta questão, foi testada a hipótese de utilizar as avaliações quantitativas destes – número de palavras e de imagens – cuja determinação não carece de avaliador e tentar, a partir destas, calcular uma avaliação qualitativa. Com base nos restantes artigos da amostra em que dispomos de ambos os tipos de avaliações, foram calculadas as matrizes de correlação entre o número de palavras (P), número de ilustrações (I) e a nota atribuída pelos avaliadores (N) para os artigos da Wikipédia (W) e da Britânica (B):

$$\begin{bmatrix} \mathbf{W} & P & I & N \\ P & 1 & & \\ I & 0,471 & 1 & \\ N & 0,345 & 0,205 & 1 \end{bmatrix} e \begin{bmatrix} \mathbf{B} & P & I & N \\ P & 1 & & \\ I & 0,345 & 1 & \\ N & 0,297 & 0,130 & 1 \end{bmatrix}$$

Infelizmente, as correlações obtidas são baixas, pelo que tentámos agrupar os dados calculando a média de palavras (Figura 54) e de ilustrações (Figura 55) para os artigos que têm a mesma nota.

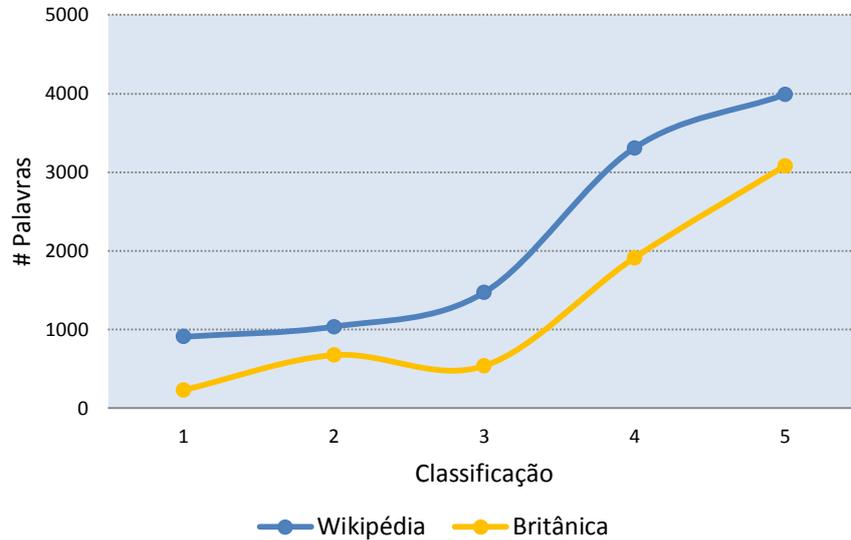


Figura 54: Média do número palavras por artigo agrupadas pela classificação obtida

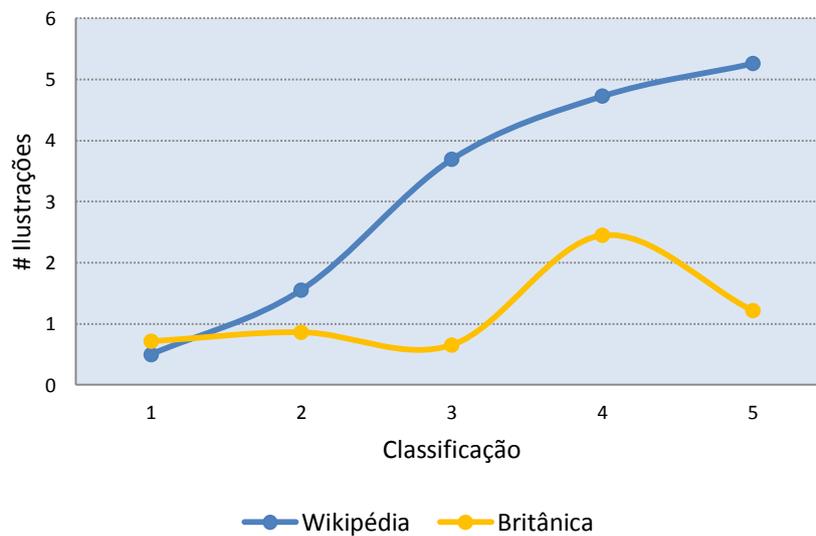


Figura 55: Média do número ilustrações por artigo agrupadas pela classificação obtida

Introduzindo esta simplificação, conseguimos aperceber-nos da tendência de as notas atribuídas crescerem com o número de palavras e de ilustrações, embora esta relação seja mais forte no caso dos artigos da Wikipédia. Tentámos, de seguida, averiguar se esta relação seria mais forte se utilizássemos a diferença ou a razão entre as avaliações, em substituição das classificações individuais.

Assim, definiram-se as variáveis: diferença entre as notas (N_{W-B}), razão entre as notas ($N_{W/B}$), diferença entre o número de palavras (P_{W-B}), razão entre o número de palavras ($P_{W/B}$) e diferença¹²⁷ entre o número de ilustrações (I_{W-B}), e calculou-se a matriz de correlação (Tabela 23):

	N_{W-B}	P_{W-B}	I_{W-B}	$P_{W/B}$	$N_{W/B}$
N_{W-B}	1				
P_{W-B}	0,395	1			
I_{W-B}	0,101	0,304	1		
$P_{W/B}$	0,347	0,418	0,147	1	
$N_{W/B}$	0,849	0,277	0,088	0,384	1

Tabela 23: Matriz Correlação 245 Artigos

A única correlação interessante é entre N_{W-B} e $N_{W/B}$ e não tem utilidade¹²⁸. Excluindo a correlação entre a diferença e a razão do número de palavras por artigo, o melhor resultado a seguir é entre a diferença entre as notas e a diferença entre as palavras (0,395), valor superior ao obtido para as correlações individuais obtidas anteriormente para as variáveis correspondentes para a Wikipédia (0,345) e Britânica (0,297).

De modo a tentar melhorar os resultados obtidos, iremos calcular a média do número de palavras e ilustrações dos artigos que têm a mesma diferença entre notas, visto termos constatado que as diferenças conduzem a melhores correlações do que as razões. Os resultados encontrados podem ser visualizados na Figura 56.

¹²⁷ Não foi também utilizado o quociente de ilustrações porque muitos artigos não as têm, o que conduz à anulação do denominador e consequente indeterminação da variável.

¹²⁸ O objectivo é encontrar uma correlação entre a diferença de notas e algo.

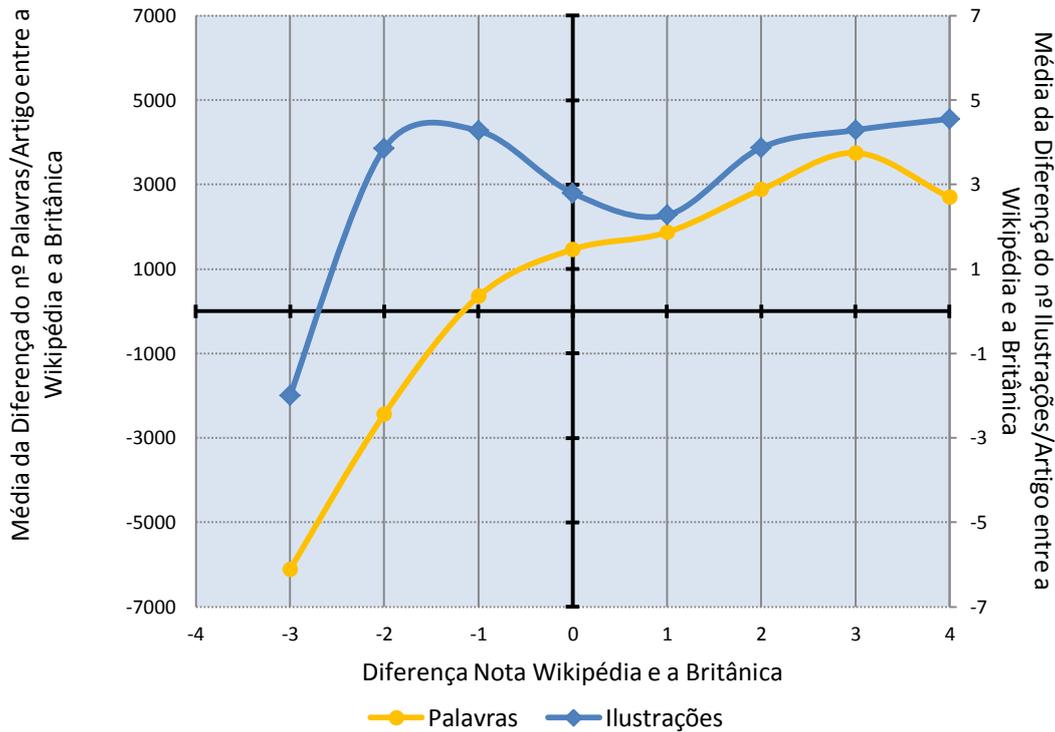


Figura 56: Nº Médio de Palavras e de Ilustrações por Artigo Agrupados pela Diferença Notas

Como se pode ver, a diferença entre o número de ilustrações não parece apresentar nenhuma correlação perceptível; no entanto, a diferença entre o número de palavras parece indicar a possibilidade de se virem a obter bons resultados.

Efectuaram-se regressões lineares, simples e múltipla, entre as diversas variáveis e obtiveram-se os seguintes coeficientes de correlação: $F(P_{W-B}, I_{W-B}) = N_{W-B}$ (0,889), $G(P_{W-B}) = N_{W-B}$ (0,895) e $H(I_{W-B}) = N_{W-B}$ (0,624). O melhor resultado obtido é aquele a que se chega utilizando apenas a diferença do número de palavras como variável dependente.

Na posse desta informação, foram ensaiadas diversas regressões não lineares, tendo o melhor resultado obtido (0,941) sido dado pela função¹²⁹:

$$F(x) = 8 \times 10^{-8}x^2 + 0,0009x - 0,8027;$$

¹²⁹ Em que $F(x) = N_{W-B}$ e $x = P_{W-B}$

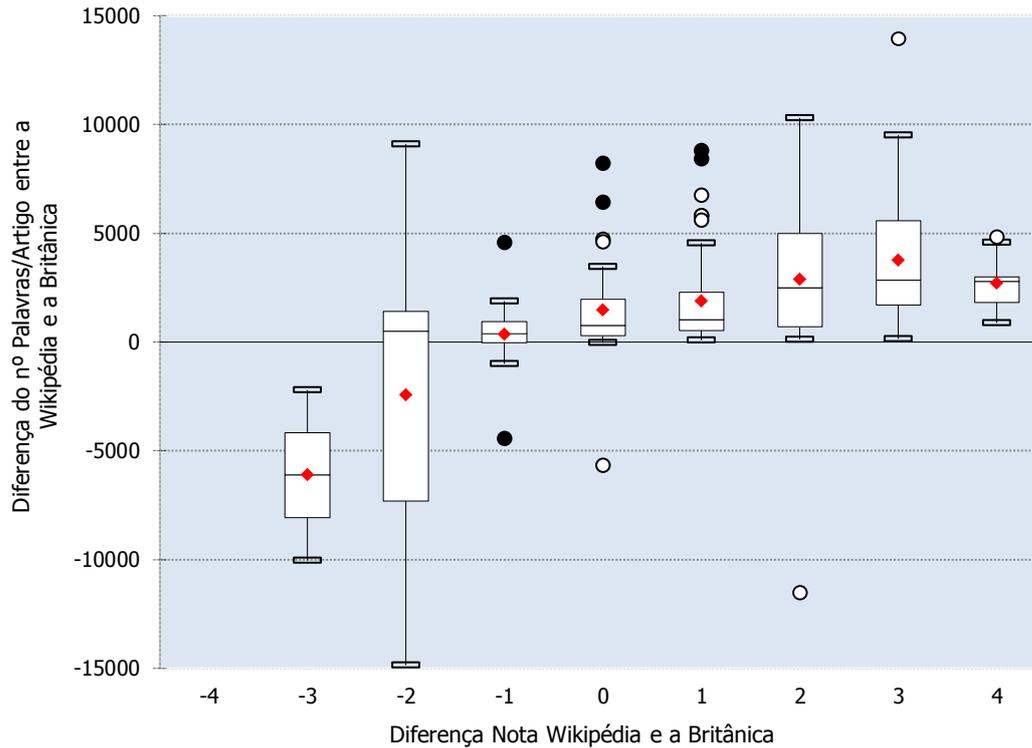


Figura 58: Box Plot diferenças entre palavras/artigo agrupadas pela diferença entre notas

No entanto, é expectável – pelos resultados obtidos – que a diferença média de notas dos artigos em questão¹³⁰ seja de sinal positivo e da mesma ordem de grandeza do que a diferença encontrada para os outros grupos; não podemos, no entanto, afirmar taxativamente que assim seja.

Apresentamos de seguida os resultados globais (Figura 59), obtidos nas avaliações dos artigos, em que se constata que em 89% dos pares de artigos avaliados a Wikipédia teve uma classificação igual ou superior à da Britânica.

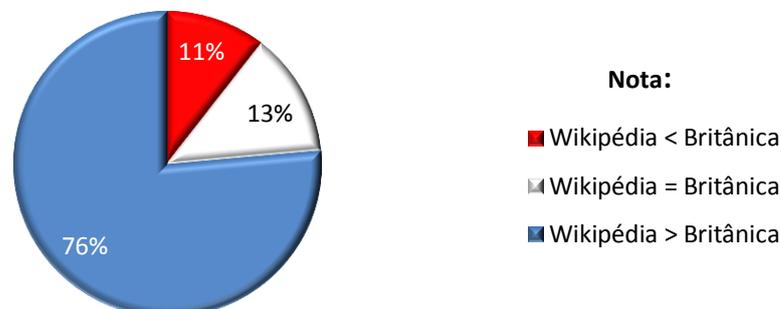


Figura 59: Diferença Avaliações Wikipédia e Britânica

¹³⁰ Os artigos de Geografia que não foram avaliados.

Na Figura 60, a área das esferas é proporcional ao número de avaliações que atribuíram o mesmo binómio de classificações (W,B)¹³¹. Das 25 combinações possíveis, apenas os binómios (1,1), (1,3) e (1,5) não foram utilizados pelos avaliadores.

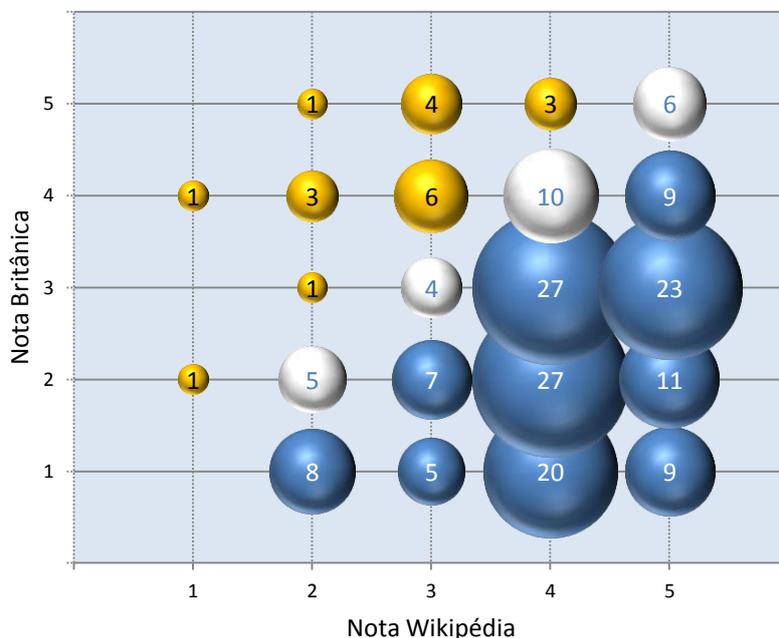


Figura 60: Avaliações Agrupadas por Binómio de Notas Iguais

Na Figura 61, podemos observar o *Box Plot* das notas atribuídas à Wikipédia e à Britânica, enquanto nas Figura 62 se encontram os histogramas correspondentes.

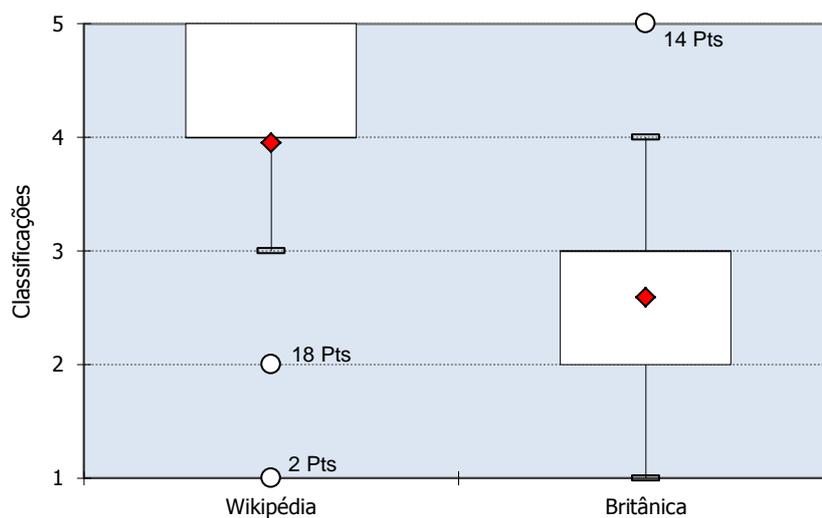


Figura 61: Box Plot – Avaliações Wikipédia e Britânica

¹³¹ A título de exemplo, o binómio (4,2) corresponde à atribuição de uma nota de 4 valores para a Wikipédia (abscissas) e de 2 valores à Britânica (ordenadas) e ocorreu 27 vezes.

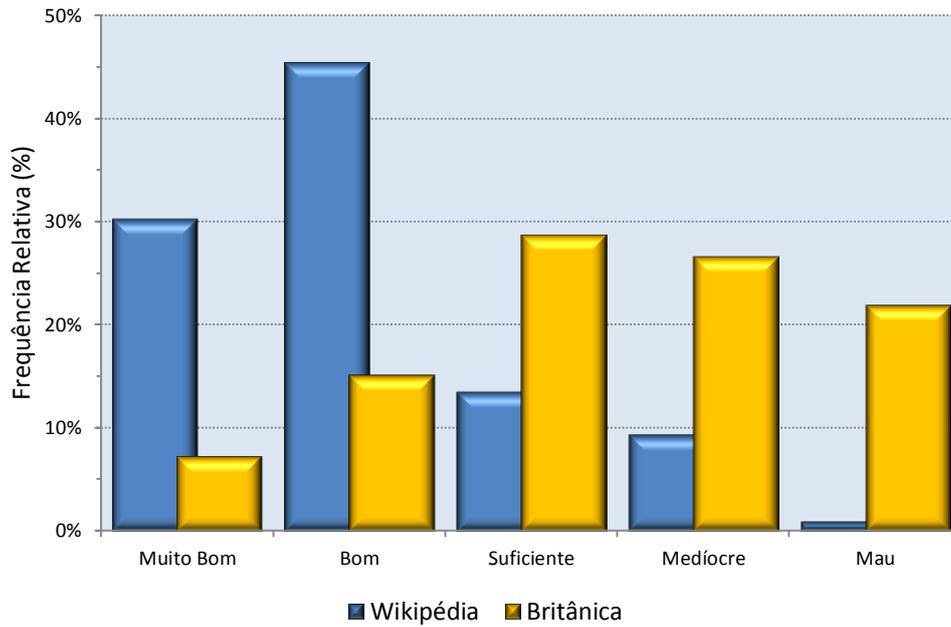


Figura 62: Histogramas das Avaliações dos Artigos

Analisemos agora a diferença entre as classificações (Figura 63 e Figura 64):

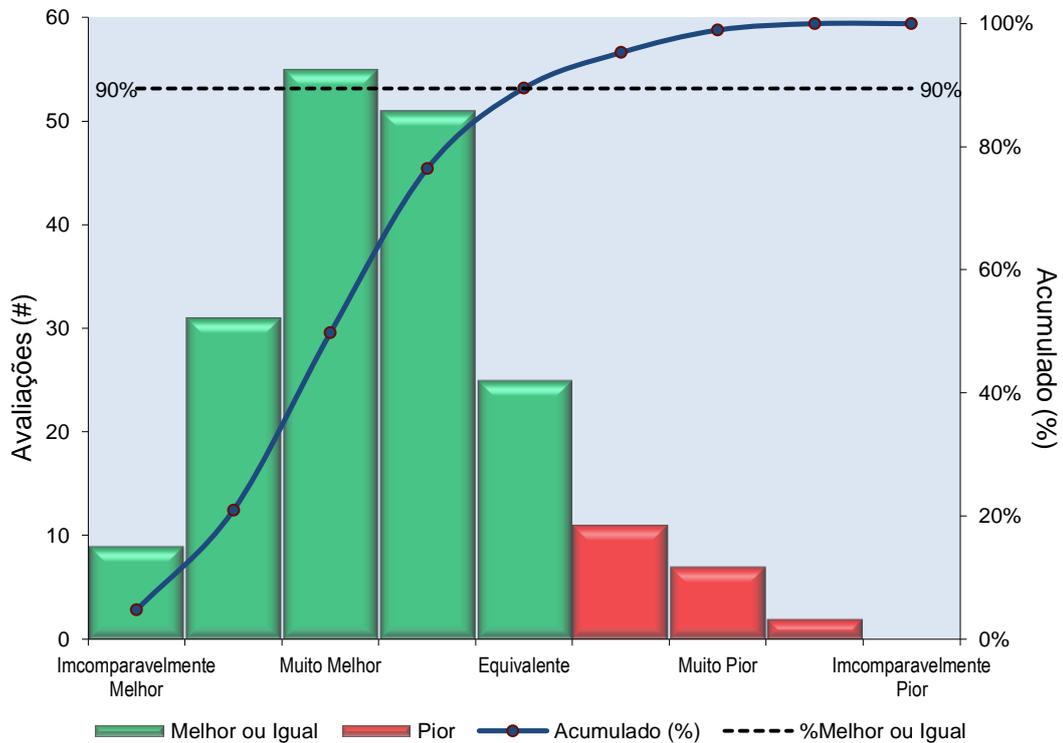


Figura 63: Qualidade Informação da Wikipédia quando comparada com a Britânica

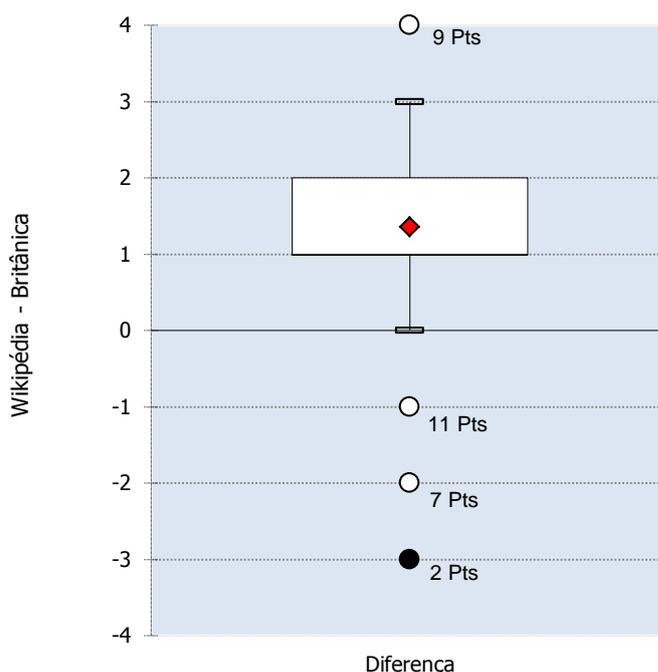


Figura 64: Box Plot – Diferença entre Avaliações

Iremos de seguida testar se as diferenças encontradas são estatisticamente significantes. Para tal, vamos começar por testar a normalidade das notas atribuídas à Wikipédia e à Britânica, testando as hipóteses:

- H_0 : a variável é normalmente distribuída;
- H_1 : a variável não é normalmente distribuída.

Foi efectuado o teste de normalidade, com os seguintes resultados (Tabela 24):

	Wikipédia		Britânica	
	Shapiro-Wilk	Probabilidade	Shapiro-Wilk	Probabilidade
Notas	0,834	0,000	0,902	0,000

Tabela 24: Teste Shapiro-Wilk – Avaliações Wikipédia e Britânica

A probabilidade associada ao teste de normalidade é inferior a 0,05 em todas as variáveis; logo, rejeita-se a hipótese nula nos dois casos, concluindo-se que nenhuma das variáveis é normalmente distribuída.

Iremos, pois, aplicar um teste não paramétrico para verificar se existem diferenças estatisticamente significativas entre as variáveis. Aplicou-se o teste de *Wilcoxon* para amostras emparelhadas, usando o seguinte teste de hipóteses:

- $H_0: \mu_1 = \mu_2$ (as médias são iguais);
- $H_1: \mu_1 \neq \mu_2$ (as médias são diferentes).

Foram obtidos os seguintes resultados (Tabela 25):

Notas	Média		Z	Probabilidade (bicaudal)
	Wikipédia	Britânica		
	4,0	2,6	-9,235	0,000

Tabela 25: Teste de Wilcoxon – Avaliações Wikipédia e Britânica

Como o valor da probabilidade é inferior a 0,05, rejeita-se a hipótese nula e conclui-se que existe diferença significativa entre as médias, ou seja, as diferenças encontradas são estatisticamente significativas. Se analisarmos os valores médios das avaliações divididas pelas categorias anteriormente apresentadas, temos (Figura 65):

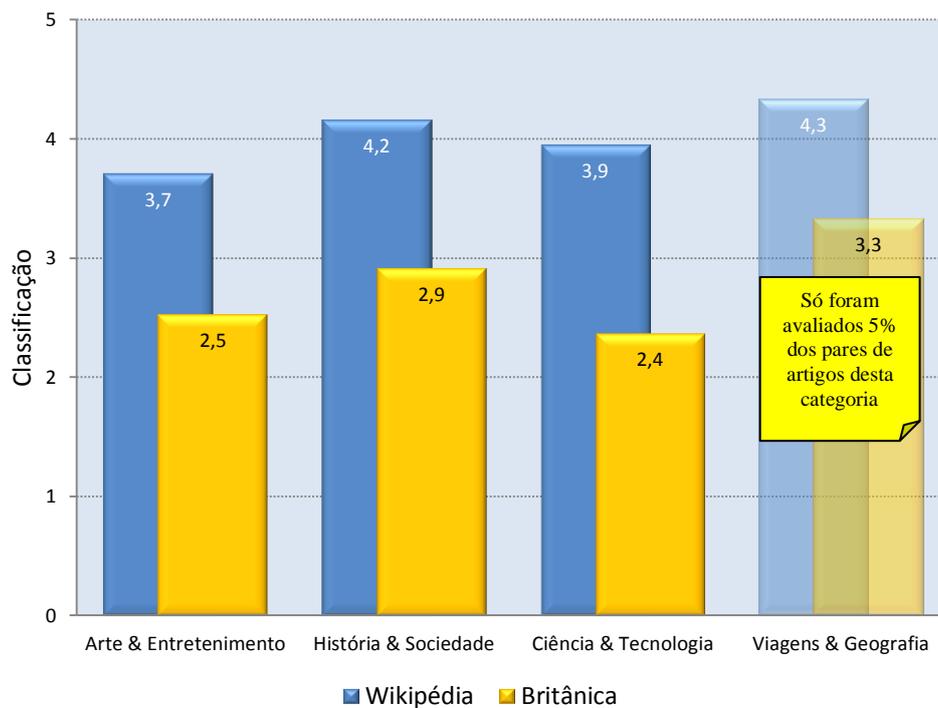


Figura 65: Média das Avaliações Wikipédia e Britânica Divididas por Categorias

Ignorando a categoria *Viagens & Geografia*, que não foi avaliada na totalidade, observa-se que as diferenças encontradas se repartem pelas restantes categorias, sendo a categoria *Ciência & Tecnologia* aquela que apresenta a maior¹³² diferença média entre as avaliações, enquanto a *Arte & Entretenimento* apresenta a menor¹³³.

Saliente-se que a avaliação negativa – menor que três – que apresentam todas as categorias da Britânica se deve seguramente ao facto de estar a ser avaliada em conjunto com outra enciclopédia e as classificações de um artigo serem feitas em comparação com o seu homólogo.

Se considerarmos a média (+1,4) e a mediana (+1,0) das diferenças entre as avaliações da Wikipédia e da Britânica e calcularmos as médias das medidas de variabilidade da diferença entre as avaliações efectuadas pelos docentes, no estudo dos 50 Princípios de Gestão poderemos ter uma aproximação do resultado que obteríamos se tivéssemos utilizado, também aqui, 12 avaliadores¹³⁴ por artigo.

Os valores absolutos das medidas de variabilidade do estudo dos 50 Princípios de Gestão foram convertidos em variações relativas em relação à média e à mediana das diferenças deste estudo, sendo em seguida transformadas de novo em variabilidades absolutas, utilizando a média e a mediana das diferenças das notas atribuídas a todos os artigos.

Esta previsão parte da hipótese de que a variabilidade das avaliações efectuadas pelos docentes de Gestão é similar à dos outros hipotéticos grupos de avaliadores de todas as outras áreas de conhecimento representadas na amostra recolhida.

¹³² 1,5 valores,

¹³³ 1,2 valores,

¹³⁴ Teria sido necessário efectuar 5.880 avaliações.

Na Figura 66, podemos ver, na escala de variação possível (-4 , 4), o modo como as diferentes medidas de variabilidade afectam o valor da diferença média entre as classificações atribuídas à Wikipédia e à Britânica:

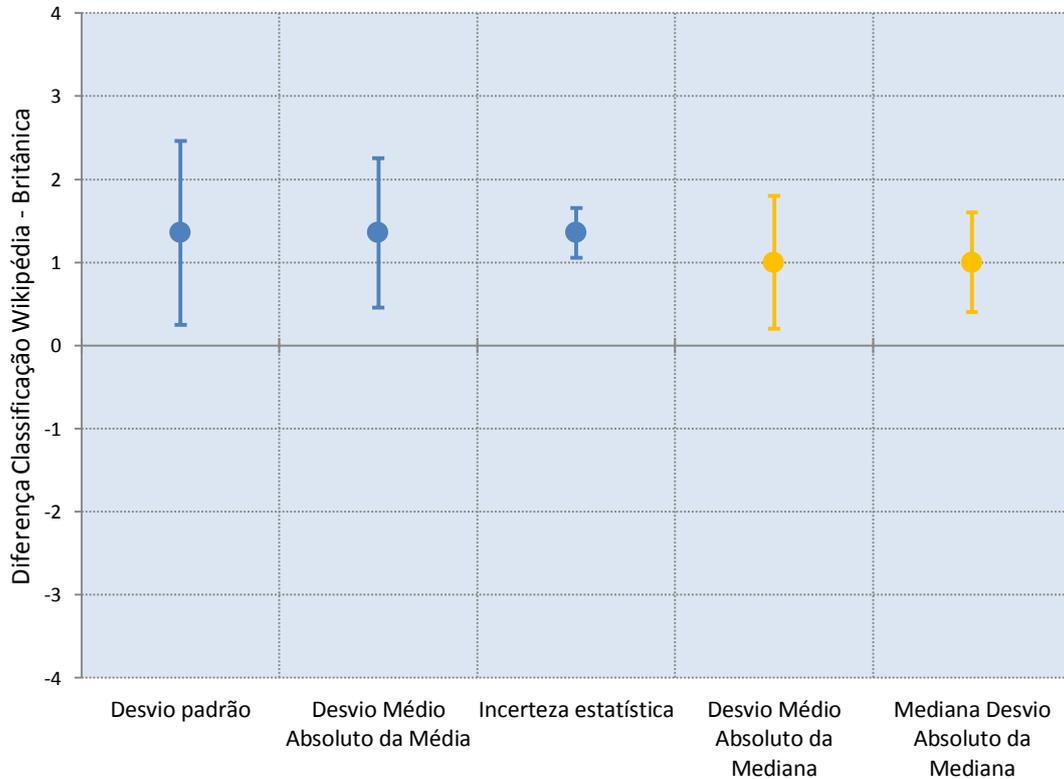


Figura 66: Variabilidade da Diferença Avaliações combinada com as Variabilidades Calculadas no Estudo dos Artigos de Gestão.

Seja qual for a medida de variabilidade escolhida, o resultado do intervalo é sempre positivo, ou seja, constata-se que, em média, as classificações atribuídas à Wikipédia são sempre superiores às da Britânica, mesmo que cada par de artigos tenha sido avaliado por múltiplos avaliadores.

*La verdad tiene muy pocos amigos y los
muy pocos amigos que tiene son suicidas.*

Antonio Porchia (1964), *Voces*

6. CONSIDERAÇÕES FINAIS

Neste capítulo, revisitam-se os temas centrais abordados na tese acrescentando-lhes a nossa interpretação pessoal, discutem-se os resultados obtidos (6.1), apresentam-se as conclusões a que chegámos e sugerem-se trabalhos futuros que poderão aumentar o conhecimento sobre o tema estudado (6.2).

6.1 REFLEXÕES E DISCUSSÃO DE RESULTADOS

A invenção do telescópio na Holanda no século XVII permitiu a Galileu Galilei observar as luas de Júpiter e o seu movimento de translação em torno deste, o que não é possível a olho nu. A descoberta de que as luas orbitam outro corpo celeste que não a Terra contribuiu decisivamente para o fim do modelo ptolemaico do universo e o abandono da teoria geocêntrica (Burke, 1995). Os avanços tecnológicos são, quase sempre, precursores do desenvolvimento científico, especialmente quando tornam possível realizar algo que antes do seu aparecimento era impossível.

No século XX, a comunicação biunívoca (um para um) era suportada sobretudo pelo correio de uma forma assíncrona e pelo telefone em tempo real. Para comunicar com muitas pessoas ao mesmo tempo (um para muitos), estavam disponíveis a televisão e a rádio em tempo real e os registos de som, imagem e palavras (livros e jornais) em diferido, sendo no entanto esta comunicação apenas realizada num único sentido (unívoca). A evolução tecnológica – em especial a massificação da internet – possibilitou a comunicação, em tempo real, de muitos para muitos, congregando ao mesmo tempo todos os tipos de *media* num único local. Outra das consequências desta tecnologia foi a diminuição substancial dos custos de produção e de distribuição de conteúdos, o que permitiu que todos nos transformássemos em potenciais criadores.

A conjugação destes dois factores – preço de criar e distribuir conteúdos e possibilidade de comunicação de muitos para muitos em tempo real – criou um terreno fértil para o

surgimento de projectos e grupos reunidos em torno de interesses comuns, em que a partilha e a discussão não estavam limitadas por barreiras geográficas, assistindo-se a uma explosão na criação de conteúdos de todos os tipos. Do mesmo modo que o arranjo cristalográfico distinto dos mesmos átomos de carbono permite obter minerais tão diferentes como a grafite e o diamante, também a qualidade das criações da *Vox populi* é extremamente variável, tendo no entanto duas características que, em termos evolucionários, geralmente produzem bons resultados – quantidade e diversidade.

Em 1995, Ward Cunningham¹³⁵ cria a primeira versão de *software* colaborativo, permitindo tirar partido da infra-estrutura da internet para agregar os esforços de criação individuais em criações de grupos que se auto-organizam em torno de um objectivo comum. As características intrínsecas deste ambiente permitem suportar as condições que James Surowiecki (2004) apresenta como necessárias ao comportamento inteligente de um grupo – diversidade de opinião, independência, descentralização e a existência de um mecanismo de agregação. Estavam criadas as condições para a colaboração em massa poder prosperar.

Uma possível caracterização das pessoas que se voluntariam para trabalhar gratuitamente neste tipo de projectos, e porque o fazem, pode emergir da análise conjunta das teorias de Daniel Pink (2005, 2009) e Clay Shirky (2008, 2010).

Daniel Pink sugere que o prazer retirado pela realização de uma tarefa, *per se*, é para muitas pessoas motivação suficiente para a sua realização, especialmente se esta não tiver um carácter de obrigatoriedade, permitir a demanda da excelência num determinado campo e tiver um propósito nobre. Para que isto aconteça, é necessário que estes indivíduos se sintam confortáveis com as recompensas externas – nomeadamente as monetárias – que auferem na realização de outras actividades.

Clay Shirky discute a existência de um excedente cognitivo na nossa sociedade devido ao aumento das regalias sociais e do nível médio de literacia, o que conduziu a um aumento do tempo livre disponível por parte de pessoas com um nível de conhecimentos substancial. Estas pessoas estarão a trocar a actividade de consumo passivo de conteúdos pela participação em projectos de criação de conteúdos que lhes

¹³⁵ Fonte: http://en.wikipedia.org/wiki/Ward_Cunningham ou <http://www.britannica.com/EBchecked/topic/1192883/Ward-Cunningham>

permitam aplicar os seus conhecimentos e desenvolver o seu potencial em actividades cujo propósito as inspire.

Assim, e em termos gerais, estamos a falar de pessoas com um nível de conhecimentos apreciável, que têm uma actividade que lhes permite satisfazer as suas necessidades materiais e que preferem despende parte do seu tempo livre a utilizar as suas potencialidades cognitivas, eventualmente não totalmente utilizadas na sua actividade profissional, apenas pelo prazer intrínseco de o fazer e não com o fito de obter recompensas externas adicionais.

Em 1993, a Microsoft criou a Encarta¹³⁶, uma enciclopédia multimédia digital disponível através da aquisição de um conjunto de discos ópticos e, mais tarde, acessível através da *Web* mediante o pagamento de uma subscrição anual. Para tal, criou uma estrutura de profissionais que asseguravam a criação, edição e gestão do projecto.

No início de 2001, entra no mercado outra enciclopédia, a Wikipédia, acessível gratuitamente *online*, cujos conteúdos são criados e editados por milhares de amadores voluntários não remunerados, não enquadrados numa estrutura hierárquica, que contribuem para o projecto quando e como muito bem entendem.

Poucos ousariam prever que, passados oito anos, a Encarta tivesse encerrado as portas e a Wikipédia fosse a fonte enciclopédica mais comumente utilizada a nível mundial. Estes factos, apesar de irrefutáveis, não nos permitem concluir que a qualidade de uma é superior à da outra: a popularidade de um produto não tem necessariamente de estar relacionado com a qualidade deste, e os factos ocorridos podem advir da relação qualidade / preço da Wikipédia – que, por ser gratuita, é infinita – bastando para tal que a sua qualidade seja não nula.

O alfa e o ómega desta tese é a tentativa de provar a credibilidade – e não a popularidade – da colaboração em massa enquanto pedra angular do modelo de negócio comumente denominado por *open source*. Para o demonstrar, realizou-se um estudo comparativo entre a qualidade dos artigos da Wikipédia e da Enciclopédia Britânica, tendo sido dado particular relevo à metodologia utilizada de modo a que os resultados obtidos fossem os mais fidedignos e abrangentes possíveis.

¹³⁶ Fonte: <http://en.wikipedia.org/wiki/Encarta>

Alguns estudos foram efectuados tentando aferir a qualidade da Wikipédia, amiudadamente efectuando comparações com outras enciclopédias, tendo no entanto sido criticados por utilizar amostras pouco representativas, metodologias de cálculo da qualidade baseados em critérios pouco abrangentes ou escolher obras especializadas dirigidas a nichos de mercado na sua comparação.

O estudo efectuado por Giles (2005), publicado na revista *Nature*, é aquele que mais se aproxima do nosso em termos de metodologia, tendo Giles chegado à conclusão de que a qualidade de ambas as enciclopédias era semelhante, com uma ligeira vantagem para a Britânica – isto em 2005. A diferença nos resultados obtidos poderá dever-se ao facto de a qualidade da Wikipédia ter vindo a aumentar e, também, à escolha de uma amostra aleatória de grande dimensão e não de um pequeno grupo de artigos sobre temas muito conhecidos.

Uma possível explicação poderá passar pelo facto de, nos temas mais comuns e relevantes, a qualidade da Britânica ser semelhante ou superior à da Wikipédia, sendo no entanto manifestamente inferior quando comparamos temas menos comuns. A utilização da amostragem aleatória foi, assim, um ponto-chave na obtenção de uma amostra representativa do universo a estudar, pois uma enciclopédia é a soma de todos os seus artigos e não apenas a daqueles com maior visibilidade.

Apesar de a amostra escolhida ser representativa do universo a estudar, os resultados obtidos dependeram da opinião pessoal dos avaliadores e dos critérios escolhidos, pelo que a generalização das conclusões deste estudo e a sua reprodutibilidade utilizando outro painel de avaliadores proveniente de outro contexto não está garantida, não sendo no entanto, *per se*, motivo suficiente para rejeitar os resultados obtidos.

A característica aberta da tecnologia, que permite que todos possam alterar os conteúdos da Wikipédia e que é o cerne do seu sucesso, é simultaneamente uma ameaça à sua credibilidade, a par da possibilidade de amadores criarem conteúdos sobre temas que não dominam. Como contraponto a estas legítimas críticas, a Wikipédia dispõe, essencialmente, da colaboração em massa na defesa contra ataques de vandalismo e na vontade de manter a qualidade do projecto por parte de quem colabora.

A percepção da qualidade de um produto baseia-se na confiança que os consumidores têm neste: a confiança não pode ser imposta ou comprada, tem de ser ganha ao longo do

tempo. Apesar do crescimento constante do número de utilizadores e de artigos disponíveis, o projecto – e a filosofia que o suporta – ainda não conquistou a mesma confiança que o poria, pelo menos, ao nível de outras enciclopédias generalistas que há muito ganharam o reconhecimento e a confiança dos consumidores.

O facto de a utilização da Wikipédia ser gratuita e a da Britânica ser paga poderá justificar, em parte, a enorme diferença entre o número de utilizadores e mesmo a percepção sobre a qualidade destas, que ficou patente nas respostas ao inquérito sobre os hábitos de consulta de informação.

A opinião preconcebida da qualidade dos conteúdos da Wikipédia e os resultados da avaliação desta numa prova cega, dentro de um mesmo grupo, contribuí para reforçar a isenção do painel de avaliadores e a conseqüente relevância dos resultados obtidos, ou seja, eventuais faltas de isenção por parte de alguns avaliadores tenderiam a favorecer a Britânica.

Se na Wikipédia excluirmos os poucos artigos que originam controvérsia e guerras de alterações, exacerbadas por paixões políticas, religiosas ou outras, a esmagadora maioria dos artigos é construída por utilizadores que dominam e se interessam pelo tema em causa e que tentam criar conteúdos de qualidade. Não nos parece plausível que alguém sem conhecimentos específicos sobre um determinado tema tenha motivação para criar e editar conteúdos em conjunto com outros que dominam o assunto em causa.

Uma explicação para o sucesso da Wikipédia poderá passar pelo altruísmo de indivíduos que se reúnem à volta de temas sobre os quais partilham a mesma paixão, sendo expectável possuírem uma dose assinalável de conhecimento sobre estes. Por outras palavras, no caso estudado, a colaboração em massa parece auto-organizar-se, resultando dessa organização uma auto-avaliação e uma autocorreção efectuada pelos pares, afastando de vez a ideia de

*Infinitos macacos a teclarem nas suas máquinas de escrever
à espera de um milagre da teoria das probabilidades¹³⁷.*

¹³⁷ Adaptado de Keen (2007)

*O Captain! my Captain! our fearful trip is done,
The ship has weather'd every rack, the prize we sought is won,
The port is near, the bells I hear, the people all exulting,
While follow eyes the steady keel, the vessel grim and daring;
But O heart! heart! heart!
O the bleeding drops of red,
Where on the deck my Captain lies,
Fallen cold and dead.*

Walt Whitman (1865), *Leaves of Grass*

6.2 CONCLUSÕES E RECOMENDAÇÕES

Esta tese é uma tentativa de ajudar a clarificar a discussão em torno do rigor e qualidade científica dos conteúdos da Wikipédia e contribuir para uma melhor percepção da qualidade de uma ferramenta utilizada por milhões diariamente e, conseqüentemente, concluir sobre a viabilidade do modelo de funcionamento que lhe está subjacente – a colaboração em massa.

Assim, e ao contrário da opinião preconcebida da maioria, mostrou-se que a Wikipédia tem uma qualidade média superior à sua concorrente directa, a Britânica, o que a transforma num caso de sucesso e permite responder à questão de investigação, concluindo que a colaboração em massa é um mecanismo que permite criar e gerir conteúdos de qualidade, pelo menos dentro do contexto e das condicionantes do caso estudado, ou seja podemos rejeitar a hipótese de a colaboração em massa não ser capaz de produzir resultados de qualidade.

As condicionantes da conclusão anterior prendem-se com o facto de não ter sido possível encontrar avaliadores para os artigos que descrevem cidades, regiões, países, rios e montanhas da área de Geografia, pelo que todas as conclusões se referem a todo o conteúdo da Wikipédia à excepção deste.

Como não podemos simplesmente baixar o nível de significância ou aumentar a margem de erro como se passássemos a ter uma amostra mais pequena, o que iria fazer com que a amostra deixasse de ser representativa da população que queremos estudar, optámos por limitar as conclusões apresentadas a todos os temas e áreas do conhecimento encontradas na Wikipédia, à excepção da Geografia.

Parece, mesmo assim, razoável afirmar que à pergunta formulada no título da tese: “Colaboração em Massa, ou Amadorismo em Massa?” se pode responder com a frase: colaboração em massa de amadores, contribuindo maioritariamente nos temas em que são peritos, embora esta só seja válida para o caso estudado e não possa ser generalizada.

O maior obstáculo encontrado na realização do estudo prendeu-se com a dificuldade em encontrar avaliadores cuja competência na área de especialização não pudesse ser posta em causa que estivessem dispostos a colaborar, uma vez que a tarefa que lhes era pedida era morosa e não trivial. Felizmente, foi possível encontrar junto da comunidade académica peritos cujo altruísmo permitiu a realização deste estudo.

Do inquérito realizado junto da comunidade académica, constatou-se que cerca de 80% dos inquiridos utiliza em primeiro lugar a Wikipédia e os resultados da pesquisa obtidos utilizando um motor de busca, numa primeira abordagem, quando necessitam de obter informação sobre um determinado tema, em alternativa às fontes tradicionais – enciclopédias, dicionários, obras de referência e afins. Esta percentagem diminui com o nível de instrução.

Numa escala de 1 (muito desconfiado) a 5 (completamente confiante), os inquiridos atribuíram um valor médio de 3,0 valores à confiança que depositam na informação obtida através da Wikipédia / Motor de Busca, e 4,5 valores à confiança que depositam na informação obtida através das fontes tradicionais.

Os subgrupos que participaram neste estudo, docentes e discentes do ensino Universitário e Secundário, não apresentaram diferenças estatisticamente significativas entre as suas respostas.

Foi avaliado um conjunto de pares de artigos sobre Gestão, extraídos da Wikipédia e da Britânica, a partir de uma amostragem por conveniência, tendo cada par sido avaliado individualmente por um grupo de professores e de alunos, numa escala de 1 (Mau) a 5 (Muito Bom), depois de terem sido previamente descaracterizados para ocultar a sua proveniência.

O grupo de docentes avaliou com uma nota média de $3,6 \pm 0,3$ os artigos provenientes da Wikipédia e com $2,4 \pm 0,2$ os seus homónimos da Britânica, enquanto os discentes

atribuíram as classificações de $4,0 \pm 0,3$ e de $2,5 \pm 0,2$. As diferenças entre as notas atribuídas à Wikipédia e à Britânica são estatisticamente significativas para cada par de artigos, e as diferenças entre os dois grupos (professores e alunos), não.

Foi recolhida uma amostra aleatória de 245 pares de artigos (490 artigos), representativa do universo composto pelos artigos que integram as duas enciclopédias, tendo cada par também sido previamente descaracterizado e avaliado por um docente proficiente na área do conhecimento em que este se enquadra. Foram avaliados todos os pares, menos os que pertencem à área de Geografia, por não ter sido possível encontrar avaliadores da área dispostos a colaborar com o estudo.

Tentou-se encontrar uma relação entre as notas das avaliações dos artigos e o número de palavras e ilustrações destes, com o intuito de simular a avaliação dos artigos de Geografia que não foram analisados. Apesar de não existir correlação para os artigos em termos individuais, existe uma correlação matematicamente significativa (0,941) entre as diferenças de notas dos pares agrupados e as diferenças entre o número de caracteres destes. A simulação das avaliações em falta, com base na equação que traduz a relação apresentada, mostra que o valor médio da diferença destas seria de +1,7, resultado da mesma ordem de grandeza ao encontrado nas avaliações dos restantes temas.

Apesar de matematicamente correcta, esta análise deve apenas ser entendida como uma aproximação que poderá estar incorrecta, pois nada garante que os artigos de Geografia sigam a mesma tendência que os restantes.

A nota média atribuída aos artigos, numa escala de 1 (Mau) a 5 (Muito Bom), foi de 4,0 para a Wikipédia e 2,6 para a Britânica. Esta diferença é estatisticamente significativa. Como cada artigo só foi avaliado por um avaliador, foi utilizada a variabilidade média das classificações dos artigos de Gestão, em que cada artigo foi avaliado por vários peritos, de modo a afectar os resultados obtidos de uma incerteza que, também aqui, iria existir se tivéssemos múltiplos avaliadores por artigo.

Admitindo que a variabilidade das respostas na área de Gestão é idêntica à de todas as outras áreas do conhecimento, conclui-se que, qualquer que seja a medida de dispersão escolhida, a classificação da Wikipédia continua a ser maior ou igual à da Britânica.

Saliente-se que a avaliação negativa – menor que três – obtida pela Britânica se deve seguramente ao facto de estar a ser avaliada em conjunto com outra enciclopédia e as classificações atribuídas a um artigo dependerem da atribuída ao seu homólogo, pelo que se recomenda que se dê uma maior relevância aos resultados das diferenças de notas entre cada par de artigos do que às classificações absolutas de cada elemento do par.

As críticas das posturas mais conservadoras em relação à Wikipédia passam muitas vezes pela afirmação de que o facto de as páginas desta serem dinâmicas – mudam ao longo do tempo – não permite garantir que o conteúdo das citações se mantenha, pois a página em questão pode vir a ser alterada no futuro. No entanto, o carácter transaccional das ferramentas Wiki e a possibilidade de gerar *links* permanentes de páginas dinâmicas – i.e. cristalizar o conteúdo de uma página num determinado instante do tempo – torna possível garantir que as referências se mantêm.

Desmontado este “problema” e demonstrado que a colaboração em massa consegue, pelo menos dentro do contexto analisado, produzir resultados comparáveis à abordagem convencional, estudos futuros que identifiquem e extrapolem as variáveis que permitem o sucesso da colaboração em massa, no caso da Wikipédia, poderão abrir caminho a projectos futuros assentes na mesma filosofia.

A própria Britânica tentou uma aproximação a esta abordagem, permitindo que os seus utilizadores proponham alterações nos seus artigos, embora estas só fiquem disponíveis para todos depois de passar pelo crivo dos seus editores. O reduzido sucesso desta iniciativa, a que não está alheio o facto de o acesso à totalidade dos conteúdos desta enciclopédia ser pago, mostra o quão importante é em projectos deste tipo criar um ecossistema em que os intervenientes sintam que estão a construir uma catedral e não apenas a partir pedra, tendo em conta que não estão a ser remunerados por isso – até por que é mais fácil aproveitar aquilo que as pessoas gostam de fazer do que obrigá-las a fazer o que não gostam...

A constatação de que a colaboração em massa, pelo menos no caso estudado, produz resultados de qualidade mostra que vale a pena realizar mais estudos nesta área. Em termos de trabalhos futuros, seria interessante voltar a utilizar esta metodologia de modo a acompanhar a evolução da qualidade dos artigos e perceber qual a tendência subjacente a projectos assentes na colaboração em massa.

Outro aspecto importante, e que foi apenas aflorado neste estudo por não ser essa a questão de investigação a que nos propusemos responder, é a identificação das características específicas da Wikipédia que levaram a que a este projecto fosse (seja) um sucesso, bem como a caracterização das pessoas que nele colaboram e a percepção das suas motivações. Estes conhecimentos levam à compreensão dos factores de sucesso e de insucesso de projectos deste tipo, o que poderá ser decisivo na sua escolha e posterior gestão.

Outra abordagem do tema que poderia ser objecto de estudos futuros seria averiguar a exequibilidade da utilização das ferramentas e princípios da colaboração em massa na migração de alguns dos aspectos da governação actualmente assentes na democracia representativa para a democracia participativa.

A publicitação, e eventual aceitação, destes resultados junto da comunidade académica poderá contribuir para o estabelecimento da colaboração em massa como instrumento credível da Gestão e o reconhecimento da Wikipédia enquanto caso de sucesso da aplicação desta; à semelhança da teoria da administração científica de Frederick Taylor, que estabeleceu os alicerces da Gestão enquanto ciência, e Henry Ford, que a aplicou na sua linha de montagem nos primórdios do século XX (Sousa, 1990).

Cabe aos gestores e líderes destas comunidades sociais, assentes na colaboração em massa, não o papel de moldar a comunidade e o modo como esta se organiza, mas sim tornar o projecto atractivo, de modo a atrair o maior número possível de contribuidores possíveis, condição necessária – embora não suficiente – ao sucesso do projecto. Os resultados deste estudo poderão também abrir portas junto da comunidade da Wikipédia, permitindo a realização de estudos futuros que clarifiquem as razões do sucesso do projecto e que permitam o desenvolvimento de abordagens teóricas sobre o tema.

A particularidade de não existir na colaboração em massa controlo directo sobre os recursos humanos, deixando quase exclusivamente a gestão dos processo de produção ao mecanismo¹³⁸ agregador das contribuições e às regras nele imbuídas, permite encontrar semelhanças entre esta e o comércio a retalho, que confia no *marketing* para

¹³⁸ O *software* Wiki.

atrair os clientes ao seu espaço pois, à semelhança da colaboração em massa em relação aos contribuidores, também não tem controla directamente a afluência de clientes.

Este tipo de projectos caracteriza-se, em termos financeiros, por um baixo rácio de custo/benefício, ou seja, o custo de falhar é baixo e os benefícios em caso de sucesso são grandes. Se tivermos em conta o esforço feito por Portugal na massificação do acesso à Internet em todos os níveis do ensino (Robalo, 2007), porque não tentar valorizar esse investimento com a criação de um repositório de materiais e conteúdos programáticos? Para tal, abrir-se-ia a edição e manipulação de um pequeno repositório embrionário inicial, previamente criado, a todos os professores e posteriormente, quiçá, também aos próprios alunos.

Caso a iniciativa fosse um sucesso, em termos de quantidade e qualidade de conteúdos, o número de potenciais beneficiários e contribuidores seriam os mais de 240 milhões¹³⁹ de falantes da língua Portuguesa, podendo, a médio prazo, ser uma alternativa aos manuais escolares em papel, aproveitando a actual tendência de substituição destes pelos seus sucedâneos digitais e acrescentando às óbvias vantagens económicas e ambientais que esta prática traz o facto de poderem ser utilizados gratuitamente, contribuindo de uma forma significativa para a igualdade de oportunidades no acesso ao conhecimento.

No entanto, para que esta utopia possa almejar transformar-se em realidade, é necessário mudar a imagem que ficou patente no inquérito realizado e provar cientificamente que existe pelo menos um (\exists^1) caso em que a colaboração em massa produz resultados melhores ou iguais ao modelo hierárquico convencional – objectivo que esperamos ter atingido com o estudo apresentado nesta tese, e que poderia ser aferido repetindo no futuro o inquérito realizado após a publicação e divulgação dos resultados obtidos.

Caso se venha a verificar uma alteração significativa na opinião do grupo inquirido, poderemos então afirmar que este estudo contribuiu para a credibilização da colaboração em massa e a concretização da visão da Wikipédia:

*Imagine a world in which every single human being
can freely share in the sum of all knowledge.*¹⁴⁰

¹³⁹ Fonte: <http://observatorio-lp.sapo.pt/pt/dados-estatisticos/falantes-de-portugues>

¹⁴⁰ Fonte: <http://wikimediafoundation.org/wiki/Vision>

BIBLIOGRAFIA

ACKOFF, R. (1974). *Redesigning the Future: a Systems Approach to Societal Problems*, Wiley-Interscience publication, Wisconsin University – Madison, USA.

ACKOFF, R. (1989). “From Data to Wisdom”, *Journal of Applied Systems Analysis*, Vol. 16, pp. 3-9.

ADAMS, J., RAESIDE, R. & KHAN, H. (2007). *Research Methods for Graduate Business and Social Science Students*, Sage Publications Inc., New Delhi, India.

ADLER, B., CHATTERJEE, K., ALFARO, L., FAELLA, M., PYE, I., & RAMAN, V. (2008). “Assigning Trust to Wikipedia Content”, *WikiSym '08: International Symposium on Wikis*, September 8–10, Porto, Portugal.

AGRANOFF, R. (2006). “Inside Collaborative Networks: Ten Lessons for Public Managers”, *Public Administration Review*, December 2006.

ALTER, S. (1996). *Information Systems: A Management Perspective*, The Benjamin - Cummings Publishing Company, Redwood City, USA.

AMARAL, L., MAGALHÃES, R., MORAIS, C., SERRANO, A., & ZORRINHO, C., Eds. (2005). *Sistemas de Informação Organizacionais*, Edições Sílabo, 1ª Edição, Lisboa, Portugal.

AMARAL, L., & VARAJÃO, J. (2007). *Planeamento de Sistemas de Informação*, FCA – Editora de Informática Lda, 4ª Edição, Lisboa, Portugal.

ANDERSON, C. (2006). *The Long Tail: Why the future of business is selling less of more*, Hyperion Books, New York, USA.

ANTUNES, F., MELO, P., & COSTA, J. (2005). “Information Management in Distributed Collaborative Systems: The Case of Collaboration Studio”, *European Journal of Operational Research*, Elsevier, 177 (2007), 1385–1399.

- ARIELY, D., GNEEZY, U., LOEWENSTEIN, G., & MAZAR, N. (2005). "Large Stakes and Big Mistakes", *Research Center for Behavioral Economics and Decision-Making*, Nº 05-11, Federal Reserve Bank of Boston.
- AUSTIN, D. (2010). "How Google Finds Your Needle in the Web's Haystack", *Monthly Essays on Mathematical Topics*, American Mathematical Society, Grand Valley State University, <http://www.ams.org/samplings/feature-column/fcarc-pagerank>, acedido em 30/09/2010.
- AXELROD, R., & HAMILTON, W. (1981). "The Evolution of Cooperation", *Science*, Vol. 211, N. 4489, 27 Mar 1981, pp. 1390-1396.
- BACON, F. (1996 [1620]). *Novum Organum*, Hanover Historical Texts Project, <http://history.hanover.edu/texts/Bacon/novorg.html>, acedido em 14/9/2011.
- BARNES, M. (2005). "Can you trust Wikipedia?", *The Guardian*, 24 October, <http://www.guardian.co.uk/technology/2005/oct/24/comment.newmedia>, acedido em 14/11/2010.
- BARTLETT, J., KOTRLIK, J., & HIGGINS, C. (2001). "Organizational Research: Determining Appropriate Sample Size in Survey Research", *Information Technology, Learning, and Performance Journal*, Vol. 19, No. 1, Spring 2001.
- BENKLER, Y. (2006). *The Wealth of Networks How Social Production Transforms Markets and Freedom*, Yale University Press, New Haven, USA.
- BERNSTEIN, S., & BERNSTEIN, R. (1998). *Schaum's Outline of Theory and Problems of Elements of Statistics I: Statistics and Probability*, McGraw-Hill, New York, USA.
- BERNSTEIN, S., & BERNSTEIN, R. (1999). *Schaum's Outline of Theory and Problems of Elements of Statistics II: Inferential Statistics*, McGraw-Hill, New York, USA.
- BLOCH, W. (2008). *The Unimaginable Mathematics of Borges' Library of Babel*, Oxford University Press, New York, USA.

- BLUMENSTOCK, J. (2008). "Automatically Assessing the Quality of Wikipedia Articles", *UC Berkeley: School of Information*, April 2008.
- BOREL, É. (1913). "Mécanique Statistique et Irréversibilité", *J. Phys.*, 5^a série, Vol. 3, pp.189-196.
- BORGES, J. L. (1999 [1939]). *The Total Library: non-fiction 1921-1986*, Penguin Press, London, UK.
- BORGES, J. L. (2000 [1941]). *The Library of Babel*, David R. Godine Publisher, New Hampshire, USA.
- BRAGUES, G. (2007). "Wiki-Philosophizing in a Marketplace of Ideas: Evaluating Wikipedia's entries on seven great minds", *University of Guelph-Humber*, <http://ssrn.com/abstract=978177> acedido em 12/03/2011.
- BROUGHTON, B. (2008). *Wikipedia: The Missing Manual*, O'Reilly Media, First Edition, Sebastopol, California, USA.
- BRUNO, F. (1997). *Not Just for the Money: An Economic Theory of Personal Motivation*, Edward Elgar Publishing, Cheltenham, UK.
- BUCKINGHAM, R., HIRSCHHEIM, R., LAND, F., & TULLY, C. (1987). *Information Systems Education: Recommendations and Implementation*, London, UK, Cambridge University Press.
- BURKE, J. (1995). *The Day the Universe Changed: How Galileo's Telescope Changed The Truth and Other Events in History That Dramatically Altered Our Understanding of the World*, Back Bay Books, London, UK.
- BURTON, R. (2007 [1621]). *The Anatomy of Melancholy, Vol. 1: Introductions and Partitions 1 and 2*, The Echo Library, Middlesex, UK.
- CALERO, C. (2004). *Handbook of Research on Web Information Systems Quality*, Information Science Reference, New York, USA.

- CAMARINHA-MATOS, L. (2004). *Virtual Enterprises and Collaborative Networks*, IFIP 18th world Computer Congress TC5/WG5.5 Working Conference on Virtual Enterprises, Toulouse, France, 22-27 August, Kluwer Academic Publishers, Massachusetts, USA.
- CAMARINHA-MATOS, L., & AFSARMANESH, H. (2005). "Collaborative Networks: a New Scientific Discipline", *Springer Journal of Intelligent Manufacturing*, 16, pp. 439–452, 2005.
- CAMARINHA-MATOS, L., AFSARMANESH, H., & ORTIZ, A. (2005). *Collaborative Networks and Their Breeding Environments*, IFIP TC5 WG 5.5 Sixth IFIP Working Conference on Virtual Enterprises, Valencia, Spain, 26-28 September 2005, Springer, New York, USA.
- CARDOSO, J. (1997). "Manual de Gestão : Guia dos Conceitos de A a Z - Executive Digest ", *Abril/Controljornal*.
- CARPENTIER, M. (1993). *Análise Numérica*, Departamento de Matemática, Instituto Superior Técnico, UTL, Lisboa, Portugal.
- CARREGAL, J. (1941). *A Evolução da Tipografia - 1440-1940 - do Prelo de Gutenberg à Rotativa*, Editora Grémio Nacional dos Industriais de Tipografia e Fotogravura, Lisboa, Portugal.
- CARVALHO, J. (2000). "Information System? Which One Do You Mean?", *Information Systems Concepts: An Integrated Discipline Emerging*, Eds. FALKENBERG, E. et al. (2000), Proceedings of the ISCO 4 Conference, Leiden, Holland.
- CHECKLAND, P., & HOLWELL, S. (1998). *Information, Systems, and Information Systems: Making sense of the Field*, John Wiley & Sons, Chichester, West Sussex, England.
- CHIAVENATO, I. (1993). *Introdução à Teoria Geral da Administração*, Mc Graw Hill, 4^a Edição, Makron Books do Brasil, São Paulo, Brasil.

- CHRISTENSEN, C. (1997). *The Innovator's Dilemma: When New Technologies Cause Great Firms to Fail*, Harvard Business School Press, Boston, USA.
- CLAUSON, K., POLEN, H., BOULOS, M., & DZENOWAGIS, J. (2008). "Scope, Completeness, and Accuracy of Drug Information in Wikipedia", *The Annals of Pharmacotherapy*, Vol. 42, N. 12, pp. 1814-1821.
- CLEVELAND, H. (1982). "Information as Resource", *The Futurist*, World Future Society, December 1982, pp. 34-39.
- COHEN, M. (2005). *The Cambridge Companion to Ancient Greek Law*, Cambridge University Press, Cambridge, UK.
- COOK, S. (2008). "The Contribution Revolution – Letting Volunteers Build Your Business", *Harvard Business Review*, October 2008.
- DALKIR, K. (2005). *knowledge Management in Theory and Practice*, Elsevier, McGill University, USA.
- DECI, E. (1971). "Effects of Externally Mediated Rewards on Intrinsic Motivation", *Journal of Personality and Social Psychology*, Vol. 18, pp. 105-115.
- DEGRACE, P., & STAHL, L. (1990). *Wicked Problems, Righteous Solutions: a Catalogue of Modern Software Engineering Paradigms*, Yourdon Press, Englewood Cliffs, USA.
- DEROSE, P., CHAI, X., GAO, B., SHEN, W., DOAN, A., BOHANNON, P., & ZHU, X. (2008). "Building Community Wikipedias: A Machine-Human Partnership Approach", *Proceedings of the 24th International Conference on Data Engineering*, ICDE 2008, April 7-12, 2008.
- DESCARTES, R. (2008 [1637]). *Discourse on the Method for Reasoning Well and for Seeking Truth in the Sciences*, Project Gutenberg EBook, http://www.gutenberg.org/catalog/world/readfile?fk_files=2585108, acedido em 12/07/2012.

- DEVGAN, L., POWE, D., BLAKEY, B., & MAKARY, M. (2007). "Wiki-Surgery? Internal validity of Wikipedia as a medical and surgical reference", *Journal of the American College of Surgeons*, Vol. 205, nº 3, pp. s76-s77.
- DUNN, D. (2001). *Statistics and Data Analysis for the Behavioral Sciences*, McGraw-Hill Higher Education, New York, USA.
- DUNN, J. (2008). *Information in Computer Science in Philosophy of Information*, Vol.8, Edited by Adriaans, P., & Benthem, J., Elsevier, Amsterdam, The Netherlands.
- EMIGH, W., & HERRING, S. (2005). "Collaborative Authoring on the Web: A Genre Analysis of Online Encyclopaedias", *Indiana University Bloomington*, <http://ella.slis.indiana.edu/~herring/wiki.pdf>, acedido em 02/03/2011
- FALLIS, D. (2010). "Toward an Epistemology of Wikipedia", *Journal of the American Society for Information Science and Technology*, 59(10), 2010, pp. 1662-1674.
- FAY, M., & PROSCHAN, M. (2010). "Wilcoxon-Mann-Whitney or t-Test? On Assumptions for Hypothesis Tests and Multiple Interpretations of Decision Rules", *Stat Surv.*, nº 4-2010, pp.1-39.
- FERSON, F., KREINOVICH, V., HAJAGOS, J., OBERKAMPF, W., & GINZBURGFETZER, L. (2007). "Experimental Uncertainty Estimation and Statistics for Data Having Interval Uncertainty", *Sandia Report*, 2007-0939, Sandia National Laboratories, Albuquerque, New Mexico.
- FETZER, J. (2004). "Information: Does it Have to be True?", *Minds and Machines Journal*, Vol. 14, Nº 2, Springer Netherlands, pp.223-229.
- FISHER, R. (1966). *The Design of Experiments*, California University, Hafner Pub. Co. (8th Edition), New York, USA.
- FLORIDI, L. (2004). "Outline of a Theory of Strongly Semantic Information", *Minds and machines Journal*, Vol. 14, Nº 2, Springer Netherlands, pp. 197- 221

- FREIXO, M. (2009). *Metodologia Científica – Fundamentos, Métodos e Técnicas*, Instituto Piaget, Lisboa, Portugal.
- GALTON, F. (1907a). "Vox Populi", *Nature*, Vol. 75, 7 March 1907, Nature Publishing Group, pp. 450-451.
- GALTON, F. (1907b). "The Ballot-Box", *Nature*, Vol. 75, 28 March 1907, Nature Publishing Group, p. 509.
- GAUR, A., & GAUR, S. (2009). *Statistical Methods for Practice and Research: A Guide to Data Analysis using SPSS*, Response Books, Business books from Sage (Second edition), New Delhi, India.
- GHAZAWNEH, A. (2008). "Managing Mass Collaboration: Toward a Process Framework", Master Thesis, School of Economics and Management, Lund University, Sweden.
- GIBBONS, J. (2008). *Monkeys with Typewriters: Myths and Realities of Social Media at Work*, Triarchy Press, Devon, United Kingdom.
- GIDE, A. (2008 [1925]). *Les Faux-Monnayeurs*, Gallimard, Paris, France.
- GILES, J. (2005). "Special Report: Internet Encyclopaedias Go Head to Head", *Nature*, Vol. 438, 14 December 2005, Nature Publishing Group, pp. 900-901.
- GIVEN, L. (2008). *The Sage Encyclopedia of Qualitative Research Methods*, Sage Publications Ltd., Thousand Oaks, California, USA.
- GOBILLOT, E. (2009). *Leadershift: Reinventing leadership for the age of mass collaboration*, Kogan Page Limited, Philadelphia, USA.
- GORARD, F. (2004). "Revisiting a 90-year-old Debate: The Advantages of the Mean Deviation", *British Journal of Educational Studies*, Volume 53, Issue 4, December 2005, pp. 417–430.

- GREENSTEIN, S., & DEVEREAUX, M. (2006). "The Crisis at Encyclopedia Britannica", *Kellogg School of Management*, Northwestern University.
- HALAVAIS, A., & LACKAFF, D. (2008). "An Analysis of Topical Coverage of Wikipedia", *Journal of Computer-Mediated Communication*, Volume 13, Number 2, January 2008 , pp. 429-440.
- HARDT, M., & NEGRI, A. (2000). *Empire*, Harvard University Press, London, England.
- HARDT, M., & NEGRI, A. (2004). *Multitude: War and Democracy in the Age of Empire*, The Penguin Press, New York, USA.
- HARLOW, H., HARLOW, M., & MEYER, M. (1950). "Learning Motivated by a Manipulation Drive", *Journal of Experimental Psychology*, Vol 40(2), pp. 228-234.
- HARLOW, H. (1953). "Motivation as a Factor in the Acquisition of New Responses", *Current Theory and Research on Motivation*, University of Nebraska Press, Lincoln, pp. 24-49.
- HAYEK, F. (1945). "The Use of Knowledge in Society", *American Economic Review*, XXXV, n° 4, pp. 519–530.
- HOBBS, T. (1957 [1651]). *Leviathan*, Basil Blackwell, Oxford, England.
- HOWARD, A. (1998). *Magna Carta - Text & Commentary*, The University Press of Virginia, USA.
- HOWE, J. (2006). " The Rise of Crowdsourcing", *Wired Magazine*, XXXV, Issue 14.06, June 2006, http://www.wired.com/wired/archive/14.06/crowds_pr.html, acedido em 3/02/2012.
- HOWE, J. (2009). *Crowdsourcing: Why the Power of the Crowd is Driving the Future of Business*, Three Rivers Press, New York, USA.

- HUFF, D. (1954). *How to Lie with Statistics*, W. W. Norton & Company, New York, USA.
- HURWICZ, L. (2007). "But who will Guard the Guardians?", *Prize Lecture*, December 8, 2007, University of Minnesota, Department of Economics, http://www.nobelprize.org/nobel_prizes/economics/laureates/2007/hurwicz_lecture.pdf, acessado em 4/08/2012.
- INÊS, P. (2010). "Avaliar a Qualidade de Dados em Sistemas de Informação", *Tese de Mestrado em gestão*, Universidade de Évora, Portugal.
- JOHNSON, S. (2002). *Emergence: The Connected Lives of Ants, Brains, Cities, and Software*, Touchstone Book, New York, USA.
- JÓLLUSKIN, G. (2009). "O Tribunal do Júri no Ordenamento Jurídico Português", *Revista da Faculdade de Ciências Humanas e Sociais*, Edições Universidade Fernando Pessoa, Portugal, pp. 116–126.
- KALOF, L., DAN A., & DIETZEEN, T. (2008). *Essentials of Social Research*, Open University Press, McGraw-Hill, Berkshire, England.
- KEEN, A. (2007). *The Cult of the Amateur: How Today's Internet is Killing our Culture*, a Currency book, published by Doubleday, New York, USA.
- KEEN, A. (2010). "Why We Must Resist the Temptation of Web 2.0", *The Next Digital Decade - Essays on the Future of the Internet*, Eds. SZOKA, B. & MARCUS, A., TechFreedom, Washington, USA, pp. 51-56
- KLARE, G. (1963). *The measurement of readability*, New York Teachers College, Ames, Iowa: University of Iowa Press.
- KRAUT, R., & KITUR, A. (2008). "Harnessing the Wisdom of Crowds in Wikipedia: Quality Through Coordination", *CSCW '08 Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work*, pp. 37-46.

- KRUSKAL, W., & WALLIS, W. (1952). "Use of Ranks in One-Criterion Variance Analysis", *Journal of the American Statistical Association*, Vol. 47, N° 260, pp. 583-621.
- LE BON, G. (2003[1896]). *The Crowd – A study of the popular mind*, Kessinger Publishing, University of Virginia, USA.
- LEE, M., & LAN, Y. (2007). "From Web 2.0 to Conversational Knowledge Management: Towards Collaborative Intelligence", *Journal of Entrepreneurship Research*, June 2007, Vol.2, No.2, p. 47-62.
- LESSING, D. (2002). *The Diaries of Jane Somers*, Flamingo, London, UK.
- LEVENE, H. (1960). "Robust Tests for Equality of Variances", *Contributions to probability and statistics: Essays in honor of Harold Hotelling*, Eds. OLKIN, I. et al. (1960), Stanford University Press, pp. 278-292.
- LEVINE, R., LOCKE, C., SEARLS, D., & WEINBERGER, D. (2000). *The Cluetrain Manifesto: The End of Business as Usual*, Basic Books, New York, USA.
- LÉVY, P. (1997). *Collective Intelligence: Mankind's Emerging World in Cyberspace*, Perseus, Cambridge, USA.
- LIH, A. (2004). "Wikipedia as Participatory Journalism: Reliable Sources? Metrics for Evaluating Collaborative Media News Resource", *5th International Symposium on Online Journalism*, University of Texas, Austin, USA.
- LIH, A. (2009). *The Wikipedia Revolution – How a Bunch of Nobodies Created the World's Greatest Encyclopedia*, Hyperion eBook, New York, USA.
- LIKERT, R. (1932). "A Technique for the Measurement of Attitudes", *Archives of Psychology*, 140, pp. 1-55.
- LIST, C., & GOODIN, R. (2001). "Epistemic Democracy: Generalizing the Condorcet Jury Theorem", *Journal of Political Philosophy*, Volume 9, pp. 277-306.

- LITTLE, R. & RUBIN, D. (1987). *Statistical Analysis with Missing Data*, John Wiley & Sons, New York, USA.
- LOUREIRO, J. (2003). *Gestão do Conhecimento*, Centro Atlântico, Lisboa, Portugal.
- MACKAY, C. (2003 [1841]). *Extraordinary Popular Delusions and the Madness of Crowds*, Harriman House Publishing, Hampshire, Great Britain.
- MALONE, T. (2006). “What is Collective Intelligence And What Will We Do About It?”, *Collective Intelligence: Creating a Prosperous World at Peace*, Ed. TOVEY, M. (2008), Earth Intelligence Network Oakton, Virginia .
- MALONE, T., LAUBACHER, R., & DELLAROCAS, C. (2010). “The Collective Intelligence Genome”, *MIT Sloan Management review*, Spring 2010, Vol. 51, Nº3.
- MARKS, H. (2009). “Collaborative Networks vs Social Networks”, *ZDNet*, <http://www.zdnet.com/blog/collaboration/collaborative-networks-vs-social-networks/621>, acedido em 12/03/2011.
- MARX, K., & ENGELS, F. (1845). *The German Ideology*, Marxists Internet Archive, <http://www.marxists.org/archive/marx/works/1845/german-ideology/ch01a.htm>, acedido em 12/08/2012.
- MASLOW, A. (1943). “A Theory of Human Motivation”, *Psychological Review*, Vol. 50(4), pp. 370-396.
- MASLOW, A. (1954). *Motivation and Personality - Psychological Review*, Harper and Row, New York, USA.
- MAYFIELD, R. (2006). “Power Law of Participation”, *Ross Mayfield's Weblog*, (27 Abril 2006), http://ross.typepad.com/blog/2006/04/power_law_of_pa.html, acedido em 14/05/2012.
- MENDES, J. (1994). *Introdução ao Estudo do Direito*, Pedro Ferreira, Cacém, Portugal.

- MILLER, J., SCOTT, E., & OKAMOTO, S. (2006). "Public Acceptance of Evolution", *Science*, Vol. 313, 11 August 2006, pp. 765-766.
- MILLER, M. (2009). *Sams Teach Yourself Wikipedia in 10 Minutes*, Pearson education, USA.
- MOODY, G. (2001). *Rebel Code: Linux And The Open Source Revolution*, Basic Books, New York, USA.
- MURUGESAN, S. (2010). *Handbook of Research on Web 2.0, 3.0, and X.0 : Technologies, Business, and Social Applications*, Information Science Reference, Hershey, New York, USA.
- NEGRI, A. (2002). "Pour une définition ontologique de la multitude ", *Multitudes*, n° 9, Mai- Juin 2012, <http://multitudes.samizdat.net/Pour-une-definition-ontologique-de>, acedido em 12/08/2012.
- NIELSON, J. (2006). *Participation Inequality: Encouraging More Users to Contribute*, http://www.useit.com/alertbox/participation_inequality.html, acedido em 10/09/2011.
- NONAKA, I., & TAKEUCHI, I. (1986). "The New New Product Development Game", *Harvard Business Review*, Jan-Feb 1986.
- NONAKA, I., & TAKEUCHI, I. (1995). *The Knowledge-Creating Company*, Oxford University Press, New York, USA.
- OLSON, G., ZIMMERMAN, A., & BOS, N. (2008). *Scientific Collaboration on the Internet*, MIT Press, Cambridge, England.
- O'NEIL, M. (2010). "Shirky and Sanger, or the Costs of Crowdsourcing", *Journal of Science Communication*, Volume 09, Issue 01, March 2010.

- O'REILLY, T. (2007). "What is Web 2.0: Design Patterns and Business Models for the Next Generation of Software", *Communications & Strategies*, No. 1, p. 17, First Quarter 2007.
- PAGE, S. (2007). *The Difference: How the Power of diversity Creates Better Groups, Firms, Schools, and Societies*, Princeton University Press, New Jersey, USA.
- PARK, H. (2008). "Univariate Analysis and Normality Test Using SAS, Stata, and SPSS. Working Paper", *The University Information Technology Services (UITS) Center for Statistical and Mathematical Computing*, Indiana University.
- PENDER, M., KAYE, L., LISA, K., CHRISTOPHER, D., & SATYAMURTHY, A. (2008). "Putting Wikipedia to the Test: A Case Study", *Special Libraries Association Annual Conference Seattle*, Washington, University of Queensland and Department of Neurology, Royal Brisbane and Women's Hospital.
- PINK, D. (2005). "The Book Stops Here", *Wired*, (13 March 2005), http://www.wired.com/wired/archive/13.03/wiki_pr.html, acessado em 16/03/2010.
- PINK, D. (2009). *Drive : The Surprising Truth About What Motivates Us*, Penguin Group, New York, USA.
- PLATÃO (1990 [380 a.C.]). *A República*, Fundação Calouste Gulbenkian, 6ª Edição, Lisboa, Portugal.
- POPPER, K. (2002 [1935]). *The Logic of Scientific Discovery*, Routledge Classics, Oxford, UK.
- POSTMAN, N. (1993). *Technopoly: The Surrender of Culture to Technology*, First Vintage Books Edition, New York, USA.
- RAMAKRISHNAN, R., BAPTIST. A., ERCEGOVAC, V., HANSELMAN, M., & SHAFT, U. (2004). "Mass Collaboration: A Case Study", *International Database Engineering and Applications Symposium (IDEAS' 04)*, pp 133-146.

- RECTOR , L. H. (2008). "Comparison of Wikipedia and Other Encyclopedias for Accuracy, Breadth, and Depth in Historical Articles", *Reference Services Review Journal*, Vol. 36, nº1, pp. 7-22.
- RETO, L., & NUNES, F. (1999). "Métodos como Estratégia de Pesquisa. Problemas Tipo numa Investigação", *Reference Services Review Journal*, ". *Revista Portuguesa de Gestão*, pp. 21-31.
- RICHARDSON, M., & DOMINGOS, P. (2003). "Building Large Knowledge Bases by Mass Collaboration", *In Proceedings of the 2nd International Conference on Knowledge Capture (K-CAP '03)*, ACM, New York, USA.
- ROBALO, M., Ed. (2007). *Sociedade da Informação: O Percurso Português – Dez Anos de Sociedade da Informação - Análise e Perspectivas*, Edições Sílabo, Lisboa, Portugal.
- ROMEIJN, J., & ATKINSON, D. (2010). "Learning Juror Competence: a Generalized Condorcet Jury Theorem", *Sage Politics Philosophy & Economics*, November 2010.
- ROSENZWEIG , R. (2006). "Can History be Open Source? Wikipedia and the Future of the Past", *The Journal of American History*, Vol. 93, nº1; pp. 117-146.
- ROUSSEAU, J. (2011 [1754]). *What is the Origin of Inequality Among Men, and is it Authorised by Natural Law?*, Constitution Society, <http://www.constitution.org/jjr/ineq.htm>, acedido em 14/08/2012.
- ROUSSEAU, J. (2012 [1762]). *The Social Contract or Principles of Political Right*, Constitution Society, <http://www.constitution.org/jjr/socon.htm>, acedido em 12/08/2012.
- RUANE, J. (2005). *Essentials of Research Methods: A Guide to Social Science Research*, Blackwell Publishing Ltd, Boston, USA.

- SAUNDERS, M., LEWIS, P., & THORNHILL, A. (2009). *Research Methods for Business Students - fifth edition*, Pearson Education Ltd, Harlow, England.
- SCHULMAN, T. (2001). *Dead Poets Society: The Screenplay*, Harvest Moon Books, USA.
- SERRANO, A., & FIALHO, C. (2003). *Gestão do Conhecimento – O Novo Paradigma das Organizações*, FCA – Editora de Informática Lda., Lisboa, Portugal.
- SHANNON, C. (1948). “A Mathematical Theory of Communication”, *The Bell System Technical Journal*, Vol. 27, pp. 379 – 423, 623 – 656.
- SHAPIRO, S., & WILK, M. (1965). “An Analysis of Variance Test for Normality”, *Biometrika*, Vol. 52, Nº 3/4), pp. 591-611.
- SHIRKY, C. (2008). *Here Comes Everybody: The Power of Organization Without Organizations*, The Penguin Press, New York, USA.
- SHIRKY, C. (2010). *Cognitive Surplus: How Technology Makes Consumers into Collaborators*, The Penguin Press, New York, USA.
- SIEGEL, L. (2008). *Against the Machine: Being Human in the Age of the Electronic Mob*, Random House Inc., New York, USA.
- SINGH, J. (2004). “Collaborative Networks as Determinants of Knowledge Diffusion Patterns”, *Management Science*, 51(5), May 2005, pp. 756-770.
- SOKAL, A. (1996). “A Physicist Experiments with Cultural Studies” *Lingua Franca*, Vol. 6, nº4 (May-June), pp.62-64. Reprinted in *The Sokal hoax: The sham that shook the Academy*, edited by the editors of *Lingua Franca*, pp. 49-53. Lincoln: University of Nebraska Press, 2000.
- SOUSA, A. (1990). *Introdução à Gestão*, Editorial Verbo., Lisboa, Portugal.

- SOYLU, F. (2009). "Academics' Views On and Uses of Wikipedia", *Journal of Communication, Culture & Technology*, Volume 9, Issue 9.2, Spring 2009.
- SPINOZA, B. (2000 [1677]). *Tractatus Politicas*, Hackett Publishers, Indianapolis, USA.
- SUNSTEIN, C. (2006). *Infotopia: How Many Minds Produce Knowledge*, Oxford University Press, New York, USA.
- SUNSTEIN, C. (2008). "Neither Hayek nor Habermas", *Springer Business and Economics*, Volume 134, Numbers 1-2 (2008).
- SUROWIECKI, J. (2004). *The Wisdom of Crowds: Why the Many Are Smarter Than the Few and How Collective Wisdom Shapes Business, Economies, Societies and Nations*, Anchor Books, New York, USA.
- SZOKA, B., & MARCUS, A., Eds. (2010). *The Next Digital Decade: Essays on the Future of The Internet*, TechFreedom, Washington, USA.
- TANCREED, H. (2004). *Aristotle: Metaphysics*. Penguin Books, London, UK.
- TAPSCOTT, D. (1998). *Growing Up Digital : The Rise of the Net Generation*, McGraw-Hill Professional, New York, USA.
- TAPSCOTT, D., & WILLIAMS, A. (2006). *Wikinomics: How Mass Collaboration Changes Everything*, Penguin Books Ltd, New York, USA.
- THIERER, A. (2010). "The Case for Internet Optimism - Saving the Net from Its Detractors", *The Next Digital Decade - Essays on the Future of the Internet*, Eds. SZOKA, B. & MARCUS, A. (2010), TechFreedom, pp. 57–88, 139-162.
- THOREAU, D. (2003 [1888]). *The Journal of Henry D. Thoreau, Volumes I-VII*, Dover Publications Inc., Vol. I, Mineola, USA.
- THORNDIKE, L. (1921). *The teacher's word book*, New York Teachers College, Columbia University, USA.

- TOUTENBURG, H., & SHALABH, S. (2009). *Statistics Analysis of Designed Experiments*, Third Edition, Springer, New York, USA.
- URDAN, T. (2005). *Statistics in Plain English*, Lawrence Erlbaum Associates Publishers, New Jersey, USA.
- URKEN, A. (2008) “Grokking Condorcet's 1785 Essai”, *Journal Électronique d'Histoire des Probabilités et de la Statistique*, Volume 4, Issue 1.
- VANDERSTOEP, S., & JOHNSTON, D. (2009). *Research Methods For Everyday Life: Blending Qualitative and Quantitative Approaches*, Published by Jossey-Bass, San Francisco, USA.
- VIÉGAS, F., WATTENBERG, M., & DAVE, K. (2004). “Studying Cooperation and Conflict Between Authors with History Flow Visualizations”, *Conference on Human Factors in Computing Systems* (Vienna, Austria 2004), ACM, pp. 575-582.
- VIITAMÄKI, S. (2008) “The FLIRT Model of Crowdsourcing: Planning and Executing Collective Customer Collaboration”, *Marketing Master's Thesis*, Department of Marketing and Management, Helsinki School of Economics, Finland.
- VIRNO, P. (2004). *A Grammar of the Multitude - For an Analysis of Contemporary Forms of Life*, Semiotext(e), Los Angeles, USA.
- WANG, R., & STRONG, D. (1996). ”Beyond Accuracy: What Data Quality Means to Data Consumers”, *Journal of Management Information Systems*, Spring, Vol. 12, nº4, pp. 5-34.
- WEAVER, W. (1982 [1963]). *Lady Luck: the theory of probability*, Dover publications inc., New York, USA.
- WHITMAN, W. (2008 [1855]). *Leaves of Grass*, Project Gutenberg, <http://www.gutenberg.org/ebooks/1322>, acedido em 3/01/2012.

- WILCOXON, F. (1945). "Individual Comparisons by Ranking Methods", *Biometrics Bulletin*, Vol. 1, N° 6, pp. 80-83.
- WILKINSON, D., & HUBERMAN B. (2007). "Assessing the Value of Cooperation in Wikipedia", *First Monday*, Vol. 12, N. 4, 2 April 2007.
- WILSON, C. (2008). "The Wisdom of the Chaperones: Digg, Wikipedia, and the Myth of Web 2.0 Democracy", *Slate*, http://www.slate.com/articles/technology/technology/2008/02/the_wisdom_of_the_chaperones.html, acedido em 12/11/2010.
- WILSON, T. (2002). "The Nonsense of Knowledge Management", *Information Research*, Vol. 8, N. 1, October 2002.
- WINCHESTER, S. (1998). *The Professor and the Madman: A Tale of Murder, Insanity, and the Making of the Oxford English Dictionary*, Harper Perennial, New York, USA.
- WINTER, R., & GREENE, E. (2007). "Juror Decision-Making", *Handbook of Applied Cognition*, Ed. DURSO, F. (2007), John Wiley & Sons, pp. 739-761.
- YUAN, W., CHEN, Y., WANG, R., & DU, Z. (2008). "Collective Intelligence in Knowledge Management", *Research and Practical Issues of Enterprise Information Systems II*, Volume 1, pp. 651-655.
- ZELENY, M. (1987). "Management Support Systems: Towards Integrated Knowledge Management", *Human Systems Management*, Elsevier science publishers, pp. 59-70.
- ZIPF, G. (1966 [1949]). *Human Behaviour and the Principle Of Least Effort*, Hafner Publishing Co Ltd, New York, USA.

ÍNDICE REMISSIVO

A

Ackoff, R., 30, 36, 187
Adams, J. et al., 35, 108, 187
Adler, B. et al., 96, 187
Alter, S., 53, 187
Amadores, 23, 26, 44, 67, 71, 105, 177, 178, 181
Amaral, L. et al., 53, 187
Anderson, C., 24, 43, 48, 70, 71, 91, 187
ANOVA, 129, 130
Ariely, D. et al., 41, 188
Austin, D., 63, 188, 196

B

Bacon, F., 136, 188
Benkler, Y., 38, 43, 46, 188
Bernstein, S. e Bernstein, R., 127, 128, 129, 188
Bloch, W., 33, 188
Borel, É., 34, 189
Borges, J. L., 33, 34, 188, 189
Bragues, G., 78, 81, 189
Britânica, 27, 28, 17, 18, 19, 21, 25, 27, 46, 77, 78,
84, 85, 87, 89, 90, 91, 92, 93, 94, 95, 97, 101, 103,
107, 109, 110, 112, 113, 117, 130, 138, 141, 149,
150, 151, 152, 153, 154, 156, 158, 159, 160, 161,
162, 164, 167, 168, 169, 170, 171, 172, 173, 177,
178, 179, 180, 181, 182, 183
Broughton, B., 88, 189
Brown-Forsythe, 129, 130
Buckingham, R. et al., 53, 189
Burke, J., 175, 189
Burton, R., 33, 189

C

Cardoso, J., 122, 190
Carpentier, M., 122, 190
Carregal, J., 23, 190
Carvalho, J., 53, 190
Chiavenato, I., 89, 190
Christensen, C., 23, 24, 191
Clauson, K. et al., 79, 81, 191
Cleveland, H., 30, 191
Cohen, M., 94, 104, 191
Colaboração em Massa, 17, 19, 24, 26, 27, 28, 29, 31,
36, 37, 38, 39, 40, 42, 44, 45, 47, 48, 49, 53, 54,
55, 62, 69, 71, 72, 73, 74, 75, 81, 83, 87, 88, 107,
137, 138, 176, 177, 178, 179, 180, 181, 183, 184,
185
Cook, S., 17, 47, 49, 191
Crowdsourcing, 28, 30, 49, 194, 198, 203

D

Dalkir, K., 53, 191
Deci, E., 41, 191
DeGrace, P. e Stahl, L., 44, 191
Descartes, R., 136, 137, 191
Devgan, L. et al., 78, 81, 192
Dunn, D., 29, 31, 127, 192

F

Fay, M. e Proschan, M., 126, 192
Fetzer, J., 29, 87, 192
Fisher, R., 125, 126, 137, 192
Floridi, L., 29, 87, 192
Freixo, M., 136, 137, 193

G

Galton, F., 56, 57, 107, 193
Gaur, A. e Gaur, S. 108, 109, 193
Gestão do Conhecimento, 44, 53
Ghazawneh, A., 38, 193
Gibbons, J. 33, 34, 193
Gide, A., 28, 193
Giles, J., 25, 77, 81, 83, 85, 178, 193
Given, L., 108, 193
Gobillot, E., 44, 45, 46, 87, 89, 193
Gorard, F., 125, 193

H

Halavais, A. e Lackaff, D., 112, 194
Hardt, M. e Negri, A., 37, 63, 194
Harlow, H. et al., 40, 194, 201
Hayek, F., 56, 57, 194, 202
Hobbes, T., 37, 55, 93, 194
Howard, A., 104, 194
Howe, J., 38, 43, 48, 49, 194
Huff, D., 114, 149, 195
Hurwicz, L., 94, 195

I

Inês, P., 99, 195
Inteligência Colectiva, 17, 28, 36, 38, 48, 49, 51, 52,
53, 54

J

Johnson, S., 90, 195
Jólluskin, G., 104, 105, 195

K

Keen, A., 25, 26, 34, 43, 67, 72, 86, 87, 195
Klare, G., 96, 195
Kruskal, W. e Wallis, W., 129, 130, 147, 196

L

Le Bom, G., 61, 196
Lessing, D., 72, 196
Levene, H., 129, 130, 196
Lévy, P., 36, 196
Lih, A., 76, 88, 90, 196
Likert, R., 100, 196
Loureiro, J., 25, 197

M

Mackay, C., 64, 65, 197
Malone, T., 17, 51, 52, 54, 90, 197
Mann-Whitney, 21, 129, 130, 160, 192
Marx, K. e Engels, F., 67, 84, 197
Maslow, A., 39, 40, 41, 42, 197
Mayfield, R., 17, 48, 197
Mendes, J., 93, 197
Miller, J. et al., 97, 198
Moody, G., 62, 198

N

Negri, A., 37, 63, 194, 198
Nielson, J., 74, 198
Nonaka, I. e Takeuchi, I., 44, 53, 198

O

O'Reilly, T., 69, 189, 199

P

Page, S., 35, 58, 59, 60, 193, 199
Park, H., 145, 199
Pender, M. et al., 79, 81, 199
Peritos, 27, 23, 24, 56, 58, 67, 77, 83, 84, 85, 87, 99,
100, 104, 105, 106, 118, 122, 181, 182
Pink, D., 26, 28, 41, 42, 78, 89, 176, 199
Platão, 29, 31, 94, 199
Popper, K., 137, 199
Postman, N., 66, 199

R

Rector, L. H., 79, 81, 200
Reto, L. e Nunes, F., 136, 200
Robalo, M., 185, 200
Rosenzweig, R., 78, 81, 200
Rousseau, J., 93, 200
Ruane, J., 108, 200

S

Saunders, M. et al., 135, 137, 201
Schulman, T., 96, 201
Serrano, A. e Fialho, C., 25, 53, 187, 201
Shannon, C., 17, 31, 32, 33, 87, 201
Shapiro, S. e Wilk, M., 21, 127, 129, 130, 145, 147,
153, 154, 158, 159, 170, 201
Shirky, C., 23, 24, 26, 27, 38, 43, 54, 76, 94, 95, 176,
198, 201
Siegel, L., 43, 66, 201
Sistemas de Informação, 28, 53, 187, 195
Sokal, A., 72, 93, 201
Sousa, A., 184, 201
Spinoza, B., 37, 55, 202
Sunstein, C., 24, 43, 57, 202
Surowiecki, J., 24, 26, 37, 43, 49, 57, 58, 61, 64, 65,
176, 202
Szoka, B. e Marcus, A., 195, 202

T

t Student, 129, 131, 190, 204
Tancred, H., 34, 202
Tapscott, D. e Williams, A., 24, 25, 26, 36, 37, 38,
39, 43, 71, 72, 86, 87, 202
Thierer, A., 17, 43, 202
Thoreau, D., 64, 202
Thorndike, L., 96, 202
Toutenburg, H. e Shalabh, S., 124, 203

U

Urdan, T., 108, 203

V

Vanderstoep, S. et Johnston, D., 82, 108, 109, 203
Viitamäki, S., 17, 50, 203
Virno, P., 37, 203
Vox Expertorum, 56, 58, 104, 117
Vox Populi, 56, 57, 58, 104, 117, 120, 193

W

Wang, R. e Strong, D., 53, 98, 99, 100, 120, 203, 204
Weaver, W., 35, 203
Whitman, W. 180, 203
Wikipédia, 27, 28, 31, 17, 18, 19, 21, 24, 25, 26, 27,
28, 35, 36, 42, 46, 47, 49, 54, 69, 73, 74, 75, 76,
77, 78, 79, 80, 81, 82, 83, 84, 85, 87, 88, 89, 90,
91, 92, 94, 95, 96, 101, 103, 107, 108, 109, 110,
112, 113, 117, 118, 130, 137, 138, 141, 142, 149,
150, 151, 152, 153, 154, 156, 158, 159, 160, 161,
162, 163, 164, 167, 168, 169, 170, 171, 172, 173,
177, 178, 179, 180, 181, 182, 183, 184, 185, 206
Wilcoxon, F., 21, 128, 129, 130, 146, 154, 158, 159,
171, 192, 204
Winchester, S., 62, 204
Winter, R. e Greene, E., 105, 204

Y

Yuan, W. et al., 53, 204

Z

Zeleny, M., 30, 204
Zipf, G., 96, 204

ANEXOS

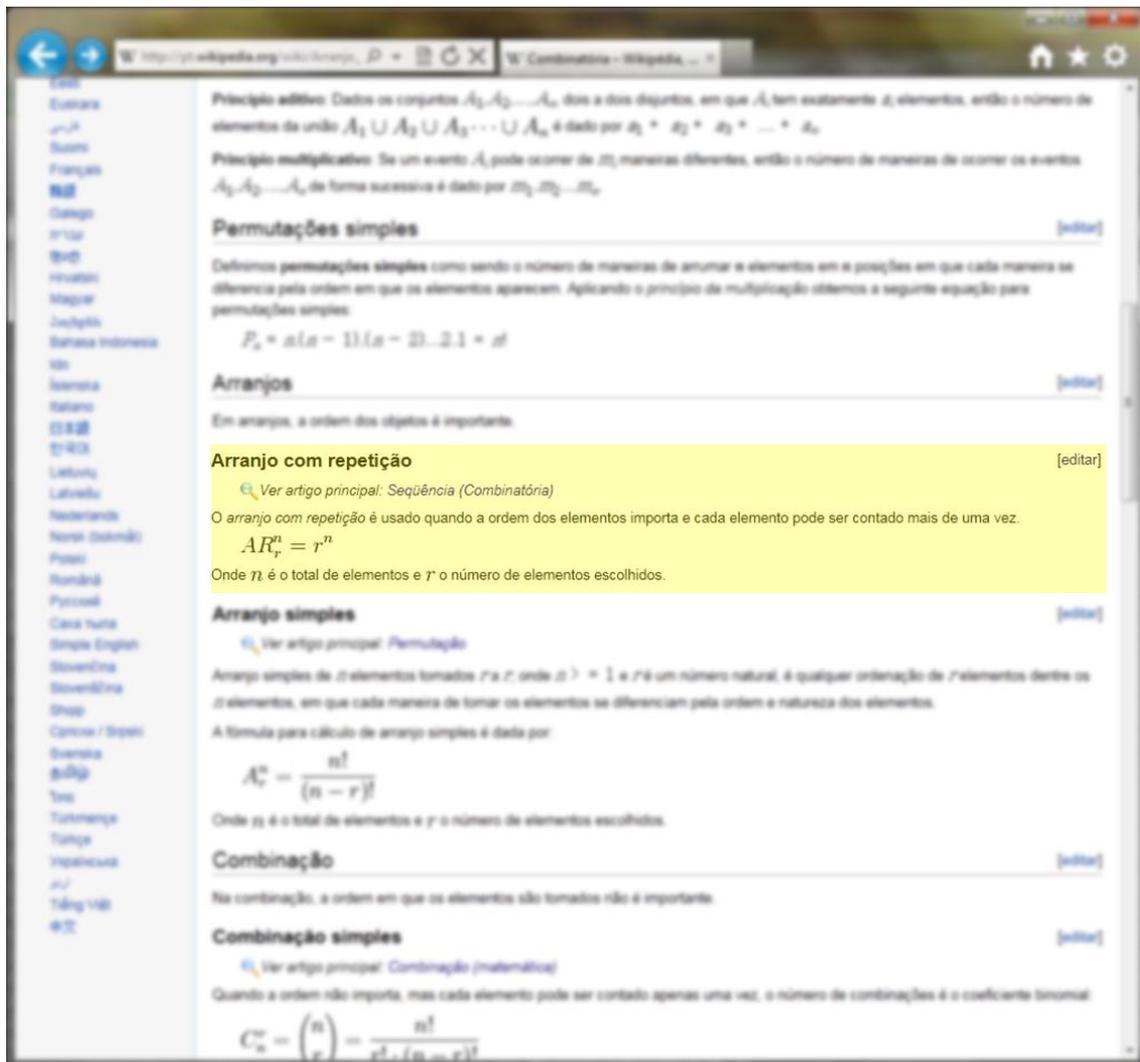
EDIÇÃO DE UM ARTIGO DA WIKIPÉDIA	I
AVALIAÇÃO DOS PARES DE ARTIGOS DA WIKIPÉDIA E DA BRITÂNICA	II
CARACTERIZAÇÃO DOS HÁBITOS DE CONSULTA DE INFORMAÇÃO	III
50 PRINCÍPIOS DE GESTÃO.....	IV

ANEXO I

EDIÇÃO DE UM ARTIGO DA WIKIPÉDIA

Exemplo da correcção de um erro encontrado na Wikipédia em Junho de 2011, que demonstra a facilidade e a ausência de necessidade de registo com que o mesmo pode ser efectuado. Apesar de potenciar o número de colaborações, esta facilidade torna a Wikipédia vulnerável à introdução propositada de erros.

A secção editada encontra-se destacada, para uma mais fácil identificação.



Wikipédia - A enciclopédia livre

Artigo **A editar Combinatória (seção)**

Origem: Wikipédia, a enciclopédia livre.

Atenção! Esta é uma página do domínio principal (artigo)!

- Não copie textos de outras fontes para a Wikipédia, pois isto pode ferir o direito autoral. Use suas próprias palavras.
- Inclua fontes fiáveis no seu texto, por meio de notas de rodapé.
- Não escreva sua autobiografia aqui. Pode utilizar sua página de usuário de acordo com as regras.
- Se as informações forem consideradas falsas ou não enciclopédicas elas poderão ser eliminadas.
- Não use "internetês"; na Wikipédia usa-se a norma-padrão da língua portuguesa, e suas edições poderão ser consideradas vandalismo.
- Não altere as versões da língua portuguesa (pt-BR para pt-PT ou pt-PT para pt-BR); se persistir nessa atitude, poderá ser bloqueado.
- Não assine este artigo nem escreva seu nome. Suas edições ficarão registradas no histórico de edições da página.
- Todo o conteúdo da Wikipédia é indexado pela maioria dos motores de busca.
- Veja também os cinco pilares da Wikipédia, nossas políticas e recomendações e a página de ajuda.

Você não está autenticado(a). Embora sua edição seja bem-vinda, o seu endereço IP será registrado no histórico desta página. Pode ocultar o seu IP criando uma conta, com a qual passará a dispor de muitos outros benefícios.

Antes de criar um novo artigo, siga o nosso guia de edição para novos usuários.

N I Edição avançada | Caracteres especiais | Ajuda

=== Arranjo com repetição ===

```
{{Ver artigo principal|[[Seqüência (Combinatória)]}}
```

O "arranjo com repetição" é usado quando a ordem dos elementos importa e cada elemento pode ser contado mais de uma vez.

`$AR^n_r = r^n$`

Onde `n` é o total de elementos e `r` é o número de elementos escolhidos.

Princípio aditivo: Dados os conjuntos A_1, A_2, \dots, A_n , dois a dois disjuntos, em que A_i tem exatamente a_i elementos, então o número de elementos da união $A_1 \cup A_2 \cup A_3 \dots \cup A_n$ é dado por $a_1 + a_2 + a_3 + \dots + a_n$.

Princípio multiplicativo: Se um evento A_i pode ocorrer de a_i maneiras diferentes, então o número de maneiras de ocorrer os eventos A_1, A_2, \dots, A_n de forma sucessiva é dado por $a_1 \cdot a_2 \cdot \dots \cdot a_n$.

Permutações simples

Definimos permutações simples como sendo o número de maneiras de arrumar n elementos em n posições em que cada maneira se diferencia pela ordem em que os elementos aparecem. Aplicando o princípio de multiplicação obtemos a seguinte equação para permutações simples:

$$P_n = n(n-1)(n-2) \dots 2 \cdot 1 = n!$$

Arranjos

Em arranjos, a ordem dos objetos é importante.

Arranjo com repetição

Ver artigo principal: Seqüência (Combinatória)

O arranjo com repetição é usado quando a ordem dos elementos importa e cada elemento pode ser contado mais de uma vez.

$$AR_r^n = n^r$$

Onde n é o total de elementos e r o número de elementos escolhidos.

Arranjo simples

Ver artigo principal: Permutação

Arranjo simples de n elementos tomados r a r , onde $n \geq r \geq 1$ e r é um número natural, é qualquer ordenação de r elementos dentre os n elementos, em que cada maneira de tomar os elementos se diferencia pela ordem e natureza dos elementos.

A fórmula para cálculo de arranjo simples é dada por:

$$A_n^r = \frac{n!}{(n-r)!}$$

Onde n é o total de elementos e r o número de elementos escolhidos.

Combinação

Na combinação, a ordem em que os elementos são tomados não é importante.

Combinação simples

Ver artigo principal: Combinação (matemática)

Quando a ordem não importa, mas cada elemento pode ser contado apenas uma vez, o número de combinações é o coeficiente binomial

$$C_n^r = \binom{n}{r} = \frac{n!}{r! \cdot (n-r)!}$$

ANEXO II

ÁVALIAÇÃO DOS PARES DE ARTIGOS DA WIKIPÉDIA E DA BRITÂNICA

Compilação das avaliações qualitativas e quantitativas efectuadas ao longo de 2010 e 2011 aos 245 pares de artigos. Os artigos foram recolhidos durante as primeiras semanas de 2010 da versão inglesa da Wikipédia e da Edição Premium da Enciclopédia Britânica Online. As avaliações foram efectuadas pelos peritos¹ cujos nomes² se encontram na figura abaixo e a quem aproveito para expressar, mais uma vez, o meu profundo agradecimento.



¹ O número de ocorrências de um nome indica a quantidade de pares de artigos que este avaliou.
² Três docentes do Departamento de História da UE optaram por manter o anonimato.

Área	Número			Linhas	URL			
	Avaliador	Afiliação	Nota			Ilustrações	Palavras	Caracteres
Henry Hobson Richardson								
Arquitetura	W	4	41	2973	17007	309	http://en.wikipedia.org/wiki/Henry_Hobson_Richardson	
11	UE	B	3	1	919	5085	92	http://www.britannica.com/EBchecked/topic/502649/Henry-Hobson-Richardson
Mozarabic art and architecture								
Arquitetura	W	3	9	1329	7065	128	http://en.wikipedia.org/wiki/Mozarabic_art_and_architecture	
11	UE	B	3	1	101	577	10	http://www.britannica.com/EBchecked/topic/1569000/Mozarabic-architecture
Norman architecture								
Arquitetura	W	4	14	1923	10967	199	http://en.wikipedia.org/wiki/Norman_architecture	
11	UE	B	4	5	787	4591	83	http://www.britannica.com/EBchecked/topic/418340/Norman-style
Black Orpheus								
Cinema	W	3	2	1256	6203	113	http://en.wikipedia.org/wiki/Black_Orpheus	
38	UB	B	2	1	194	1025	19	http://www.britannica.com/EBchecked/topic/68125/Black-Orpheus
Michael Redgrave								
Cinema	W	5	1	3277	16757	305	http://en.wikipedia.org/wiki/Sir_Michael_Redgrave	
38	UB	B	1	0	278	1486	27	http://www.britannica.com/EBchecked/topic/494679/Sir-Michael-Scudamore-Redgrave
Midnight Cowboy								
Cinema	W	4	2	1914	9239	168	http://en.wikipedia.org/wiki/Midnight_Cowboy	
38	UB	B	3	1	174	873	16	http://www.britannica.com/EBchecked/topic/381606/Midnight-Cowboy
Richard Burton								
Cinema	W	4	3	6774	34572	629	http://en.wikipedia.org/wiki/Richard_Burton	
38	UB	B	4	1	337	1757	32	http://www.britannica.com/EBchecked/topic/85898/Richard-Burton
The Sound of Music								
Cinema	W	4	5	6049	30661	557	http://en.wikipedia.org/wiki/The_Sound_of_Music_%28film%29	
38	UB	B	2	1	236	1312	24	http://www.britannica.com/EBchecked/topic/555356/The-Sound-of-Music
The Three Stooges								
Cinema	W	5	10	9228	48039	873	http://en.wikipedia.org/wiki/Three_Stooges	
38	UB	B	5	1	1018	5186	94	http://www.britannica.com/EBchecked/topic/593914/the-Three-Stooges
Akebono Tarō								
Desporto	W	4	2	2537	12722	231	http://en.wikipedia.org/wiki/Akebono_Tar%C5%8D	
31	UE	B	3	0	592	3054	56	http://www.britannica.com/EBchecked/topic/11487/Akebono
Dallas Mavericks								
Desporto	W	5	3	4026	20237	368	http://en.wikipedia.org/wiki/Dallas_Mavericks	
31	UE	B	3	0	310	1596	29	http://www.britannica.com/EBchecked/topic/1271184/Dallas-Mavericks
Di Stéfano								
Desporto	W	5	2	2161	10961	199	http://en.wikipedia.org/wiki/Alfredo_Di_St%C3%A9fano	
31	UE	B	2	1	311	1632	30	http://www.britannica.com/EBchecked/topic/936749/Alfredo-Di-Stefano

	Área		Número					URL	
	Avaliador	Afiliação	Nota	Ilustrações	Palavras	Caracteres	Linhas		
Artes & Entretenimento	Doc Rivers								
	Desporto	W	4	2	1096	5636	102	http://en.wikipedia.org/wiki/Glenn_Anton_Rivers	
	31	UE	B	4	0	470	2372	43	http://www.britannica.com/EBchecked/topic/1459452/Glenn-Rivers
	high jump								
	Desporto	W	5	6	3312	17397	316	http://en.wikipedia.org/wiki/High_jump	
	31	UE	B	2	4	315	1514	28	http://www.britannica.com/EBchecked/topic/265250/high-jump
	Irina Rodnina								
	Desporto	W	5	1	929	4937	90	http://en.wikipedia.org/wiki/Irina_Rodnina	
	31	UE	B	3	1	213	1164	21	http://www.britannica.com/EBchecked/topic/506649/Irina-Rodnina
	New York Rangers								
	Desporto	W	5	5	8262	44275	805	http://en.wikipedia.org/wiki/New_York_Rangers	
	31	UE	B	3	0	473	2348	43	http://www.britannica.com/EBchecked/topic/665363/New-York-Rangers
	Richard Petty								
	Desporto	W	5	7	5088	25890	471	http://en.wikipedia.org/wiki/Richard_Petty	
	31	UE	B	3	1	367	1843	34	http://www.britannica.com/EBchecked/topic/758273/Richard-Petty
	Joan Lluís Vives								
	Filosofia	W	2	2	1007	5702	104	http://en.wikipedia.org/wiki/Juan_Luis_Vives	
	37	IPB	B	2	1	306	1598	29	http://www.britannica.com/EBchecked/topic/631418/Juan-Luis-Vives
	codex								
	História	W	4	5	1511	7695	140	http://en.wikipedia.org/wiki/Codex	
8	UE	B	2	1	366	1876	34	http://www.britannica.com/EBchecked/topic/123916/codex	
Don Juan									
História	W	3	2	1735	8890	162	http://en.wikipedia.org/wiki/Don_Juan		
4	II	B	4	0	460	2399	44	http://www.britannica.com/EBchecked/topic/168882/Don-Juan	
Hambletonian 10									
História	W	5	1	732	3714	68	http://en.wikipedia.org/wiki/Hambletonian_10		
4	II	B	4	0	130	697	13	http://www.britannica.com/EBchecked/topic/253220/Hambletonian	
Suzhou embroidery									
História	W	3	0	135	767	14	http://en.wikipedia.org/wiki/Suzhou_embroidery		
3	UE	B	4	0	107	567	10	http://www.britannica.com/EBchecked/topic/570408/Suzhou-embroidery	
Pong									
Informática	W	5	6	4535	25012	455	http://en.wikipedia.org/wiki/Pong		
22	UE	B	3	0	253	1245	23	http://www.britannica.com/EBchecked/topic/932625/Pong	
Booth Tarkington									
Literatura	W	2	2	1348	7209	131	http://en.wikipedia.org/wiki/Booth_Tarkington		
17	UL	B	1	1	258	1498	27	http://www.britannica.com/EBchecked/topic/583605/Booth-Tarkington	

	Área		Número					URL	
	Avaliador	Afiliação	Nota	Ilustrações	Palavras	Caracteres	Linhas		
Artes & Entretenimento	C. S. Lewis								
	Literatura	W	4	8	9933	51083	929	http://en.wikipedia.org/wiki/C._S._Lewis	
	17	UL	B	1	3	422	2143	39	http://www.britannica.com/EBchecked/topic/338121/C-S-Lewis
	Decadent movement								
	Literatura	W	2	1	532	3295	60	http://en.wikipedia.org/wiki/Decadent_movement	
	17	UL	B	1	0	54	262	5	http://www.britannica.com/EBchecked/topic/154853/decadence
	Gísla saga								
	Literatura	W	2	1	1167	5895	107	http://en.wikipedia.org/wiki/G%C3%ADsla_saga	
	17	UL	B	1	0	73	372	7	http://www.britannica.com/EBchecked/topic/234354/Gisla-saga
	Johann Jakob Bodmer								
	Literatura	W	2	1	299	1617	29	http://en.wikipedia.org/wiki/Johann_Jakob_Bodmer	
	17	UL	B	1	1	191	1068	19	http://www.britannica.com/EBchecked/topic/71053/Johann-Jakob-Bodmer
	John le Carré								
	Literatura	W	4	1	3294	17570	319	http://en.wikipedia.org/wiki/John_le_Carr%C3%A9	
	17	UL	B	2	1	398	2030	37	http://www.britannica.com/EBchecked/topic/333310/John-le-Carre
	Jonathan Franzen								
	Literatura	W	4	1	2537	14199	258	http://en.wikipedia.org/wiki/Jonathan_Franzen	
	17	UL	B	1	1	376	2055	37	http://www.britannica.com/EBchecked/topic/1284424/Jonathan-Franzen
	Kavya								
	Literatura	W	1	0	146	785	14	http://en.wikipedia.org/wiki/K%C4%81vya	
17	UL	B	2	0	220	1195	22	http://www.britannica.com/EBchecked/topic/313574/kavya	
Malcolm Lowry									
Literatura	W	2	1	1121	5879	107	http://en.wikipedia.org/wiki/Malcolm_Lowry		
17	UL	B	2	0	354	1779	32	http://www.britannica.com/EBchecked/topic/350045/Malcolm-Lowry#ref=ref263620	
Miguel Ángel Asturias									
Literatura	W	4	5	6642	34818	633	http://en.wikipedia.org/wiki/Miguel_%C3%81ngel_Asturias		
17	UL	B	1	2	340	1816	33	http://www.britannica.com/EBchecked/topic/40122/Miguel-Angel-Asturias	
Nikos Kazantzakis									
Literatura	W	3	4	2809	15123	275	http://en.wikipedia.org/wiki/Nikos_Kazantzakis		
17	UL	B	1	1	319	1715	31	http://www.britannica.com/EBchecked/topic/313888/Nikos-Kazantzakis#ref=ref116509	
Sawako Ariyoshi									
Literatura	W	2	0	421	2233	41	http://en.wikipedia.org/wiki/Ariyoshi_Sawako		
17	UL	B	1	0	263	1472	27	http://www.britannica.com/EBchecked/topic/34806/Ariyoshi-Sawako	
Stendhal									
Literatura	W	3	3	1510	7882	143	http://en.wikipedia.org/wiki/Stendhal		
17	UL	B	4	1	2509	13278	241	http://www.britannica.com/EBchecked/topic/565258/Stendhal	

	Área		Número					URL	
	Avaliador	Afiliação	Nota	Ilustrações	Palavras	Caracteres	Linhas		
Artes & Entretenimento	Twilight								
	Literatura	W	4	5	6780	38236	695	http://en.wikipedia.org/wiki/Twilight_Saga	
	17	UL	B	2	0	303	1619	29	http://www.britannica.com/EBchecked/topic/1458382/Twilight-Saga
	V. S. Naipaul								
	Literatura	W	2	0	1805	9602	175	http://en.wikipedia.org/wiki/V._S._Naipaul	
	17	UL	B	1	1	437	2353	43	http://www.britannica.com/EBchecked/topic/401894/Sir-V-S-Naipaul#ref=ref749949
	Vittorio Sereni								
	Literatura	W	2	0	252	1330	24	http://en.wikipedia.org/wiki/Vittorio_Sereni	
	17	UL	B	2	0	269	1418	26	http://www.britannica.com/EBchecked/topic/535475/Vittorio-Sereni
	William Blake								
	Literatura	W	4	15	10667	56590	1029	http://en.wikipedia.org/wiki/William_Blake	
	17	UL	B	4	5	7471	37236	677	http://www.britannica.com/EBchecked/topic/68793/William-Blake/261262/Blake-as-a-poet#ref=ref211229
	barrel piano								
	Música	W	3	1	626	3128	57	http://en.wikipedia.org/wiki/Barrel_piano	
	18	UE	B	2	0	99	499	9	http://www.britannica.com/EBchecked/topic/54036/barrel-piano
	Bhangra								
	Música	W	4	3	5724	29793	542	http://en.wikipedia.org/wiki/Bhangra	
	18	UE	B	4	1	986	5313	97	http://www.britannica.com/EBchecked/topic/63998/bhangra
	Black Flag								
	Música	W	5	3	5982	30995	564	http://en.wikipedia.org/wiki/Black_Flag_%28band%29	
18	UE	B	3	0	366	1879	34	http://www.britannica.com/EBchecked/topic/67829/Black-Flag	
Giovanni Carlo Maria Clari									
Música	W	3	0	197	997	18	http://en.wikipedia.org/wiki/Giovanni_Carlo_Maria_Clari		
18	UE	B	3	0	126	668	12	http://www.britannica.com/EBchecked/topic/119827/Giovanni-Carlo-Maria-Clari	
Hamilton Harty									
Música	W	3	1	921	4785	87	http://en.wikipedia.org/wiki/Hamilton_Harty		
18	UE	B	2	0	116	608	11	http://www.britannica.com/EBchecked/topic/256239/Sir-Hamilton-Harty	
Little Willie John									
Música	W	4	2	1054	5249	95	http://en.wikipedia.org/wiki/Little_Willie_John		
18	UE	B	3	0	155	758	14	http://www.britannica.com/EBchecked/topic/304638/Little-Willie-John	
Outlaw country									
Música	W	5	4	1702	8637	157	http://en.wikipedia.org/wiki/Outlaw_music		
18	UE	B	3	0	281	1557	28	http://www.britannica.com/EBchecked/topic/435612/outlaw-music	
Quincy Jones									
Música	W	4	4	3644	19134	348	http://en.wikipedia.org/wiki/Quincy_Jones		
18	UE	B	3	1	442	2308	42	http://www.britannica.com/EBchecked/topic/709663/Quincy-Jones	

	Área					Número			URL
	Avaliador	Afiliação	Nota	Ilustrações	Palavras	Caracteres	Linhas		
Artes & Entretenimento	R.E.M.								
		Música	W	5	3	7554	39838	724	http://en.wikipedia.org/wiki/R.e.m.
	18	UE	B	4	2	795	4237	77	http://www.britannica.com/EBchecked/topic/497516/REM
	Talking Heads								
		Música	W	4	3	1979	10319	188	http://en.wikipedia.org/wiki/Talking_Heads
	18	UE	B	3	1	558	3014	55	http://www.britannica.com/EBchecked/topic/581530/Talking-Heads
	The Andrews Sisters								
		Música	W	5	2	5908	30295	551	http://en.wikipedia.org/wiki/The_Andrews_Sisters
	18	UE	B	3	1	656	3402	62	http://www.britannica.com/EBchecked/topic/1299246/The-Andrew-Sisters
	The Cleveland Orchestra								
		Música	W	4	2	834	5075	92	http://en.wikipedia.org/wiki/Cleveland_Orchestra
	18	UE	B	3	0	206	1204	22	http://www.britannica.com/EBchecked/topic/121437/Cleveland-Orchestra
	The Mekons								
		Música	W	4	0	1073	5342	97	http://en.wikipedia.org/wiki/The_Mekons
	18	UE	B	3	0	232	1286	23	http://www.britannica.com/EBchecked/topic/373586/The-Mekons
	Fernand Léger								
		Pintura	W	4	6	1970	10160	185	http://en.wikipedia.org/wiki/Fernand_L%C3%A9ger
	15	MJV	B	2	1	963	5011	91	http://www.britannica.com/EBchecked/topic/334999/Fernand-Leger
George Catlin									
	Pintura	W	4	7	1253	6763	123	http://en.wikipedia.org/wiki/George_Catlin	
15	MJV	B	1	2	291	1545	28	http://www.britannica.com/EBchecked/topic/99958/George-Catlin	

Área	Avaliador	Afiliação	Nota	Ilustrações	Número			URL
					Palavras	Caracteres	Linhas	
Managerial economics								
Economia	7	W	2	0	514	3298	60	http://en.wikipedia.org/wiki/Managerial_economics
	UE	B	3	0	110	727	13	http://www.britannica.com/EBchecked/topic/361242/managerial-economics
abstract object								
Filosofia	37	W	3	0	590	3229	59	http://en.wikipedia.org/wiki/Abstract_object
	IPB	B	2	0	231	1324	24	http://www.britannica.com/EBchecked/topic/131303/concrete
John Dewey								
Filosofia	37	W	4	1	6317	35203	640	http://en.wikipedia.org/wiki/John_Dewey
	IPB	B	3	2	1758	9529	173	http://www.britannica.com/EBchecked/topic/160445/John-Dewey
Juvenal of Jerusalem								
Filosofia	37	W	4	0	224	1079	20	http://en.wikipedia.org/wiki/Juvenal_of_Jerusalem
	IPB	B	3	0	118	594	11	http://www.britannica.com/EBchecked/topic/308980/Saint-Juvenal
Normative ethics								
Filosofia	37	W	4	0	918	5053	92	http://en.wikipedia.org/wiki/Normative_ethics
	IPB	B	3	0	272	1611	29	http://www.britannica.com/EBchecked/topic/418412/normative-ethics
Acts of Union								
História	4	W	4	6	3796	19892	362	http://en.wikipedia.org/wiki/Acts_of_Union,_in_1707
	II	B	4	0	325	1781	32	http://www.britannica.com/EBchecked/topic/614670/Act-of-Union
Ahmad Sirhindi								
História	8	W	5	0	2416	12228	222	http://en.wikipedia.org/wiki/Ahmad_Sirhindi
	UE	B	5	0	437	2291	42	http://www.britannica.com/EBchecked/topic/10170/Shaykh-Ahmad-Sirhindi
American Indian Movement								
História	9	W	4	1	6363	35619	648	http://en.wikipedia.org/wiki/American_Indian_Movement
	UE	B	1	0	240	1319	24	http://www.britannica.com/EBchecked/topic/19799/American-Indian-Movement
Arab–Israeli conflict								
História	8	W	5	5	7117	39541	719	http://en.wikipedia.org/wiki/Arab-Israeli_Conflict
	UE	B	3	4	1065	5504	100	http://www.britannica.com/EBchecked/topic/31439/Arab-Israeli-wars
Battle of Adrianople								
História	4	W	4	3	2646	13842	252	http://en.wikipedia.org/wiki/Battle_of_Adrianople
	II	B	3	0	147	799	15	http://www.britannica.com/EBchecked/topic/6534/Battle-of-Adrianople
Charles II, Duke of Lorraine								
História	8	W	3	1	757	3763	68	http://en.wikipedia.org/wiki/Charles_II,_Duke_of_Lorraine
	UE	B	4	0	178	878	16	http://www.britannica.com/EBchecked/topic/106755/Charles-I
Christian Magnus Falsen								
História	3	W	4	1	341	1743	32	http://en.wikipedia.org/wiki/Christian_Magnus_Falsen
	UE	B	3	0	261	1386	25	http://www.britannica.com/EBchecked/topic/201085/Christian-Magnus-Falsen

Área	Número		URL				
	Avaliador	Afiliação	Nota	Ilustrações	Palavras	Caracteres	Linhas
Cimbri							
História	W	5	3	2961	15070	274	http://en.wikipedia.org/wiki/Cimbri
4	II	B	2	0	113	595	11 http://www.britannica.com/EBchecked/topic/117886/Cimbri
Civil Constitution of the Clergy							
História	W	5	3	2177	11356	206	http://en.wikipedia.org/wiki/Civil_Constitution_of_the_Clergy
4	II	B	4	0	352	1868	34 http://www.britannica.com/EBchecked/topic/119210/Civil-Constitution-of-the-Clergy
civil township							
História	W	3	3	1248	6982	127	http://en.wikipedia.org/wiki/Civil_township
4	II	B	5	0	262	1439	26 http://www.britannica.com/EBchecked/topic/601134/township
Cleomenes III							
História	W	5	0	1674	8617	157	http://en.wikipedia.org/wiki/Cleomenes_III
4	II	B	3	0	297	1579	29 http://www.britannica.com/EBchecked/topic/121201/Cleomenes-III
Cynthia Ann McKinney							
História	W	5	4	8111	45833	833	http://en.wikipedia.org/wiki/Cynthia_Ann_McKinney
4	II	B	3	1	379	2080	38 http://www.britannica.com/EBchecked/topic/1471371/Cynthia-McKinney
Dôn							
História	W	4	1	259	1454	26	http://en.wikipedia.org/wiki/D%C3%B4n
3	UE	B	2	0	141	715	13 http://www.britannica.com/EBchecked/topic/168849/Don
Elihu Root							
História	W	3	4	1682	8636	157	http://en.wikipedia.org/wiki/Elihu_Root
9	UE	B	2	1	451	2391	43 http://www.britannica.com/EBchecked/topic/509465/Elihu-Root
Evagoras							
História	W	5	0	560	2790	51	http://en.wikipedia.org/wiki/Evagoras
4	II	B	3	0	231	1160	21 http://www.britannica.com/EBchecked/topic/196796/Evagoras
Frederick III, Elector of Saxony							
História	W	5	2	489	2501	45	http://en.wikipedia.org/wiki/Frederick_III,_Elector_of_Saxony
3	UE	B	4	1	222	1162	21 http://www.britannica.com/EBchecked/topic/217946/Frederick-III
Galla Placidia							
História	W	4	3	4101	21595	393	http://en.wikipedia.org/wiki/Aelia_Galla_Placidia
4	II	B	2	0	131	670	12 http://www.britannica.com/EBchecked/topic/462612/Aelia-Galla-Placidia
Ghadar Party							
História	W	2	3	1004	5438	99	http://en.wikipedia.org/wiki/Ghadar_Party
9	UE	B	1	0	105	598	11 http://www.britannica.com/EBchecked/topic/232350/Ghadar
Herodotus							
História	W	5	6	5386	29152	530	http://en.wikipedia.org/wiki/Herodotus
4	II	B	3	2	2070	10095	184 http://www.britannica.com/EBchecked/topic/263507/Herodotus

Área	Número		URL				
	Avaliador	Afiliação	Nota	Ilustrações	Palavras	Caracteres	Linhas
Howard Hughes							
História	W	4	9	8532	45111	820	http://en.wikipedia.org/wiki/Howard_Hughes
9	UE	B	1	2	563	2842	http://www.britannica.com/EBchecked/topic/274919/Howard-Hughes#
James Douglas, 9th Earl of Douglas							
História	W	4	0	434	2111	38	http://en.wikipedia.org/wiki/James_Douglas,_9th_Earl_of_Douglas
4	II	B	3	0	233	1104	http://www.britannica.com/EBchecked/topic/170173/James-Douglas-9th-earl-of-Douglas
Kazakhs							
História	W	5	16	3828	21515	391	http://en.wikipedia.org/wiki/Kazakhs
8	UE	B	3	1	407	2180	http://www.britannica.com/EBchecked/topic/313741/Kazakh
Lord George Murray							
História	W	5	1	1381	6939	126	http://en.wikipedia.org/wiki/Lord_George_Murray
4	II	B	2	0	187	1028	http://www.britannica.com/EBchecked/topic/398232/Lord-George-Murray
Minamoto no Yoshiie							
História	W	4	1	736	3703	67	http://en.wikipedia.org/wiki/Minamoto_Yoshiie
3	UE	B	3	0	211	1117	http://www.britannica.com/EBchecked/topic/383446/Minamoto-Yoshiie
Muromachi period							
História	W	5	7	1951	10923	199	http://en.wikipedia.org/wiki/Muromachi_period
3	UE	B	2	0	197	1085	http://www.britannica.com/EBchecked/topic/398169/Muromachi-period
Oakes Ames							
História	W	2	2	744	3969	72	http://en.wikipedia.org/wiki/Oakes_Ames
9	UE	B	2	1	294	1573	http://www.britannica.com/EBchecked/topic/20447/Oakes-Ames
Order of the Garter							
História	W	4	15	5351	27147	494	http://en.wikipedia.org/wiki/The_Most_Noble_Order_of_the_Garter
4	II	B	5	0	771	3741	http://www.britannica.com/EBchecked/topic/226255/The-Most-Noble-Order-of-the-Garter
Quechuan languages							
História	W	5	3	4279	23893	434	http://en.wikipedia.org/wiki/Quechuan_languages
4	II	B	3	0	144	775	http://www.britannica.com/EBchecked/topic/486747/Quechuan-languages
Satrap							
História	W	4	0	1240	6700	122	http://en.wikipedia.org/wiki/Satrap
8	UE	B	4	0	192	1040	http://www.britannica.com/EBchecked/topic/525034/satrap
Siouan languages							
História	W	4	1	684	4059	74	http://en.wikipedia.org/wiki/Siouan
4	II	B	3	0	96	578	http://www.britannica.com/EBchecked/topic/546400/Siouan-languages
Stasi							
História	W	4	11	4952	26935	490	http://en.wikipedia.org/wiki/Stasi
9	UE	B	2	0	628	3486	http://www.britannica.com/EBchecked/topic/563751/Stasi

Área	Número		URL				
	Avaliador	Afiliação	Nota	Ilustrações	Palavras	Caracteres	Linhas
Stephen Arnold Douglas							
História	W	3	8	3496	18379	334	http://en.wikipedia.org/wiki/Stephen_A._Douglas
9	UE	B	1	695	3677	67	http://www.britannica.com/EBchecked/topic/170192/Stephen-A-Douglas
The University of Western Ontario							
História	W	5	9	3407	19380	352	http://en.wikipedia.org/wiki/University_of_Western_Ontario
4	II	B	2	51	340	6	http://www.britannica.com/EBchecked/topic/640756/University-of-Western-Ontario
Tokugawa Ieyasu							
História	W	5	8	5441	27497	500	http://en.wikipedia.org/wiki/Tokugawa_Ieyasu
3	UE	B	3	1772	9306	169	http://www.britannica.com/EBchecked/topic/598292/Tokugawa-Ieyasu
trade union							
História	W	5	4	7136	38626	702	http://en.wikipedia.org/wiki/Trade_union
3	UE	B	3	654	3730	68	http://www.britannica.com/EBchecked/topic/1357701/trade-union
Tuscarora							
História	W	4	2	3124	17062	310	http://en.wikipedia.org/wiki/Tuscarora_%28tribe%29
4	II	B	3	234	1317	24	http://www.britannica.com/EBchecked/topic/610553/Tuscarora
Ulster Unionist Party							
História	W	3	4	4374	24090	438	http://en.wikipedia.org/wiki/Ulster_Unionist_Party
9	UE	B	2	1292	6933	126	http://www.britannica.com/EBchecked/topic/689836/Ulster-Unionist-Party
Union of Orthodox Rabbis of the United States and Canada							
História	W	4	0	1015	5472	99	http://en.wikipedia.org/wiki/Union_of_Orthodox_Rabbis
8	UE	B	4	119	714	13	http://www.britannica.com/EBchecked/topic/1582056/Union-of-Orthodox-Rabbis-of-the-United-States-
Uralic languages							
História	W	5	4	4348	25266	459	http://en.wikipedia.org/wiki/Uralic_languages
8	UE	B	5	10018	53840	979	http://www.britannica.com/EBchecked/topic/619069/Uralic-languages
USS Constitution							
História	W	4	13	14297	74815	1360	http://en.wikipedia.org/wiki/USS_Constitution
9	UE	B	1	352	1783	32	http://www.britannica.com/EBchecked/topic/134174/Constitution
William Hastings, 1st Baron Hastings							
História	W	4	2	739	3676	67	http://en.wikipedia.org/wiki/William_Hastings,_1st_Baron_Hastings
8	UE	B	2	136	694	13	http://www.britannica.com/EBchecked/topic/256761/William-Hastings-Baron-Hastings
Battle of Leyte Gulf							
Militar	W	4	11	9539	48261	877	http://en.wikipedia.org/wiki/Battle_of_Leyte_Gulf
16	MP	B	1	286	1475	27	http://www.britannica.com/EBchecked/topic/338453/Battle-of-Leyte-Gulf
Military logistics							
Militar	W	3	4	1548	8875	161	http://en.wikipedia.org/wiki/Military_logistics
16	MP	B	5	12122	65875	1198	http://www.britannica.com/EBchecked/topic/346423/logistics

Avaliador	Área		Nota	Ilustrações	Número			URL
	Afiliação				Palavras	Caracteres	Linhas	
Abiathar								
	Religião	W	5	0	594	2910	53	http://en.wikipedia.org/wiki/Abiathar
36	ME	B	4	0	83	417	8	http://www.britannica.com/EBchecked/topic/1274/Abiathar
Covenant (biblical)								
	Religião	W	2	3	2722	13835	252	http://en.wikipedia.org/wiki/Covenant_%28biblical%29
36	ME	B	5	2	4922	25946	472	http://www.britannica.com/EBchecked/topic/141085/covenant
Joachim of Fiore								
	Religião	W	5	1	1307	6571	119	http://en.wikipedia.org/wiki/Joachim_of_Fiore
36	ME	B	5	2	1084	5468	99	http://www.britannica.com/EBchecked/topic/304176/Joachim-of-Fiore
Pope Leo VII								
	Religião	W	5	0	279	1329	24	http://en.wikipedia.org/wiki/Leo_VII
36	ME	B	5	0	113	513	9	http://www.britannica.com/EBchecked/topic/336244/Leo-VII
Saint Martin of Tours								
	Religião	W	5	11	5665	29151	530	http://en.wikipedia.org/wiki/Saint_Martin_of_Tours
36	ME	B	3	1	395	2020	37	http://www.britannica.com/EBchecked/topic/366997/Saint-Martin-of-Tours#
synod								
	Religião	W	4	0	1791	9507	173	http://en.wikipedia.org/wiki/Synod
36	ME	B	2	0	188	1031	19	http://www.britannica.com/EBchecked/topic/578481/synod
cargo cult								
	Sociologia	W	5	1	2211	11718	213	http://en.wikipedia.org/wiki/Cargo_cult
35	UE	B	1	0	308	1718	31	http://www.britannica.com/EBchecked/topic/95744/cargo-cult
Max Weber								
	Sociologia	W	5	5	6709	39681	721	http://en.wikipedia.org/wiki/Max_Weber
35	UE	B	4	1	2631	14697	267	http://www.britannica.com/EBchecked/topic/638565/Max-Weber
Social norms								
	Sociologia	W	5	1	1911	10795	196	http://en.wikipedia.org/wiki/Norm_%28sociology%29
35	UE	B	1	0	266	1389	25	http://www.britannica.com/EBchecked/topic/418203/norm

Área	Avaliador	Afiliação	Nota	Ilustrações	Número			URL
					Palavras	Caracteres	Linhas	
Herbicide								
Agronomia		W	4	1	4336	24548	446	http://en.wikipedia.org/wiki/Herbicide
30	UE	B	2	1	288	1609	29	http://www.britannica.com/EBchecked/topic/262755/herbicide
Plant pathology								
Agronomia		W	2	4	1552	8771	159	http://en.wikipedia.org/wiki/Plant_pathology
29	UE	B	4	6	16422	91994	1673	http://www.britannica.com/EBchecked/topic/463327/plant-disease
Coma Berenices								
Astronomia		W	4	3	1501	7373	134	http://en.wikipedia.org/wiki/Coma_Berenices
5	UE	B	1	1	143	736	13	http://www.britannica.com/EBchecked/topic/127214/Coma-Berenices
Dumitru Prunariu								
Astronomia		W	4	3	685	3771	69	http://en.wikipedia.org/wiki/Dumitru_Prunariu
5	UE	B	3	0	277	1431	26	http://www.britannica.com/EBchecked/topic/1512631/Dumitru-Prunariu
Infrared astronomy								
Astronomia		W	4	2	1424	8062	147	http://en.wikipedia.org/wiki/Infrared_astronomy
5	UE	B	2	1	643	3713	68	http://www.britannica.com/EBchecked/topic/287941/infrared-astronomy
Konstantin Tsiolkovsky								
Astronomia		W	4	5	1388	7918	144	http://en.wikipedia.org/wiki/Konstantin_Tsiolkovsky
5	UE	B	2	1	819	4401	80	http://www.britannica.com/EBchecked/topic/607781/Konstantin-Eduardovich-Tsiolkovsky
Orion								
Astronomia		W	4	10	3559	18338	333	http://en.wikipedia.org/wiki/Orion_constellation
5	UE	B	1	1	148	771	14	http://www.britannica.com/EBchecked/topic/432659/Orion
hormone								
Biologia		W	1	1	1670	9901	180	http://en.wikipedia.org/wiki/Hormone
6	UE	B	4	6	11704	65403	1189	http://www.britannica.com/EBchecked/topic/271826/hormone
Balatá								
Botânica		W	3	2	516	2808	51	http://en.wikipedia.org/wiki/Balat%C3%A1
1	UE	B	1	0	88	461	8	http://www.britannica.com/EBchecked/topic/49876/balata
Bulbophyllum								
Botânica		W	5	6	944	5800	105	http://en.wikipedia.org/wiki/Bulbophyllum
1	UE	B	1	1	64	342	6	http://www.britannica.com/EBchecked/topic/84037/Bulbophyllum
Cyperaceae								
Botânica		W	3	2	328	2269	41	http://en.wikipedia.org/wiki/Cyperaceae
1	UE	B	5	1	4389	24086	438	http://www.britannica.com/EBchecked/topic/148485/Cyperaceae
Lycopene								
Botânica		W	5	4	2855	16823	306	http://en.wikipedia.org/wiki/Lycopene
1	UE	B	1	1	75	378	7	http://www.britannica.com/EBchecked/topic/352572/lycopene

Área	Número		URL				
	Avaliador	Afiliação	Nota	Ilustrações	Palavras	Caracteres	Linhas
Michelia champaca							
Botânica	W	2	1	449	2366	43	http://en.wikipedia.org/wiki/Michelia_champaca
1	UE	B	2	1	150	813	http://www.britannica.com/EBchecked/topic/105128/champac
Pennisetum							
Botânica	W	3	2	347	2256	41	http://en.wikipedia.org/wiki/Pennisetum
1	UE	B	1	1	116	679	http://www.britannica.com/EBchecked/topic/450083/Pennisetum
Poaceae							
Botânica	W	4	22	2065	11945	217	http://en.wikipedia.org/wiki/Poaceae
1	UE	B	5	2	6493	34818	http://www.britannica.com/EBchecked/topic/465603/Poaceae
Electronic commerce							
Economia	W	4	0	1697	10261	187	http://en.wikipedia.org/wiki/E-commerce
32	UE	B	3	0	682	4020	http://www.britannica.com/EBchecked/topic/183748/e-commerce
John Tyndall							
Física	W	4	8	10682	58903	1071	http://en.wikipedia.org/wiki/John_Tyndall
5	UE	B	2	1	387	2205	http://www.britannica.com/EBchecked/topic/611577/John-Tyndall
Lift (force)							
Física	W	4	7	5652	29545	537	http://en.wikipedia.org/wiki/Lift_%28force%29
5	UE	B	2	1	17178	81207	http://www.britannica.com/EBchecked/topic/211272/fluid-mechanics/77497/Lift#ref=ref611779
magnetic susceptibility							
Física	W	4	0	1628	9187	167	http://en.wikipedia.org/wiki/Magnetic_susceptibility
5	UE	B	2	0	319	1905	http://www.britannica.com/EBchecked/topic/357313/magnetic-susceptibility
meson							
Física	W	4	3	3804	18871	343	http://en.wikipedia.org/wiki/Meson
5	UE	B	1	0	494	2591	http://www.britannica.com/EBchecked/topic/376781/meson#ref258563
moiré pattern							
Física	W	4	10	2054	10106	184	http://en.wikipedia.org/wiki/Moir%C3%A9_pattern
5	UE	B	1	1	208	1068	http://www.britannica.com/EBchecked/topic/387754/moire-pattern
Particle physics							
Física	W	4	2	1764	9880	180	http://en.wikipedia.org/wiki/Particle_physics
5	UE	B	1	0	88	562	http://www.britannica.com/EBchecked/topic/445074/particle-physics
Radiant energy							
Física	W	4	2	1468	8259	150	http://en.wikipedia.org/wiki/Radiant_energy
5	UE	B	1	1	59	332	http://www.britannica.com/EBchecked/topic/488482/radiant-energy
String theory							
Física	W	4	3	10741	60694	1104	http://en.wikipedia.org/wiki/String_theory
5	UE	B	2	0	2002	11245	http://www.britannica.com/EBchecked/topic/574576/string-theory

Área	Avaliador		Nota		Número		Linhas	URL
	Avaliador	Afiliação	Nota	Ilustrações	Palavras	Caracteres		
Miocene								
Geologia	W	5	3	1425	8120	148	http://en.wikipedia.org/wiki/Miocene_epoch	
12	UL	B	5	2	661	3690	http://www.britannica.com/EBchecked/topic/384700/Miocene-Epoch	
Ornithopods								
Geologia	W	3	4	723	4747	86	http://en.wikipedia.org/wiki/Ornithopod	
13	UL	B	4	3	476	2757	http://www.britannica.com/EBchecked/topic/432993/ornithopod	
Paleontology								
Geologia	W	3	19	9323	55848	1015	http://en.wikipedia.org/wiki/Palaeontology	
12	UL	B	5	1	230	1377	http://www.britannica.com/EBchecked/topic/439548/paleontology	
Permian–Triassic extinction event								
Geologia	W	5	7	10040	60157	1094	http://en.wikipedia.org/wiki/Permian%E2%80%93Triassic_extinction_event	
12	UL	B	4	1	1225	6851	http://www.britannica.com/EBchecked/topic/1481207/Permian-extinction	
Rhodochrosite								
Geologia	W	4	3	701	3976	72	http://en.wikipedia.org/wiki/Rhodochrosite	
14	UL	B	4	2	111	706	http://www.britannica.com/EBchecked/topic/501692/rhodochrosite	
salt marsh								
Geologia	W	4	9	2757	15109	275	http://en.wikipedia.org/wiki/Salt_marsh	
14	UL	B	2	1	112	570	http://www.britannica.com/EBchecked/topic/519827/salt-marsh	
Therapsida								
Geologia	W	3	8	810	5860	107	http://en.wikipedia.org/wiki/Therapsida	
13	UL	B	3	0	317	1709	http://www.britannica.com/EBchecked/topic/591228/therapsid	
Toarcian								
Geologia	W	3	4	551	3350	61	http://en.wikipedia.org/wiki/Toarcian_Stage	
13	UL	B	4	0	169	960	http://www.britannica.com/EBchecked/topic/597672/Toarcian-Stage	
Subway								
História	W	4	20	4433	24367	443	http://en.wikipedia.org/wiki/Rapid_transit	
4	II	B	3	5	823	4404	http://www.britannica.com/EBchecked/topic/571195/subway	
C++								
Informática	W	5	3	5478	30525	555	http://en.wikipedia.org/wiki/C_Plus_Plus	
22	UE	B	2	0	45	312	http://www.britannica.com/EBchecked/topic/688345/C	
Computer-aided engineering								
Informática	W	2	1	646	3807	69	http://en.wikipedia.org/wiki/Computer-aided_engineering	
24	UE	B	4	0	155	965	http://www.britannica.com/EBchecked/topic/130566/computer-aided-engineering#ref=ref6418	
Linux								
Informática	W	5	13	7391	42883	780	http://en.wikipedia.org/wiki/Linux	
22	UE	B	3	0	263	1430	http://www.britannica.com/EBchecked/topic/686713/Linux	

Área	Número		URL				
	Avaliador	Afiliação	Nota	Ilustrações	Palavras	Caracteres	Linhas
Malware							
Informática	W	4	1	6167	34301	624	http://en.wikipedia.org/wiki/Malware
26	UE	B	3	0	363	2049	37 http://www.britannica.com/EBchecked/topic/1477142/malware
relational database							
Informática	W	2	3	1990	10904	198	http://en.wikipedia.org/wiki/Relational_database
28	UE	B	4	0	143	780	14 http://www.britannica.com/EBchecked/topic/496814/relational-database
Robert Elliot Kahn							
Informática	W	5	2	1200	6554	119	http://en.wikipedia.org/wiki/Robert_Elliot_Kahn
25	UE	B	4	0	525	2997	54 http://www.britannica.com/EBchecked/topic/932776/Robert-Elliot-Kahn
Sun Microsystems							
Informática	W	5	10	7691	44259	805	http://en.wikipedia.org/wiki/Sun_Microsystems
23	UE	B	3	1	1312	7488	136 http://www.britannica.com/EBchecked/topic/573631/Sun-Microsystems-Inc
XML							
Informática	W	4	1	5894	32245	586	http://en.wikipedia.org/wiki/XML
27	UE	B	1	0	184	899	16 http://www.britannica.com/EBchecked/topic/1086474/XML
Ernst Eduard Kummer							
Matemática	W	3	1	503	2937	53	http://en.wikipedia.org/wiki/Ernst_Eduard_Kummer
33	UE	B	3	0	287	1473	27 http://www.britannica.com/EBchecked/topic/324800/Ernst-Eduard-Kummer
hyperbola							
Matemática	W	4	7	5881	28113	511	http://en.wikipedia.org/wiki/Hyperbola
33	UE	B	3	0	275	1297	24 http://www.britannica.com/EBchecked/topic/279494/hyperbola
Intuitionistic logic							
Matemática	W	4	1	1949	10024	182	http://en.wikipedia.org/wiki/Intuitionistic_logic
33	UE	B	3	0	536	2722	49 http://www.britannica.com/EBchecked/topic/369221/foundations-of-
Matrix							
Matemática	W	4	11	9633	51651	939	http://en.wikipedia.org/wiki/Matrix_%28mathematics%29
33	UE	B	3	0	1216	4879	89 http://www.britannica.com/EBchecked/topic/369516/matrix
Möbius strip							
Matemática	W	4	6	2979	14353	261	http://en.wikipedia.org/wiki/Mobius_strip
33	UE	B	1	1	78	422	8 http://www.britannica.com/EBchecked/topic/386825/Mobius-strip
variance							
Matemática	W	4	0	4107	19433	353	http://en.wikipedia.org/wiki/Variance
33	UE	B	1	0	37	194	4 http://www.britannica.com/EBchecked/topic/623377/variance
Fuel injection							
Mecânica	W	4	3	6885	37616	684	http://en.wikipedia.org/wiki/Fuel_injection
10	UE	B	1	0	139	765	14 http://www.britannica.com/EBchecked/topic/221388/fuel-injection

Área	Número		URL					
	Avaliador	Afiliação	Nota	Ilustrações	Palavras	Caracteres	Linhas	URL
ratchet								
Mecânica	W	4	3	415	2137	39	http://en.wikipedia.org/wiki/Ratchet_%28device%29	
10	UE	B	2	1	298	1372	25	http://www.britannica.com/EBchecked/topic/491900/ratchet
Morning sickness								
Medicina	W	4	1	803	4893	89	http://en.wikipedia.org/wiki/Morning_sickness	
19	UE	B	3	0	204	1039	19	http://www.britannica.com/EBchecked/topic/392562/morning-sickness
Tooth enamel								
Medicina	W	4	11	4804	25535	464	http://en.wikipedia.org/wiki/Tooth_enamel	
21	UE	B	4	1	187	927	17	http://www.britannica.com/EBchecked/topic/186360/enamel
Coastal artillery								
Militar	W	4	9	1581	8462	154	http://en.wikipedia.org/wiki/Coast_Artillery	
16	MP	B	2	1	232	1280	23	http://www.britannica.com/EBchecked/topic/123129/coastal-artillery
Carrack								
Naval	W	4	2	2234	10751	195	http://en.wikipedia.org/wiki/Carrack	
16	MP	B	1	0	198	966	18	http://www.britannica.com/EBchecked/topic/96864/carrack
clipper								
Naval	W	2	3	1095	5522	100	http://en.wikipedia.org/wiki/Clipper_ship	
16	MP	B	1	1	30	168	3	http://www.britannica.com/EBchecked/topic/121871/clipper-ship
destroyer								
Naval	W	5	20	5035	26608	484	http://en.wikipedia.org/wiki/Destroyer	
16	MP	B	1	1	211	1177	21	http://www.britannica.com/EBchecked/topic/159394/destroyer
adhesive								
Química	W	5	5	3183	16978	309	http://en.wikipedia.org/wiki/Adhesive	
34	UE	B	3	0	2460	14604	266	http://www.britannica.com/EBchecked/topic/5823/adhesive
carbon monoxide								
Química	W	5	6	4938	28306	515	http://en.wikipedia.org/wiki/Carbon_monoxide	
34	UE	B	1	0	372	2037	37	http://www.britannica.com/EBchecked/topic/95021/carbon-monoxide
Malonic acid								
Química	W	3	3	373	2142	39	http://en.wikipedia.org/wiki/Malonic_acid	
34	UE	B	1	0	180	930	17	http://www.britannica.com/EBchecked/topic/360443/malonic-acid#ref=ref280513
RDX								
Química	W	5	2	3018	16109	293	http://en.wikipedia.org/wiki/RDX	
34	UE	B	1	0	155	815	15	http://www.britannica.com/EBchecked/topic/492672/RDX
Zinc								
Química	W	5	21	12044	67692	1231	http://en.wikipedia.org/wiki/Zinc	
34	UE	B	3	1	1753	8786	160	http://www.britannica.com/EBchecked/topic/657264/zinc

Área	Número		URL					
	Avaliador	Afiliação	Nota	Ilustrações	Palavras	Caracteres	Linhas	URL
Lungworm								
Veterinária	W	4	0	786	4224	77	http://en.wikipedia.org/wiki/Lungworm	
20	UE	B	2	0	118	647	12	http://www.britannica.com/EBchecked/topic/351620/lungworm
Acanthuridae								
Zoologia	W	5	11	1166	7943	144	http://en.wikipedia.org/wiki/Surgeonfish	
2	UE	B	2	1	170	891	16	http://www.britannica.com/EBchecked/topic/575146/surgeonfish
bony fish								
Zoologia	W	4	1	525	3080	56	http://en.wikipedia.org/wiki/Bony_fish	
2	UE	B	2	1	240	1286	23	http://www.britannica.com/EBchecked/topic/73261/bony-fish
Crotalinae								
Zoologia	W	5	3	1908	10501	191	http://en.wikipedia.org/wiki/Crotalinae	
2	UE	B	1	0	74	408	7	http://www.britannica.com/EBchecked/topic/461927/pit-viper
dragonfly								
Zoologia	W	5	9	2114	12539	228	http://en.wikipedia.org/wiki/Dragonfly	
2	UE	B	3	7	737	3872	70	http://www.britannica.com/EBchecked/topic/170810/dragonfly
horse								
Zoologia	W	5	15	11972	66569	1210	http://en.wikipedia.org/wiki/Horse	
2	UE	B	4	35	6151	31203	567	http://www.britannica.com/EBchecked/topic/272156/horse
Labeo								
Zoologia	W	4	1	945	5557	101	http://en.wikipedia.org/wiki/Labeo	
2	UE	B	2	1	153	732	13	http://www.britannica.com/EBchecked/topic/326696/labeo
Mallard								
Zoologia	W	5	11	2439	13179	240	http://en.wikipedia.org/wiki/Mallard	
2	UE	B	2	3	485	2494	45	http://www.britannica.com/EBchecked/topic/360302/mallard
moth								
Zoologia	W	4	7	2185	12028	219	http://en.wikipedia.org/wiki/Moth	
2	UE	B	2	4	315	1751	32	http://www.britannica.com/EBchecked/topic/393915/moth
owlet moth								
Zoologia	W	4	0	749	4820	88	http://en.wikipedia.org/wiki/Owlet_moth	
2	UE	B	2	3	336	1833	33	http://www.britannica.com/EBchecked/topic/436332/owlet-moth
Sardine								
Zoologia	W	4	8	2193	12242	223	http://en.wikipedia.org/wiki/Sardine	
2	UE	B	1	1	445	2265	41	http://www.britannica.com/EBchecked/topic/524126/sardine
scorpionfish								
Zoologia	W	4	3	468	2793	51	http://en.wikipedia.org/wiki/Scorpion_fish	
2	UE	B	2	2	149	806	15	http://www.britannica.com/EBchecked/topic/529367/scorpion-fish

Área	Avaliador		Afiliação		Número			URL
			Nota	Ilustrações	Palavras	Caracteres	Linhas	
Delaware Aqueduct								
Arquitetura	W	4	1	353	2116	38	http://en.wikipedia.org/wiki/Delaware_Aqueduct	
11	UE	B	3	0	125	581	11	http://www.britannica.com/EBchecked/topic/156400/Delaware-Aqueduct
Ernestine duchies								
História	W	4	2	2474	13466	245	http://en.wikipedia.org/wiki/Saxon_duchies	
4	II	B	5	0	583	3203	58	http://www.britannica.com/EBchecked/topic/526039/Saxon-duchies#ref=ref119293
National Portrait Gallery								
Pintura	W	5	3	1606	9025	164	http://en.wikipedia.org/wiki/National_Portrait_Gallery_%28London%29	
15	MJV	B	2	1	302	1710	31	http://www.britannica.com/EBchecked/topic/405261/National-Portrait-Gallery
Bayonne								
Geografia	W		8	5362	29473	536	http://en.wikipedia.org/wiki/Bayonne,_New_Jersey	
		B	0	175	895	16	http://www.britannica.com/EBchecked/topic/56921/Bayonne	
Belarus								
Geografia	W		26	11143	63963	1163	http://en.wikipedia.org/wiki/Belarus	
		B	15	9661	53548	974	http://www.britannica.com/EBchecked/topic/59081/Belarus	
Benue								
Geografia	W		1	419	2325	42	http://en.wikipedia.org/wiki/Benue_State	
		B	0	230	1161	21	http://www.britannica.com/EBchecked/topic/61231/Benue	
Berkshire								
Geografia	W		11	4320	24671	449	http://en.wikipedia.org/wiki/Berkshire,_England	
		B	1	608	3234	59	http://www.britannica.com/EBchecked/topic/62029/Berkshire	
Bielefeld								
Geografia	W		7	1502	8407	153	http://en.wikipedia.org/wiki/Bielefeld	
		B	1	283	1546	28	http://www.britannica.com/EBchecked/topic/64818/Bielefeld	
Black Sea								
Geografia	W		31	5441	30533	555	http://en.wikipedia.org/wiki/Black_Sea	
		B	7	2973	14972	272	http://www.britannica.com/EBchecked/topic/68234/Black-Sea	
British Rail								
Geografia	W		12	3251	18068	329	http://en.wikipedia.org/wiki/British_Rail	
		B	0	425	2276	41	http://www.britannica.com/EBchecked/topic/80334/British-Railways	
Brooklyn								
Geografia	W		23	7739	43508	791	http://en.wikipedia.org/wiki/Brooklyn	
		B	2	370	2030	37	http://www.britannica.com/EBchecked/topic/81247/Brooklyn	
Buryatia								
Geografia	W		7	1641	9143	166	http://en.wikipedia.org/wiki/Buryatia	
		B	1	515	2897	53	http://www.britannica.com/EBchecked/topic/86012/Buryatia	

Área	Número		Palavras	Caracteres	Linhas	URL
	Avaliador	Afiliação				
Canadian Prairies						
Geografia	W	5	1174	6351	115	http://en.wikipedia.org/wiki/Prairie_Provinces
	B	1	48	280	5	http://www.britannica.com/EBchecked/topic/473851/Prairie-Provinces
County Tyrone						
Geografia	W	2	1233	7027	128	http://en.wikipedia.org/wiki/County_Tyrone
	B	1	293	1518	28	http://www.britannica.com/EBchecked/topic/611953/Tyrone
Duars						
Geografia	W	3	1714	9053	165	http://en.wikipedia.org/wiki/Duars
	B	1	130	635	12	http://cache.britannica.com/EBchecked/topic/172688/Duars
Fukuyama, Hiroshima						
Geografia	W	5	969	5331	97	http://en.wikipedia.org/wiki/Fukuyama,_Hiroshima
	B	1	102	570	10	http://www.britannica.com/EBchecked/topic/221684/Fukuyama
Gascony						
Geografia	W	4	1498	7915	144	http://en.wikipedia.org/wiki/Gascony
	B	0	487	2633	48	http://www.britannica.com/EBchecked/topic/226518/Gascony
Görlitz						
Geografia	W	10	1332	7017	128	http://en.wikipedia.org/wiki/G%C3%B6rlitz
	B	0	199	1040	19	http://www.britannica.com/EBchecked/topic/239365/Gorlitz
Helena, Montana						
Geografia	W	9	3056	16761	305	http://en.wikipedia.org/wiki/Helena,_Montana
	B	2	332	1709	31	http://www.britannica.com/EBchecked/topic/259894/Helena
Highland						
Geografia	W	3	1243	7134	130	http://en.wikipedia.org/wiki/Highland_%28council_area%29
	B	2	677	3678	67	http://www.britannica.com/EBchecked/topic/682996/Highland
Huntingdon County, Pennsylvania						
Geografia	W	5	1087	6310	115	http://en.wikipedia.org/wiki/Huntingdon_County,_Pennsylvania
	B	1	146	849	15	http://www.britannica.com/EBchecked/topic/277124/Huntingdon
Ithaca, New York						
Geografia	W	16	6865	37034	673	http://en.wikipedia.org/wiki/Ithaca,_New_York
	B	1	193	1024	19	http://www.britannica.com/EBchecked/topic/297971/Ithaca
Jackson, Michigan						
Geografia	W	4	2543	13879	252	http://en.wikipedia.org/wiki/Jackson,_Michigan
	B	1	254	1348	25	http://www.britannica.com/EBchecked/topic/298754/Jackson
Jos						
Geografia	W	2	1669	8497	154	http://en.wikipedia.org/wiki/Jos
	B	0	603	3034	55	http://www.britannica.com/EBchecked/topic/306282/Jos

Área	Avaliador	Afiliação	Nota	Número			URL
				Ilustrações	Palavras	Caracteres	
Kaliningrad							
Geografia	W	B	23	3286	18381	334	http://en.wikipedia.org/wiki/Kaliningrad
			1	711	3652	66	http://www.britannica.com/EBchecked/topic/310209/Kaliningrad
Kantō Plain							
Geografia	W	B	2	248	1337	24	http://en.wikipedia.org/wiki/Kant%C5%8D_plain
			0	520	2574	47	http://www.britannica.com/EBchecked/topic/311537/Kanto-Plain
Kerala							
Geografia	W	B	31	12455	74155	1348	http://en.wikipedia.org/wiki/Kerala
			7	2441	13590	247	http://www.britannica.com/EBchecked/topic/315300/Kerala
León							
Geografia	W	B	4	755	3809	69	http://en.wikipedia.org/wiki/Le%C3%B3n_province%29
			0	200	1074	20	http://www.britannica.com/EBchecked/topic/336359/Leon
Lisburn							
Geografia	W	B	4	2723	14042	255	http://en.wikipedia.org/wiki/Lisburn
			0	204	1092	20	http://www.britannica.com/EBchecked/topic/343264/Lisburn
Málaga							
Geografia	W	B	27	5002	26569	483	http://en.wikipedia.org/wiki/M%C3%A1laga
			2	459	2383	43	http://www.britannica.com/EBchecked/topic/359455/Malaga
Marche							
Geografia	W	B	6	1535	8041	146	http://en.wikipedia.org/wiki/Marche
			1	458	2517	46	http://www.britannica.com/EBchecked/topic/364212/Marche
Mecklenburg-Vorpommern							
Geografia	W	B	45	3374	19467	354	http://en.wikipedia.org/wiki/Mecklenburg-Vorpommern
			6	1253	7046	128	http://www.britannica.com/EBchecked/topic/372034/Mecklenburg-West-Pomerania
Meiningen							
Geografia	W	B	8	1218	6349	115	http://en.wikipedia.org/wiki/Meiningen
			1	151	770	14	http://www.britannica.com/EBchecked/topic/373389/Meiningen
Nepal							
Geografia	W	B	19	8991	49393	898	http://en.wikipedia.org/wiki/Nepal
			4	7471	40514	737	http://www.britannica.com/EBchecked/topic/409152/Nepal
Oise River							
Geografia	W	B	3	306	1675	30	http://en.wikipedia.org/wiki/Oise_%28river%29
			1	213	1059	19	http://www.britannica.com/EBchecked/topic/426304/Oise-River
Owensboro							
Geografia	W	B	10	3197	17342	315	http://en.wikipedia.org/wiki/Owensboro
			0	196	1053	19	http://www.britannica.com/EBchecked/topic/436301/Owensboro

Área	Número		Número			URL	
	Avaliador	Afiliação	Nota	Ilustrações	Palavras		Caracteres
Oxfordshire							
Geografia	W		3	1748	9854	179	http://en.wikipedia.org/wiki/Oxfordshire
	B		1	752	4089	74	http://www.britannica.com/EBchecked/topic/436567/Oxfordshire
Pisa							
Geografia	W		12	4981	25452	463	http://en.wikipedia.org/wiki/Pisa
	B		1	998	5233	95	http://www.britannica.com/EBchecked/topic/461669/Pisa
Pra River							
Geografia	W		1	246	1265	23	http://en.wikipedia.org/wiki/Pra_River_%28Ghana%29
	B		0	109	537	10	http://www.britannica.com/EBchecked/topic/473530/Pra-River
Prahova							
Geografia	W		2	741	4247	77	http://en.wikipedia.org/wiki/Prahova_County
	B		0	244	1353	25	http://www.britannica.com/EBchecked/topic/473817/Prahova
Queensland							
Geografia	W		15	3593	20810	378	http://en.wikipedia.org/wiki/Queensland
	B		18	7711	43052	783	http://www.britannica.com/EBchecked/topic/486932/Queensland
Renmark							
Geografia	W		4	738	3821	69	http://en.wikipedia.org/wiki/Renmark,_South_Australia
	B		0	158	884	16	http://www.britannica.com/EBchecked/topic/498146/Renmark
Richmond upon Thames							
Geografia	W		4	1349	7032	128	http://en.wikipedia.org/wiki/Richmond_upon_Thames
	B		1	433	2311	42	http://www.britannica.com/EBchecked/topic/502846/Richmond-upon-Thames
Rio Grande do Sul							
Geografia	W		21	7047	37466	681	http://en.wikipedia.org/wiki/Rio_Grande_do_Sul
	B		2	730	3758	68	http://www.britannica.com/EBchecked/topic/504264/Rio-Grande-do-Sul
Saône River							
Geografia	W		2	309	1753	32	http://en.wikipedia.org/wiki/Sa%C3%B4ne
	B		1	149	774	14	http://www.britannica.com/EBchecked/topic/523619/Saone-River
Saskatchewan River							
Geografia	W		2	775	4235	77	http://en.wikipedia.org/wiki/Saskatchewan_River
	B		1	241	1315	24	http://www.britannica.com/EBchecked/topic/524737/Saskatchewan-River
Sneeuberge							
Geografia	W		0	123	701	13	http://en.wikipedia.org/wiki/Sneeuberg
	B		0	72	370	7	http://www.britannica.com/EBchecked/topic/550414/Sneeuberg
Surakarta							
Geografia	W		11	2533	13885	252	http://en.wikipedia.org/wiki/Surakarta
	B		0	226	1229	22	http://www.britannica.com/EBchecked/topic/574943/Surakarta

Área	Número		Nota	Ilustrações	Palavras	Caracteres	Linhas	URL
	Avaliador	Afiliação						
Swansea								
Geografia	W		23	10235	56623	1030	http://en.wikipedia.org/wiki/Swansea	
	B		1	475	2495	45	http://www.britannica.com/EBchecked/topic/576322/Swansea	
Wahiba Sands								
Geografia	W		5	563	3035	55	http://en.wikipedia.org/wiki/Ramlat_al-Wahiba	
	B		0	101	488	9	http://www.britannica.com/EBchecked/topic/11913/Al-Wahibah-Dunes	
Wallonia								
Geografia	W		29	7509	41597	756	http://en.wikipedia.org/wiki/Wallonia	
	B		1	107	656	12	http://www.britannica.com/EBchecked/topic/634954/Wallonia	
West Bank								
Geografia	W		9	9797	55318	1006	http://en.wikipedia.org/wiki/West_Bank	
	B		1	1786	9537	173	http://www.britannica.com/EBchecked/topic/640076/West-Bank	
Woodward								
Geografia	W		4	3234	17503	318	http://en.wikipedia.org/wiki/Woodward,_Oklahoma	
	B		1	136	706	13	http://www.britannica.com/EBchecked/topic/647688/Woodward	
Zielona								
Geografia	W		9	1779	9603	175	http://en.wikipedia.org/wiki/Zielona_G%C3%B3ra	
	B		0	166	890	16	http://www.britannica.com/EBchecked/topic/657103/Zielona-Gora	
Zlín								
Geografia	W		7	2050	10725	195	http://en.wikipedia.org/wiki/Zl%C3%ADn	
	B		1	166	929	17	http://www.britannica.com/EBchecked/topic/657652/Zlin	
Alentejo								
Geografia	W		4	1037	5405	98	http://en.wikipedia.org/wiki/Alentejo	
	B		1	277	1522	28	http://www.britannica.com/EBchecked/topic/13813/Alentejo	
Barnstable								
Geografia	W		14	2736	14785	269	http://en.wikipedia.org/wiki/Barnstable,_Massachusetts	
	B		1	234	1270	23	http://www.britannica.com/EBchecked/topic/53701/Barnstable	

COLABORAÇÃO EM MASSA OU AMADORISMO EM MASSA?

UM ESTUDO COMPARATIVO DA QUALIDADE DA INFORMAÇÃO CIENTÍFICA
PRODUZIDA UTILIZANDO OS CONCEITOS E FERRAMENTAS WIKI.

Doutoramento em Gestão

Sistemas de Informação

Fernando Silvério Nifrário Rodrigues

Orientador:

Professor Doutor Paulo Resende Silva



A massificação do acesso à Internet tornou-nos a todos em potenciais criadores de conteúdos que não passam pela teia de crivos que os editores conceberam para se assegurar da qualidade dos conteúdos que comercializam.

Apesar de o bom senso aconselhar a não confiarmos em informações disponibilizadas por fontes cuja credibilidade desconhecemos, eventualmente a maioria de nós já escolheu um produto com base nas opiniões publicadas na *Amazon* por outros compradores, criou uma imagem de uma determinada empresa/produto/pessoa ao ler artigos de opinião em *blogs* escritos por *amadores*, utilizou a *Wikipédia* cujos conteúdos são livremente editados por anónimos sem a supervisão de um editor, etc.

Este estudo pretende aferir o grau de confiabilidade da informação científica criada através da colaboração em massa, anónima e gratuita, por contraponto à abordagem tradicional, através de um estudo comparativo da qualidade dos conteúdos das enciclopédias generalistas Wikipédia e Britânica. O objectivo não é a comparação, *tout court*, dos produtos *per se*, mas descobrir se, e em que condições, a colaboração em massa funciona.

Solicitamos assim a sua colaboração, na qualidade de perito na área de Gestão, através da atribuição de uma classificação a quatro pares de artigos, provenientes das enciclopédias acima referidas, tendo em atenção os seguintes aspectos:

Cada *pdf* enviado em anexo contém as versões de cada enciclopédia sobre o tema descrito no nome do ficheiro, estando as páginas correspondentes a cada uma delas assinaladas pelas letras **A** e **B** em marca de água. Os números antes do nome do tema servem para identificar a área de conhecimento e o artigo em causa.

Deve classificar ambas as versões de cada tema de 1 a 5 (1: Mau, 2:Medíocre, 3: Suficiente, 4:Bom e 5: Muito Bom)

Pode utilizar os critérios que entender, desde que sejam os mesmos, em todos os artigos que avaliar. Sugerimos que proceda como se estivesse a cotar uma pergunta de um exame, ao nível de licenciatura, de dois alunos e que aplique os critérios que normalmente utilizaria nessa situação.

Os artigos foram descaracterizados (de modo a evitar que o conhecimento da sua proveniência influencie a sua classificação) e contêm apenas o corpo principal do artigo e as figuras a este associado. Estas alterações levaram a perdas de formatação e de qualidade das ilustrações, pelo que solicitamos que apenas tenha em conta os conteúdos.

Caso entenda que alguns dos artigos que lhe foram confiados extravasam a sua área de conhecimento e conheça alguém que seja perito nessa área, agradecemos que lhe faça chegar os artigos em questão.

Para nos fazer chegar as suas avaliações, basta indicar o número do artigo e as classificações que atribuiu às versões A e B.

Caso não pretenda que o seu nome seja incluído na lista de peritos que colaboraram neste estudo e/ou receber as conclusões a que chegarmos, por favor indique-o na sua resposta. Note que não serão publicadas as classificações individuais que nos enviar, mas sim o valor médio das respostas dos peritos convidados, pelo que está garantido o anonimato das suas respostas.

Se achar necessário, a classificação dos artigos pode ser efectuado por mais professores que sejam, também, peritos na área em causa. Desde já agradeço a sua colaboração.

MASS COLLABORATION OR MASS AMATEURISM?

A COMPARATIVE STUDY OF THE QUALITY OF SCIENTIFIC INFORMATION
PRODUCED USING THE WIKI TOOLS AND CONCEPTS

Management of Information Systems

Fernando Silvério Nifrário Rodrigues
(PhD candidate)

Supervisor:

Professor Paulo Resende da Silva



The widespread mass access to the Internet has turned us all into potential makers of contents which do not go through the examination and selection sieves that editors have created so as to rest assured that the products they sell are quality ones.

Although common sense advises us not to trust information made available by sources whose credibility is unknown, most of us have possibly already chosen a product based on buyer reviews on Amazon; eventually, we have also created an image of a certain company/ product/ person by reading opinion articles on blogs written by *amateurs*, or have used *Wikipedia*, whose contents are freely edited without the supervision of an editor by anonymous contributors.

This study intends to check to what extent information generated through anonymous and freely bestowed mass collaboration is reliable as opposed to the traditional approach. In order to achieve that goal, I develop a comparative study between *Wikipedia* and *Encyclopaedia Britannica* with regard to accuracy, depth and detail of information in both, in order to confront the quality of the knowledge repository produced by them. My goal is not the mere comparison of the products *per se*: I want to perceive whether, and under which conditions, mass collaboration works.

Thus, I would appreciate your collaboration as an expert in your chosen field of knowledge. To do so, I will ask you to rate a number of articles obtained through performing a random sampling of contents in the above mentioned encyclopaedias. While classifying the articles, the following criteria should be taken into account:

- Each attached *pdf* document contains the versions of each encyclopaedia on the theme referred to in the filename. The pages which correspond to each encyclopaedia are watermarked by the letters A and B. The numbers that come before the name of the theme are used to identify the field of knowledge and the article at stake.
- You should rate both versions of each theme using a 1 to 5 scale (1: Poor, 2: Mediocre, 3: Fair, 4: Good, 5: Very Good)
- You may use the criteria you think are the most suitable, provided they are always the same in every article you rate. I would suggest you approached this classification in the way you do when are you are grading the same question written by two undergraduate students in an exam, thus applying here the same criteria you use in that situation.
- The articles have been reformatted in order that the rating might not be influenced by the knowledge of their source, so they have just the main body of the article and the figures associated to it. These changes have led to formatting and illustrations quality losses, so I ask you just to bear in mind the contents, not the look.
- Should you consider that some of the articles you are working with were not in your field of expertise and should you know someone who is an expert in that area, I would appreciate if you forwarded them the articles at stake.
- To forward your classifications to me, you should just indicate the number of the article and the ratings you have given versions A and B.

Should you prefer your name was not included in the list of experts that have collaborated on this study and/ or not to be informed about the conclusions of this study, please say so in your answer. The individual ratings you will send me will not be published; only the average value of the answers of the invited experts will be published, in order to guarantee the anonymity of your answer. Thank you in anticipation.

ANEXO III

CARACTERIZAÇÃO DOS HÁBITOS DE CONSULTA DE INFORMAÇÃO

Compilação das respostas ao inquérito efectuado junto dos Docentes da Universidade de Évora, durante o mês de Maio de 2011. O pedido de colaboração foi efectuado através de correio electrónico e as respostas foram dadas através de um *website* que utiliza *cookies* e o registo do IP, de modo a garantir que cada respondente só o faz uma vez.

Docentes Universitários														
Quando necessita de obter informação - aprofundar um determinado tema, onde procura primeiro? (0: Motor de Busca / Wikipédia - 1: Enciclopédia / Dicionário / Obra de Referência)	#1	0	#41	0	Qual o grau de confiança que têm na informação obtida através de um Motor de Busca / Wikipédia? [1: Desconfio Totalmente, 5: Confio Completamente]	#1	4	#41	2	Qual o grau de confiança que têm na informação obtida através de uma Enciclopédia / Dicionário / Obra de Referência? [1: Desconfio Totalmente, 5: Confio Completamente]	#1	4	#41	4
	#2	0	#42	0		#2	4	#42	3		#2	5	#42	5
	#3	1	#43	1		#3	3	#43	1		#3	5	#43	4
	#4	0	#44	0		#4	3	#44	3		#4	4	#44	4
	#5	0	#45	0		#5	4	#45	2		#5	5	#45	5
	#6	0	#46	0		#6	3	#46	4		#6	4	#46	5
	#7	0	#47	1		#7	3	#47	3		#7	5	#47	5
	#8	1	#48	0		#8	1	#48	2		#8	4	#48	5
	#9	0	#49	0		#9	3	#49	2		#9	4	#49	4
	#10	0	#50	1		#10	2	#50	2		#10	4	#50	4
	#11	1	#51	0		#11	3	#51	2		#11	4	#51	4
	#12	1	#52	0		#12	2	#52	3		#12	5	#52	5
	#13	1	#53	0		#13	2	#53	3		#13	5	#53	5
	#14	0	#54	1		#14	2	#54	2		#14	4	#54	4
	#15	1	#55	0		#15	4	#55	3		#15	4	#55	5
	#16	0	#56	0		#16	3	#56	2		#16	4	#56	5
	#17	1	#57	0		#17	2	#57	3		#17	4	#57	5
	#18	1	#58	0		#18	1	#58	3		#18	4	#58	4
	#19	0	#59	1		#19	3	#59	1		#19	3	#59	3
	#20	1	#60	0		#20	2	#60	4		#20	5	#60	4
	#21	0	#61	1		#21	4	#61	2		#21	4	#61	5
	#22	0	#62	0		#22	4	#62	3		#22	5	#62	4
	#23	0	#63	1		#23	2	#63	1		#23	2	#63	5
	#24	0				#24	4				#24	3		
	#25	0				#25	3				#25	4		
	#26	0				#26	3				#26	5		
	#27	0				#27	4				#27	3		
	#28	0				#28	3				#28	5		
	#29	0				#29	3				#29	4		
	#30	0				#30	4				#30	5		
	#31	0				#31	1				#31	5		
	#32	0				#32	2				#32	5		
	#33	0				#33	4				#33	5		
	#34	0				#34	3				#34	4		
	#35	0				#35	3				#35	4		
	#36	0				#36	3				#36	5		
	#37	0				#37	3				#37	4		
	#38	0				#38	4				#38	2		
	#39	0				#39	3				#39	5		
	#40	0				#40	3				#40	5		

Compilação das respostas ao inquérito efectuado junto dos Alunos da Universidade de Évora, durante o mês de Maio de 2011. O pedido de colaboração foi efectuado através de correio electrónico e as respostas foram dadas através de um *website* que utiliza *cookies* e o registo do IP, de modo a garantir que cada respondente só o faz uma vez.

Discentes Universitários														
Quando necessita de obter informação - aprofundar um determinado tema, onde procura primeiro? (0: Motor de Busca / Wikipédia - 1: Enciclopédia / Dicionário / Obra de Referência)	#1	0	#41	1	Qual o grau de confiança que têm na informação obtida através de um Motor de Busca / Wikipédia? [1: Desconfio Totalmente, 5: Confio Completamente]	#1	3	#41	3	Qual o grau de confiança que têm na informação obtida através de uma Enciclopédia / Dicionário / Obra de Referência? [1: Desconfio Totalmente, 5: Confio Completamente]	#1	4	#41	4
	#2	0	#42	0		#2	2	#42	2		#2	4	#42	5
	#3	0	#43	0		#3	4	#43	1		#3	4	#43	4
	#4	0	#44	0		#4	3	#44	3		#4	5	#44	5
	#5	0	#45	0		#5	3	#45	1		#5	5	#45	5
	#6	0	#46	1		#6	2	#46	2		#6	5	#46	4
	#7	0	#47	0		#7	2	#47	3		#7	5	#47	4
	#8	0	#48	1		#8	4	#48	3		#8	5	#48	4
	#9	0	#49	0		#9	3	#49	3		#9	5	#49	4
	#10	0	#50	1		#10	4	#50	3		#10	4	#50	5
	#11	0	#51	0		#11	3	#51	4		#11	4	#51	5
	#12	0	#52	0		#12	2	#52	4		#12	4	#52	5
	#13	1	#53	0		#13	2	#53	4		#13	4	#53	5
	#14	0	#54	0		#14	3	#54	4		#14	4	#54	5
	#15	0	#55	0		#15	4	#55	3		#15	4	#55	4
	#16	0	#56	0		#16	4	#56	4		#16	5	#56	5
	#17	0	#57	0		#17	3	#57	3		#17	4	#57	5
	#18	0	#58	0		#18	4	#58	5		#18	5	#58	5
	#19	0	#59	0		#19	4	#59	2		#19	5	#59	4
	#20	0	#60	0		#20	3	#60	3		#20	4	#60	5
	#21	1	#61	0		#21	2	#61	3		#21	4	#61	5
	#22	0	#62	0		#22	3	#62	3		#22	5	#62	5
	#23	0	#63	0		#23	2	#63	3		#23	5	#63	4
	#24	1	#64	0		#24	3	#64	3		#24	5	#64	4
	#25	0	#65	0		#25	4	#65	3		#25	4	#65	5
	#26	0	#66	0		#26	4	#66	4		#26	4	#66	5
	#27	0	#67	0		#27	4	#67	3		#27	5	#67	5
	#28	1	#68	0		#28	2	#68	4		#28	4	#68	5
	#29	0	#69	0		#29	5	#69	4		#29	5	#69	4
	#30	0	#70	1		#30	3	#70	3		#30	3	#70	5
	#31	0	#71	0		#31	4	#71	3		#31	4	#71	4
	#32	0	#72	0		#32	3	#72	3		#32	4	#72	5
	#33	1	#73	1		#33	3	#73	2		#33	5	#73	4
	#34	0	#74	1		#34	3	#74	4		#34	5	#74	5
	#35	0	#75	0		#35	3	#75	4		#35	5	#75	5
	#36	1	#76	1		#36	2	#76	3		#36	5	#76	4
	#37	0	#77	0		#37	5	#77	3		#37	5	#77	5
	#38	1	#78	0		#38	2	#78	3		#38	4	#78	5
	#39	0				#39	3				#39	4		
	#40	0				#40	2				#40	5		

Compilação das respostas ao inquérito efectuado junto dos Docentes da Escola Secundária Diogo Gouveia - Beja, durante o mês de Abril de 2011. O pedido de colaboração foi efectuado presencialmente, sendo as respostas efectuadas através do preenchimento de um boletim em papel.

Docentes Secundário														
Quando necessita de obter informação - aprofundar um determinado tema, onde procura primeiro? (0: Motor de Busca / Wikipédia - 1: Enciclopédia / Dicionário / Obra de Referência)	#1	0			Qual o grau de confiança que têm na Informação obtida através de um Motor de Busca / Wikipédia? [1: Desconfio Totalmente, 5: Confio Completamente]	#1	4			Qual o grau de confiança que têm na Informação obtida através de uma Enciclopédia / Dicionário / Obra de Referência? [1: Desconfio Totalmente, 5: Confio Completamente]	#1	5		
	#2	0				#2	4				#2	5		
	#3	0				#3	3				#3	5		
	#4	0				#4	3				#4	5		
	#5	0				#5	5				#5	5		
	#6	0				#6	3				#6	5		
	#7	0				#7	3				#7	4		
	#8	0				#8	3				#8	5		
	#9	0				#9	2				#9	5		
	#10	1				#10	3				#10	5		
	#11	0				#11	3				#11	4		
	#12	1				#12	2				#12	4		
	#13	1				#13	3				#13	5		
	#14	0				#14	4				#14	5		
	#15	0				#15	3				#15	4		
	#16	1				#16	2				#16	5		
	#17	0				#17	2				#17	5		
	#18	0				#18	4				#18	4		
	#19	0				#19	3				#19	4		
	#20	1				#20	2				#20	4		
	#21	1				#21	3				#21	4		
	#22	0				#22	3				#22	4		
	#23	1				#23	4				#23	5		
	#24	0				#24	4				#24	5		
	#25	0				#25	3				#25	4		
	#26	0				#26	2				#26	4		
	#27	0				#27	3				#27	4		
	#28	1				#28	3				#28	5		
	#29	0				#29	4				#29	4		
	#30	0				#30	4				#30	4		
	#31	1				#31	3				#31	5		
	#32	1				#32	1				#32	5		
	#33	0				#33	4				#33	5		
	#34	0				#34	4				#34	4		
	#35	0				#35	3				#35	3		
	#36	0				#36	3				#36	5		
	#37	1				#37	3				#37	5		
	#38	0				#38	2				#38	4		

Compilação das respostas ao inquérito efectuado junto dos Alunos dos 10º, 11º e 12º anos da Escola Secundária Diogo Gouveia - Beja, durante o mês de Abril de 2011. O pedido de colaboração foi efectuado presencialmente, sendo as respostas efectuadas através do preenchimento de um boletim em papel.

Discentes Secundário														
Quando necessita de obter informação - aprofundar um determinado tema, onde procura primeiro? (0: Motor de Busca / Wikipédia - 1: Enciclopédia / Dicionário / Obra de Referência)	#1	0	#41	0	Qual o grau de confiança que têm na informação obtida através de um Motor de Busca / Wikipédia? [1: Desconfio Totalmente, 5: Confio Completamente]	#1	4	#41	5	Qual o grau de confiança que têm na informação obtida através de uma Enciclopédia / Dicionário / Obra de Referência? [1: Desconfio Totalmente, 5: Confio Completamente]	#1	5	#41	5
	#2	0	#42	0		#2	2	#42	5		#2	4	#42	5
	#3	0	#43	0		#3	3	#43	3		#3	4	#43	5
	#4	0	#44	0		#4	5	#44	4		#4	5	#44	5
	#5	0	#45	0		#5	3	#45	3		#5	4	#45	4
	#6	0	#46	0		#6	4	#46	3		#6	5	#46	4
	#7	0	#47	0		#7	3	#47	3		#7	4	#47	4
	#8	0	#48	0		#8	4	#48	3		#8	4	#48	5
	#9	0	#49	0		#9	3	#49	3		#9	5	#49	5
	#10	0	#50	0		#10	3	#50	3		#10	5	#50	5
	#11	0	#51	0		#11	3	#51	3		#11	5	#51	4
	#12	0	#52	0		#12	3	#52	4		#12	5	#52	4
	#13	0	#53	0		#13	4	#53	2		#13	5	#53	4
	#14	0	#54	0		#14	3	#54	3		#14	5	#54	5
	#15	0	#55	0		#15	3	#55	4		#15	5	#55	4
	#16	0	#56	1		#16	3	#56	3		#16	4	#56	4
	#17	0	#57	0		#17	4	#57	4		#17	5	#57	5
	#18	0	#58	0		#18	4	#58	3		#18	5	#58	4
	#19	0	#59	0		#19	3	#59	2		#19	4	#59	4
	#20	0	#60	0		#20	2	#60	3		#20	3	#60	4
	#21	0	#61	0		#21	2	#61	4		#21	5	#61	5
	#22	0	#62	0		#22	4	#62	4		#22	5	#62	5
	#23	0	#63	0		#23	4	#63	3		#23	5	#63	4
	#24	0	#64	0		#24	5	#64	3		#24	5	#64	4
	#25	0	#65	0		#25	4	#65	3		#25	5	#65	4
	#26	1	#66	0		#26	3	#66	3		#26	5	#66	5
	#27	0	#67	0		#27	4	#67	3		#27	5	#67	5
	#28	0	#68	0		#28	3	#68	3		#28	4	#68	5
	#29	0	#69	0		#29	4	#69	2		#29	4	#69	4
	#30	1				#30	3				#30	5		
	#31	1				#31	3				#31	4		
	#32	0				#32	3				#32	5		
	#33	0				#33	3				#33	5		
	#34	0				#34	2				#34	4		
	#35	0				#35	3				#35	5		
	#36	0				#36	4				#36	5		
	#37	0				#37	3				#37	5		
	#38	0				#38	3				#38	5		
	#39	0				#39	3				#39	5		
	#40	0				#40	1				#40	5		

ANEXO IV

50 PRINCÍPIOS DE GESTÃO

Resultado da procura efectuada na Enciclopédia Britânica e na Wikipédia, em Outubro de 2010, a partir das palavras-chave retiradas do Manual de Gestão: guia dos conceitos de A a Z da Executive Digest.

TEMA:	WIKIPÉDIA	BRITÂNICA	TEMA:	WIKIPÉDIA	BRITÂNICA
Activity Based Costing	✓	✓	Management by Objectives	✓	
BCG Matrix	✓		Marketing-Mix	✓	
Benchmarking	✓		Motivation	✓	
Brainstorming	✓		Organizational Culture	✓	
Brand Management	✓		Outplacement	✓	
Break-Even	✓		Outsourcing	✓	✓
Business Ethics	✓		Product Life Cycle	✓	
Excellence	✓		Project Management	✓	
Business Networking	✓		Risk Management	✓	
Reengineering	✓		Scenario Planning	✓	
Change Management	✓		Strategic Alliance	✓	
Competitive Advantage	✓		Strategic Thinking	✓	
Core Competence	✓		SWOT Analysis	✓	
Downsizing	✓		Synergy	✓	
Empowerment	✓		X, Y, Z	✓	
Eco-Management	✓		Total Quality Management	✓	✓
Franchising	✓		Trade Marketing	✓	
Globalization	✓		Value Analysis	✓	
Horizontal Organization	✓		Value Chain	✓	
Just-in-Time	✓	✓	Industry Structural Analysis		
Lateral Thinking	✓		Self Management Teams		
Leadership	✓		Internet and Intranet		
Lean Production	✓		Principles of Common Sense		
Learning Organization	✓		Time Based Competition		
Loyalty Marketing	✓		Urgent vs Important		

Encontrados: Wikipédia 44/50 (88%) - Britânica 4/50 (8%) - Não Encontrados 6/50 (12%) - Pares Artigos 4/50 (8%)

URLS DOS QUATRO PARES DE ARTIGOS ENCONTRADOS

TEMA	WIKIPÉDIA	BRITÂNICA
Activity Based Costing	http://en.wikipedia.org/wiki/Activity_Based_Costing	www.britannica.com/EBchecked/topic/4631/activity-based-costing
Just-in-Time	http://en.wikipedia.org/wiki/Just-in-time_(business)	www.britannica.com/EBchecked/topic/308705/just-in-time-manufacturing
Outsourcing	http://en.wikipedia.org/wiki/Outsourcing	www.britannica.com/EBchecked/topic/1016357/outsourcing
Total Quality Management	http://en.wikipedia.org/wiki/Total_quality_management	www.britannica.com/EBchecked/topic/1387320/Total-Quality-Management

AVALIAÇÕES QUANTITATIVAS¹ DOS QUATRO PARES DE ARTIGOS DE GESTÃO

Tema:	Wikipédia			Britânica		
	Palavras	Caracteres	Linhas	Palavras	Caracteres	Linhas
Activity Based Costing	1500	8570	156	810	4134	75
Just-in-time	3057	16353	297	130	762	14
Outsourcing	2679	16123	293	104	557	10
Total Quality Management	397	2326	42	67	467	8

AVALIADORES

As avaliações foram efectuadas pelos peritos cujos nomes² se encontram na tabela abaixo e por um grupo de 16 alunos do 3º ano da Licenciatura em Gestão 2010/11, a quem aproveito para expressar, mais uma vez, o meu profundo agradecimento.

Universidade de Évora	Instituto Politécnico de Tomar
Jorge Casas Novas	Docente Gestão #1
José Roberto	Docente Gestão #2
José Ventura	Docente Gestão #3
Paulo Resende da Silva	Docente Gestão #4
Rui Fragoso	José Ramos
Rui Quaresma	Nuno Madeira

¹ Nenhum dos artigos em causa continha ilustrações, pelo que a tabela não faz referência a estas.

² Quatro docentes do Instituto Politécnico de Tomar optaram por manter o anonimato.

AVALIAÇÕES QUALITATIVAS DOS QUATRO PARES DE ARTIGOS DE GESTÃO

AVALIADOR		JUST IN TIME		ACTIVITY BASED COSTING		OUTSOURCING		TOTAL QUALITY MANAGEMENT	
		W	B	W	B	W	B	W	B
Alunos 3º Ano Gestão da Universidade de Évora	#1	5	2	4	3	4	2	5	2
	#2	5	2	4	3	4	2	5	2
	#3	4	2	4	3	4	2	5	2
	#4	4	2	3	3	4	2	3	2
	#5	5	2	4	3	5	2	4	2
	#6	3	4	2	4	3	3	4	3
	#7	4	2	4	3	4	2	5	2
	#8	5	1	5	3	5	2	5	2
	#9	4	3	3	2	3	1	2	1
	#10	5	2	4	3	4	2	5	2
	#11	5	3	4	2	4	3	5	2
	#12	4	1	4	1	3	2	2	1
	#13	4	3	4	4	3	3	5	4
	#14	4	3	5	3	4	4	3	5
	#15	3	4	2	3	4	3	5	3
	#16	5	3	4	2	3	1	4	4
Docentes de Gestão da Universidade de Évora e do Instituto Politécnico de Tomar*	#1	5	2	4	3	5	1	3	1
	#2	5	1	5	2	5	1	5	1
	#3	3	4			3	4	3	3
	#4	4	2	4	3	1	3	3	1
	#5	4	1	4	3	1	3	2	1
	#6*	3	2	4	4	4	2	2	2
	#7*	4	2	4	3	5	3	3	2
	#8*	4	3	4	4	5	3	2	2
	#9*	4	3	4	3	5	2	3	3
	#10*	4	3	4	4	5	3	3	2
	#11*	3	2	3	3	5	2	3	3
	#12	4	2	4	3	1	3	3	1

(AS AVALIAÇÕES ESTÃO ORDENADAS PELA DATA DE RECEPÇÃO DAS MESMAS)

COLABORAÇÃO EM MASSA OU AMADORISMO EM MASSA?

UM ESTUDO COMPARATIVO DA QUALIDADE DA INFORMAÇÃO CIENTÍFICA
PRODUZIDA UTILIZANDO OS CONCEITOS E FERRAMENTAS *WIKI*.

**Doutoramento em Gestão
Sistemas de Informação**

Fernando Silvério Nifrário Rodrigues

Orientador:

Professor Doutor Paulo Resende Silva



A massificação do acesso à Internet tornou-nos a todos em potenciais criadores de conteúdos que não passam pela teia de crivos que os editores conceberam para se assegurar da qualidade dos conteúdos que comercializam.

Apesar de o bom senso aconselhar a não confiarmos em informações disponibilizadas por fontes cuja credibilidade desconhecemos, eventualmente a maioria de nós já escolheu um produto com base nas opiniões publicadas na *Amazon* por outros compradores, criou uma imagem de uma determinada empresa/produto/pessoa ao ler artigos de opinião em *blogs* escritos por *amadores*, utilizou a Wikipédia cujos conteúdos são livremente editados por anónimos sem a supervisão de um editor, etc.

Este estudo pretende aferir o grau de confiabilidade da informação científica criada através da colaboração em massa, anónima e gratuita, por contraponto à abordagem tradicional, através de um estudo comparativo da qualidade dos conteúdos das enciclopédias generalistas Wikipédia e Britânica. O objectivo não é a comparação, *tout court*, dos produtos *per si*, mas descobrir se, e em que condições, a colaboração em massa funciona.

Solicitamos assim a sua colaboração, na qualidade de perito na área de Gestão, através da atribuição de uma classificação a quatro pares de artigos, provenientes das enciclopédias acima referidas, tendo em atenção os seguintes aspectos:

- Cada *pdf* enviado em anexo contém as versões de cada enciclopédia sobre o tema descrito no nome do ficheiro, estando as páginas correspondentes a cada uma delas assinaladas pelas letras **A** e **B** em marca de água.
- Deve classificar ambas as versões de cada tema de 1 a 5 (1: Mau, 2:Medíocre, 3: Suficiente, 4:Bom e 5: Muito Bom)
- Pode utilizar os critérios que entender, desde que sejam os mesmos, em todos os artigos que avaliar. Pedimos que proceda como se estivesse a cotar uma pergunta de um exame, ao nível de licenciatura, de dois alunos e que aplique os critérios que normalmente utilizaria nessa situação.
- Os artigos foram descaracterizados (de modo a evitar que o conhecimento da sua proveniência influencie a sua classificação) e contêm apenas o corpo principal do artigo e as figuras a este associado. Estas alterações levaram a perdas de formatação e de qualidade das ilustrações, pelo que solicitamos que apenas tenha em conta os conteúdos.

Para nos fazer chegar as suas avaliações, basta completar o quadro seguinte:

Artigo	Versão	Classificação				
		1	2	3	4	5
JUST IN TIME	A	1	2	3	4	5
	B	1	2	3	4	5
ACTIVITY BASED COSTING	A	1	2	3	4	5
	B	1	2	3	4	5
OUTSOURCING	A	1	2	3	4	5
	B	1	2	3	4	5
TOTAL QUALITY MANAGEMENT	A	1	2	3	4	5
	B	1	2	3	4	5

Caso não pretenda que o seu nome seja incluído na lista de peritos que colaboraram neste estudo e/ou receber as conclusões a que chegarmos, por favor indique-o na sua resposta. Note que não serão publicadas as classificações individuais que nos enviar, pelo que está garantido o anonimato das suas respostas.

COLABORAÇÃO EM MASSA OU AMADORISMO EM MASSA?

UM ESTUDO COMPARATIVO DA QUALIDADE DA INFORMAÇÃO CIENTÍFICA
PRODUZIDA UTILIZANDO OS CONCEITOS E FERRAMENTAS WIKI.

Apesar de o bom senso aconselhar a não confiarmos em informações disponibilizadas por fontes cuja credibilidade desconhecemos, eventualmente a maioria de nós já escolheu um produto com base nas opiniões publicadas na *Amazon* por outros compradores, criou uma imagem de uma determinada empresa/produto/pessoa ao ler artigos de opinião em *blogs* escritos por *amadores*, utilizou a Wikipédia cujos conteúdos são livremente editados por anónimos sem a supervisão de um editor, etc.

Este estudo pretende aferir o grau de confiabilidade da informação científica criada através da colaboração em massa, anónima e gratuita, por contraponto à abordagem tradicional, através de um estudo comparativo da qualidade dos conteúdos das enciclopédias generalistas Wikipédia e Britânica. Solicitamos assim a sua colaboração, na qualidade de Aluno na área de Gestão, através da atribuição de uma classificação a quatro pares de artigos sobre Gestão, provenientes das enciclopédias acima referidas, tendo em atenção os seguintes aspectos:

- Cada artigo contém duas (2) versões de cada enciclopédia sobre o mesmo tema, estando as páginas correspondentes a cada uma delas assinaladas pelas letras **A** e **B** em marca de água.
- Os artigos foram descaracterizados (de modo a evitar que o conhecimento da sua proveniência influencie a sua classificação) e contém apenas o corpo principal do artigo e as figuras a este associado. Estas alterações levaram a perdas de formatação e de qualidade das ilustrações, pelo que solicitamos que apenas tenha em conta os conteúdos
- Deve classificar ambas as versões de cada tema de 1 a 5 (1: Mau, 2:Medíocre, 3: Suficiente, 4:Bom e 5: Muito Bom). Pode utilizar os critérios que entender, desde que sejam os mesmos, em todos os artigos. Pedimos que utilize os critérios pelos quais é normalmente avaliado.

Artigo	Versão	Classificação				
JUST IN TIME	A	1	2	3	4	5
	B	1	2	3	4	5
ACTIVITY BASED COSTING	A	1	2	3	4	5
	B	1	2	3	4	5
OUTSOURCING	A	1	2	3	4	5
	B	1	2	3	4	5
TOTAL QUALITY MANAGEMENT	A	1	2	3	4	5
	B	1	2	3	4	5

Doutorando:
Fernando Nifrário Rodrigues

DOUTORAMENTO EM GESTÃO
SISTEMAS DE INFORMAÇÃO

Orientador:
Professor Doutor Paulo Resende Silva

PAR DE ARTIGOS DESCARACTERIZADOS

JUST IN TIME

Just-in-time (business)

Just-in-time (JIT) is an inventory strategy that strives to improve a business's return on investment by reducing in-process inventory and associated carrying costs. Just In Time production method is also called the Toyota Production System. To meet JIT objectives, the process relies on signals or Kanban (看板 *Kanban*) between different points in the process, which tell production when to make the next part. Kanban are usually 'tickets' but can be simple visual signals, such as the presence or absence of a part on a shelf. Implemented correctly, JIT focuses on continuous improvement and can improve a manufacturing organization's return on investment, quality, and efficiency. To achieve continuous improvement key areas of focus could be flow, employee involvement and quality.

Quick notice that stock depletion requires personnel to order new stock is critical to the inventory reduction at the center of JIT. This saves warehouse space and costs. However, the complete mechanism for making this work is often misunderstood.

For instance, its effective application cannot be independent of other key components of a lean manufacturing system or it can "...end up with the opposite of the desired result."^[1] In recent years manufacturers have continued to try to hone forecasting methods (such as applying a trailing 13 week average as a better predictor for JIT planning,^[2] however some research demonstrates that basing JIT on the presumption of stability is inherently flawed.^[3]

Philosophy

Philosophy of JIT is simple: inventory is waste. JIT inventory systems expose hidden causes of inventory keeping, and are therefore not a simple solution for a company to adopt. The company must follow an array of new methods to manage the consequences of the change. The ideas in this way of working come from many different disciplines including statistics, industrial engineering, production management, and behavioral science. The JIT inventory philosophy defines how inventory is viewed and how it relates to management.

Inventory is seen as incurring costs, or waste, instead of adding and storing value, contrary to traditional accounting. This does not mean to say JIT is implemented without an awareness that removing inventory exposes pre-existing manufacturing issues. This way of working encourages businesses to eliminate inventory that does not compensate for manufacturing process issues, and to constantly improve those processes to require less inventory. Secondly, allowing any stock habituates management to stock keeping. Management may be tempted to keep stock to hide production problems. These problems include backups at work centers, machine reliability, process variability, lack of flexibility of employees and equipment, and inadequate capacity.

In short, the just-in-time inventory system focus is having "the right material, at the right time, at the right place, and in the exact amount"-Ryan Grabosky, without the safety net of inventory. The JIT system has broad implications for implementers.

Transaction cost approach

JIT reduces inventory in a firm. However, a firm may simply be outsourcing their input inventory to suppliers, even if those suppliers don't use Just in time (Naj 1993). Newman (1994) investigated this effect and found that suppliers in Japan charged JIT customers, on average, a 5% price premium.

Environmental concerns

During the birth of JIT, multiple daily deliveries were often made by bicycle. Increased scale has required a move to vans and lorries (trucks). Cusumano (1994) highlighted the potential and actual problems this causes with regard to gridlock and burning of fossil fuels. This violates three JIT waste guidelines:

1. Time—wasted in traffic jams

2. Inventory—specifically pipeline (in transport) inventory
3. Scrap—fuel burned while not physically moving

Price volatility

JIT implicitly assumes a level of input price stability that obviates the need to buy parts in advance of price rises. Where input prices are expected to rise, storing inventory may be desirable.

Quality volatility

JIT implicitly assumes that input parts quality remains constant over time. If not, firms may hoard high quality inputs. As with price volatility, a solution is to work with selected suppliers to help them improve their processes to reduce variation and costs. Longer term price agreements can then be negotiated and agreed-upon quality standards made the responsibility of the supplier. Fixing up of standards for volatility of quality according to the quality circle

Demand stability

Karmarker (1989) highlights the importance of relatively stable demand, which helps ensure efficient capital utilization rates. Karmarker argues that without significantly stable demand, JIT becomes untenable in high capital cost production.

Supply Stability

In the U.S., the 1992 railway strikes caused General Motors to idle a 75,000-worker plant because they had no supply.

JIT Implementation Design

Based on a diagram modeled after the one used by Hewlett-Packard's Boise plant to accomplish its JIT program.

1) F Design Flow Process

- **F** Redesign/relayout for flow
 - **L** Reduce lot sizes
 - **O** Link operations
 - **W** Balance workstation capacity
 - **M** Preventive maintenance
 - **S** Reduce setup Times

2) Q Total Quality Control

- **C** worker compliance
 - **I** Automatic inspection
 - **M** quality measures
 - **M** fail-safe methods
 - **W** Worker participation

3) S Stabilize Schedule

- **S** Level schedule
 - **W** Establish freeze windows
 - **UC** Underutilize Capacity

4) K Kanban Pull System

- **D** Demand pull
 - **B** Backflush
 - **L** Reduce lot sizes

5) **V** Work with Vendors

- **L** Reduce lead time
 - **D** Frequent deliveries
 - **U** Project usage requirements
 - **Q** Quality expectations

6) **I** Further Reduce Inventory in Other Areas

- **S** Stores
 - **T** Transit
 - **C** Implement carrousel to reduce motion waste
 - **C** Implement conveyor belts to reduce motion waste

7) **P** Improve Product Design

- **P** Standard production configuration
 - **P** Standardize and reduce the number of parts
 - **P** Process design with product design
 - **Q** Quality expectations

Effects

A surprising effect was that factory response time fell to about a day. This improved customer satisfaction by providing vehicles within a day or two of the minimum economic shipping delay.

Also, the factory began building many vehicles to order, eliminating the risk they would not be sold. This improved the company's return on equity.

Since assemblers no longer had a choice of which part to use, every part had to fit perfectly. This caused a quality assurance crisis, which led to a dramatic improvement in product quality. Eventually, Toyota redesigned every part of its vehicles to widen tolerances, while simultaneously implementing careful statistical controls for quality control. Toyota had to test and train parts suppliers to assure quality and delivery. In some cases, the company eliminated multiple suppliers.

When a process or parts quality problem surfaced on the production line, the entire production line had to be slowed or even stopped. No inventory meant a line could not operate from in-process inventory while a production problem was fixed. Many people in Toyota predicted that the initiative would be abandoned for this reason. In the first week, line stops occurred almost hourly. But by the end of the first month, the rate had fallen to a few line stops per day. After six months, line stops had so little economic effect that Toyota installed an overhead pull-line, similar to a bus bell-pull, that let any worker on the line order a line stop for a process or quality problem. Even with this, line stops fell to a few per week.

The result was a factory that has been studied worldwide. It has been widely emulated, but not always with the expected results, as many firms fail to adopt the full system.^[4]

The just-in-time philosophy was also applied to other segments of the supply chain in several types of industries. In the commercial sector, it meant eliminating one or all of the warehouses in the link between a factory and a retail establishment. Examples in sales, marketing, and customer service involve applying information systems and mobile hardware to deliver customer information as needed, and reducing waste by video conferencing to cut travel time^[5].

Benefits

Main benefits of JIT include:

- *Reduced setup time.* Cutting setup time allows the company to reduce or eliminate inventory for "changeover" time. The tool used here is SMED (single-minute exchange of dies).
- *The flow of goods from warehouse to shelves improves.* Small or individual piece lot sizes reduce lot delay inventories, which simplifies inventory flow and its management.
- *Employees with multiple skills are used more efficiently.* Having employees trained to work on different parts of the process allows companies to move workers where they are needed.
- *Production scheduling and work hour consistency synchronized with demand.* If there is no demand for a product at the time, it is not made. This saves the company money, either by not having to pay workers overtime or by having them focus on other work or participate in training.
- *Increased emphasis on supplier relationships.* A company without inventory does not want a supply system problem that creates a part shortage. This makes supplier relationships extremely important.
- *Supplies come in at regular intervals throughout the production day.* Supply is synchronized with production demand and the optimal amount of inventory is on hand at any time. When parts move directly from the truck to the point of assembly, the need for storage facilities is reduced.

Problems

Within a JIT system

Just-in-time operation leaves suppliers and downstream consumers open to supply shocks and large supply or demand changes. For internal reasons, Ohno saw this as a feature rather than a bug. He used an analogy of lowering the water level in a river to expose the rocks to explain how removing inventory showed where production flow was interrupted. Once barriers were exposed, they could be removed. Since one of the main barriers was rework, lowering inventory forced each shop to improve its own quality or cause a holdup downstream. A key tool to manage this weakness is production levelling to remove these variations. Just-in-time is a means to improving performance of the system, not an end.

Very low stock levels means shipments of the same part can come in several times per day. This means Toyota is especially susceptible to flow interruption. For that reason, Toyota uses two suppliers for most assemblies. As noted in Liker (2003), there was an exception to this rule that put the entire company at risk because of the 1997 Aisin fire. However, since Toyota also makes a point of maintaining high quality relations with its entire supplier network, several other suppliers immediately took up production of the Aisin-built parts by using existing capability and documentation. Thus, a strong, long-term relationship with a few suppliers is better than short-term, price-based relationships with many competing suppliers. Toyota uses this long-term relationship to send Toyota staff to help suppliers improve their processes. These interventions have been going on for twenty years and have created a more reliable supply chain, improved margins for Toyota and suppliers, and lowered prices for customers. Toyota encourages their suppliers to use JIT with their own suppliers.

Within a raw material stream

As noted by Liker (2003) and Womack and Jones (2003), it ultimately would be desirable to introduce synchronised flow and link JIT through the entire supply stream. However, none followed this in detail all the way back through the processes to the raw materials. With present technology, for example, an ear of corn cannot be grown and delivered to order. The same is true of most raw materials, which must be discovered and/or grown through natural processes that require time and must account for natural variability in weather and discovery. The part of this currently viewed as impossible is the *synchronised* part of flow and the *linked* part of JIT. It is for the reasons stated raw materials companies decouple their supply chain from their clients' demand by carrying large 'finished goods' stocks. Both flow and JIT can be implemented in isolated process islands within the raw materials stream. The

challenge becomes to achieve that isolation by some means other than carrying huge stocks, as most do today.

Because of this, almost all value chains are split into a part made-to-forecast and a part that could, by using JIT, become make-to-order. Historically, the make-to-order part has often been within the retailer portion of the value chain. Toyota took Piggly Wiggly's supermarket replenishment system and drove it at least half way through their automobile factories. Their challenge today is to drive it all the way back to their goods-inwards dock. Of course, the mining of iron and making of steel is still not connected to an order for a particular car. Recognising JIT could be driven back up the supply chain has reaped Toyota huge benefits and a dominant position in the auto industry.

Note that the advent of the mini mill steelmaking facility is starting to challenge how far back JIT can be implemented, as the electric arc furnaces at the heart of many mini-mills can be started and stopped quickly, and steel grades changed rapidly.

Oil

It has been frequently charged that the oil industry has been influenced by JIT.^{[6] [7] [8]}

The argument is presented as follows:

The number of refineries in the United States has fallen from 279 in 1975 to 205 in 1990 and further to 149 in 2004. As a result, the industry is susceptible to supply shocks, which cause spikes in prices and subsequently reduction in domestic manufacturing output. The effects of hurricanes Katrina and Rita are given as an example: in 2005, Katrina caused the shutdown of 9 refineries in Louisiana and 6 more in Mississippi, and a large number of oil production and transfer facilities, resulting in the loss of 20% of the US domestic refinery output. Rita subsequently shut down refineries in Texas, further reducing output. The GDP figures for the third and fourth quarters showed a slowdown from 3.5% to 1.2% growth. Similar arguments were made in earlier crises.

Beside the obvious point that prices went up because of the reduction in supply and not for anything to do with the practice of JIT, JIT students and even oil & gas industry analysts question whether JIT as it has been developed by Ohno, Goldratt, and others is used by the petroleum industry. Companies routinely shut down facilities for reasons other than the application of JIT. One of those reasons may be economic rationalization: when the benefits of operating no longer outweigh the costs, including opportunity costs, the plant may be economically inefficient. JIT has never subscribed to such considerations directly; following Waddel and Bodek (2005), this ROI-based thinking conforms more to Brown-style accounting and Sloan management. Further, and more significantly, JIT calls for a reduction in inventory capacity, not production capacity. From 1975 to 1990 to 2005, the annual average stocks of gasoline have fallen by only 8.5% from 228,331 to 222,903 bbls to 208,986 (Energy Information Administration data). Stocks fluctuate seasonally by as much as 20,000 bbls. During the 2005 hurricane season, stocks never fell below 194,000,000 bbl (m^3), while the low for the period 1990 to 2006 was 187,017,000 bbl (m^3) in 1997. This shows that while industry storage capacity has decreased in the last 30 years, it hasn't been drastically reduced as JIT practitioners would prefer.

Finally, as shown in a pair of articles in the Oil & Gas Journal, JIT does not seem to have been a goal of the industry. In Waguespack and Cantor (1996), the authors point out that JIT would require a significant change in the supplier/refiner relationship, but the changes in inventories in the oil industry exhibit none of those tendencies. Specifically, the relationships remain cost-driven among many competing suppliers rather than quality-based among a select few long-term relationships. They find that a large part of the shift came about because of the availability of short-haul crudes from Latin America. In the follow-up editorial, the Oil & Gas Journal claimed that "casually adopting popular business terminology that doesn't apply" had provided a "rhetorical bogey" to industry critics. Confessing that they had been as guilty as other media sources, they confirmed that "It also happens not to be accurate."

Business models following similar approach

Vendor-managed inventory

Vendor-managed inventory (VMI) employs the same principles as those of JIT inventory, however, the responsibilities of managing inventory is placed with the vendor in a vendor/customer relationship. Whether it's a manufacturer managing inventory for a distributor, or a distributor managing inventory for their customers, the management role goes to the vendor.

An advantage of this business model is that the vendor may have industry experience and expertise that lets them better anticipate demand and inventory needs. The inventory planning and controlling is facilitated by applications that allow vendors access to their customer's inventory data.

Another advantage to the customer is that inventory cost usually remains on the vendor's books until used by the customer, even if parts or materials are on the customer's site.

Customer-managed inventory

With customer-managed inventory (CMI), the customer, as opposed to the vendor in a VMI model, has responsibility for all inventory decisions. This is similar to JIT inventory concepts. With a clear picture of their inventory and that of their supplier's, the customer can anticipate fluctuations in demand and make inventory replenishment decisions accordingly.

References

- [1] A study of the Toyota Production System, Shigeo Shingo, Productivity Press, 1989, p 187
- [2] Gilliland, Michael. "Is Forecasting a Waste of Time?" (<http://www.scmr.com/article/CA232251.html>), *Supply Chain Management Review*, July/August 2002.
- [3] Ruffa, Stephen A., (2008). *Going Lean: How the Best Companies Apply Lean Manufacturing Principles to Shatter Uncertainty, Drive Innovation, and Maximize Profits*, AMACOM (American Management Association)
- [4] Alan Pilkington, "Manufacturing Strategy Regained: Evidence for the Demise of Best-Practice" (<http://cmr.berkeley.edu/search/articleDetail.aspx?article=4323>), *California Management Review*, (1998) Vol. 41, No. 1, pp.31-42.
- [5] Paul H. Selden (1997). *Sales Process Engineering: A Personal Workshop*. Milwaukee, WI: ASQ Quality Press. pp. 113–117.
- [6] Bongiorno, Sara. "All in the timing" (<http://web.archive.org/web/20040909200301/http://www.businessreport.com/newsDetail.cfm?aid=127>), *The Greater Baton Rouge Business Report*, 19 July 2004.
- [7] "Online NewsHour: Rising gas prices -- April 30, 1996" (http://www.pbs.org/newshour/bb/economy/gas_hike_4-30.html). . Retrieved 2007-09-24.
- [8] "Story taken from Time magazine May 13, 1996 Volume 147, No. 20" (<http://www.econ.ucsb.edu/~tedb/ceep/news/gas.html>). . Retrieved 2007-09-24.

just-in-time manufacturing (JIT), Production-control system, developed by Toyota Motor Corp. and imported to the West, that has revolutionized manufacturing methods in some industries. By relying on daily deliveries of most supplies, it eliminates waste due to overproduction and lowers warehousing costs. Supplies are closely monitored and quickly altered to meet changing demands, and small and accurate resupply deliveries must be made just as they are needed. Because there are no spares, the components must be free of defects. Plants wholly dedicated to the JIT concept require a logistics staff to schedule production, balancing product demand with plant capacity and availability of inputs. JIT has worked most effectively for large automobile manufacturers, which may have several thousand suppliers feeding parts into 100 factories that assemble components for 20 assembly lines.



PAR DE ARTIGOS DESCARACTERIZADOS

ACTIVITY BASED COST

Activity-based costing

Activity-based costing (ABC) is a costing model that identifies activities in an organization and assigns the cost of each activity resource to all products and services according to the actual consumption by each: it assigns more indirect costs (overhead) into direct costs.

In this way, an organization can precisely estimate the cost of individual products and services so they can identify and eliminate those that are unprofitable and lower the prices of those that are overpriced.[]

In a business organization, the ABC methodology assigns an organization's resource costs through activities to the[] products and services provided to its customers. It is generally used as a tool for understanding product and customer[] cost and profitability. As such, ABC has predominantly been used to support strategic decisions such as pricing,[] outsourcing, identification and measurement of process improvement initiatives.[]

Historical development

Traditionally cost accountants had arbitrarily added a broad percentage of expenses into the indirect cost. In addition, activities include actions that are performed both by people and machine. However, as the percentages of indirect or overhead costs rose, this technique became increasingly inaccurate, because indirect costs were not caused equally by all products. For example, one product might take more time in one expensive machine than another product—but since the amount of direct labor and materials might be the same, additional cost for use of the machine is not being recognized when the same broad 'on-cost' percentage is added to all products. Consequently, when multiple products share common costs, there is a danger of one product subsidizing another.

ABC is based on George Staubus' Activity Costing and Input-Output Accounting.^[1] The concepts of ABC were developed in the manufacturing sector of the United States during the 1970s and 1980s. During this time, the Consortium for Advanced Management-International, now known simply as CAM-I, provided a formative role for studying and formalizing the principles that have become more formally known as Activity-Based Costing.^[2]

Robin Cooper and Robert S. Kaplan, proponents of the Balanced Scorecard, brought notice to these concepts in a number of articles published in *Harvard Business Review* beginning in 1988. Cooper and Kaplan described ABC as an approach to solve the problems of traditional cost management systems. These traditional costing systems are often unable to determine accurately the actual costs of production and of the costs of related services. Consequently managers were making decisions based on inaccurate data especially where there are multiple products.

Instead of using broad arbitrary percentages to allocate costs, ABC seeks to identify cause and effect relationships to objectively assign costs. Once costs of the activities have been identified, the cost of each activity is attributed to each product to the extent that the product uses the activity. In this way ABC often identifies areas of high overhead costs per unit and so directs attention to finding ways to reduce the costs or to charge more for costly products.

Activity-based costing was first clearly defined in 1987 by Robert S. Kaplan and W. Bruns as a chapter in their book *Accounting and Management: A Field Study Perspective*.^[3] They initially focused on manufacturing industry where increasing technology and productivity improvements have reduced the relative proportion of the direct costs of labor and materials, but have increased relative proportion of indirect costs. For example, increased automation has reduced labor, which is a direct cost, but has increased depreciation, which is an indirect cost.

Like manufacturing industries, financial institutions have diverse products and customers, which can cause cross-product, cross-customer subsidies. Since personnel expenses represent the largest single component of non-interest expense in financial institutions, these costs must also be attributed more accurately to products and customers. Activity based costing, even though originally developed for manufacturing, may even be a more useful tool for doing this.^[4] ^[5]

Activity-based costing was later explained in 1999 by Peter F. Drucker in the book *Management Challenges of the 21st Century*.^[6] He states that traditional cost accounting focuses on what it costs to *do something*, for example, to

cut a screw thread; activity-based costing also records the cost of *not doing*, such as the cost of waiting for a needed part. Activity-based costing records the costs that traditional cost accounting does not do.

The overhead costs assigned to each activity comprise an activity cost pool.

Methodology

- Cost allocation
- Fixed cost
- Variable cost
- Cost driver

Direct labor and materials are relatively easy to trace directly to products, but it is more difficult to directly allocate indirect costs to products. Where products use common resources differently, some sort of weighting is needed in the cost allocation process. The **cost driver** is a factor that creates or drives the cost of the activity. For example, the cost of the activity of bank tellers can be ascribed to each product by measuring how long each product's transactions (**cost driver**) takes at the counter and then by measuring the number of each type of transaction. Other example, for the running machinery activity, the driver is likely to be machine operating hours. That is, machine operating hours drive labour, maintenance, and power cost during the running machinery activity.

Uses

- It helps to identify inefficient products, departments and activities
- It helps to allocate more resources on profitable products, departments and activities
- It helps to control the costs at an individual level and on a departmental level
- It helps to find unnecessary costs
- It helps fixing the price of a product or service scientifically

Limitations

Even in activity-based costing, some overhead costs are difficult to assign to products and customers, such as the chief executive's salary. These costs are termed 'business sustaining' and are not assigned to products and customers because there is no meaningful method. This lump of unallocated overhead costs must nevertheless be met by contributions from each of the products, but it is not as large as the overhead costs before ABC is employed.

Although some may argue that costs untraceable to activities should be "arbitrarily allocated" to products, it is important to realize that the only purpose of ABC is to provide information to management. Therefore, there is no reason to assign any cost in an arbitrary manner.

Cost

ABC is considered a relatively costly accounting methodology.^[7]

Lean accounting methods have been developed in recent years to provide relevant and thorough accounting, control, and measurement systems without the complex and costly methods of ABC. Lean Accounting takes an opposite direction from ABC by working to eliminate cost allocations rather than find complicated methods of allocation.

While lean accounting is primarily used within lean manufacturing, the approach has proven useful in many other areas including healthcare, construction, financial services, governments, and other industries.

Prevalence

Following initial enthusiasm, ABC lost ground in the 1990s, to alternative metrics, such as Kaplan's balanced scorecard and economic value added.^[7]

ABC has stagnated over the last five to seven years,

— Kaplan, 1998

However, application of an activity based recording may be applied without change to incremental **activity based accounting** not replacing any synoptic and retrospective modeling process with *costing*, but to transform concurrent process accounting to a most authentic approach.

Public sector use

ABC is widely used in the public sector, including by the United States Marine Corps.^[7]

Its use by the UK Police has been mandated since the 2003-04 UK tax year as part of England and Wales' National Policing Plan, specifically the Policing Performance Assessment Framework.^[8] An independent 2008 report concluded that ABC was an inefficient use of resources: it was expensive and difficult to implement for small gains, and a poor value, and that alternative methods should be used.^[9]

Furthermore, the South African government, specifically the National Treasury has tasked specialists to craft a master guideline for local government in South Africa, for purposes of cost optimisation, tariff setting, balanced budget setting and equitable fund allocations

References

- [1] Staubus, George J. *Activity Costing and Input-Output Accounting* (Richard D. Irwin, Inc., 1971).
- [2] Consortium for Advanced Manufacturing-International (<http://www.cam-i.org>)
- [3] Kaplan, Robert S. and Bruns, W. *Accounting and Management: A Field Study Perspective* (Harvard Business School Press, 1987) ISBN 0-87584-186-4
- [4] Sapp, Richard, David Crawford and Steven Rebishcke "Article title?" *Journal of Bank Cost and Management Accounting* (Volume 3, Number 2), 1990.
- [5] Author(s)? "Article title?" *Journal of Bank Cost and Management Accounting* (Volume 4, Number 1), 1991.
- [6] Drucker Peter F. *Management Challenges of the 21st Century*. New York:Harper Business, 1999.
- [7] Activity-Based Costing (ABC): In recent years, ABC has lost ground in the metric wars. But it may be set for a resurgence (<http://www.cfo.com/article.cfm/3007694>), by David M. Katz
- [8] Police Service National ABC Model Manual of Guidance Version 2.3 June 2007 (http://police.homeoffice.gov.uk/publications/finance-and-business-planning/ABC_Manual_of_Guidance2835.pdf?view=Binary)
- [9] The Review of Policing Final Report by Sir Ronnie Flanagan February 2008 (http://police.homeoffice.gov.uk/publications/police-reform/Review_of_policing_final_report/flanagan-final-report2835.pdf?view=Binary)

Drucker, Peter F.. *Management Challenges of the 21st Century*. New York:Harper Business, 1999.

Activity-based costing

Activity-based costing, is based on the fact that many costs are driven by factors other than product volume. The first task is to identify the activities that drive costs. The next step is to estimate the costs that are driven by each activity and to state them as averages per unit of activity. Management can use these averages to guide its efforts to reduce costs. In addition, if management wants an estimate of the cost of a specific product, the accountant can estimate how many of the activity units are associated with that product and multiply those numbers by the average costs per activity unit.

For example, suppose that costs driven by the number of machine hours average \$12 per machine hour, costs driven by the number of production batches average \$100 a batch, and the costs of keeping a product in the line average \$100 a year for each kind of material or component part used. Keeping in the line a product that is assembled from six component parts thus incurs costs of $6 \times \$100 = \600 a year, irrespective of volume and even if the product is not made at all during the period. If the plant manufactures 10,000 of these units in a year, the unit cost of product maintenance is $\$600/10,000 = \$.06$ a unit. If this product is manufactured in batches of 1,000 units, then batch-driven costs average $\$100/1,000 = \$.10$ a unit. And, if a batch requires 15 machine hours, hour-driven costs average $15 \times \$12/1,000 = \$.18$ a unit. At the 10,000-unit volume, then, the cost of this product is $\$.06 + \$.10 + \$.18 = \$.34$ a unit plus the cost of materials.

Product cost finding under activity-based costing is almost always a process of estimating costs before production takes place. The method of process costing and job-order costing can be used either in preparing estimates before the fact or in assigning costs to products as production proceeds. Even when job-order costing is used to tally the costs actually incurred on individual jobs, the overhead rates are usually predetermined—that is, they represent the average planned overhead cost at some production volume. The main reason for this is that actual overhead cost averages depend on the total volume and efficiency of operations and not on any one job alone. The relevance of job-order cost information will be impaired if these external fluctuations are allowed to change the amount of overhead cost assigned to a particular job.

Many systems go even farther than this. Estimates of the average costs of each type of material, each operation, and each product are prepared routinely and identified as standard costs. These are then readily available whenever estimates are needed and can also serve as an important element in the company's performance-reporting system, as described below.

Similar methods of cost finding can be used to determine or estimate the cost of providing services rather than physical goods. Most advertising agencies and consulting firms, for example, maintain some form of job cost records, either as a basis for billing their clients or as a means of estimating the profitability of individual jobs or accounts.

The methods of cost finding described in the preceding paragraphs are known as full, or absorption, costing methods, in that the overhead rates are intended to include provisions for all manufacturing costs. Both process and job-order costing methods can also be adapted to

variable costing in which only variable manufacturing costs are included in product cost. Variable costs rise or fall in proportion to the quantity of output. Total fixed costs, in contrast, are the same at all volume levels within the normal range.

Unit cost under variable costing represents the average variable cost of making the product. Compared to the average full cost, the average variable cost is more useful when making short-term managerial decisions. In deciding whether to manufacture goods in large lots, for example, management needs to estimate the cost of carrying larger amounts of finished goods in inventory. More variable costs will have to be incurred to build the inventory to a higher level; fixed manufacturing costs presumably will be unaffected.

Furthermore, when a management decision changes the company's fixed costs, the change is unlikely to be proportional to the change in volume; therefore, average fixed cost is seldom a valid basis for estimating the cost effects of such decisions. Variable costing eliminates the temptation to use average fixed cost in estimating changes in the total fixed cost. When variable costing is used, supplemental rates for fixed overhead production costs must be provided to measure the costs to be assigned to end-of-year inventories. This practice is followed because generally accepted accounting principles (GAAP) in the United States and in most other countries require that inventories be measured at full product cost for external financial reporting.

PAR DE ARTIGOS DESCARACTERIZADOS

OUTSOURCING

Outsourcing

Outsourcing or **sub-servicing** often refers to the process of contracting to a third-party.^[1]

Overview

A precise definition of outsourcing has yet to be agreed upon. Thus, the term is used inconsistently. However, outsourcing is often viewed as involving the contracting out of a business function - commonly one previously performed in-house - to an external provider.^[2] In this sense, two organizations may enter into a contractual agreement involving an exchange of services and payments. Of recent concern is the ability of businesses to outsource to suppliers outside the nation, sometimes referred to as offshoring or offshore outsourcing (which are odd terms because doing business with another country does not mean you have to go offshore^{[3] [4] [5] [6] [7]}) In addition, several related terms have emerged to grasp various aspects of the complex relationship between economic organizations or networks, such as nearshoring, multisourcing^{[8] [9]} and strategic outsourcing.^[10]

Reasons

Organizations that outsource are seeking to realize benefits or address the following issues:^{[11] [12] [13] [14]}

- **Cost savings** — The lowering of the overall cost of the service to the business. This will involve reducing the scope, defining quality levels, re-pricing, re-negotiation, and cost re-structuring. Access to lower cost economies through offshoring called "labor arbitrage" generated by the wage gap between industrialized and developing nations.^[15]
- **Focus on Core Business** — Resources (for example investment, people, infrastructure) are focused on developing the core business. For example often organizations outsource their IT support to specialised IT services companies.
- **Cost restructuring** — Operating leverage is a measure that compares fixed costs to variable costs. Outsourcing changes the balance of this ratio by offering a move from fixed to variable cost and also by making variable costs more predictable.
- **Improve quality** — Achieve a steep change in quality through contracting out the service with a new service level agreement.
- **Knowledge** — Access to intellectual property and wider experience and knowledge.^[16]
- **Contract** — Services will be provided to a legally binding contract with financial penalties and legal redress. This is not the case with internal services.^[17]
- **Operational expertise** — Access to operational best practice that would be too difficult or time consuming to develop in-house.
- **Access to talent** — Access to a larger talent pool and a sustainable source of skills, in particular in science and engineering.^{[3] [18]}
- **Capacity management** — An improved method of capacity management of services and technology where the risk in providing the excess capacity is borne by the supplier.
- **Catalyst for change** — An organization can use an outsourcing agreement as a catalyst for major step change that can not be achieved alone. The outsourcer becomes a Change agent in the process.
- **Enhance capacity for innovation** — Companies increasingly use external knowledge service providers to supplement limited in-house capacity for product innovation.^{[19] [20]}
- **Reduce time to market** — The acceleration of the development or production of a product through the additional capability brought by the supplier.^[21]
- **Commodification** — The trend of standardizing business processes, IT Services, and application services which enable to buy at the right price, allows businesses access to services which were only available to large

corporations.

- Risk management — An approach to risk management for some types of risks is to partner with an outsourcer who is better able to provide the mitigation.^[22]
- Venture Capital — Some countries match government funds venture capital with private venture capital for start-ups that start businesses in their country.^[23]
- Tax Benefit — Countries offer tax incentives to move manufacturing operations to counter high corporate taxes within another country.
- Scalability — The outsourced company will usually be prepared to manage a temporary or permanent increase or decrease in production.
- Creating leisure time — Individuals may wish to outsource their work in order to optimise their work-leisure balance.^[24]

Implications

Management, the corporation and consumers

Management processes

Greater physical distance between higher management and the production floor employees often require a change in management methodologies. As inspection and feedback may not be as direct and frequent as in internal processes. This often requires the assimilation of new communication methods such as Voice over ip, Instant messaging, and Issue Tracking Systems, new Time management methods such as Time Tracking Software, and new cost and schedule assessment tools such as Cost Estimation Software.

Quality of service

Quality in terms of end-user-experience is best measured through customer satisfaction questionnaires which are professionally designed to capture an unbiased view of quality. Surveys can be one of research.^[25]

Language skills

In the area of call centers end-user-experience is deemed to be of lower quality when a service is outsourced. This is exacerbated when outsourcing is combined with off-shoring to regions where the first language and culture are different.^[26] The questionable quality is particularly evident when call centers that service the public are outsourced and offshored.

Call center agents may speak a variety of the language with different linguistic features such as accents, word use and phraseology, which may make them more difficult to understand for the clients. The visual clues that are present in face-to-face encounters are missing from the call center interactions and this also may lead to misunderstandings and difficulties.^[27] In addition to language and accent differences, a lack of local social and geographic knowledge is often present, leading to misunderstandings or mis-communications.

Security

Before outsourcing an organization is responsible for the actions of all their staff and liable for their actions. When these same people are transferred to an outsourcer they may not change desk but their legal status has changed. They are no longer directly employed or responsible to the organization. This causes legal, security and compliance issues that need to be addressed through the contract between the client and the suppliers. This is one of the most complex areas of outsourcing and requires a specialist third party adviser.

Fraud is a specific security issue that is criminal activity whether it is by employees or the supplier staff. However, it can be disputed that the fraud is more likely when outsourcers are involved, for example credit card theft when there is scope for fraud by credit card cloning. In April 2005, a high-profile case involving the theft of \$350,000 from four

Citibank customers occurred when call center workers acquired the passwords to customer accounts and transferred the money to their own accounts opened under fictitious names. Citibank did not find out about the problem until the American customers noticed discrepancies with their accounts and notified the bank.^[28]

Qualifications of outsourcers

The outsourcer may replace staff with less qualified people or with people with different non-equivalent qualifications.^[29]

In the engineering discipline there has been a debate about the number of engineers being produced by the major economies of the United States, India and China. The argument centers around the definition of an engineering graduate and also disputed numbers. The closest comparable numbers of annual graduates of four-year degrees are United States (137,437) India (112,000) and China (351,537).^{[30] [31]}

Public opinion

In the United States there is a strong public opinion against outsourcing (especially when combined with offshoring) because it leads to job displacement. However, outsourcing supporters draw on mainstream economics to argue that outsourcing should bring down prices, providing greater economic benefit to all.

Standpoint of labor

From the standpoint of labor outsourcing may represent a new threat, contributing to worker insecurity, and reflective of the general process of globalization.^[32]

On June 26, 2009, Jeff Immelt, the CEO of General Electric, called for the United States to increase its manufacturing base employment to 20% of the workforce commenting that the U.S. has outsourced too much and can no longer rely on consumer spending to drive demand.^[33]

Standpoint of government

Western governments may attempt to compensate workers affected by outsourcing through various forms of legislation. In Europe, the Acquired Rights Directive attempts to address the issue. The Directive is implemented differently in different nations. In the United States, the Trade Adjustment Assistance Act is meant to provide compensation for workers directly affected by international trade agreements. Whether or not these policies provide the security and fair compensation they promise is debatable.

By country

United States

'Outsourcing' became a popular political issue in the United States during the 2004 U.S. presidential election. The political debate centered on outsourcing's consequences for the domestic U.S. workforce. Democratic U.S. presidential candidate John Kerry criticized U.S. firms that outsource jobs abroad or that incorporate overseas in tax havens to avoid paying their "fair share" of U.S. taxes during his 2004 campaign, calling such firms "Benedict Arnold corporations".

Criticism of outsourcing, from the perspective of U.S. citizens, generally revolves around the costs associated with transferring control of the labor process to an external entity in another country. A Zogby International poll conducted in August 2004 found that 71% of American voters believed that "outsourcing jobs overseas" hurt the economy while another 62% believed that the U.S. government should impose some legislative action against companies that transfer domestic jobs overseas, possibly in the form of increased taxes on companies that outsource.^[34]

Union busting is one cause of outsourcing. As unions are disadvantaged by union busting legislation, workers lose bargaining power and it becomes easier for corporations to fire them and ship their job overseas.^[35]

Another given rationale is the high corporate income tax rate in the U.S. relative to other OECD nations,^{[36] [37] [38]} and the practice of taxing revenues earned outside of U.S. jurisdiction, a very uncommon practice. However, outsourcing is not solely a U.S. phenomenon as corporations in various nations with low tax rates outsource as well, which means that high taxation can only partially, if at all, explain US outsourcing. For example, the amount of corporate outsourcing in 1950 would be considerably lower than today, yet the tax rate was actually higher in 1950.^[39]

It is argued that lowering the corporate income tax and ending the double-taxation of foreign-derived revenue (taxed once in the nation where the revenue was raised, and once from the U.S.) will alleviate corporate outsourcing and make the U.S. more attractive to foreign companies. However, while the US has a high official tax rate, the actual taxes paid by US corporations may be considerably lower due to the use of tax loopholes, tax havens, and attempts to "game the system".^[40] Rather than avoiding taxes, outsourcing may be mostly driven by the desire to lower labor costs (see standpoint of labor above). Sarbanes-Oxley has also been cited as a factor for corporate flight from U.S. jurisdiction.

References

- [1] "Terms and Definitions" (<http://www.ventureoutsource.com/node/18/print>). *ventureoutsource.com*. Retrieved 2007-10-05.
- [2] Overby, S (2007) ABC: An Introduction to Outsourcing (http://www.cio.com/article/40380/ABC_An_Introduction_to_Outsourcing). CIO.com.
- [3] Manning et al. (2008) A Dynamic Perspective on Next-Generation Offshoring: The Global Sourcing of Science and Engineering Talent (http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1287369) *Academy of Management Perspectives* 22.3: 35-54.
- [4] Norwood et al. (2006) Off-Shoring: An Elusive Phenomenon (<http://www.napawash.org/Pubs/Off-ShoringJan06.pdf>). National Academy of Public Administration
- [5] Babu, M. (2005) Myth: All Outsourcing Is Offshoring (<http://www.computerworld.com/managementtopics/outsourcing/story/0,10801,105828,00.html?SKC=management-105828>) www.computerworld.com
- [6] McCue, A. (2006) Indian outsourcingers to launch European invasion (<http://www.silicon.com/publicsector/0,3800010403,39155869,00.htm>) www.silicon.com.
- [7] Gibson (2006) India 2.0 Aims to Sustain Its Global IT Influence (<http://www.eweek.com/article/2/0,1895,1931483,00.asp?kc=ewnws022706dix1k0000599>) eWeek
- [8] (Q4 2006) Mandatory Multisourcing Discipline (<http://www.btquarterly.com/?mc=mandatory-multisourcing-discipline&page=ss-viewresearch>) *Business Trends Quarterly*
- [9] (2006) Mandatory Multisourcing Discipline (http://www.businessweek.com/globalbiz/content/nov2006/gb20061108_738883.htm?chan=top+news_top+news+index_global+business)
- [10] see Holcomb & Hiitt, 2007
- [11] Gareiss, R (2002, 18 Nov) *Analyzing The Outsourcers* (ftp://ftp.hp.com/pubs/services/spotlight/info/info_week.pdf). Information Week.
- [12] Drezner, D.W. (2004) The Outsourcing Boogeyman (<http://www.foreignaffairs.org/20040501faessay83304/daniel-w-drezner/the-outsourcing-boogeyman.html>) www.foreignaffairs.org
- [13] Engardio, P. (2006) Outsourcing: Job Killer or Innovation Boost? (http://www.businessweek.com/globalbiz/content/nov2006/gb20061108_738883.htm?chan=top+news_top+news+index_global+business) *Business Week*
- [14] "Is it Virtuous to be Virtual? The VC Viewpoint" (<http://www.nature.com/nbt/journal/v27/n10/full/nbt1009-886.html>), Justin Chakma, Jeff L. Calcagno, Ali Behbahani and Shawn Mojtabedian, *Nature Biotechnology* Volume 27, Number 10, October 2009.
- [15] Engardio, P. & Arndt, M. & Foust, D. (2006) The Future Of Outsourcing (http://www.businessweek.com/globalbiz/content/nov2006/gb20061108_738883.htm?chan=top+news_top+news+index_global+business) *Business Week*
- [16] Engardio, P. & Kripalani, M. (2006) The Rise Of India (<http://www.nasscom.in/Nasscom/templates/NormalPage.aspx?id=44614>) *Business Week*
- [17] Rothman, J. (2003) 11 Steps to Successful Outsourcing: A Contrarian's View (<http://www.computerworld.com/developmenttopics/development/story/0,10801,84847,00.html>) www.computerworld.com
- [18] Lewin, A.Y. & Couto, V. Next Generation Offshoring: The Globalization of Innovation (https://offshoring.fuqua.duke.edu/om_report.pdf) Offshoring Research Network (<https://offshoring.fuqua.duke.edu/>) 2006 Survey Report
- [19] Lewin, A.Y. & Couto, V. Next Generation Offshoring: The Globalization of Innovation (https://offshoring.fuqua.duke.edu/om_report.pdf) Offshoring Research Network (<https://offshoring.fuqua.duke.edu/>) 2006 Survey Report

- [20] Couto et al. *Offshoring 2.0: Contracting Knowledge and Innovation to Expand Global Capabilities* (https://offshoring.fuqua.duke.edu/pdfs/DukeServiceProviderReport_web.pdf) Offshoring Research Network (<https://offshoring.fuqua.duke.edu/>) 2007 Service Provider Report
- [21] Nadeem, S (2009) "The Uses and Abuses of Time: Globalization and Time Arbitrage in India's Outsourcing Industries" (http://nadeem.commons.gc.cuny.edu/files/2010/04/Nadeem_Global-Networks.pdf). Global Networks
- [22] Roehrig, P (2006) Bet On Governance To Manage Outsourcing Risk (<http://www.biquarterly.com/?mc=bet-governance&page=ss-viewresearch>). Business Trends Quarterly
- [23] "Russia finally gets serious about venture capital" (<http://venturebeat.com/2007/06/01/russia-finally-gets-serious-about-venture-capital/>). VentureBeat. 2007-06-01. Retrieved 2010-03-15.
- [24] Gamerman, Ellen (2007-06-02). "Outsourcing Your Life" (<http://online.wsj.com/article/SB118073815238422013.html>). Wall Street Journal. Retrieved 2010-07-23.
- [25] Maddock, B. & Warren, C. & Worsley A. (2005) Survey of canteens and food services in Victorian schools (<http://www.blackwell-synergy.com/doi/abs/10.1111/j.1747-0080.2005.00008.x>)
- [26] Nadeem, S (2009) Macaulay's (Cyber) Children: The Cultural Politics of Outsourcing in India (http://nadeem.commons.gc.cuny.edu/files/2010/04/CS_Nadeem.pdf). Cultural Sociology
- [27] Alster, N (2005) Customer Disservice. (http://www.cfo.com/article.cfm/4390954/c_2984406?f=archives) www.CFO.com.
- [28] Ribeiro, J (2005) Indian call center workers charged with Citibank fraud (http://www.infoworld.com/article/05/04/07/HNcitibankfraud_1.html). www.infoworld.com
- [29] Stein, R (2005) Hospital Services Performed Overseas (<http://www.washingtonpost.com/wp-dyn/articles/A12392-2005Apr23.html>). www.washingtonpost.com
- [30] Wadhwa, V (2005) About That Engineering Gap (http://www.businessweek.com/smallbiz/content/dec2005/sb20051212_623922.htm). www.businessweek.com
- [31] Gereffi, G. & Wadhwa, V. Framing the Engineering Outsourcing Debate: Placing the United States on a Level Playing Field with China and India. (<http://memp.pratt.duke.edu/outsourcing/>) Duke University.
- [32] Krugman, Paul (2006). "Feeling No Pain." *New York Times*, March 6, 2006
- [33] Bailey, David and Soyoun Kim (June 26, 2009). GE's Immelt says U.S. economy needs industrial renewal (<http://www.guardian.co.uk/business/feedarticle/8578904>). *UK Guardian*. Retrieved on June 28, 2009.
- [34] Zogby International survey results online at zogby.com (<http://www.zogby.com/news/ReadNews.dbm?ID=870>)
- [35] http://act.americanrightswork.org/p/dia/action/public/?action_KEY=576
- [36] "Veronique de Rugy on Corporate Flight & Taxes on NRO Financial" (http://www.nationalreview.com/nrof_comment/comment-rugy041802.asp). Nationalreview.com. 2002-04-18. Retrieved 2010-03-15.
- [37] "U.S. Lagging Behind OECD Corporate Tax Trends" (<http://www.taxfoundation.org/news/show/1466.html>). The Tax Foundation. 2006-05-05. Retrieved 2010-03-15.
- [38] John Tamny. "John Tamny on Hillary Clinton Economics on NRO Financial" (<http://article.nationalreview.com/?q=ODZjNj14ZTNjZmNiOGMxYjAwOTg1ZGI0NmFiOWFjZj1=>). Article.nationalreview.com. Retrieved 2010-03-15.
- [39] "Microsoft Word - 10-16-03tax-rev.doc" (<http://www.cbpp.org/files/10-16-03tax.pdf>) (PDF). Retrieved 2010-03-15.
- [40] "High Corporate Tax Rate Is Misleading at" (<http://www.smartmoney.com/investing/economy/high-corporate-tax-rate-is-misleading-22463/>). Smartmoney.com. Retrieved 2010-03-15.

outsourcing, work arrangement made by an employer who hires an outside contractor to perform work that could be done by company personnel. Outsourcing has been a frequent point of dispute for organized labour. If, for example, an employer has a labour contract with a union, and the outsourced work could be performed by union members, then the union will typically object to such a practice because it takes work away from the union's members. Management favours outsourcing, or subcontracting, often to nonunion providers, because these activities can often reduce costs. Outsourcing can also reduce the number of employees in a collective bargaining unit.

B

PAR DE ARTIGOS DESCARACTERIZADOS
TOTAL QUALITY MANAGEMENT

Total quality management

Since the late 1980s, firms around the world have launched **Total Quality Management (TQM)** programs in an attempt to retain or regain competitiveness in order to achieve customer satisfaction in the face of increasing competition from around the world in this era of globalization. TQM is an integrative philosophy of management for continuously improving the quality of products and processes.^[1]

TQM functions on the premise that the quality of the products and processes is the responsibility of everyone who is involved with the creation or consumption of the products or services offered by the organization. In other words, TQM capitalizes on the involvement of management, workforce, suppliers, and even customers, in order to meet or exceed customer expectations. Considering the practices of TQM as discussed in six empirical studies, Cua, McKone, and Schroeder (2001) identified the nine common TQM practices as cross-functional product design, process management, supplier quality management, customer involvement, information and feedback, committed leadership, strategic planning, cross-functional training, and employee involvement.^[2]

TQM and Six Sigma

The Six Sigma process improvement originated in 1986 from Motorola's drive towards reducing defects by minimizing variation in processes through metrics measurement.^[3] Applications of the Six Sigma project execution methodology have since expanded to include practices common in Total Quality Management and Supply Chain Management, such as increasing customer satisfaction, and developing closer supplier relationships.^[4]

The main difference between TQM and Six Sigma (a newer concept) is the approach. TQM tries to improve quality by ensuring conformance to internal requirements, while Six Sigma focuses on improving quality by reducing the number of defects and impurities.^[5]

References

- [1] Ahire, S. L. 1997. Management Science- Total Quality Management interfaces: An integrative framework. *Interfaces* 27 (6) 91-105.
- [2] Cua, K. O., K. E. McKone, and R. G. Schroeder. 2001. Relationships between implementation of TQM, JIT, and TPM and manufacturing performance. *Journal of Operations Management* 19 (6) 675-694.
- [3] Anand, G., P. T. Ward, and M. V. Tatikonda. 2010. Role of explicit and tacit knowledge in six sigma projects: An empirical examination of differential project success. *Journal of Operations Management* 28 (4) 303-315.
- [4] Hoerl, R. W. 2001. Six Sigma black belts: what do they need to know? *Journal of Quality Technology* (4) 391-406.
- [5] "Six Sigma vs. Total Quality Management" (<http://www.pmiut.com/six-sigma-vs-total-quality-management>). Retrieved April 19, 2010.

Total Quality Management (TQM), Management practices designed to improve the performance of organizational processes in business and industry. Based on concepts developed by statistician and management theorist W. Edwards Deming, TQM includes techniques for achieving efficiency, solving problems, imposing standardization and statistical control, and regulating design, housekeeping, and other aspects of business or production processes. *See also* International Organization for Standardization (ISO); Total Quality Control (TQC).

