

ARTIGO

Recebido em:
21/06/2015

Aceito em:
22/09/2015

Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação, v. 20, n. 44, p. 127-144, set./dez., 2015. ISSN 1518-2924. DOI: 10.5007/1518-2924.2015v20n44p127

A literatura internacional sobre *e-Science* nas bases de dados LISA e LISTA

The international literature on e-Science in the database LISA and LISTA

Maira Murrieta **COSTA**

Doutoranda em Ciência da Informação pela Universidade de Brasília. Analista de Ciência e Tecnologia do Ministério da Ciência, Tecnologia e Inovação - mairamurrieta@gmail.com

Murilo Bastos da **CUNHA**

Professor do Departamento de Ciência da Informação e Documentação da Universidade de Brasília - murilobc@unb.br

Resumo

Apresenta brevemente o conceito de *e-Science*. Trata-se de um estudo quantitativo que analisa as publicações sobre o tema *e-Science* na área da ciência da informação. O universo da pesquisa é formado por 202 registros de documentos das bases de dados LISA e LISTA. A análise dos dados revela um crescimento do volume de publicações sobre o tema e demonstra que dentre os 202 registros analisados, 172 são de artigos de periódicos. Ainda, confirma o fenômeno de que poucos autores publicam muito. A Inglaterra foi o país que mais se destacou na análise dos dados, visto que demonstrou pioneirismo em programas de governo para *e-Science*. Dentre as associações profissionais da ciência da informação, destaca-se a atuação da Association of Research Libraries. Ressalta-se que foram encontrados 3 artigos de autores brasileiros indexados sobre o tema.

Palavras-chave: *E-Science*. Ciberinfraestrutura. Dilúvio de dados. Quarto paradigma. Big data. Bibliometria. Cientometria.

Abstract

It briefly introduces the concept of e-Science. This is a quantitative study that analyzes the publications in Information Science dealing with the topic e-Science. The research sample consists of 202 records of documents retrieved from the "Library and Information Science" (LISA) and "Library and Information Science & Technology Abstracts" (LISTA) databases. Data analysis reveals a growth of literature dealing with the subject; and shows that among the 202 records analyzed, 172 of them were journal articles. The result also confirms the phenomenon that few authors publish much. Among the countries that are working with e-Science the United Kingdom is playing a major role. In terms of professional library and information associations highlight the Association of Research Libraries. The paper concludes that the UK stands out in the treatment of e-science by information science. It is noteworthy that were found three articles by Brazilian authors indexed on the subject.

Keywords: *E-Science. cyber infrastructure. Data deluge. Fourth paradigm. Data-intensive scientific discovery. Big data. Bibliometric study.*



v. 20, n. 44, 2015
p. 127-144
ISSN 1518-2924



Esta obra está licenciada sob uma [Licença Creative Commons](https://creativecommons.org/licenses/by/4.0/).

1 INTRODUÇÃO

A quantidade de publicações em cada campo do conhecimento, após a Segunda Guerra Mundial, cresceu exponencialmente, duplicando a cada dez ou quinze anos (SOLLA PRICE, 1976). Esse fenômeno deu origem à chamada *Big Science*. No Século XXI, a pesquisa colaborativa que produz grande volume de dados deu origem aos termos: *Big Data*, *cyberscience*, *cyber infrastructure* e *e-Science*.

Este artigo contextualiza o surgimento da *e-Science* e sua relação com a ciência da informação. Por se tratar de um tema novo para a ciência da informação, este artigo tem como objetivo geral: *conhecer a comunidade científica que produz literatura sobre e-Science na ciência da informação*. Seus objetivos específicos são sintetizados a seguir.

- Identificar quantitativamente os autores em ciência da informação que mais publicaram sobre *e-Science*.
- Identificar quantitativamente os periódicos que mais publicaram sobre a *e-Science*.
- Identificar as associações de classe em ciência da informação que já publicaram sobre *e-Science*.
- Identificar os eventos relacionados com a ciência da informação e *e-Science*.

Após buscas exaustivas nas bases de dados internacionais ERIC, LISA, LISTA e Web of Science, foi verificado que ainda não há na literatura um trabalho com características bibliométricas que identifique qual o autor da ciência da informação que mais tem publicado sobre o fenômeno da *e-science*, qual o periódico em ciência da informação que mais publicou sobre o tema, qual país tem despontado nessa área de pesquisa dentre outros aspectos de mapeamento da literatura sobre a *e-Science*.

A partir do exposto, infere-se que não há trabalho na literatura sobre ciência da informação atrelado ao fenômeno da *e-Science* que identifique a sua comunidade, os autores que mais publicam, bem como os países que produzem essa literatura. Ou seja, não há um mapeamento a respeito do tema que facilite a identificação dos canais de comunicação utilizados pela comunidade.

Compreender a comunidade de determinada área do conhecimento é essencial para identificar os grandes autores, as instituições que têm fomentado a pesquisa, a maturidade da área, as questões relevantes da comunidade, dentre outros. Para Vilan Filho (2010, p. 19), *“é importante entender a natureza específica de cada grupo produtor do conhecimento científico para entender os processos de produção do conhecimento destes grupos, possibilitando ações de incentivo e promoção da eficiência na pesquisa científica.”*

A comunidade de uma determinada área do conhecimento é formada por cursos de formação, associações de classe e a literatura sobre essa área. Nos cursos de formação encontram-se os professores e pesquisadores. Esses pesquisadores precisam publicar seus resultados de pesquisa. Apesar de diferentes áreas do conhecimento darem preferência a canais distintos para divulgarem seus resultados, a quantidade de artigos de periódicos é

normalmente usada para avaliar o pesquisador (MEADOWS, 1999; MUELLER, 2007). Na visão de Tenopir e King (2001, p. 15), “os periódicos científicos são geralmente tidos como de grande valor para os cientistas.”

A respeito do assunto, Ziman (1976) argumenta que o ato de publicar artigos em revistas científicas é exigido pelos pares como prova definitiva de efetiva atividade em pesquisa científica.

Ao considerar-se as observações de Ziman (1976), Meadows (1999) e Tenopir e King (2001), entende-se que a publicação de periódicos é relevante para o pesquisador independentemente de sua área de conhecimento. Nesse sentido, o presente artigo se propôs a estudar a literatura internacional sobre o fenômeno da *e-Science* na ciência da informação a partir de duas bases de dados relacionadas à ciência da informação: a *Library and Information Science Abstracts* (LISA) e a *Library and Information Science & Technology Abstracts* (LISTA).

2 A E-SCIENCE

São várias as denominações utilizadas para tratar as mudanças ocorridas na condução da ciência contemporânea, tais como a colaboração *on-line* da ciência, a infraestrutura de tecnologia de informação necessária para suportar essa colaboração, o grande volume de dados produzido, bem como a demanda de computação em *grid* para processar a análise desse grande volume de dados (JANKOWSKI, 2007). Dentre essas denominações, destacam-se os termos *e-Science* e *cyberinfrastructure*.

De acordo com Jankowski (2007), o termo *cyberinfrastructure* está enraizado com a iniciativa, dos cientistas americanos, no ano de 2003, em obter fontes de financiamento na National Science Foundation. Essa iniciativa culminou, em 2003, com a publicação do Atkins Report. Nas palavras do relatório: “(...) se a infraestrutura é necessária para a economia industrial, então... *cyberinfrastructure* é necessária para a economia do conhecimento.” (ATKINS, 2003, p. 5)

Por outro lado, o termo *e-Science* surge de iniciativas que emergem na Europa, particularmente no Reino Unido, quando John Taylor, então diretor geral do Escritório de Ciência e Tecnologia do Reino Unido, cunha o termo, no ano de 1999, durante o lançamento de um grande programa de financiamento (JANKOWSKI, 2007).

Para Hey e Trefethen (2003), Marcum e George (2010) e Gray (2007) a *e-Science* refere-se à coleção de instrumentos e tecnologias necessárias para apoiar a pesquisa científica do Século XXI – intrínseca à natureza colaborativa e multidisciplinar, bem como pelo grande volume de dados produzidos que precisam estar disponibilizados em rede.

Nesse sentido, Gray (2007, p. 17) afirmou que “*e-science* é o ponto onde a TI (tecnologia da informação) encontra os cientistas”. O fato é que autores como Hey e Trefethen (2003), bem como Gray (2007), Mayer-Schonberger e Cukier (2013) têm destacado a importância da tecnologia da informação na forma de se fazer ciência. Para esses autores, os desafios tecnológicos incluem a necessidade de melhor captar, analisar, modelar,

visualizar e preservar as informações científicas, tornando os sistemas de computação vitais para o moderno ambiente de pesquisa.

3 METODOLOGIA

Trata-se de uma pesquisa descritiva e de levantamento (*survey*), que utilizou a bibliometria, um método quantitativo baseado em análises estatísticas, para análise de dados. Portanto, a análise de dados assume um caráter quantitativo.

A pesquisa teve como fonte de dados os metadados dos artigos publicados nas bases de dados: *Library and Information Science Abstracts* (LISA) e *Library e Information Science & Technology Abstracts* (LISTA). O período de coleta desses metadados foi de 19/03/2013 a 19/06/2013.

A Base de Dados LISA foi selecionada por permitir acesso a artigos de periódicos, desde 1969 até a presente data, de mais 440 títulos de periódicos relacionados com Biblioteconomia e Ciência da Informação, conferindo exaustividade à busca.

A Base de Dados LISTA foi selecionada por indexar mais de 675 das principais revistas científicas, além de livros, relatórios de pesquisas e protocolos. O período de cobertura dessa base de dados remonta a meados dos anos 1960, o que propicia uma pesquisa exaustiva sobre o tema.

Em função dos índices de revocação e precisão no resultados das buscas realizadas aleatoriamente, optou-se na busca estruturada por utilizar apenas o termo *e-Science*, aumentando assim o índice de revocação. O período de cobertura do levantamento bibliográfico compreendeu os anos de 2003 a 2013. O resultado desse levantamento é observado na Tabela 1.

Tabela 1: Levantamento sobre e-Science

Nome da Base	2003/2004	2005/2006	2007/2008	2009/2010	2011/2012	Total de estudos
LISA	11	18	30	35	20	114
LISTA	10	22	44	51	19	148*

Fonte: Os autores - Dados coletados até 10/06/2013. *Dois registros são de 2013

O universo desta pesquisa é representado pelos artigos indexados nas bases de dados LISA e LISTA. Como no total de 262 registros bibliográficos havia 60 registros duplicados, optou-se por trabalhar com o universo da pesquisa, o que permitiu uma maior compreensão sobre a literatura publicada até maio de 2013. Assim, o universo desta pesquisa é representado por um total de 202 registros bibliográficos ($n = 202$), conforme ilustra a Figura 1.

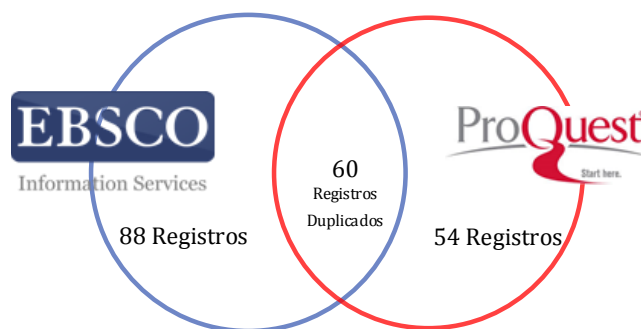


Figura 1: Representação do Universo de Pesquisa
Fonte: Os autores

3.1 Procedimentos metodológicos na análise bibliométrica

Após o download dos arquivos de dados (extensão .txt) do resultado da busca bibliográfica executada na LISA e na LISTA foi realizada a limpeza dos arquivos .txt separadamente, padronizando as *tags* de registro (REG), título do artigo (TI), autor (AU), periódico (PR), data (DT), palavras-chave (PC), organização (AD). Após a limpeza, os arquivos foram salvos em formato .csv, – um arquivo para os registros da LISTA e outro para os registros da LISA.

A etapa seguinte consistiu em copiar os dados do arquivo .csv da LISA para o arquivo .csv da LISTA. Conseqüentemente, os registros originários da LISA precisaram ser renumerados sequencialmente a partir do registro 148 (último da LISTA). Após a renumeração, salvou-se o arquivo unificado em formato .xls para permitir a ordenação pelo título de forma a identificar-se os registros duplicados. Após a conferência do título do artigo, foi conferido o nome do periódico e o do autor. Uma vez confirmada a semelhança, um dos registros foi considerado duplicado e eliminado do arquivo unificado. Eliminaram-se, assim, 60 registros. As etapas operacionais acima relatadas podem ser observadas na Figura 2.

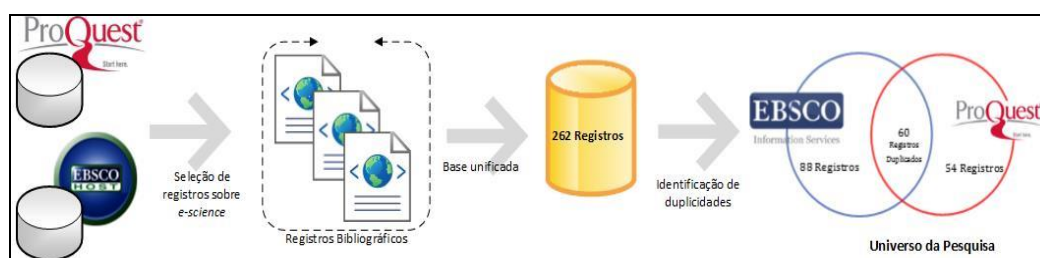


Figura 2: Procedimentos operacionais para a análise bibliométrica – Fase 1
Fonte: Os autores.

A quarta etapa teve início com a organização de uma tabela matriz que tivesse os dados de número de registro, título do artigo, autor (es), periódico, data e organização. Também foi elaborada uma planilha só com o número de registro e o autor, uma vez que um artigo pode ter mais de um autor. Conseqüentemente, foi elaborada mais uma planilha que continha o número do registro e a instituição de filiação do autor. Assim, para analisar os dados,

foram elaboradas inicialmente três planilhas – uma matriz, uma de autor(es) e outra de instituição do(s) autor(es). A estrutura completa dos metadados analisados é observada na Figura 3.

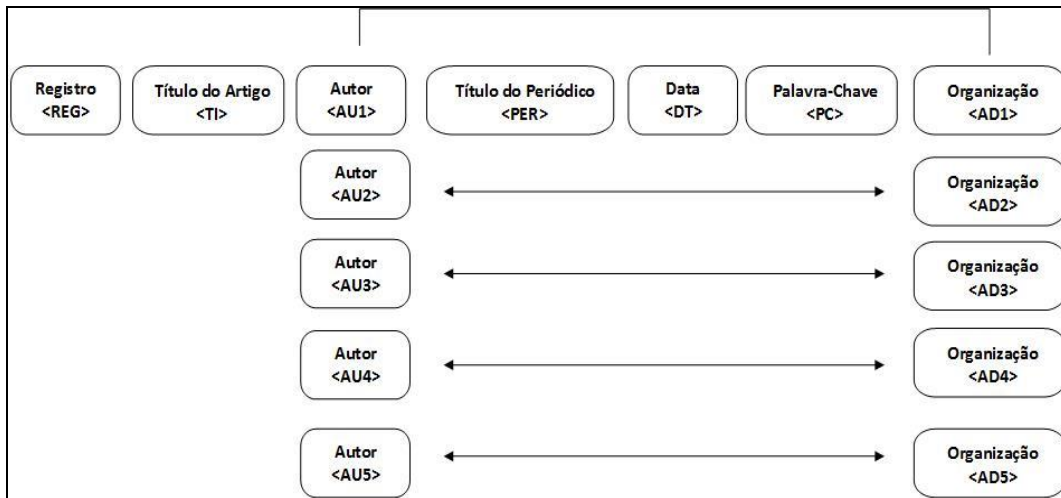


Figura 3: Estrutura dos metadados analisados

Fonte: Os autores.

A etapa seguinte foi realizada com o *software* SPSS que importou os dados da planilha matriz em formato .xls. Com o auxílio do SPSS foi possível identificar o número de publicações por ano e o periódico que mais publicou sobre o tema. A planilha de autores permitiu identificar os autores que mais produziram artigos sobre *e-Science*, enquanto a planilha de instituições permitiu identificar as instituições com destaque no tema. Todas as etapas até aqui mencionadas foram feitas com o auxílio do SPSS, porém, os gráficos foram elaborados no Excel.

A identificação das palavras-chaves mais usadas na indexação das bases foi realizada com o apoio do Excel e ilustrada em uma nuvem de *tags* desenvolvida no aplicativo *Wordle*,¹ disponível gratuitamente na internet.

Foram recuperados nas buscas 31 registros referentes a artigos publicados no período entre 2011 e 2013. Esses registros não foram levados em consideração no Gráfico 3 – referente a publicações sobre o tema por ano. Tais artigos foram desprezados em função dos respectivos anos ainda não terem sido completamente indexados pelas bases de dados quando da época do estudo. Merece ser comentado que a eliminação desses registros não interfere no quadro de autores que mais publicaram sobre o tema, tão pouco no quadro de periódicos que mais publicaram sobre *e-Science*.

3.2 Limitações da Pesquisa

Uma limitação da pesquisa é o fato de não ter sido considerada uma base de dados francófona como fonte de dados. Assim, autores franceses e seus respectivos periódicos certamente não serão encontrados no resultado deste artigo.

¹ Wordle – <http://wordle.net>

Também constitui uma limitação da pesquisa só terem sido analisadas as bases de dados relacionadas com a ciência da informação (LISA e LISTA), portanto os resultados desta pesquisa compreendem apenas uma pequena faceta do tema *e-Science*. Em face do exposto, a extrapolação de resultados deve ser feita com cautela no que diz respeito ao tema *e-Science* na ciência da informação. Apesar de ser um estudo quantitativo, os resultados aqui apresentados não podem ser extrapolados para o tema *e-Science* e sua relação com outras áreas do conhecimento.

4 ANÁLISE DOS DADOS

Os dados foram analisados com o apoio dos softwares Excel e SPSS. O SPSS foi utilizado para realizar as análises descritivas, calculando seus percentuais, percentuais válidos e valores acumulados. O Excel, por sua vez, foi utilizado para elaborar os gráficos. Já as nuvens de *tags* foram criadas com o aplicativo *Wordle*.

O Gráfico 1 ilustra os tipos de publicações indexados pelas bases de dados LISA e LISTA. Estão agrupados em **conference** os *proceedings*, *conferences reports* e *conferences*. O agrupamento **outros** é formado por editoriais, opiniões (*opinion*), cartas (*letters*) e um estudo de caso (*case study*) que não foi indexado como artigo.

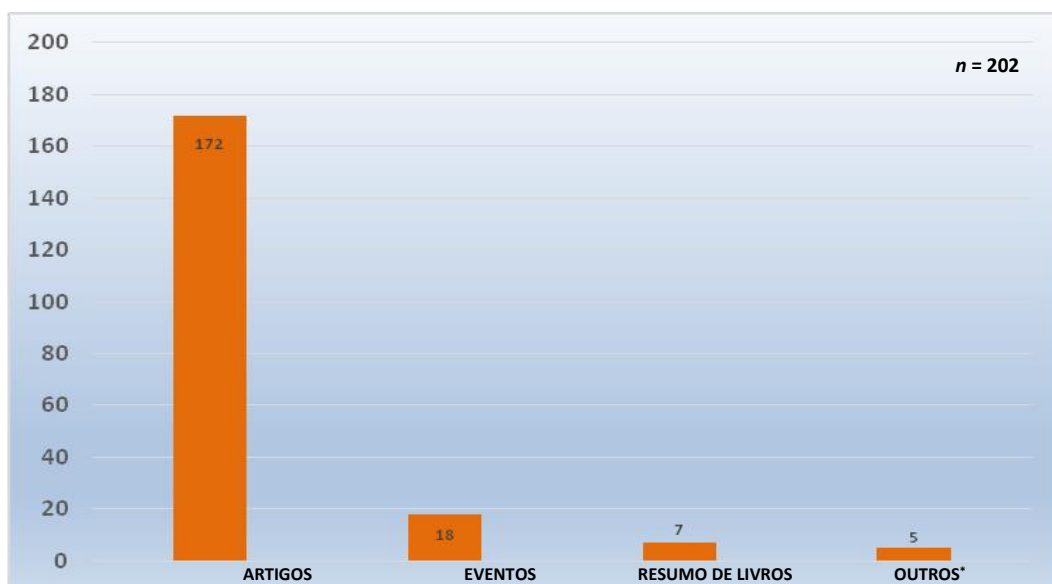


Gráfico 1: Tipos de documentos indexados pela LISA e LISTA (2003 - 2013)

Fonte: Os autores. *Outros (editorial, opinião, cartas)

A análise do Gráfico 1 indica que há uma preponderância de artigos científicos no conteúdo das bases de dados. Merece ser comentado que dos 7 resumos de livros, 6 tratavam do livro *"The data deluge: can libraries cope with e-science?"* - publicado em 2010 pelos editores Denna B. Marcum e Gerald George.

O Gráfico 2 demonstra que apenas 10% dos autores publicaram mais de um artigo sobre *e-Science* no período de 2003 a 2013. Ressalta-se que foram levados em consideração nesse gráfico apenas os artigos científicos, ou seja, foram excluídos da análise os documentos agrupados como *conference*, *book review* e outros.

O Gráfico confirma o fenômeno da **Lei de Lotka**, também conhecida como lei de frequência da produtividade científica, que considera que poucos autores produzem muito e, por outro lado, muitos autores produzem pouco.

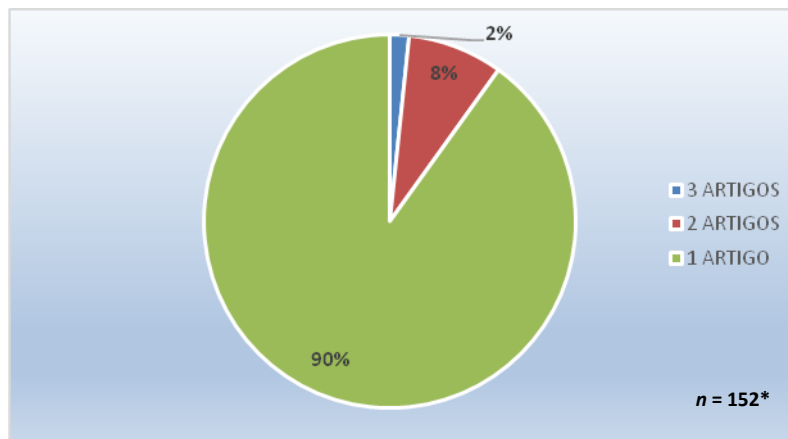


Gráfico 2: Distribuição da quantidade de autores versus artigos publicados sobre e-Science - Período (2003/2013)

Fonte: Os autores. *Dos 172 artigos, 20 não possuíam informação sobre o autor, logo, n = 152.

Com fundamento nos autores que informaram sua instituição de origem no artigo, é possível afirmar que apesar do termo *e-Science* ter surgido na Inglaterra, autores vinculados a instituições nos Estados Unidos já produziram mais sobre o tema, conforme pode ser observado no Gráfico 3, onde na sequência aparece a Inglaterra.

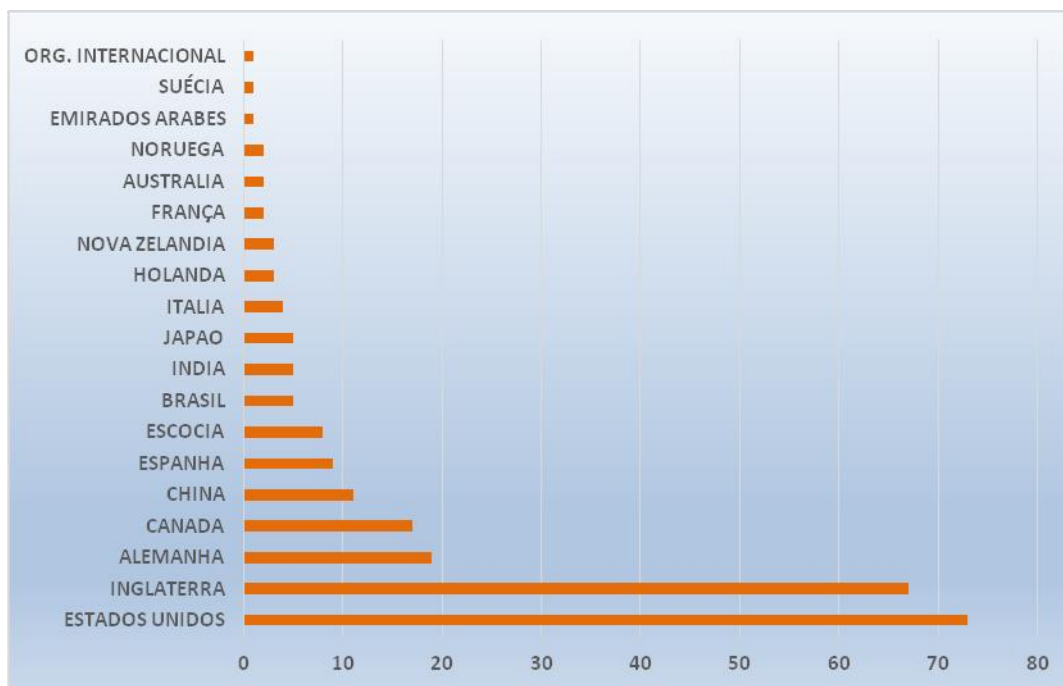


Gráfico 3: Distribuição dos autores por País (com fundamento na instituição informada no artigo)

Fonte: Os autores.

Há que se ressaltar, na leitura do Gráfico 3, que nem todos os autores informaram a sua instituição (um total de 73 registros), o que impossibilitou identificar o País de origem da produção. Por outro lado, uma minoria de registros (20 registros) não possuíam informação de autoria, logo também não possuíam informação da instituição. Incluem-se nestes casos alguns resumos de livros, cartas, dentre outros.

Ainda sobre o Gráfico 3, constata-se que os três maiores produtores de documentos sobre *e-Science* são membros do G8. Por outro lado, a produção da França, apesar de membro do G8, é menor que a do Brasil. Nesse aspecto, é importante lembrar que as bases de dados americanas pouco indexam documentos franceses. Assim, o Gráfico 3 deve ser analisado cuidadosamente, pois para evitar essas disparidades, essa pesquisa precisava ter incluído os registros da base de dados Francis.

Dentre os demais países, percebe-se a presença da produção de artigos entre os membros do BRICS², representando assim a emergência desses países tanto em termos industriais quanto no desenvolvimento científico.

Dentre as instituições, obtiveram destaque pelo número de autores e publicações, as relacionadas na Tabela 2.

² São membros do BRICS: Brasil, Rússia, Índia, China e África do Sul.

Tabela 2: Instituições em destaque

PAIS	INSTITUIÇÃO	OCORRÊNCIAS (n = 202)
Inglaterra	University of Oxford	13
EUA	Purdue University	10
Inglaterra	University of California	7
Inglaterra	University of Manchester	7
China	Chinese Academy of Sciences	6
Escócia	University of Edinburgh	6
Inglaterra	Arts and Humanities e-Science Support Centre	5
Inglaterra	Oxford eResearch Centre	5
EUA	University of Washington	5
EUA	Microsoft Corporation	4

Fonte: Os autores.

A partir da análise da Tabela 2, fica evidente a maturidade da Inglaterra no tratamento de dados oriundos da *e-Science*, pois na Tabela figuraram não apenas universidades, mas também centros especializados em *e-Science* e *e-Research*. O resultado não causa surpresa, afinal no ano 2000 a Inglaterra já possuía o *National e-Science Center*.

Outro centro da Inglaterra que apareceu na análise, mas em menor evidência, foi o *e-Science Centre*. Também apareceu o *e-Science Core Programme*. Esse programa é gerido pelo Conselho de Pesquisa em Ciências da Engenharia e Física, em nome das comunidades de todos os Conselhos de Pesquisa. Tem apoiado o desenvolvimento de tecnologias genéricas, como o software conhecido como *middleware* - necessário para permitir que diferentes recursos trabalhem de forma integrada por meio de redes, bem como criem *grids* computacionais.

O fato da Inglaterra, já em 2001, ter lançado, de forma pioneira, um programa, conferiu-lhe certa maturidade no tratamento de dados oriundos da *e-Science*. Tal afirmação se justifica pelo volume de trabalhos vinculados ao tema pelas Universidades de Oxford e Manchester. Também apresentam trabalhos significativos sobre o tema o Arts and Humanities e-Science Support Centre, o Oxford eResearch Centre, o National e-Science Center, o e-Science Centre, bem como o e-Science Core Programme.

Já no contexto dos Estados Unidos destacam-se os trabalhos vinculados às universidades de Purdue e de Washington. Além disso, percebe-se um interesse de grandes corporações, como a Microsoft, pelo tema, destacando-se o fato de Tony Hey, atual vice-presidente da área de pesquisa da Microsoft, ter sido o diretor do e-Science Core Programme no Reino Unido.

A única associação de classe que foi recuperada pelos registros da LISA e LISTA foi a Association of Research Libraries (ARL). Foi constatado que 7 publicações envolviam a ARL, são elas: 1) ARL launches e-Science agenda; 2) e-Science and data support service: a study of ARL members institutions; 3) e-Science in research libraries: new ARL reports to research libraries; 4) Reinventing librarianship: themes from ARL-CNI Forum; 5) ARL activities; 6) ARL e-science survey e 7) Libraries and changing research

practice: a report of the ARL-CNI Forum on e-research and cyberinfrastructure.

O número de instituições e sua distribuição por alguns dos países que se destacaram na análise dos dados pode ser observado no Gráfico 4.



Gráfico 4: Distribuição das instituições por países

Fonte: Os autores.

A análise do campo referente ao local de publicação permitiu fazer um ranking dos periódicos que mais publicaram sobre o tema *e-Science*. O ranking pode ser observado na Tabela 3.

Tabela 3: Periódicos que mais publicaram sobre e-Science

PERIÓDICO	n = 202			
	FREQ	%	% VÁLIDO	% ACUMULADO
Ariadne	10	5,0	5,0	5,0
IEEE Intelligent Systems	7	3,5	3,5	8,4
D-Lib Magazine	6	3,0	3,0	11,4
Information Services & Use	6	3,0	3,0	14,4
Library Hi Tech News	6	3,0	3,0	17,3
Information Today	5	2,5	2,5	19,8
International Journal of Web Services Research	5	2,5	2,5	22,3
Issues in Science & Technology Librarianship	5	2,5	2,5	24,8
Sci-Tech News	5	2,5	2,5	27,2
Electronic Library	4	2,0	2,0	29,2
Information, Communication & Society	4	2,0	2,0	31,2
Journal of Information Processing & Management	4	2,0	2,0	33,2
ABI Technik	3	1,5	1,5	34,7
ARL: A Bimonthly Report on Research Library Issues & Actions	3	1,5	1,5	36,1
Communications of the ACM	3	1,5	1,5	37,6
International Journal on Digital Libraries	3	1,5	1,5	39,1
Journal of Library Administration	3	1,5	1,5	40,6
Journal of the Medical Library Association	3	1,5	1,5	42,1

Library Hi Tech	3	1,5	1,5	43,6
-----------------	---	-----	-----	------

Fonte: Os autores com fundamento em dados da LISA e LISTA em junho de 2013.

A Tabela 3 revela que o periódico que mais publicou sobre o tema foi o *Ariadne*, periódico de origem inglesa voltado para profissionais da informação. Sua área de concentração é sobre bibliotecas digitais, abordando os temas de evolução dos serviços de informação e redes de informação.

Em segundo lugar aparece o periódico *IEEE Intelligent Systems*, periódico acadêmico voltado para engenheiros de software, designers, gerentes de sistemas de informação, engenheiros de conhecimento, pesquisadores e profissionais em áreas como finanças, manufatura, medicina, defesa e as ciências. A publicação é produzida e patrocinada pela *IEEE Computer Society*, seu patrocínio é oferecido pela *Association for the Advancement of Artificial Intelligence* (Associação para o Avanço da Inteligência Artificial).

Os demais periódicos estão voltados para a ciência da informação. Sendo que 5 deles são periódicos 'profissionais' – *Library Hi Tech News*, *Information Today*, *Communication of the ACM*, *Journal of the Medical Library Association*, *Library Hi Tech* e *ARL: Bi-monthly Report on Research Library Issues & Actions*, este último publicado pela *ARL*.

No que diz respeito aos eventos, a *IATUL³ Annual Conference Proceedings* foi a que mais teve ocorrência nas bases de dados. O último evento da *IATUL* foi realizado em junho de 2014 na Finlândia, sendo que o próximo será realizado em julho de 2015 na cidade de Hannover, na Alemanha. Além deste evento, foram identificados os seguintes: *eSciDoc*, *International Bielefeld Conference*, *NFAIS⁴ Annual Conference* e a *ICSTI⁵ Annual Conference*.

Para saber se o tema *e-Science* está crescendo dentro da Ciência da Informação, os registros da base de dados foram organizados por ano, de forma a identificar se ocorre um crescimento anual da produção literária da área. O Gráfico 5 ilustra o crescimento da produção até o ano de 2010.

³ Internacional Association of Scientific and Technological Libraries.

⁴ National Federation of Advanced Information Services.

⁵ National Federation of Advanced Information Services.

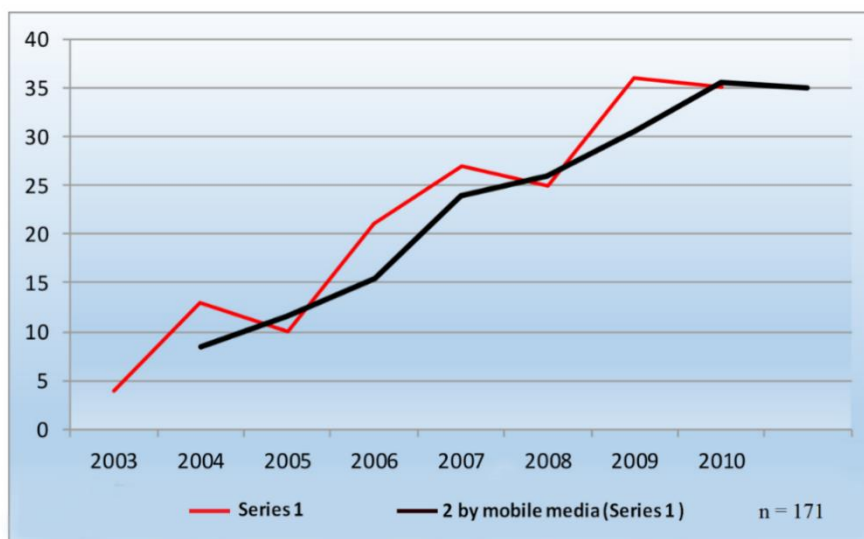


Gráfico 5: Produção de Artigos por Ano

Fonte: Os autores com fundamento em dados da LISA e LISTA em junho de 2013.

Foram desconsiderados para o Gráfico 5 a produção referente a 2011⁶ e 2012⁷ em função do período de atualização das bases de dados. Ou seja, há um entendimento de que o volume de publicações indexados em 2011 e 2012 ainda não reflete a realidade em função do gap que ocorre entre o período de publicação do volume de um periódico e o tempo que as bases de dados levam para incluí-lo e indexá-lo.

Merece ser ressaltado que o Gráfico 5 demonstra um leve decréscimo no ano de 2010 em relação a 2009 e uma posterior estagnação no volume de publicações. Uma possível explicação para a leve queda do gráfico é a criação de periódicos específicos sobre *e-Science*, que, dada a sua recente criação, ainda não foram indexados pelas bases de dados LISA e LISTA. Como exemplo, cita-se os periódicos *International Journal of Digital Curation* e o *Journal of eScience Librarianship*, criados em 2006 e 2012, respectivamente.

Para identificar qual seria o melhor termo para se realizar uma busca sobre o tema, foi realizada uma análise das palavras-chave indexadas nos 202 registros. A Tabela 4 ilustra os 20 termos mais indexados.

⁶ Em junho de 2013, as bases de dados possuíam apenas 17 registros referentes a 2011.

⁷ Em junho de 2013, as bases de dados possuíam apenas 12 registros referentes a 2012.

Tabela 4: Vinte palavras-chave mais indexadas

PALAVRAS-CHAVE	FREQ	%	% VALIDA	% ACUMULADA
Science	30	3,4	3,4	3,4
Digital-libraries	22	2,5	2,5	5,9
Electronic publishing	22	2,5	2,5	8,4
Research	21	2,4	2,4	10,8
Electronic-media	13	1,5	1,5	12,3
Librarians	12	1,4	1,4	13,7
Information resources management	10	1,1	1,1	14,8
Computer Network Resources	9	1	1	15,8
Conferences	9	1	1	16,9
Libraries	9	1	1	17,9
Open access	9	1	1	18,9
Scholarly publishing	9	1	1	20
UK	9	1	1	21
e-Science	8	0,9	0,9	21,9
Research libraries	8	0,9	0,9	22,8
Scholarly publications	8	0,9	0,9	23,7
Collaboration	7	0,8	0,8	24,5
Information resources	7	0,8	0,8	25,3
Information science	7	0,8	0,8	26,1
Information services	7	0,8	0,8	26,9

Fonte: Os autores com fundamento em dados da LISA e LISTA em junho de 2013.

Interessante destacar que a análise dos 114 registros recuperados na base de dados LISA revela que o termo *e-Science* não foi usado como indexador (palavra-chave) em nenhum dos registros, conforme ilustrado na Figura 4.



Figura 4: Nuvem de Tags dos termos indexados na base de dados LISA

Fonte: Os autores.

Percebe-se que as palavras-chave utilizadas na LISA são mais genéricas, por exemplo: ciência, publicação eletrônica, bibliotecas digitais,

publicação acadêmica dentre outros. Esses termos certamente aumentariam a revocação da busca, mais diminuiriam sua precisão. Infere-se que há uma possibilidade do termo não ter sido utilizado por ser muito novo e não constar ainda nos vocabulários controlados.

Interessante observar o uso do termo UK (United Kingdom) como palavra-chave. Infere-se que esse termo revela o lastro da Inglaterra em projetos na área de *e-Science*.

Por outro lado, em um mesmo período de tempo, verificou-se que a base de dados LISTA usou o termo *e-Science* como indexador, conforme ilustra a Figura 5.

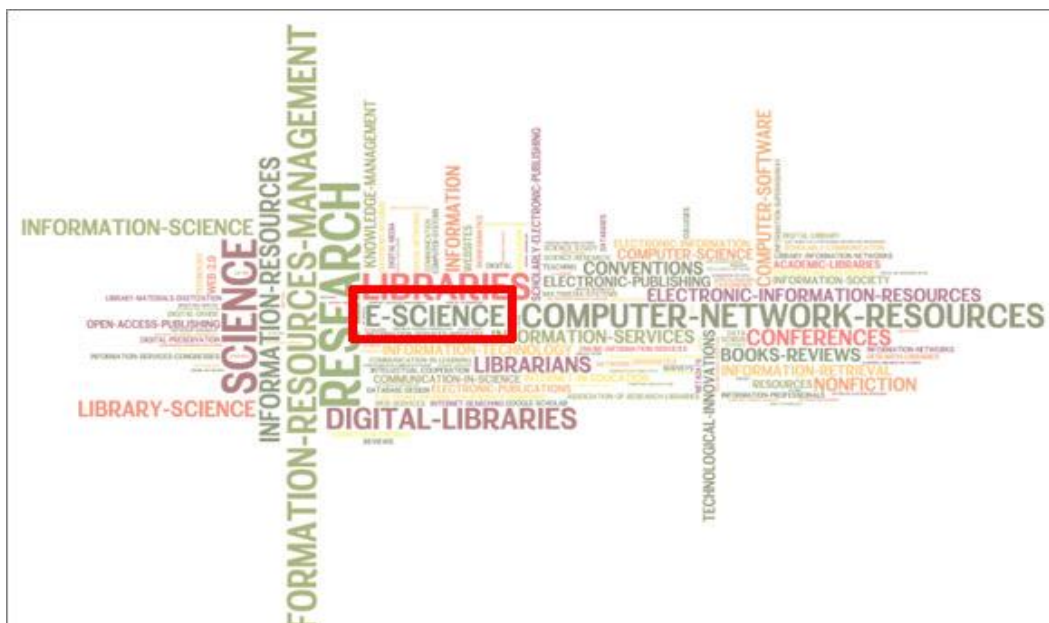


Figura 5: Nuvem de *Tags* dos termos indexados na base de dados LISTA

Fonte: Os autores.

Na LISTA também é observado o uso de termos mais genéricos como palavras-chave. A exemplo cita-se: ciência da informação, bibliotecas digitais, bibliotecários. Porém, o termo *e-Science* aparece com razoável destaque na nuvem, o que indica a quantidade de vezes que foi utilizado.

Para compreender de fato porque a LISA não utilizou o termo e a LISTA sim, é necessário fazer uma entrevista com os responsáveis pela política de indexação das respectivas bases de dados. Qualquer inferência mais precisa pode ser leviana.

O que se pode afirmar, com fundamento nos dados, é que a busca realizada pela pesquisadora teve um retorno de 262 registros (202 excluindo os registros duplicados) pelo fato de não ter se delimitado o campo para a busca. Se a delimitação tivesse sido feita na palavra-chave, o resultado seria frustrante, pois o termo *e-Science* não foi indexado pela LISA e não teve absoluta expressão na busca realizada na LISTA. Entende-se que os registros foram recuperados pelo termo estar no título ou no resumo.

Merece ser comentado que os termos: dilúvio de dados, ciberinfraestrutura e *big data* que seriam mais precisos, também não foram

utilizados nas respectivas bases de dados. Nesse cenário, é interessante lembrar que o termo *cyberinfrastructure* foi cunhado pela National Science Foundation. Assim, é possível inferir que o termo criado no Reino Unido (*e-Science*) tenha ganhado mais adeptos, levando-o a ser um termo indexador.

Conclui-se, portanto, que ao se realizar buscas relacionadas ao fenômeno da *e-Science* é preciso aumentar o índice de revocação, pois uma busca precisa certamente levará a um resultado frustrante.

A análise dos dados revelou que dois são os autores que mais publicaram sobre *e-Science*, sendo que todos eles publicaram 3 artigos no período de 2003-2011. A lista dos autores pode ser observada no Tabela 5.

Tabela 5: Autores que mais publicaram sobre *e-Science*

AUTOR	QTD. TOTAL	QTD. ARTIGOS	QTD. OUTROS
Blanke, Tobias	5	3	2
Mullins, James L.	4	3	1
Aschenbrenner, Andreas	3	3	--
Dovey, Matthew J.	3	3	--
Gore, Sally A.	3	3	--
Guo Jing	3	3	--
Hai Zhuge	3	3	--
Paterson, Lorraine	3	3	--
Becker, Carolin	3	2	1
Kallerbon, Reiner	3	2	1
Rusnak, Ute	3	2	1

Fonte: Os autores.

Interessante observar que em pesquisa no Google Acadêmico sobre *e-Science* e bibliotecas, os artigos em destaque são de Tony Hey, o primeiro *The data deluge: an e-science perspective*, citado por 367 e; o segundo: *E-science and its implications for the library community*, citado por 64. Esse segundo artigo também é recomendado na bibliografia sobre o tema elaborada por Szigeti e Wheeler (2011). Porém, na busca realizada na LISA e no LISTA o Tony Hey aparece como autor de apenas 2⁸ artigos, por isso não foi representado na Tabela 5.

O único autor que figura na lista dos que mais publicaram e está relacionado na bibliografia elaborada por Szigeti e Wheeler (2011) é o doutor James L. Mullins⁹, professor da Purdue University. Os resultados apresentados na Tabela 5 confirmam que nem sempre o autor que mais publica é o mais citado.

⁸ (i) HEY, Tony. Why engage in e-science? Library & Information Update. 2004.; (ii) HEY, Tony; HEY, Jessie. E-science and its implications for the library community. Library Hi-Tech. 2006.

⁹ Enabling international access to scientific data sets: creation of the distributed data curation center.

5 CONCLUSÃO

As contribuições da Ciência da Informação para a *e-Science* ainda são incipientes. Tal afirmativa se faz possível em função da quantidade de trabalhos publicados na base de dados *Library Information Science Technology Abstracts* (LISTA), bem como na base de dados *Library Information Science Abstracts* (LISA). No entanto, os dados revelam uma curva crescente no volume de trabalhos sobre o tema.

O Reino Unido se destaca no que diz respeito a centros de tratamento de dados da *e-Science* em relação aos outros países. Atribui-se esse destaque ao pioneirismo da Inglaterra no tratamento desses dados.

A única associação profissional que emergiu da análise de dados foi a Association of Research Libraries, o que denota um campo ainda desconhecido para os profissionais da informação. Para atualizar-se recomenda-se a participação no IATUL, bem como a leitura dos títulos de periódicos *Ariadne* e *IEEE Intelligent Systems*.

Os autores que se destacam em termos de publicação sobre o tema são Tobias Blanke e James L. Mullins. No entanto, recomenda-se a leitura dos trabalhos de Tony Hey, autor pioneiro na área de *e-Science*.

REFERÊNCIAS

ATKINS, Daniel E. *et al.* **Revolutionizing science and engineering through cyber infrastructure**: report of the National Science Foundation Blue-Ribbon Advisory Panel on Cyber infrastructure". Washington, DC, January, 2003, <http://www.nsf.gov/cise/sci/reports/atkins.pdf>>. Acesso em: 30 ago. 2012.

GRAY, Jim. **eScience: a transformed scientific method**. Palestra apresentada no Conselho Nacional de Pesquisa dos Estados Unidos (NRC-CSTB), Mountain View, Califórnia, 11 janeiro de 2007. Arquivo disponível em <http://research.microsoft.com/en-us/um/people/gray/talks/NRC-CSTB_eScience.ppt>. Acesso em 30 ago. 2012.

HEY, Tony; TREFETHEN, Anne. E-Science and its implications. **Philosophical Transactions of the Royal Society (A)**, v. 15, n. 361, p. 1809-1825, June, 2003.

JANKOWSKI, Nicholas W. Exploring e-science: an introduction. **Journal of Computer-Mediated Communication**, v. 12, n. 2, p. 549-562, 2007.

MARCUM, Deanna B.; GEORGE, Gerald (Ed.). **The data deluge**: can libraries cope with e-science? Santa Barbara, California: Libraries Unlimited, 2010. 139 p.

MAYER-SCHÖNBERGER; Viktor; CUKIER, Kenneth. **Big data**: como extrair volume, variedade, velocidade e valor da avalanche de informação cotidiana. Rio de Janeiro, Elsevier, 2013. 256 p.

MEADOWS, A. J. **A comunicação científica**. Brasília: Briquet de Lemos, 1999. 268 p.

MUELLER, Suzana Pinheiro Machado. Literatura científica, comunicação científica e ciência da informação. In: TOUTAIN, Lídia Maria Batista Brandão (Org.). **Para entender a ciência da informação**. Salvador: EDUFBA, 2007. p. 125-144.

SOLLA PRICE, Derek J. **O desenvolvimento da ciência**: análise histórica, filosófica, sociológica e econômica. Rio de Janeiro: Livros Técnicos e Científicos, 1976.

SZIGETI, Kathy; WHEELER, Kathy. Science and technology resources on the Internet: essential readings in e-Science. **Issues in Science and Technology Librarianship**, v. 64, 2011. Disponível em:

<<http://dx.doi.org/10.5062/F400001J>>. Acesso em : 30 abr. 2014.

TENOPIR, Carol; KING, Donald W. A importância dos periódicos para o trabalho científico. **Revista de Biblioteconomia de Brasília**, v.25, n. 1, p. 15-26, 2001.

VILLAN FILHO, Jayme Leiro. **Autoria múltipla em artigos de periódicos científicos nas áreas de informação do Brasil**. Brasília, Universidade de Brasília, Programa de Pós Graduação em Ciência da Informação, 2010. [Tese].

ZIMAN, John. **O conhecimento confiável**. Campinas: Papyrus, 1976.

Editor do artigo: Adilson Luiz Pinto