ADADOSDEPESQUISACURADORIA DIGITALESCIENCECIÊNCIAABERTA REPOSITÓRIOSCIENTÍFICOSLIMPE ZADEDADOSQUALIDADEDEDADO SMÉTODOSDELIMPEZADEDADOSI NTERNETDASCOISASEFICIÊNCIAE NERGÉTICATECNOLOGIASDEINFO RMAÇÃOECOMUNICAÇÃOANÁLISE DEREDESSOCIAISTOINTERNACION ALIZAÇÃOACADÊMICACOMUNICA ÇÃOCIENTÍFICAAULAINVERTIDATE CNOLOGIAINOVAÇÃOEMEDUCAÇ ÃOPROCESSOSDEENSINOAPRENDI ZAGEMIDIOMAINGLÊSINOVAÇÃOE MAPRENDIZAGEMOBJETOSDEAPR ENZDIZAGEMMETODOLOGÍADECZ

NOVASPRÁTICASEMINFORMAÇÃOECONHECIMENTO

VOLUME5NÚMERO2JULDEZ2016

ISSN:2237-826X

www.revistas.ufpr.br/atoz

Publicação

Universidade Federal do Paraná
Mestrado Interdisciplinar em Ciência, Gestão e Tecnologia da

AtoZ: novas práticas em informação e conhecimento
revistas.ufpr.br/atoz

Universidade Federal do Paraná
Setor de Ciências Sociais Aplicadas
Programa de Pós-Graduação em Ciência, Gestão e Tecnologia da Informação
Av. Prefeito Lothário Meissner, 632 - Campus III
Jardim Botânico
Curitiba – PR, Brasil
80210-170
Fone: *+55*(41)3360-4389
Fax: *+55*(41)3336-4471
E-mail: revistaatoz@ufpr.br
URL: http://revistas.ufpr.br/atoz

Periodicidade: Semestral
ISSN: 2237-826X
Diretrizes para autores: http://revistas.ufpr.br/atoz/about/submissions#authorGuidelines

Qualis/Capes
B4 - Planejamento Urbano e Regional/Demografia; Saúde Coletiva / B5 - Ciências Sociais Aplicadas I; Arquitetura e Urbanismo

Indexada/registrada em
Directory of Open Access Journals (DOAJ); Sumários.org; Google Acadêmico; LivRe! Portal para periódicos de livre acesso na Internet; InfoBCI; Latindex Catálogo; Bielefeld Academic Search Engine; INFOBILA: Información Bibliotecológica Latinoamericana; REDIB (Red Iberoamericana de Innovación y Conocimiento Científico)

Todo o conteúdo da Revista (incluindo-se instruções, política editorial e modelos) está sob uma licença Creative Commons Atribuição-NãoComercial-Compartilhalgual 3.0 Não Adaptada.

Ao serem publicados por esta Revista, os artigos são de livre uso em ambientes educacionais, de pesquisa e não comerciais, com atribuição de autoria obrigatória.

O ©***copyright*** dos artigos e da entrevista pertence aos respectivos autores/entrevistados com cessão de direitos para a AtoZ no que diz respeito à inclusão do material publicado (revisado por pares/pós-print) em sistemas/ferramentas de indexação, agregadores ou curadores de conteúdo. Os autores têm permissão e são encorajados a depositar seus artigos em páginas pessoais, repositórios e/ou portais institucionais antes (pré-print) e após (pós-print) a publicação na AtoZ. Solicita-se apenas que, quando possível, a referência bibliográfica (incluindo o link/URL do artigo) seja elaborada com base na publicação na AtoZ: novas práticas em informação e conhecimento.

Comitê Editorial
Dra. Patrícia Zeni Marchiori, Universidade Federal do Paraná (UFPR), Grupo Metodologias para Gestão da Informação UFPR/CNPq, Brasil
Msc. Eduardo Michelotti Bettoni, Observatórios Sesi/Senai/IEL, Grupo Metodologias para Gestão da Informação UFPR/CNPq, Brasil
Msc. Andre Luiz Appel, Universidade Federal do Rio de Janeiro (UFRJ), Grupo Metodologias para Gestão da Informação UFPR/CNPq, Brasil
Dra. Helena Nunes Silva, Universidade Federal do Paraná (UFPR), Grupo Metodologias para Gestão da Informação UFPR/CNPq, Brasil
Dra. Denise Fukumi Tsunoda, Universidade Federal do Paraná (UFPR), Grupo Metodologias para Gestão da Informação UFPR/CNPq, Brasil

Conselho Consultivo
Dra. Ana Esmeralda Carelli, Universidade Estadual de Londrina - UEL, Brasil
Msc. Augusto José Waszczynskyj Antunes das Neves, Universidade Federal do Paraná - UFPR, Brasil
Dra. Avanilde Kemczinski, Universidade do Estado de Santa Catarina - UDESC, Brasil
Dr. Carlos Olavo Quandt, Pontifícia Universidade Católica do Paraná - PUC PR, Brasil
Dra. Cassandra Ribeiro Joye, Instituto Federal de Educação, Ciência e Tecnologia do Ceará - IFCE, Brasil
Dra. Cláudia Regina Z. Bomfá, Universidade Federal de Santa Maria - UFSM, Brasil
Dr. Claudio Cesar de Sá, Universidade do Estado de Santa Catarina - UDESC, Brasil
Dr. Daniel Cebrian Robles, Consultor independente, Espanha
Dra. Deborah Ribeiro Carvalho, Pontifícia Universidade Católica do Parana - PUCPR, Brasil
Dra. Faimara do Rocio Strauhs, Universidade Tecnológica Federal do Paraná - UTFPR, Brasil
Dr. Filiberto Felipe Martínez Arellano, Universidad Nacional Autónoma de México - UNAM, México
Dr. Francisco José Ruiz Rey, Universidad de Málaga - UMA, Espanha
Msc. Frank Coelho de Alcântara, Universidade Positivo - UP, Brasil
Dr. Francisco José Ruiz Rey, Universidad de Málaga - UMA, Espanha
Msc. Frank Coelho de Alcântara, Universidade Positivo - UP, Brasil
Dra. Isabela Gasparini, Universidade do Estado de Santa Catarina - UDESC, Brasil
Dr. Jamerson Viegas Queiroz, Universidade Federal do Rio Grande do Norte - UFRN, Brasil
Dra. Janine Kniess, Universidade do Estado de Santa Catarina - UDESC, Brasil
Dr. José Barata Oliveira, Instituto de Desenvolvimento de Novas Tecnologias - UNINOVA, Portugal
Dr. Juan José Monedero Moya, Universidad de Málaga - UMA, Espanha
Dra. Lídia de Jesus Oliveira Loureiro da Silva, Universidade de Aveiro - UA, Portugal
Dra. Lucila Pérez Cascante, Universidad Casa Grande - UCG, Equador
Dra. Maria Cristina Vieira de Freitas, Universidade de Coimbra - UC, Portugal
Dra. Maria da Graça de Melo Simões, Universidade de Coimbra - UC, Portugal
Dr. Maria do Carmo Duarte Freitas, Universidade Federal do Paraná, Brasil

NOVAS PRÁTICAS EM INFORMAÇÃO E CONHECIMENTO

Dra. María Gladys Ceretta Soria, Universidad de la República - UdelaR, Uruguai
Dra. Maria Salet Ferreira Novellino, Escola Nacional de Ciências Estatísticas - IBGE, Brasil
Dr. Mauro José Belli, Universidade Federal do Paraná - UFPR, Brasil
Msc. Murilo Artur Araújo da Silveira, Universidade Federal de Pernambuco - UFPE, Brasil
Msc. Victor Marcos Ferracutti, Universidad Nacional del Sur - UNS, Argentina

Editores de Seção - Expediente
Msc. Andre Luiz Appel, Grupo de Pesquisa UFPR/CNPq - Metodologias para Gestão da Informação, Brasil
Msc. Eduardo Michelotti Bettoni, Grupo de Pesquisa UFPR/CNPq - Metodologias para Gestão da Informação, Brasil

Editores de Seção - Editorial
Dra. Patricia Zeni Marchiori, Universidade Federal do Paraná - UFPR, Brasil

Editores de Seção - Entrevistas
Msc. Eduardo Michelotti Bettoni, Grupo de Pesquisa UFPR/CNPq - Metodologias para Gestão da Informação, Brasil

Editores de Seção - Artigos
Dra. Patricia Zeni Marchiori, Universidade Federal do Paraná - UFPR, Brasil
Msc. Eduardo Michelotti Bettoni, Grupo de Pesquisa UFPR/CNPq - Metodologias para Gestão da Informação, Brasil

Editores de Seção - Short Papers
Dra. Patricia Zeni Marchiori, Universidade Federal do Paraná - UFPR, Brasil
Msc. Andre Luiz Appel, Grupo de Pesquisa UFPR/CNPq - Metodologias para Gestão da Informação, Brasil

Editores de Leiaute
Msc. Andre Luiz Appel, Grupo de Pesquisa UFPR/CNPq - Metodologias para Gestão da Informação, Brasil
Msc. Eduardo Michelotti Bettoni, Grupo de Pesquisa UFPR/CNPq - Metodologias para Gestão da Informação, Brasil
Marcelo Batista de Carvalho, Bolsista UFPR/TN, Grupo de Pesquisa UFPR/CNPq - Metodologias para Gestão da Informação, Brasil

Apoio técnico
Biblioteca Digital de Periódicos (BDP), UFPR, Brasil

Capa
Marcelo Batista de Carvalho, Bolsista UFPR/TN, Grupo de Pesquisa UFPR/CNPq - Metodologias para Gestão da Informação, Brasil

Projeto gráfico
Grupo de Pesquisa UFPR/CNPq - Metodologias para Gestão da Informação

Avaliadores da edição (v. 5 n. 2)
Dr. Adriano Heemann, Universidade Federal do Paraná - UFPR, Brasil
Msc. Andre Appel, Grupo de Pesquisa UFPR/CNPq - Metodologias para Gestão da Informação, Brasil
Msc. Augusto José Antunes das Neves, Universidade Federal do Paraná - UFPR, Brasil
Dr. Cícero Bezerra, Universidade Federal do Paraná - UFPR, Brasil
Dr. Daniel Cebrian Robles, Consultor independente, Espanha
Dra. Deborah Carvalho, Pontifícia Universidade Católica do Parana - PUCPR, Brasil
Dra. Denise Fukumi Tsunoda, Universidade Federal do Paraná - UFPR, Brasil
Msc. Eduardo Bettoni, Grupo de Pesquisa UFPR/CNPq - Metodologias para Gestão da Informação, Brasil
Msc. Frank Coelho de Alcântara, Universidade Positivo - UP, Brasil
Dra. Janine Kniess, Universidade do Estado de Santa Catarina - UDESC, Brasil
Profa. Lídia Oliveira Loureiro da Silva, Universidade de Aveiro - UA, Portugal
Dra. Patricia Marchiori, Universidade Federal do Paraná - UFPR, Brasil

DOI da edição (v. 5 n. 2)
10.5380/atoz/v5i2

AtoZ : Novas Práticas em Informação e Conhecimento. — Vol. 5, n. 2 (jul./dez. 2016)- . — Curitiba : Universidade Federal do Paraná, Programa de Pós-Graduação em Ciência, Gestão e Tecnologia da Informaçãoo, 2016- .

v.

Semestral.
Publicação online: <http://revistas.ufpr.br/atoz>
ISSN 2237-826X

1. Comunicação científica – Periódico. 2. Informação – Periódico. 3. Conhecimento – Periódico.
I. Programa de Pós-Graduação em Ciência, Gestão e Tecnologia da Informação. II. Universidade Federal do Paraná.

CDD 001(8162)

ISSN: 2237-826X
EDITORIAL

Dados e informação: significados e engajamento

Data and information: meanings and engagement

Patricia Zeni Marchiori[1], Andre Luiz Appel[2], Eduardo Michelotti Bettoni[2]

[1]*Universidade Federal do Paraná - UFPR, Curitiba, PR, Brasil*
[2]*Grupo de Pesquisa UFPR/CNPq - Metodologias para Gestão da Informação, Curitiba, PR, Brasil*

Autor para correspondência/*Mail to:* Patricia Zeni Marchiori (editoratoz@gmail.com)

Copyright © 2016 Marchiori, Appel & Bettoni. Todo o conteúdo da Revista (incluindo-se instruções, política editorial e modelos) está sob uma licença Creative Commons Atribuição-NãoComercial-CompartilhaIgual 3.0 Não Adaptada. Ao serem publicados por esta Revista, os artigos são de livre uso em ambientes educacionais, de pesquisa e não comerciais, com atribuição de autoria obrigatória. Mais informações em http://revistas.ufpr.br/atoz/about/submissions#copyrightNotice.

Este volume 5/número 2 da AtoZ apresenta quatro artigos e um *short paper*, além da tradicional entrevista com especialistas. A tônica dos trabalhos publicados segue uma tendência já percebida na dinâmica de recebimento, avaliação e eventual publicação de investigações: em maior ou menor grau os autores buscam extrair sentido e testar a efetividade de estratégias de engajamento de públicos-alvo diferenciados em distintos processos ligados à dados e à informação. A pesquisa de Silva e Martins, no âmbito da Universidade Federal de Goiás (UFG), descreve e analisa a realidade e a potencialidade da coautoria e internacionalização da produção científica na Instituição, que representa - em última instância - o impacto das pesquisas da UFG em audiências e parcerias externas. O engajamento, via tecnologia educativa, está presente na experiência de motivação, baseada no uso da "aula invertida", relatada por Ruiz-Robles, junto a estudantes de inglês; e, no desenho, implantação e avaliação de cursos online baseados em objetos de aprendizagem, cuja experiência é relatada no *short paper* assinado por Zambrano e demais coautores.

Quanto ao papel das tecnologias para auxiliar na exploração de significados, Almeida e colaboradores exploram condições tecnológicas para aumentar a qualidade de volumes robustos de dados, assim como Sonego, Marcelino e Gruber verificam - via conexões da literatura - relações pervasivas da "Internet das Coisas" (IoT) com uma área de potencial aplicação (eficiência energética).

Todos os trabalhos publicados nesta edição (e em outras) da AtoZ apresentam resultados provenientes de coleta e tratamento de dados em distintos vieses metodológicos experimentais, quali-quantitativos ou de análise de redes sociais, por exemplo. O desafio de se armazenar, tratar e, especialmente, disponibilizar ao público os dados de pesquisa é o tema da entrevista gentilmente cedida pelos doutores Luis Fernando Sayão e Luana Farias Sales, os quais apresentam e detalham o estado da arte neste tema e os possíveis caminhos para os autores, gestores/curadores dos repositórios e os usuários interessados em reutilizar dados.

Foi também no decorrer da elaboração desse número, entre julho e dezembro de 2016, que o Projeto elaborado pela equipe gestora da AtoZ - e aprovado no Edital de Apoio à Editoração e Publicação de Periódicos Científicos (PRPPG/UFPR, 2016) - finalizou seus objetivos. Constituiu-se uma base de referências do material publicado na AtoZ (desde o número de estréia até o v. 5 n.1.), definindo-se e gerando-se indicadores de desempenho. Os resultados permitiram verificar que a base de dados de citações tem efetivo potencial para subsidiar indicadores e melhor compreender as relações da Revista com outras fontes de informação na área, o que permitirá - a médio prazo - avaliar a situação da AtoZ no macro cenário das revistas acadêmico-científicas. Ao se definir um conjunto de métricas de avaliação e testá-las resulta - de imediato - em apoio ao processo decisório do Corpo Editorial, e, potencialmente, na exploração da metodologia proposta por outros editores. O Projeto está na fase de submissão de relatório final e prestação de contas, motivo pelo qual o detalhamento dos resultados será objeto do próximo Editorial da Revista. Contudo, a trajetória metodológica, que inclui a especificação das fases de execução, se encontra disponível para consulta na base e-Lis, no seguinte link: http://eprints.rclis.org/30525/.

ISSN: 2237-826X
ENTREVISTA | INTERVIEW

Curadoria digital e dados de pesquisa

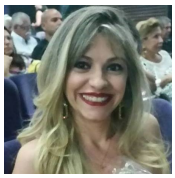
Digital curation and research data

Luís Fernando Sayão[1], Luana Farias Sales[1]

[1]*Comissão Nacional de Energia Nuclear - CNEN, Rio de Janeiro, RJ, Brasil*

Autor para correspondência/*Mail to:* Luana Ferreira Sales (lsales@ien.gov.br)

Luís Fernando Sayão possui graduação em Física pela Universidade Federal do Rio de Janeiro (1978), mestrado em Ciência da Informação pela Universidade Federal do Rio de Janeiro/Instituto Brasileiro de Informação em Ciência e Tecnologia (UFRJ/IBICT) e doutorado em Ciência da Informação pela UFRJ/IBICT (1994). Trabalha desde 1980 na Comissão Nacional de Energia Nuclear onde já exerceu os cargos de: chefe do Centro de Informações Nucleares (CIN); chefe da Divisão de Tecnologia da Informação; representante do Brasil no INIS - International Nuclear Information System (AIEA/ONU); coordenador-geral da RRIAN - Red Regional de Información en el Área Nuclear. É conselheiro do CONARQ - Conselho Nacional de Arquivos, membro do Câmara Técnica de Documentos Eletrônicos do CONARQ; docente do Programa de Pós-Graduação em Biblioteconomia da UNIRIO - Universidade Federal do Estado do Rio de Janeiro e do Programa de Pós-Graduação em Memória e Acervos da Fundação Casa de Rui Barbosa. Foi membro do Comitê Técnico-Científico do IBICT e da Comissão de Ensino da CNEN. É coautor do livro "Guia de gestão de dados de pesquisa para bibliotecários e pesquisadores" e organizador/autor dos livros "Bibliotecas digitais: saberes e práticas" e "Implantação e gestão de repositórios institucionais". Tem como áreas de interesse: bibliotecas e arquivos digitais, publicações eletrônicas, interoperabilidade de sistemas de informação, acervos culturais digitais, curadoria de dados de pesquisa e preservação digital.



Luana Farias Sales possui Graduação em Biblioteconomia e Documentação pela Universidade Federal Fluminense (2003). Mestre em Ciência da Informação pelo convênio UFF/IBICT (2004-2006). Doutora em Ciência da Informação pelo Programa de Pós-Graduação do IBICT/UFRJ (2011-2014)., Atualmente é Analista em C & T da CNEN, atuando como pesquisadora do Instituto de Engenharia Nuclear na linha de Gestão do Conhecimento Nuclear e coordenando o Repositório Carpe dIEN : repositório de dados e informações em Energia Nuclear. Atua ainda como professora da Universidade Federal do Estado do Rio de Janeiro – UNIRIO, ministrando disciplinas relacionadas à Organização do Conhecimento. Tem experiência na área de Ciência da Informação, com ênfase em Organização e Representação do Conhecimento e Recuperação de Informação. Possui interesse em tópicos ligados à Comunicação Científica, Tecnologia de Informação e Gestão do Conhecimento, desenvolvendo pesquisas especificamente nas temáticas de e-Science, curadoria digital de dados de pesquisa, biblioteca digital, metadados, repositórios institucionais, repositórios de dados, Sistemas CRIS, objetos digitais e publicações ampliadas, sendo coautora de diversos artigos sobre estas temáticas e ainda do "Guia de gestão de dados de pesquisa para bibliotecários e pesquisadores".

Copyright © 2016 Sayão & Sales. Todo o conteúdo da Revista (incluindo-se instruções, política editorial e modelos) está sob uma licença Creative Commons Atribuição-NãoComercial-CompartilhaIgual 3.0 Não Adaptada. Ao serem publicados por esta Revista, os artigos são de livre uso em ambientes educacionais, de pesquisa e não comerciais, com atribuição de autoria obrigatória. Mais informações em http://revistas.ufpr.br/atoz/about/submissions#copyrightNotice.

Resumo

Os Doutores Luis Fernando Sayão e Luana Farias Sales apresentam as condições que sustentam o desenvolvimento da área de curadoria digital e os emergentes repositórios de dados de pesquisa, assim como as controvérsias, desafios e responsabilidades éticas para os depositantes, gestores dos repositórios, usuários e profissionais da informação.

Palavras-chave: Dados de pesquisa; Curadoria digital; e-Science; Ciência aberta; Repositórios científicos.

Abstract

Luis Fernando Sayão, PhD and Luana Farias Sales, PhD present the conditions underpinning the development of the digital curation area and the emerging research data repositories, as well as the controversies, challenges and ethical responsibilities for depositants, repository managers, users and information professionals.

Keywords: *Research data; Digital curation; E-Science; Open science; Scientific repositories.*

1. Que elementos/condições caracterizam a emergência do desenvolvimento da área de curadoria digital?

A sociedade contemporânea – apoiada pelo poder pervasivo das tecnologias digitais e da *web* – produz e consome um volume extraordinário de informações em formatos digitais. Esses registros digitais são criados e aplicados em todo espectro social, mudando comportamento, negócios, formas de governar, de ensinar, inaugurando padrões inéditos de socialização e dando margem ao surgimento de novos fenômenos como é o *Big Data*.

Na esfera científica, as transformações são mais contundentes: o desenvolvimento de novos aparatos científicos, instrumentos, sensores, escalas e o uso intensivo de modelos de simulação da natureza geram uma quantidade imensa de dados, delineiam as fronteiras de um novo paradigma científico – conhecida como *eScience* -, criam novas metodologias de produção de conhecimento científico, formas de compartilhamento e de socialização entre os pesquisadores e alteram o fluxo tradicional de comunicação científica e de revisão por pares. Aliado a isso, os pressupostos da ciência aberta, que considera o conhecimento científico como um bem da humanidade, demandam que metodologias, instrumentos, softwares e dados estejam abertos para garantir o princípio da

reprodutibilidade e de autocorreção da ciência e a transparência de seus fluxos e de sua trajetória de erros e acertos.

Neste cenário de mudanças, os dados deixam de ser um mero subproduto da atividade de pesquisa e se tornam protagonistas na geração de novos conhecimentos e descobertas. Para tal, precisam estar disponíveis e interpretáveis para que possam transmitir conhecimento no tempo e no espaço e para que sejam reusados em diversos contextos alimentando a pesquisa interdisciplinar.

O uso e a geração intensiva de dados pelas atividades acadêmicas e de pesquisa criam a necessidade urgente de infraestruturas gerencias e tecnológicas que tratem de forma dinâmica o ciclo de vida dos dados – do seu planejamento até a seu arquivamento confiável – e insira esses ativos na infraestrutura mundial de informação para a pesquisa. Neste ponto que os procedimentos práticos e teóricos, que coletivamente chamamos de curadoria digital, se inserem como essencial para a captura, tratamento, preservação, arquivamento e acesso, pelo tempo que for preciso, dos dados de pesquisa considerados de valor continuo.

2. Caso existam, quais são as correntes/motivações contrárias ao depósito e compartilhamento de dados brutos associados aos manuscritos de pesquisa?

Muitos pesquisadores ainda estão presos na ideia de que os dados gerados por suas pesquisas são propriedades deles e não precisam ser compartilhados ou só precisam ser compartilhados de forma restrita entre seus colegas mais próximos. Porém, grande parte desses dados é produzida ou coletada com financiamento público e podem ser consideradas, portanto, um bem público. Nessa perspectiva, esses dados precisam que o seu potencial de reuso se realize em novas pesquisas e que se amplie seu alcance como inspirador de pesquisas interdisciplinares; aliados a isso é preciso também otimizar os recursos investidos na pesquisa pelas agências de fomento e minimizar a duplicação de esforços dispendidos em ciclos de geração de dados. Subjacente a essa discussão está o pressuposto fundamental de que a ciência avança mais rapidamente quando os seus estoques informacionais são compartilhados.

Porém, o padrão de compartilhamento varia enormemente no mundo heterogêneo e complexo da ciência. Há domínios científicos em que o compartilhamento de dados é determinante para o avanço da área e faz parte das suas metodologias, das suas formas de socialização e da sua cultura, como acontece, por exemplo, na Astronomia. Outros domínios disciplinares são tradicionalmente mais fechados e precisam de mudanças comportamentais e políticas mandatórias para disseminarem abertamente os seus dados.

Contudo, é preciso observar que uma parcela das coleções de dados é proveniente de pesquisas que precisam que seus produtos de pesquisa sejam arquivados e acessados de forma restrita. Isto acontece por vários motivos, por exemplo, devido à possibilidade de patentes, por interesses comerciais, por segurança ou por se tratarem de dados sensíveis que precisam de tratamentos específicos, como processos de anonimização. No entanto, as políticas de gestão de dados, preveem um intervalo de tempo para que esses dados possam ser utilizados de forma privilegiada por seus autores – conhecido como tempo de embargo – e dispõe de instrumentos para tratamento de dados que apresentem elementos mais críticos. Portanto, mesmo esses dados que não podem ser compartilhados em um primeiro momento precisam ser preservados e geridos para reuso futuro. A própria seleção do que vai ser ou não preservado e quando vai ser compartilhado já é uma das etapas do processo de curadoria digital de dados de pesquisa.

É preciso lembrar que um fator determinante para o crescimento da cultura de compartilhamento de dados é a formalização dos mecanismos de autoria e de recompensa por parte das instituições de pesquisa, agências de fomento e editores científicos. Os pesquisadores que geram, coletam e organizam as coleções de dados demandam que a autoria sobre esses dados seja identificada e reconhecida, e a partir daí possam ser citados, avaliados por pares e recompensados por seu trabalho, como são por um artigo de periódico.

3. No caso de um pesquisador desejar ter seus dados de pesquisa amplamente disponíveis e acessíveis a longo prazo, quais as alternativas disponíveis?

Existem várias opções, como disponibilizar no *website* do projeto ou da página pessoal, mas publicar as coleções de dados na *web* é só uma das etapas de um fluxo mais longo e complexo. Os dados para transmitirem conhecimento precisam de tratamentos específicos, de catalogação e de documentação que inclui metadados descritivos e disciplinares e documentos que garantam sua interpretação, como cadernos de laboratório, roteiro de entrevistas, manuais, dicionário de dados etc.; além do mais, os dados precisam ser identificados por meio de identificadores persistentes como o DOI (Digital Object Indentifier). Nessa direção, a melhor opção é depositar os dados em repositórios de dados, sejam institucionais, disciplinares, multidisciplinares, orientados para projetos ou os ligados aos periódicos científicos. Há centenas de repositórios com tecnologias, gestão, licenças, fluxos, submissão, custos, níveis de certificação e estratégias de preservação que variam muito. A maneira mais fácil de localizá-los e de identificar as suas características é consultando os diretórios de repositório como o re3data.org (www.re3data.org). Além do mais, os repositórios de dados dão maior visibilidade aos dados, o que implica em mais citações para estes e para os artigos correspondentes; permitem um maior grau de reuso e

compartilhamento; oferecem estruturas de preservação – como migração - e armazenamento seguro; oferecem ainda possibilidade de interação entre os pesquisadores, anotações e administração de licenças e de tempo de embargo. Outro fator de grande importância é o estabelecimento de *links* entre dados e artigos de periódicos que possibilita a formulação de novas concepções de publicações científicas e de um alto grau de contextualização para os dados, artigos e projetos.

4. O que é uma "publicação ampliada"/*enhanced publication*? Já existem boas práticas nesse sentido no Brasil? E, especificamente, no campo da Ciência da Informação?

Uma publicação ampliada é uma instância de um objeto digital complexo que se caracteriza pela vinculação de artigos aos dados que subsidiaram a sua criação, bem como à outras informações que auxiliem na contextualização da pesquisa, por exemplo, informações sobre os autores, sobre os equipamentos utilizados, projetos, etc. Essas vinculações podem ter valores semânticos e estarem associados à diversas ontologias. No Brasil, no campo da Ciência da Informação, a revista Informação & Tecnologia, do GT-8 da ANCIB, se configura como uma tentativa de publicação ampliada, no entanto, muitos artigos ainda aparecem no modelo tradicional porque os dados das pesquisas não foram cedidos pelos seus autores.

5. Quais os procedimentos éticos utilizados na criação, disponibilidade e armazenamento dos *datasets*, em relação a dados sensíveis (área de Saúde, segmento de negócios, por exemplo)? Qual o papel de cada ator neste sentido? (depositante, repositório, usuário, por exemplo)?

As infraestruturas de gestão de coleções de dados de pesquisa precisam estar ancoradas em uma política formal que defina todos os procedimentos do fluxo de curadoria: procedimentos práticos, formatos, metadados, arquivamento, preservação e segurança, licenças e, sobretudo, o tratamento que deve ser conferido ao compartilhamento de dados sensíveis, pessoais, confidenciais, de informações de interesse comercial ou que serão importantes nos processos de patenteamento, além das questões associadas à proteção de direitos de propriedade intelectual.

O pesquisador tem um papel importante nessas questões, pois é cada vez mais exigido pelas instituições de pesquisa e pelas agências de fomento que o próprio pesquisador formalize em um documento - conhecido como "Plano de Gestão de Dados de Pesquisa" - entre outras coisas, seu compromisso sobre como as questões éticas e de privacidade e os procedimentos para garantir a segurança dos dados, especialmente dos dados sensíveis, na fase de compartilhamento serão endereçados. Por exemplo: as estratégias de anonimização e de criptografia, no caso de transmissões por rede dos *datasets* das áreas de saúde que identificam pacientes; e o cuidado com o compartilhamento de dados que envolvam a localização de espécimes em riscos e de habitats sensíveis.

Pelo lado do usuário, os termos de uso e as licenças associadas aos dados – ambas estabelecidas ou adotadas pelas políticas de cada repositório – definem as responsabilidades e os limites de uso que os usuários devem respeitar, incluindo nesse escopo o reconhecimento da autoria dos dados por meio de citação padronizada.

6. Sob sua perspectiva, em que aspectos as discussões sobre curadoria digital - notadamente orientadas para a comunicação científica - extrapolam este ambiente e se dirigem ao "mercado" (por exemplo, a já reconhecida importância da inserção do cientista de dados em *fintechs* (ex: Nubank) e alguns meios de comunicação (ex.: NexoJornal)?

O conceito de reuso é determinante nessa questão. A possibilidade de as coleções de dados de pesquisa serem analisadas e reinterpretadas em outros contextos fora dos limites do domínio científico, ou de suas metodologias serem usadas em empreendimentos comerciais ou culturais e artísticos ou ainda em processos de inovação é plenamente viável. Isto é especialmente importante e tecnologicamente factível quando caminhamos de uma *web* de documentos para uma *web* de dados, em que as tecnologias digitais se associam às tecnologias semânticas e criam ambientes de informação mais inteligentes, contextualizados e convergentes, baseados em sistemas de informação interoperáveis

No âmbito de dados culturais, a ideia de reuso e reinterpretação de conteúdos culturais digitais começa a se ampliar e se tornar também um novo nicho de negócios para a indústria de conteúdos.

Por exemplo, a Europeana, apoiada na ontologia Europeana Data Model (EDM), disponibiliza os recursos que ela agrega, para reuso em diversos setores. O Projeto Europeana Space (http://www.europeana-space.eu), cujo lema é "um espaço de possibilidades para o reuso criativo de conteúdo cultural" ilustra bem esse novo conceito de reinterpretação de informações culturais. O objetivo do Projeto – conforme informa seu *website* - é a criação de novas oportunidades de emprego e de crescimento econômico no setor das indústrias criativas europeias com base nos recursos culturais digitais. Como resultado final, o projeto espera gerar produtos e serviços inéditos prontos e testados para serem distribuídos no mercado.

Esse tipo de reuso implica numa curadoria mais sofisticada e com perspectivas mais abertas. Porém, antes de tudo é preciso compreender que os dados de pesquisa são, em grande parte, gerados/coletados com recursos

públicos; são disseminados baseados em licenças que não permitem o uso comercial; e existe uma preocupação forte com questões éticas, de privacidade e de propriedade intelectual. Isto significa que teremos pela frente debates interessantes ao redor dessa questão.

7. Quais competências devem ser desenvolvidas por profissionais para o trabalho com curadoria digital?

O ponto crucial da curadoria de dados de pesquisa é que ela precisa estar inserida organicamente nos fluxos de geração de conhecimento científico para ter validade. De uma forma prática, isto significa que os profissionais de informação e as infraestruturas informacionais e tecnológicas subjacentes às bibliotecas de pesquisa devem estar imbricadas nas atividades dos laboratórios e das outras atividades acadêmicas e de pesquisas da instituição. Isto demanda uma reformulação nos perfis profissionais dos bibliotecários e arquivistas que agora se tornam profissionais de dados. Além disso, a interlocução necessária à curadoria exige ainda novos profissionais provenientes dos domínios específicos – como pesquisadores - e profissionais da área de computação. Vejamos:

A biblioteca de pesquisa – alinhadas a ciência voltada para produtos finais: teses, artigos, inventos, patentes - sempre se concentrou na pós-publicação, na custodia e disseminação de artefatos informacionais acabados, como periódicos e livros; o ciclo de vida dos dados de pesquisa, entretanto, é mais complexo e pressupõe a captura de dados em diferentes estágios de seu processamento e ainda considera linhagens e versões que variam no tempo.

A curadoria se inicia ainda no planejamento dos dados e não se encerra com o fim dos projetos, pois os dados continuam a evoluir. Como desdobramento, as bibliotecas agora têm que se preocupar com os estágios de pré-publicação e com uma gestão contínua. O bibliotecário tem que se preocupar com a gênese dos dados e com toda a documentação necessária à interpretação e à contextualização dos dados ao longo do tempo. Isto implica que ele precisar conhecer as peculiaridades da área em que atua e os fluxos de trabalho dos laboratórios e seus produtos de pesquisa, precisa conhecer também todo o ciclo de vida dos dados de sua instituição e como isso se relaciona com a atividade de curadoria; os metadados gerais e os do domínio disciplinar; os padrões de catalogação; e as teorias que subsidiam a organização do conhecimento e seus instrumentos como taxonomias tesauro e ontologias.

Porém, a curadoria de dados de pesquisa exige uma equipe com muitas expertises que vão além da biblioteconomia de dados:

Para começar, os dados – especialmente os observacionais - precisam ser preservados e arquivados de modo que propriedades arquivísticas, como proveniência, confiabilidade e autenticidade sejam mantidas. Isto acontece porque os pesquisadores do futuro vão basear suas pesquisas na confiança que têm sobre esses dados, portanto eles devem estar hospedados em ambientes estáveis e confiáveis, pois a *web*, por si só, não tem memória permanente, não foi projetada para isso, e não garante que essas propriedades sejam asseguradas. Endereçar essas questões é o papel dos arquivistas de dados.

A curadoria implica, em grande escala, em adicionar valor aos dados. Nessa direção, os dados precisam ser avaliados, analisados, enriquecidos com anotações, ligados por *hiperlinks* com outros recursos, comentados, agregados, "regerados" e descartados. Este papel, que exige conhecimento profundo da área, é atribuído ao pesquisador e a outros especialistas da área.

Por fim, os dados precisam ser analisados, processados por programas e equipamentos específicos que identificam padrões que ajudem na formulação de hipóteses, como *data mining*; precisam ser visualizados e experimentados por sistemas interativos que exibem suas diversas faces. Isto fica por conta dos cientistas da computação, que no contexto da curadoria são chamados de cientistas de dados.

Como citar esta entrevista (APA):
Sayão, L. F. & Sales, L. F. (2016). Curadoria digital e dados de pesquisa. *AtoZ: novas práticas em informação e conhecimento, 5*(2), 67 – 71. Recuperado de: http://dx.doi.org/10.5380/atoz.v5i2.49708

AtoZ
novas práticas em informação e conhecimento

ISSN: 2237-826X
ARTIGO | PAPER

Melhoria na qualidade de dados com a aplicação de "data cleaning" na base de dados de acidentes aeronáuticos da aviação civil brasileira

Application of data cleaning to improve data quality at the Brazilian Civil Aviation Aircraft Accidents Database

Cleibson Aparecido de Almeida[1], Leonardo Derckan Rodrigues Silva[2], Elaine Cristina da Silva Schilipack[3], Nivaldo Aparecido Minervi[3]

[1]*Universidade Aberta de Portugal – UAb, Lisboa, Portugal*
[2]*Instituto de Tecnologia da Aeronáutica – ITA, São José dos Campos, SP, Brasil*
[3]*Universidade Federal do Paraná – UFPR, Curitiba, PR, Brasil*

Autor para correspondência/*Mail to:* Cleibson Aparecido de Almeida (contato@cleibsonalmeida.blog.br)

Financiamento/*Funding:* Centro de Investigação e Prevenção de Acidentes Aeronáuticos (CENIPA/FAB)

Recebido/*Submitted*: 21 Jun. 2016; **Aceito/*Approved*:** 26 Ago. 2016

Copyright © 2016 Almeida et al.. Todo o conteúdo da Revista (incluindo-se instruções, política editorial e modelos) está sob uma licença Creative Commons Atribuição-NãoComercial-CompartilhaIgual 3.0 Não Adaptada. Ao serem publicados por esta Revista, os artigos são de livre uso em ambientes educacionais, de pesquisa e não comerciais, com atribuição de autoria obrigatória. Mais informações em http://revistas.ufpr.br/atoz/about/submissions#copyrightNotice.

Resumo

Introdução: Apresenta a aplicação de técnicas de *data cleaning* na base de dados de acidentes aeronáuticos da aviação civil brasileira com o objetivo de mensurar o grau de melhoria na qualidade dos dados.
Método: inicialmente realizou-se uma revisão de literatura sobre os conceitos de *data cleaning* e qualidade de dados e, em seguida, aplicaram-se as técnicas de *data cleaning* em uma base de dados composta por 4601 registros, referentes aos acidentes aeronáuticos ocorridos entre os anos de 1979 e 2014 na aviação civil brasileira. A medição da melhoria na qualidade dos dados foi realizada por meio da métrica "percentual de melhoria dos dados".
Resultados: Observando-se o contexto geral todos os atributos da base de dados houve uma melhoria de 9% quanto à qualidade dos dados, com atributos, como por exemplo o peso, fabricante e modelo das aeronaves, que apresentaram um grau de melhoria acima de 55% após a aplicação da metodologia.
Conclusão: A técnica de *data cleaning* pode ser utilizada para definir políticas para a melhoria contínua em bases de dados e melhorar os processos de decisão nas organizações que tratam sobre aviação, em especial na área de segurança de voo.

Palavras-chave: Limpeza de dados; Qualidade de dados; Métodos de limpeza de dados

Abstract

***Introduction**: It shows the application of techniques of data cleaning in the aeronautical accidents of brazilian civil aviation with the aim of measuring the degree of improvement in the quality of the data.*
***Method**: Initially, there was a literature review on the concepts of data cleaning and data quality, and then applied the techniques of data cleaning in a database composed of 4601 records, relating to aviation accidents that occurred between the years of 1979 and 2014 in brazilian civil aviation. The measurement of the improvement in the quality of the data was performed using the metric "percent of improvement of data".*
***Results**: Observing the general context all the attributes of the database there was a 9% improvement on the quality of the data, with attributes, such as weight, manufacturer and model of the aircraft, which had a degree of improvement over 55% after application of the methodology.*
***Conclusion**: The data cleaning technique can be used to define policies for continuous improvement in data bases and improve decision-making processes in organizations that deal with aviation, particularly in the area of flight safety.*

***Keywords:** Data cleansing; Data quality; Data cleansing methods*

INTRODUÇÃO

Nos últimos 50 anos as organizações vêm transformando seus processos baseados na economia industrial para uma economia de informação. Desde então, o mundo tem visto uma transformação da competição que se baseava na excelência produtiva, marketing e distribuição de produtos para uma atividade baseada na captação e processamento de dados como pré-requisito para definir estratégias de produção, marketing e distribuição de produtos e serviços.

Como fator fundamental para essa transformação, a popularização dos computadores, que vem ocorrendo desde meados dos anos 1990 e, mais recentemente, os dispositivos móveis, têm sido utilizados pelas empresas como o principal meio para a captura, armazenamento, visualização e descoberta de oportunidades baseadas nas informações produzidas.

Lopes (2006) ressalta que:

> Se a informação é a moeda da nova economia, então os dados são a matéria prima essencial necessária para alcançar o sucesso. Eles são os bens que formam a base dos planos estratégicos e ações que determinam o bom desempenho de um empreendimento.

Acompanhando esta tendência, a aviação civil brasileira tem coletado e armazenado diversas informações. São informações sobre rotas e horários de voos, disponibilidade de aeroportos, registros de aeronaves e tripulantes, dentre outras.

No contexto da aviação civil brasileira, a área de segurança de voo tem sido alvo de interesse da sociedade, principalmente após os dois últimos acidentes aeronáuticos que causaram grande comoção nacional – o A-022/CENIPA/2008 (Centro de Investigação e Prevenção de Acidentes [CENIPA], 2008) e o A-Nº67/CENIPA/2009 (Centro de Investigação e Prevenção de Acidentes [CENIPA], 2009), nos quais morreram 154 e 199 pessoas, respectivamente.

São apenas dois casos entre os vários acidentes aeronáuticos que ocorrem anualmente no Brasil. Somente no ano de 2014 foram registrados 145 acidentes com um total de 70 fatalidades.

De fato, os acidentes aeronáuticos acontecem em diversas condições e são causados por diferentes fatores, porém uma base de dados histórica poderia auxiliar nos trabalhos de identificação dos riscos, prevenção, redução e um melhor entendimento sobre esses acidentes.

Entretanto nem sempre uma base de dados histórica apresenta condições de qualidade para utilização em processos decisórios, visto que podem estar cercadas de problemas concernentes a anomalias de dados. Estes problemas podem ser provenientes da falta de padrão/validação na entrada de dados; da agregação de sistemas legados; da ruptura na coleta de informação em determinados períodos; entre outros.

De acordo com Kanki e Seamster (2002), uma aviação segura depende de padronizações operacionais seguindo taxonomias internacionais e isso deveria garantir uma boa qualidade nos dados gerados nas atividades aeronáuticas.

Desta forma, este trabalho tem como objetivo mensurar a melhoria na qualidade dos dados após a aplicação de um *data cleaning* (processo de limpeza de dados) na base de dados de acidentes aeronáuticos ocorridos na aviação civil brasileira entre 1979 e 2014.

Para isso, este documento está estruturado de forma a apresentar brevemente os conceitos sobre *data cleaning* e qualidade de dados, os procedimentos metodológicos adotados, a apresentação e análise dos resultados e conclusão.

REFERENCIAL TEÓRICO

Data cleaning

Data cleaning, ou limpeza de dados - também conhecida como *data cleansing* ou *scrubbing* - é um conjunto de técnicas utilizadas para detectar e remover anomalias em bases de dados (Rahm & Do, 2000).

De acordo com Vasco (2013), as anomalias existentes em bancos de dados podem ser divididas em três categorias: a) anomalias de sintaxe; b) anomalias de semântica e; c) anomalias de cobertura.

As anomalias de sintaxe condizem ao formato e valores adotados para a representação do dado. Podem conter erros lexicais, erros no formato do domínio e irregularidades. As anomalias de semântica condizem ao não entendimento do dado registrado. Podem conter violações das restrições de integridade, contradições, duplicidade de registros e registros inválidos. As anomalias de cobertura condizem à ausência de informação quando esta for uma premissa do dado. Podem conter valores omissos e registros omissos (o Quadro 1 apresenta alguns exemplos de anomalias em bases de dados).

Quanto às técnicas utilizadas para a limpeza de dados, Oliveira, Rodrigues, e Henriques (2004) destacam que elas podem ser divididas em abordagens especializadas e abordagens genéricas.

As abordagens especializadas são utilizadas sobre um determinado campo de domínio da base de dados, ou seja, foca em um problema concreto já conhecido como correção de nomes, endereços, registros duplicados.

As abordagens genéricas são utilizadas de forma sistemática na base de dados, ou seja, cobrem um campo mais vasto de problemas e adaptam-se a diferentes domínios que sejam de desconhecimento do analista. Tanto a abordagem específica, quanto a abordagem genérica, depende de técnicas (algoritmos/regras) que automatizem, de forma total ou parcial, as tarefas de detecção e correção das anomalias (Rahm & Do, 2000).

De acordo com Lopes (2006) e Vasco (2013), algumas técnicas gerais que podem ser utilizadas para a detecção de erros em *data cleaning* são: a) técnicas estatísticas; b) clusterização; c) baseado em padrão; d) regras associativas; e) análise sintática (*Parsing*); f) transformação de dados; g) aplicação de restrições de integridade e; h) eliminação de duplicatas.

Apesar do *data cleaning* ter como objetivo uma base de dados isenta de erros, isso nem sempre é alcançado. Por isso, é salutar a adoção de medidas preventivas que visem evitar futuros trabalhos com dados anômalos. Além de ser mais eficaz, a prevenção tem um custo menor do que futuras correções.

Atributo	Valor registrado	Valor esperado	Tipo de anomalia
gênero	40	'Masculino' ou 'Feminino'	Erro Lexical
localidade	São Paulo BR	'São Paulo - BR'	Erro no Formato do Domínio
gênero	Masculino | M	'Masculino' | 'Masculino'	Irregularidade
gênero	Homossexual	'Masculino' ou 'Feminino'	Violação da restrição de integridade
Gênero – gravidez?	M – SIM	'F - SIM' ou 'M-NÃO'	Contradição
localidade	São Paulo – BR | S. Paulo - BR	'São Paulo – BR' | 'São Paulo – BR'	Duplicidade de Registros
gênero	NULL	'Masculino' ou 'Feminino'	Valor Omisso
Id – gênero - localidade	23 – NULL - NULL	'23 - Masculino - São Paulo'	Registro Omisso

Quadro 1. Exemplos de anomalias de dados.
Fonte: Adaptado de Vasco (2013).

Qualidade de dados

Com foco na menor quantidade possível de anomalias, a qualidade de dados é um fator importante para determinar a precisão da tomada de decisão quando se utiliza informações provenientes de bancos de dados.

De acordo com Orr (1998), "nenhum sistema de informação tem uma qualidade dados de 100%". No entanto, a principal preocupação com a qualidade de dados não é a garantia de perfeição, mas garantir um nível de qualidade em que seja possível tomar decisões razoáveis de forma rápida e que garanta a sobrevivência de uma organização.

O nível de qualidade de um banco de dados pode ser medido de forma específica, utilizando os conceitos apresentados no trabalho de Strong, Lee, e Wang (1997). Em sua pesquisa, os autores definiram categorias e dimensões para a qualidade de dados (Quadro 2).

Categoria	Dimensão	Definição
Intrínseca	Acuracidade	A informação é correta e confiável?
	Reputabilidade	A informação espelha a sua fonte ou conteúdo?
	Credibilidade	A informação é verdadeira?
	Objetividade	A informação é imparcial?
Acessibilidade	Acessibilidade	A informação está disponível de forma fácil?
	Segurança	Até quanto a informação é restrita para resguardar a sua segurança?
Contextualidade	Relevância	A informação é útil para a tarefa em questão?
	Completude	A informação é completa e suficiente ao nível de largura e profundidade?
	Temporalidade	A informação está atualizada para a tarefa em questão?
	Valor agregado	A informação é benéfica e há vantagens em sua utilização?
	Quantidade	O volume de informação é adequado para a tarefa em questão?
Representatividade	Interpretação	A informação está em linguagem, símbolo e unidade apropriada para a tarefa em questão? Sua definição é clara?
	Facilidade de uso	A informação é facilmente manipulada e aplicada em diferentes tarefas?
	Facilidade de entendimento	A informação é facilmente entendida?
	Representação concisa	A informação está representada de forma compacta?
	Representação consistente	A informação está representada no mesmo formato para um mesmo atributo em diferentes registros?

Quadro 2. Categorias e dimensões na qualidade de dados.
Fonte: Adaptado de Strong et al. (1997) e Pipino, Lee, e Wang (2002).

Outra forma específica para medir a qualidade dos dados foi descrita por Vasco (2013), conforme os itens a seguir: a) precisão: é o quociente entre o número de valores corretos e o número total de valores existentes no conjunto de dados; b) integridade: este critério é satisfeito quando não existem registros inválidos, nem violações de restrição de integridade ou valores omissos; c) completude: é o percentual de preenchimento em um registro, considerando que todos os atributos considerados de preenchimento obrigatório devam estar alimentados; d) validade: é o percentual de registros que são entidades válidas. Registros que violem as condições de integridade são considerados inválidos; e) consistência: é o percentual de registros que possuem anomalias de sintaxe ou contradições; f) conformidade de esquema: é o percentual registros que estão conforme a estrutura sintática definida pelo esquema relacional; g) uniformidade: é o percentual de registros que não contém irregularidades nos seus valores; h) densidade: é o percentual de valores não omissos nos registros e; i) unicidade: percentual de registros únicos, ou seja, registros sem duplicidades.

MÉTODO

Descrição do conjunto de dados

O conjunto de dados utilizado na pesquisa corresponde ao número total de acidentes aeronáuticos ocorridos com aeronaves brasileiras, entre 1979 e 2014, resultando em 4601 registros (eventos).

Foram utilizadas as duas principais entidades do banco de dados, com um total de 19 (dezenove) atributos, conforme ilustrado na Figura 1.

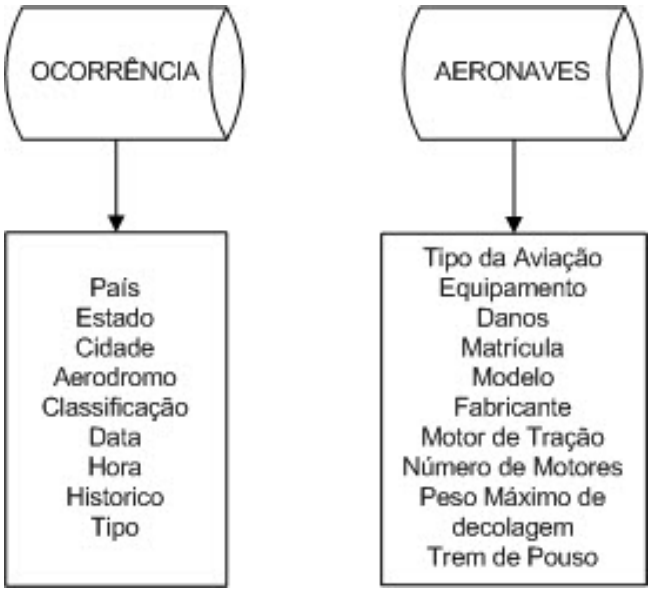


Figura 1. Entidades Ocorrência e Aeronaves.
Fonte: Elaborado pelos autores.

Detalhamento dos atributos

O Quadro 3 apresenta a descrição e exemplos para cada um dos 19 atributos utilizados no estudo.

Aplicação do Data Cleaning no banco de dados

A aplicação do *data cleaning* foi realizada mediante um processo controlado e mensurado após a atualização do banco de dados (Figura 2).

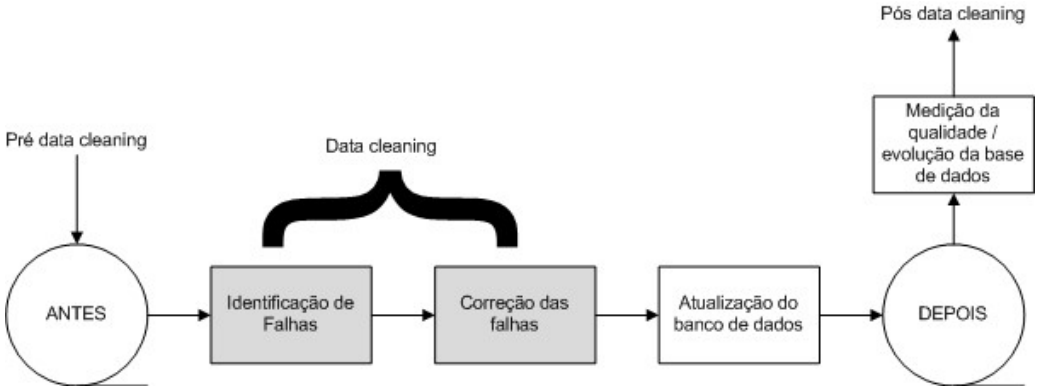


Figura 2. Processo para aplicação do *data cleaning*.
Fonte: Elaborado pelos autores.

Utilizando a abordagem especializada, o *data cleaning* foi voltado à erros lexicais, erros no formato de domínio, irregularidades nos dados, violações de restrição de integridade, duplicidade de registros, valores omissos e registros omissos.

Para isso foram adotadas as técnicas estatísticas para correção dos atributos numéricos, baseado em padrão. Nos atributos do tipo data, hora e texto foram aplicadas regras associativas.

A execução das atividades foi parcialmente automatizada com o apoio dos seguintes softwares: a) QLIKVIEW 11; b) MS EXCEL 2007 e, c) R PROJECT 3.0.

Atributo (tipo de dado)	Descrição	Exemplo
País (texto)	Nome do país onde ocorreu o acidente.	'BRASIL', 'PERU', 'CHILE'...
Estado (texto)	Sigla do Estado/Província onde ocorreu o acidente.	'SP', 'EX', 'EX'...
Cidade (texto)	Nome da cidade onde ocorreu o acidente.	'SÃO PAULO', 'LIMA', 'SANTIAGO'...
Aeródromo (texto)	Código ICAO* da pista de pouso/decolagem onde ocorreu o acidente (composto por 4 caracteres).	'SBGR', 'SBTY', 'KLOP'...
Classificação (texto)	Classificação da ocorrência. Neste caso apenas acidentes.	'ACIDENTE'
Data (data)	Data do acidente no formato dia/mês/ano no padrão numérico.	'01/09/2011', '15/02/1979'...
Hora (hora)	Horário local do acidente no formato hora:minuto no padrão 24h.	'11:20', '18:59', 23:45'...
Histórico (texto)	Texto resumo que descreve a ocorrência. Pode ser qualquer conjunto de frases formadas por qualquer tipo de caracteres.	'A aeronave PTRFG estava sobrevoando o aeroporto SBGR quando colidiu com um urubu'...
Tipo (texto)	Taxonomia de tipo de ocorrência aeronáutica prevista no MCA 3-6 .	'COLISÃO COM OBSTÁCULO NO SOLO'...
Tipo da Aviação (texto)	Define se a aeronave é CIVIL ou MILITAR. Nesta pesquisa será considerada apenas a CIVIL.	'CIVIL'
Equipamento (texto)	Tipo de aeronave que se acidentou. Pode ser um AVIÃO, HELICÓPTERO ou outro conforme o RAB**.	'AVIÃO', 'PLANADOR', 'HELICÓPTERO'...
Danos (texto)	Nível do dano da aeronave naquele acidente. Pode assumir um entre quatro diferentes valores.	'NENHUM', 'LEVE', 'SUBSTANCIAL', 'DESTRUÍDA'
Matrícula (texto)	Marca da aeronave registrada no RAB**. É composta por 5 ou mais caracteres.	'PTOLK', 'PUGHT', 'N600XL'...
Modelo (texto)	Modelo de fabricação da aeronave.	'BOEING 757', 'EMB 190'...
Fabricante (texto)	Nome do fabricante da aeronave.	'EMBRAER', 'BOEING', 'BOMBARDIER'...
Motor de Tração (texto)	Tipo do motor da aeronave.	'JATO', 'TURBOÉLICE'...
Número de Motores (número)	Quantidade de motores presente na aeronave. Podendo ser 0 ou mais.	'0', '1'...'4'...
Peso Máximo de Decolagem (número)	Peso máximo de decolagem de uma aeronave.	'1263', '123890'...
Trem de Pouso (texto)	Tipo do trem de pouso da aeronave.	'CONVENCIONAL', 'TRICICLO'...

Quadro 3. Descrição dos atributos utilizados na pesquisa.
Fonte: Elaborado pelos autores.
Notas: *ICAO: International Civil Aviation Organization. **RAB: Registro Aeronáutico Brasileiro (http://www2.anac.gov.br/aeronaves/cons_rab.asp).

O QLIKVIEW foi utilizado para fazer a extração dos dados disponíveis no Sistema Gerenciador de Banco de Dados (SGBD), em ORACLE 10G DATABASE, da Organização. O MS EXCEL foi utilizado como ferramenta auxiliar nas atividades de visualização dos dados em formato tabular e preparação de dados para reinserção no SGBD após o *data cleaning*. O R-PROJECT foi utilizado para identificação e correção das falhas identificadas.

Métrica utilizada para a qualidade dos dados

A métrica utilizada para a medição da qualidade dos dados foi a do percentual de melhoria dos dados, ou seja, $(QEA \div N) \times 100$, sendo que o QEA representa a quantidade de erros encontrados na situação pré *data cleaning* em comparação com a situação pós *data cleaning* e N representa o total de registros na base dados.

Para exibição dos resultados foram utilizados gráficos estatísticos apresentando o percentual de melhoria dos dados em cada um dos atributos considerados no estudo.

RESULTADOS E DISCUSSÃO

Os resultados apresentados nesta seção mostram a melhoria da base de dados após aplicação do *data cleaning* quanto às seguintes anomalias: a) correção de erros lexicais; b) correção de erros de formato no domínio; c) eliminação de irregularidades; d) melhoria nas restrições de integridade; e) eliminação de registros duplicados; f) preenchimento de valores omissos e g) eliminação de registros omissos.

Em relação à melhoria dos dados da entidade Ocorrência, pode-se observar na Figura 3 o percentual de melhoria em cada um dos atributos que passaram pelo *data cleaning*. De forma geral, o erro médio encontrado nesta entidade foi de 6% (seis por cento).

Os atributos Aeródromo e Cidade apresentaram os maiores percentuais de erros, ou seja, 49% e 38.4%, respectivamente. Dentre as anomalias identificadas nestes dois atributos, concentraram-se esforços para melhorar as restrições de integridade, corrigir erros no formato do domínio e o preenchimento de valores omissos.

Os demais atributos da entidade Ocorrência apresentaram um percentual abaixo do erro médio (6%), sendo que o atributo País teve um percentual de melhoria igual a 0, ou seja, não continha erros antes da aplicação do *data cleaning*.

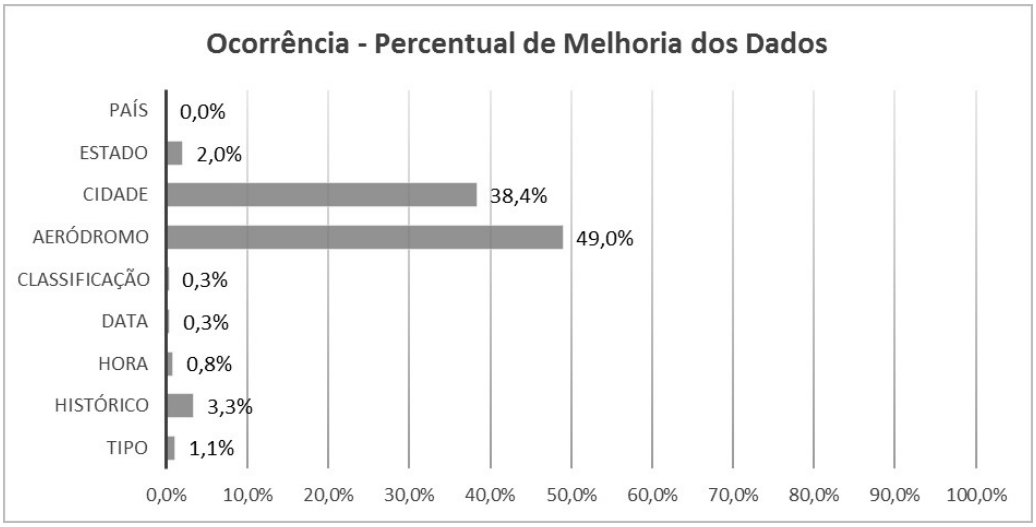


Figura 3. Percentual de melhoria na entidade Ocorrência – Base de dados de acidentes aeronáuticos da aviação civil brasileira.
Fonte: Elaborado pelos autores, via análise dos dados.

Quanto à entidade Aeronave, o percentual de erros encontrados e corrigidos foi maior quando comparado com a entidade Ocorrência. Enquanto a Ocorrência apresentou um erro médio de 6%, a entidade Aeronave teve uma média de erro igual a 12%.

A Figura 4 apresenta o percentual de melhoria realizada na entidade aeronave. Os atributos que apresentaram maior incidência de erros foram o Peso Máximo de Decolagem (98.4%), Trem de Pouso (84%), Modelo (77.2%) e Fabricante (56.1%).

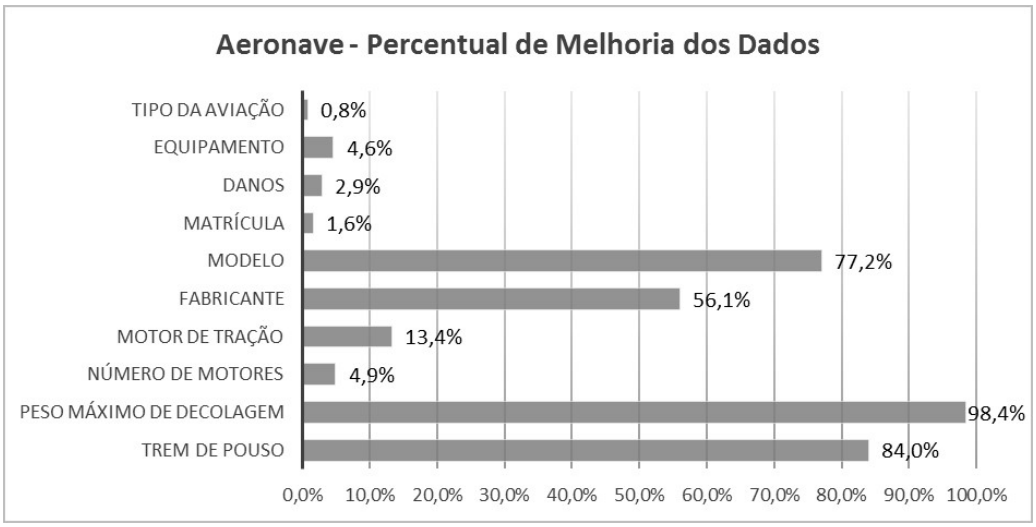


Figura 4. Percentual de melhoria na entidade Aeronave - Base de dados de acidentes aeronáuticos da aviação civil brasileira.
Fonte: Elaborado pelos autores, via análise dos dados.

Uma simples analogia com os valores brutos dessa amostra levaria a considerar que esta base de dados, composta por 4601 acidentes aeronáuticos, tinha mais de 4500 aeronaves com problemas relativos a informação do Peso Máximo de Decolagem.

Com isso, é possível inferir que a tomada de decisão com esses dados antes da aplicação do *data cleaning*, aqui proposto, proporcionava um erro de decisão de 98.4% ao gestor que utilizava esta informação. Este valor significativo quanto ao erro de decisão no Peso Máximo de Decolagem revelou três principais problemas existentes no gerenciamento no SGBD, como:

a) por ser um atributo de preenchimento não obrigatório no SGBD, na maioria das vezes o usuário não se atentava com preenchimento desta informação no banco de dados. Conforme literatura apresentada, isso gerou erros do tipo "valor omisso";

b) por ser uma informação proveniente de terceiros, o usuário necessitava consultar e confiar na credibilidade de uma base de dados externa, o que nem sempre garantia que a informação foi corretamente consultada pelo usuário. Conforme literatura apresentada, isso gerou "erros no formato do domínio";

c) por requerer um grau de conhecimento *a priori* quanto aos modelos de aeronaves existentes, muitas vezes

o usuário preenchia este campo sem realizar uma pesquisa mais aprofundada sobre o assunto. Com isso, o campo era preenchido com valores baseados em conhecimentos prejulgados pelos usuários, baseados em aeronaves similares e distorcidos da realidade especifica daquela aeronave. Conforme literatura apresentada, isso gerou erros do tipo "violação da restrição de integridade".

Estes três tipos de anomalias também foram identificadas em outros três atributos (Trem de pouso, Modelo e Fabricante), que apresentaram um percentual de erros em 84%, 77.2% e 56.1%, respectivamente.

Em contrapartida, os atributos Tipo da Aviação, Equipamento, Danos, Matrícula e Número de Motores apresentaram um percentual de melhoria abaixo do erro médio desta entidade (12%). Os erros encontrados nesses atributos estavam ligados a duplicidade de registros e erros no formato do domínio.

No contexto geral, a melhoria da base de dados utilizado neste estudo foi de 9%. Isso quer dizer que a cada 100 acidentes aeronáuticos contidos nesta base de dados, 9 (nove) possuíam erros que poderiam afetar uma tomada de decisão.

Em consulta verbal aos especialistas que trabalham há mais de dez anos com essa base de dados, foi percebido que a integração ineficiente de sistemas legados, sistemas com falhas na entrada de dados, processos operacionais instáveis e a falta de comprometimento com a qualidade dos dados foram alguns dos fatores que corroboraram para os problemas de qualidade dos dados utilizados nesta pesquisa. Este fato vai ao encontro do proposto na pesquisa realizada por Oliveira et al. (2004) e ao mesmo tempo comprova a afirmação de Orr (1998), sobre a dificuldade em obter uma qualidade de dados 100% em um sistema de informação.

CONCLUSÃO

O presente estudo apresentou a aplicação da técnica de *data cleaning* em uma base de dados de acidentes aeronáuticos ocorridos entre 1979 e 2014, no Brasil. Nesse cenário, vale ressaltar que os especialistas que trabalham há mais de dez anos com a base de dados utilizada nesta pesquisa afirmaram que a utilização de sistemas com falhas na entrada dos dados, problemas com integração de sistemas legados, instabilidade nos processos operacionais e a falta de comprometimento com a qualidade dos dados foram os fatores causadores da baixa qualidade dos dados.

Dentre as anomalias identificadas na base de dados utilizada, destaca-se que o preenchimento de atributos com valores omissos e correções de erros no formato dos domínios e violações nas restrições de integridade são aquelas que mais influenciaram na melhoria da qualidade dos dados.

Com isso, sugere-se que antes de iniciar algum procedimento de *data cleaning* em uma base de dados, sejam observadas a existência desses três tipos de anomalias, visto que este pequeno esforço inicial poderá melhorar significativamente a qualidade dos dados.

Desta forma, a pesquisa atingiu o seu objetivo, ou seja, mensurar o grau de melhoria na qualidade dos dados após a aplicação da técnica de *data cleaning* na base de dados de acidentes aeronáuticos da aviação civil brasileira. Fica conclusivo que a aplicação da técnica de *data cleaning* é uma boa prática para a definição de políticas que foquem na melhoria contínua em bases de dados utilizadas para a tomada de decisão, neste caso específico, auxiliar no processo decisório durante a identificação de riscos na prevenção de acidentes aeronáuticos.

Os autores propõem que se realizem pesquisas adicionais em áreas fora do escopo da aviação civil brasileira, para identificar limitações ou lacunas onde a técnica de *data cleaning* seja relevante para a melhoria da qualidade dos dados. Esta proposta também ajudará na generalização dos resultados até aqui alcançados.

REFERÊNCIAS

Centro de Investigação e Prevenção de Acidentes. (2008). *Relatório final a-022/cenipa/2008*.

Centro de Investigação e Prevenção de Acidentes. (2009). *Relatório final a-nº67/cenipa/2009*.

Kanki, B. G., & Seamster, T. L. (2002). *Aviation information management: From documents to data*. Burlington: Ashgate.

Lopes, F. P. (2006). *Administração de dados: Técnicas, metodologias e ferramentas para garantir a qualidade dos dados*. Recife: Universidade Federal de Pernambuco.

Oliveira, P. J., Rodrigues, F., & Henriques, P. R. (2004). *Limpeza de dados: Uma visão geral.* Recuperado de http://wiki.di.uminho.pt/twiki/pub/Research/Doutoramentos/SDDI2004/ArtigoOliveira.pdf

Orr, K. (1998, Feb.). Data quality and systems theory. *Communications of the ACM*, *41*(2), 66–71. doi: 10.1145/269012.269023

Pipino, L. L., Lee, Y. W., & Wang, R. Y. (2002, Apr.). Data quality assessment. *Communications of the ACM*, *45*(4), 211–218. doi: 10.1145/505248.506010

Rahm, E., & Do, H. H. (2000). Data cleaning: Problems and current approaches. *IEEE Data Engineering Bulletin*, *23*(4), 3–13. Recuperado de http://sites.computer.org/debull/A00dec/issue1.htm

Strong, D. M., Lee, Y. W., & Wang, R. Y. (1997, May). Data quality in context. *Communications of the ACM*, *40*(5), 103–110. doi: 10.1145/253769.253804

Vasco, D. O. (2013). *Identificação de anomalias contextuais*. Porto: Universidade do Porto.

Como citar este artigo (APA):
Almeida, C. A., Silva, L. D. R., Schilipack, E. C. S. & Minervi, N. A. (2016). Melhoria na qualidade de dados com a aplicação de "data cleaning" na base de dados de acidentes aeronáuticos da aviação civil brasileira. *AtoZ: novas práticas em informação e conhecimento, 5*(2), 72 – 79. Recuperado de: http://dx.doi.org/10.5380/atoz.v5i2.47303

AtoZ
novas práticas em informação e conhecimento

ISSN: 2237-826X
ARTIGO | PAPER

A Internet das Coisas aplicada ao conceito de eficiência energética: uma análise quantitativo-qualitativa do estado da arte da literatura

The Internet of Things applied to the energy efficiency concept: a quantitative-qualitative analysis of the state-of-the-art

Arildo Antônio Sônego[1], Roderval Marcelino[1], Vilson Gruber[1]

[1] *Universidade Federal de Santa Catarina - UFSC, Florianópolis, SC, Brasil*

Autor para correspondência/*Mail to:* Arildo Antônio Sônego (asonego@gmail.com)

Recebido/*Submitted*: 28 Jul. 2016; **Aceito/*Approved*:** 06 Set. 2016

Copyright © 2016 Sônego, Marcelino & Gruber. Todo o conteúdo da Revista (incluindo-se instruções, política editorial e modelos) está sob uma licença Creative Commons Atribuição-NãoComercial-CompartilhaIgual 3.0 Não Adaptada. Ao serem publicados por esta Revista, os artigos são de livre uso em ambientes educacionais, de pesquisa e não comerciais, com atribuição de autoria obrigatória. Mais informações em http://revistas.ufpr.br/atoz/about/submissions#copyrightNotice.

Resumo

Introdução: A Internet das Coisas corresponde a um novo paradigma, que imputa aos objetos a capacidade de disponibilizar informações a respeito de seu funcionamento e ao ambiente no qual tais objetos estão inseridos, caracterizando-se como abrangentes e diversificados os universos de prováveis aplicações a ela relacionadas. Em uma época em que se evidenciam ações e discussões dedicadas à sustentabilidade e ao meio ambiente, este artigo objetiva apurar o estado atual das pesquisas envolvendo a Internet das Coisas associado ao conceito de "eficiência energética".
Método: Tendo como referência a base de dados Scopus procedeu-se uma revisão sistemática da literatura, acompanhada de uma análise quantitativo-qualitativa. Foram aferidas informações no período compreendido entre os anos de 2001 a 2016, além da análise das cinquenta publicações com maiores citações.
Resultados: Constatou-se a existência de 895 documentos relacionados aos assuntos pesquisados. Contudo, poucos (foram identificados sete documentos) discorrem sobre a efetiva aplicação da Internet das Coisas à "eficiência energética".
Conclusão: Além de se determinar o estado da arte das pesquisas a nível mundial dos temas propostos, foi possível ratificar os conceitos a eles associados, bem como identificar as publicações mais relevantes, para que possam servir de referência e ponto de partida a outros trabalhos do gênero.

Palavras-chave: Internet das Coisas (IoT); Eficiência energética; Tecnologias de informação e comunicação (TIC)

Abstract

***Introduction**: The Internet of Things represents a new paradigm, which assign to objects the ability to provide information about its operations and the environment in which they live, characterized as a comprehensive and diverse universe and applications related to it. Nowadays, due to the recognition of actions and discussions devoted to sustainability and the environment, this article aims to assess the current state of research involving the Internet of Things, linked to the concept of "energy efficiency".*
***Method**: Starting with reference to the database Scopus, it was conducted a systematic review of the literature, accompanied by a quantitative and qualitative analysis. The registers were collected using a 2001-2016 window, as well as it was made an analysis of fifty publications with higher citations.*
***Results**: It was found 895 documents related to the subjects studied. However, few (seven documents were identified) discussed the effective application of Internet of Things in "energy efficiency".*
***Conclusion**: In addition to determining the state of the art research in the themes proposed, it was possible to ratify the concepts associated with them, as well as identify the most relevant publications which can provide a baseline for other investigations.*

***Keywords:** Internet of Things (IoT); Energy efficiency; Information and communication technologies (ICT)*

INTRODUÇÃO

Ao se analisar a adoção das Tecnologias de Informação e Comunicação (TIC) constata-se um crescimento considerável em recente período da sociedade moderna. Dispositivos como telefones celulares, *tablets*, *smartphones* e *notebooks* são frequentemente empregados em atividades cotidianas de profissionais das mais diferentes áreas, servindo também como fonte de diversão em momentos de relaxamento e lazer.

Conforme Barros (2011, p. 30), "[o]bserva-se, como em nenhum outro momento, o impacto da inclusão da tecnologia na vida das pessoas, atingindo a todos, de todas as maneiras, mesmo aqueles excluídos, vivendo em condições as mais adversas."e, de acordo com Nascimento (2011), a cada dia mais e mais pessoas, empresas, governos e outras organizações sociais tornam-se dependentes do uso contínuo de novas tecnologias, principalmente das Tecnologias de Informação e Comunicação. Entende-se, que as TIC tornaram-se elementos onipresentes ao ser humano.

A popularização do uso da Internet é considerada um dos motivadores deste processo. Na ótica de Pinho (2011, p. 98), "[p]arece que a Internet pode ser colocada como um marco civilizatório: a vida antes e depois da Internet, pois ela tem criado expectativas elevadas de mudanças, algumas até revolucionárias.".

A Internet é provavelmente o maior sistema de engenharia já criado pela humanidade, com centenas de milhões de computadores conectados, enlaces de comunicação e comutadores, bilhões de usuários que se conectam por meio de *laptops* e *tablets*, e com uma série de dispositivos como sensores, *webcams*, consoles para jogos, quadros de imagens e até mesmo máquinas de lavar sendo conectadas (Kurose & Ross, 2013).

Na concepção de Sato (2015), a rápida popularização da Internet - e seu acesso a partir de dispositivos móveis - pode ser apontada como a parte mais visível das transformações pelas quais passa o cenário dos meios de comunicação, principalmente nas últimas décadas. Este cenário promissor apresenta-se favorável à manifestação de ideias, ao surgimento de conceitos e teorias, bem como à proliferação de novas tecnologias, entre elas, a Internet das Coisas (IoT - *Internet of Things*).

Em uma definição mais ampla, a Internet das Coisas relaciona-se à capacidade dos objetos, conectados em rede, disponibilizarem informações a respeito de seu funcionamento. Tal tecnologia, descreve Ferreira (2014), tem a finalidade de proporcionar inteligência para objetos, de modo a permitir seu controle e a notificação de alterações em seu estado.

Lacerda e Lima-Marques (2015, p. 159), afirmam que, "[i]nterligados em rede, os objetos são capazes de realizar ações de forma independente e gerar dados em quantidade e variedade exponenciais, como produto das interações".

Nesta conjuntura, entende-se que a dimensão, contexto e escopo de adoção de tais aplicações, apresentam-se como um dos desafios desta tecnologia. Segundo Wu, Talwar, K., Himayat, e Johnson (2011), não é possível representar de maneira exaustiva todas as alternativas de soluções atribuídas à Internet das Coisas, uma vez que o processo de compreensão do seu potencial ainda encontra-se em fase inicial, visto a variedade de aplicações que podem ser endereçadas em diversos segmentos do mercado. A Figura 1 contém exemplos de concentração proporcional por área (1-Insignificante, 5= Muito importante) de aplicações para a Internet das Coisas.

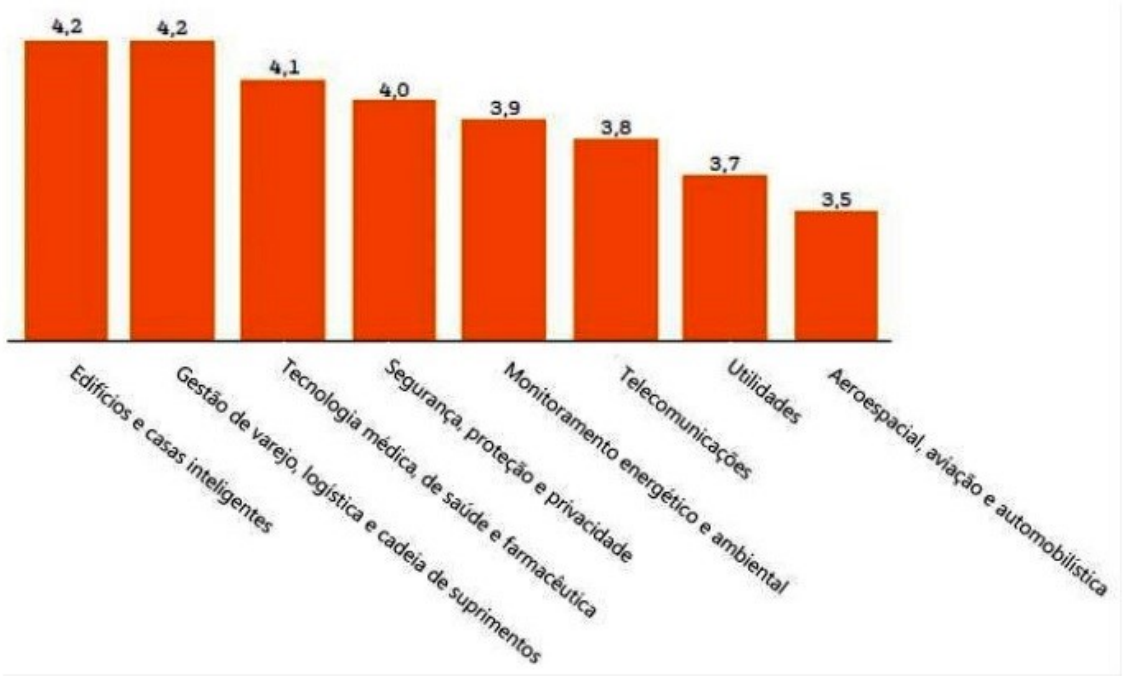


Figura 1. Áreas de concentração para a Internet das Coisas.
Fonte: Lacerda (2015, p. 72).

Uma das áreas elencadas anteriormente se relaciona à aplicação da Internet das Coisas no monitoramento energético e ambiental, sugerindo um uso adequado e consciente dos recursos naturais, contemplando a "eficiência energética".

O desenvolvimento econômico e os altos padrões de vida, destacam Hinrichs, Kleinbach, e Reis (2015), são processos complexos que necessitam de um abastecimento adequado e confiável de energia. Para Yu, Cecati, Dillon, e Simões (2011), a necessidade de minimizar ou mesmo equilibrar o impacto ao meio ambiente criado pelas atividades que fazem parte do estilo de vida do século XXI, promovendo o uso consciente dos recursos naturais, apresentam-se como um dos desafios da humanidade.

Desta maneira, de acordo com Wu et al. (2011), o processo de eficiência energética está condicionado à integração de novas tecnologias, como por exemplo, as TIC, sendo a Internet das Coisas considerada ideal para este ambiente. Assim, este artigo busca investigar a evolução das pesquisas no campo da Internet das Coisas objetivando compreender quais contribuições este paradigma pode oferecer ao conceito de "eficiência

energética".

INTERNET DAS COISAS

O conceito primordial associado à Internet das Coisas (IoT - *Internet of Things*) relaciona-se à capacidade que os objetos possuem de se comunicar, reportando informações acerca de seu estado e funcionamento. Segundo Serafim (2014), esta tecnologia consiste em interligar os objetos de uso cotidiano do ambiente real com a Internet, tornando-os então objetos inteligentes.

O termo Internet das Coisas, destaca Gogliano Sobrinho (2013), refere-se a um novo paradigma, que tem por premissa a integração entre objetos de uso cotidiano e a Internet. Contudo, sustenta o autor, para muitos esse conceito demonstra-se abstrato, e muitas vezes de difícil compreensão, no que tange à maneira de como se procede a essa integração.

Em 1999, o pesquisador do MIT (Massachusetts Institute of Technology) Kevin Ashton fez uso do termo Internet das Coisas pela primeira vez, em uma apresentação direcionada à empresa Procter & Gamble (Serafim, 2014).

O significado do termo, explana Ferreira (2014), ampliou-se e passou a abranger a área de sensores e atuadores sem fio, de objetos conectáveis às redes que utilizam o protocolo TCP/IP (*Transmission Control Protocol / Internet Protocol*), assim como as tecnologias de semântica de dados, concebendo desta maneira, uma visão orientada às coisas, uma visão orientada à internet e uma visão orientada à semântica, conforme apresentado na Figura 2.

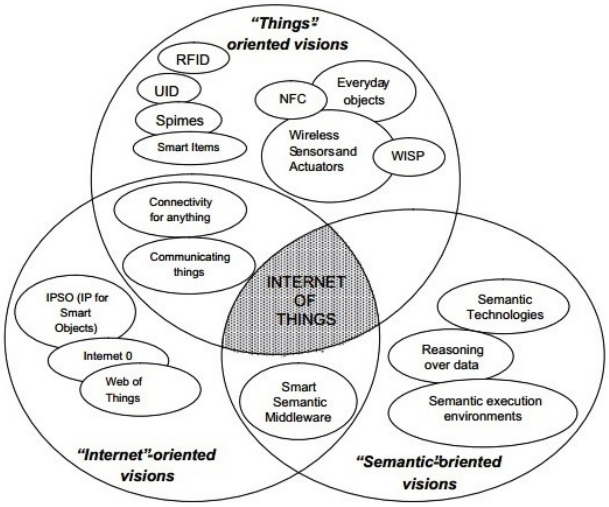


Figura 2. Diferentes visões da Internet das Coisas.
Fonte: Atzori, Iera, e Morabito (2010, p. 2789).

Ao referirem-se às diferentes visões, Atzori et al. (2010) detalham que a visão orientada às coisas objetiva demonstrar propostas que assegurem o melhor aproveitamento dos recursos dos dispositivos e sua comunicação; a visão orientada à semântica foca na representação, armazenamento, pesquisa e organização da informação gerada, procurando soluções para a modelagem das descrições que permitam um tratamento adequado para os dados produzidos pelos objetos; enquanto que a visão orientada à internet tem o intuito de conceber modelos e técnicas destinadas a interoperabilidade dos dispositivos em rede.

A Internet das Coisas, preconiza Lacerda (2015), provê vários benefícios para a sociedade, concebendo efeitos significativos nas áreas de meio ambiente, saúde, comunicação, segurança, comodidade e urbanismo, uma vez que as aplicações são tantas quantas forem possíveis de se imaginar ao associar-se objetos com informações. Desta maneira, uma das aplicações sugeridas e investigadas neste artigo, diz respeito à eficiência energética.

EFICIÊNCIA ENERGÉTICA

De acordo com Roméro e Reis (2012), a energia está presente em todas as atividades humanas e, como não se encontra disponível de maneira direta na natureza - sendo obtida por meio de transformações de recursos naturais - seu consumo de forma adequada manifesta-se como um dos requisitos essenciais para a construção de um modelo de desenvolvimento sustentável. Burattini (2008) defende que a busca pela eficiência energética deve fazer parte consciente de todas as ações do ser humano moderno.

Para Pereira (2010), a utilização de energia tem sido intensiva e crescente desde a Revolução Industrial, uma vez que sua adoção é essencial para o funcionamento dos mais diversos setores e atividades da sociedade. Como

exemplo, a autora cita a energia elétrica, responsável pelo funcionamento da maior parte dos equipamentos de edificações residenciais, comerciais e públicas. Na atualidade, complementa Oliveira (2015), emprega-se a energia elétrica para aquecer, refrescar, iluminar, preparar e conservar alimentos, gerenciar informações, entre outros.

Em termos de suprimento energético, sustenta Palz (2002), a eletricidade é o alicerce sobre o qual se encontra edificada a indústria eletrônica, a indústria dos aparelhos elétricos, uma considerável parte dos transportes públicos, a informática, bem como a indústria automobilística. Neste contexto, Burattini (2008, p. 47), reforça que, "[o] crescente desenvolvimento tecnológico envolvendo dispositivos que utilizam a eletricidade para seu funcionamento fez aumentar a demanda de energia elétrica".

Diversos são os meios de produzir energia elétrica, a partir de fontes renováveis ou não renováveis, conforme exemplificado na Figura 3, baseada na matriz elétrica brasileira, cada qual com suas vantagens e desvantagens econômicas e ambientais. Conforme Palz (2002), qualquer forma de energia pode ser transformada em eletricidade.

Uma fonte de energia é denominada renovável, destacam Goldemberg e Lucon (2012), quando as condições naturais possibilitam sua reposição em um curto espaço de tempo como, por exemplo: a energia solar, a energia eólica, a energia hidráulica e a biomassa (lenha, carvão vegetal, resíduos orgânicos). Em contrapartida, complementam os autores, as fontes não renováveis de energia são aquelas que a natureza não tem condições de repor em um horizonte de tempo compatível com seu consumo pelos seres humanos, como o carvão mineral, petróleo, gás natural e urânio.

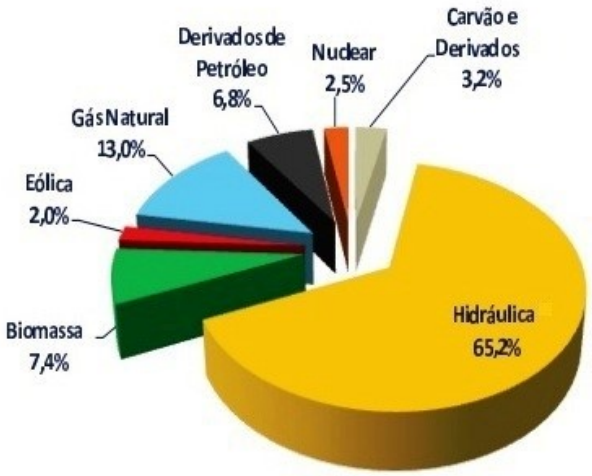


Figura 3. Matriz elétrica brasileira.
Fonte: Empresa de Pesquisa Energética [EPE] (2015, p. 35).

Ressalta-se a colaboração de Roméro e Reis (2012), que entendem que as fontes energéticas impactam com maior ou menor intensidade o meio ambiente, intensificando a necessidade de adoção de tecnologias menos poluentes sob qualquer aspecto. Neste cenário, Trsic e Fresqui (2012) ponderam que, o planeta Terra é um organismo vivo, sendo difícil introduzir novos corpos (tecnologias e afins) sem provocar algum tipo de reação.

Além do exposto, novas tecnologias, salientam Lamberts, Ghisi, Pereira, e Batista (2010) e Oliveira (2015), também contribuem para a Eficiência Energética. Dentre estas tecnologias, sugere-se a aplicação do paradigma de Internet das Coisas ao conceito de “eficiência energética”.

METODOLOGIA

Uma pesquisa consiste de uma investigação sistemática, que objetiva desenvolver teorias, estabelecer evidências e resolver problemas. É fundamental que o pesquisador esteja suficientemente informado do que foi pesquisado, como foi pesquisado, que resultados foram encontrados e o que ainda não foi pesquisado (Morandi & Camargo, 2015).

Assim, para determinar a confiabilidade da pesquisa quanto aos seus resultados e análises, optou-se pela adoção da revisão sistemática da literatura e da análise quantitativo-qualitativa. Neste primeiro aspecto, Sampieri, Collado, e Lucio (2013, p. 76) ressaltam que a revisão sistemática, "[c]onsiste em detectar, consultar e obter a bibliografia e outros materiais úteis para os propósitos do estudo, dos quais extraímos e sintetizamos a informação relevante e necessária para o problema de pesquisa.". Os autores salientam que este procedimento

deve ser seletivo, uma vez que todo ano, em diversas partes do mundo são publicados milhares de artigos em revistas acadêmicas, periódicos, livros e outros tipos de materiais nas diferentes áreas do conhecimento.

De acordo com Freire (2013), trata-se de um processo de levantamento de dados onde são exigidas revisões rigorosas de publicações acadêmicas em busca de indícios que possam levar à identificação de evidências a respeito de um tema de pesquisa ou mesmo um tópico em uma área desejada. Este método, ratificam Morandi e Camargo (2015), proporciona uma visão robusta e abrangente, permitindo que os pesquisadores se mantenham a par do que tem sido estudado em suas áreas de interesse.

Para a elaboração desta pesquisa, adotou-se a plataforma Scopus (www.scopus.com), base internacional responsável por publicações científicas de caráter multidisciplinar. Inicialmente realizou-se a análise estatística dos documentos, definindo-se como ponto de corte o período compreendido a partir de 2001 até 2016, uma vez que corresponde ao início do século XXI, englobando um universo de pesquisa de 15 anos, considerado como satisfatório para análises desta natureza. Para a realização da análise descritiva dos documentos, tendo como base este contexto, efetuou-se a escolha de 50 publicações utilizando como critério o número de citações que cada uma recebeu, em ordem decrescente. Por conseguinte, com base na leitura do resumo de cada obra selecionada, elencaram-se os documentos que apresentaram maior aderência ao assunto, excluindo-se aqueles que continham informações voltadas a universos diferentes do objeto de pesquisa proposto.

RESULTADOS E DISCUSSÕES

Como procedimento inicial, optou-se pela análise estatística dos documentos inserindo-se, na janela de busca da plataforma Scopus, o termo IoT, abreviatura para *Internet of Things*. A pesquisa reportou a ocorrência de 8.902 publicações, distribuídas segundo a Figura 4.

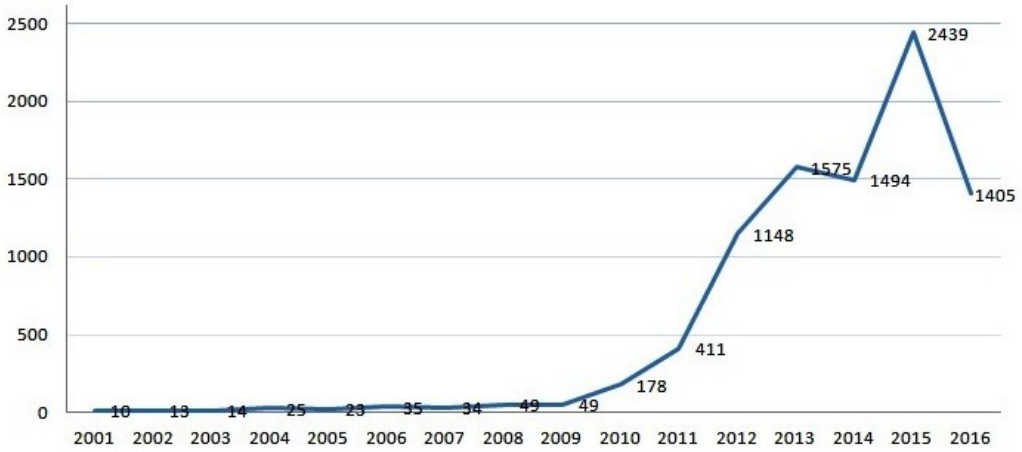


Figura 4. Evolução das pesquisas relacionadas à Internet das Coisas – Base Scopus (2001-junho/2016).
Fonte: os autores.

Observa-se, apesar do termo ter sido utilizado pela primeira vez no ano de 1999, que apenas em 2010 iniciou-se o efetivo interesse pelo assunto, ocorrendo uma ligeira queda em 2014, sendo que o ápice ocorre em 2015. Por 2016 ser o ano em curso, ainda não há dados conclusivos sobre a pesquisa. Porém pela evolução do gráfico, acredita-se que esse número possa superar o de 2015.

Na seguinte pesquisa na base de dados Scopus, inseriram-se os termos *energy efficiency*, *light efficiency* e *house efficiency*, todos relacionados à "eficiência energética". Ressalta-se que essa busca ainda não está associada à Internet das Coisas e sim aos assuntos e aplicações de uma maneira genérica. Como resultado, obteve-se a quantia de 346.761 documentos, conforme apresenta a Figura 5. Estes dados exibem valores bem mais expressivos quando comparados aos anteriores, relacionados à Internet das Coisas, indicando que a "eficiência energética" desperta considerável interesse por parte dos pesquisadores, pelo menos quantitativamente, devido, provavelmente, a ser um tema mais antigo. Todavia, conforme a Figura 5, em termos percentuais, as pesquisas voltadas diretamente à Internet das Coisas, apresentam evolução no crescimento a cada ano.

Nota-se também, um aumento constante no número de documentos publicados a cada ano e, tal qual a pesquisa anterior, o ano de 2015 contém uma maior quantidade quando comparada com os demais. Da mesma forma, como 2016 é o ano em curso, a quantidade de documentos ainda é reduzida, devido a não totalidade dos dados.

Posteriormente, visando filtrar as publicações associadas ao tema proposto, realizou-se a pesquisa associando-se o termo *Internet of Things* aos termos *energy efficiency*, *light efficiency* e *house efficiency*. A pesquisa culminou com a exibição de 895 documentos, distribuídos de acordo com a Figura 6. Tal qual a Figura 4, relacionado aos

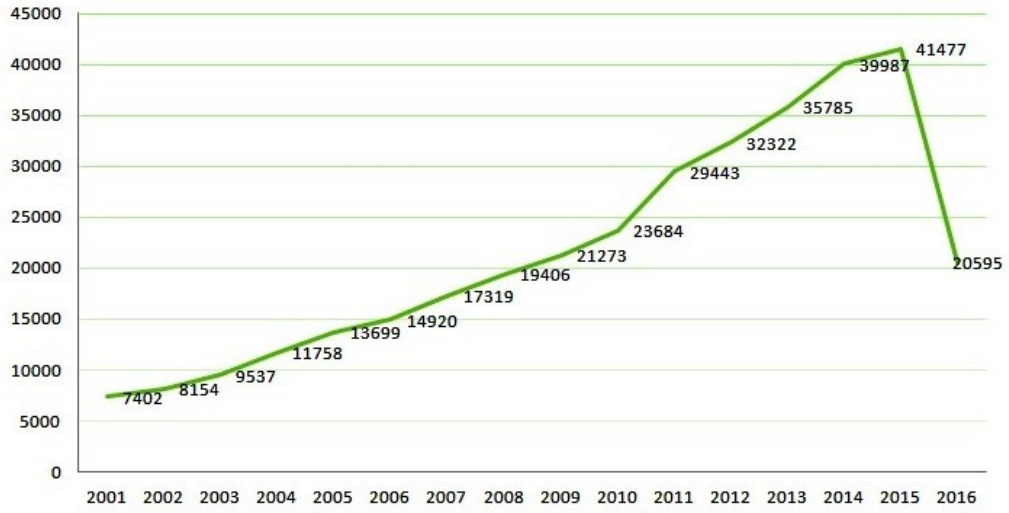


Figura 5. Evolução das pesquisas relativas à "eficiência energética" – Base Scopus (2001-junho/2016).
Fonte: os autores.

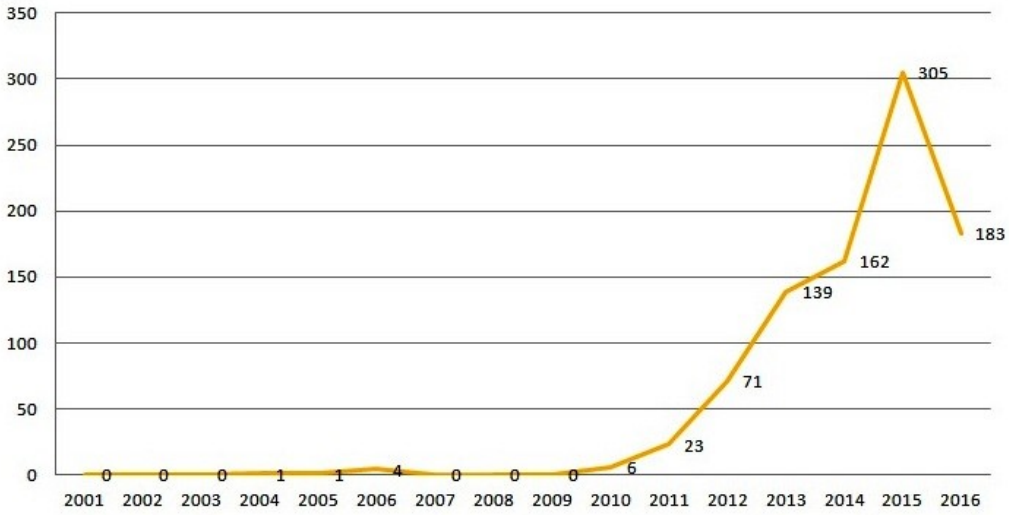


Figura 6. Publicações relacionadas à Internet das Coisas associada à "eficiência energética" – Base Scopus (2001-junho/2016).
Fonte: os autores.

estudos a respeito da Internet das Coisas, observa-se uma evolução das pesquisas a partir do ano de 2010, sendo que em 2015 ocorrem as maiores quantidades.

A Figura 7 apresenta a distribuição de publicações por países, considerando os dez que mais contribuíram.

Constata-se uma predominância da China, com quase o dobro em relação aos Estados Unidos, segundo colocado. É interessante frisar que os dois, em conjunto, são responsáveis por quase metade (cerca de 43%) de todas as publicações reportadas. Em relação ao Brasil, foram encontradas ao total, onze publicações, fazendo com que o país ocupe a vigésima posição.

A Figura 8 contém dados das publicações de acordo com seu tipo. Os *papers*, classificados como documentos publicados em conferências e seminários internacionais, são responsáveis pela maior quantidade de documentos, com 53,85% das publicações, seguidos pelos artigos, documentos publicados em revistas e sites especializados, com 38,66%. A soma dos dois tipos de documento totaliza 92,51%, sendo que o restante é distribuído entre as demais formas.

Em se tratando de áreas de conhecimento, a Figura 9 apresenta os resultados obtidos.

A maioria dos documentos está associada às áreas de Computação, Engenharia e Matemática, sendo que apenas uma pequena parte demonstra-se atrelada especificamente à área energética. Atenta-se ao fato, que cada documento pode estar associado a mais de uma área específica.

O procedimento seguinte diz respeito à análise descritiva das publicações elencadas, conforme a quantidade de citações recebidas. Em termos de datas, o primeiro documento publicado intitula-se *An improved concept for a Higher-Order Mode IOT* e foi escrito por Bohlen e Wright (2004). No artigo, que não recebeu nenhuma citação,

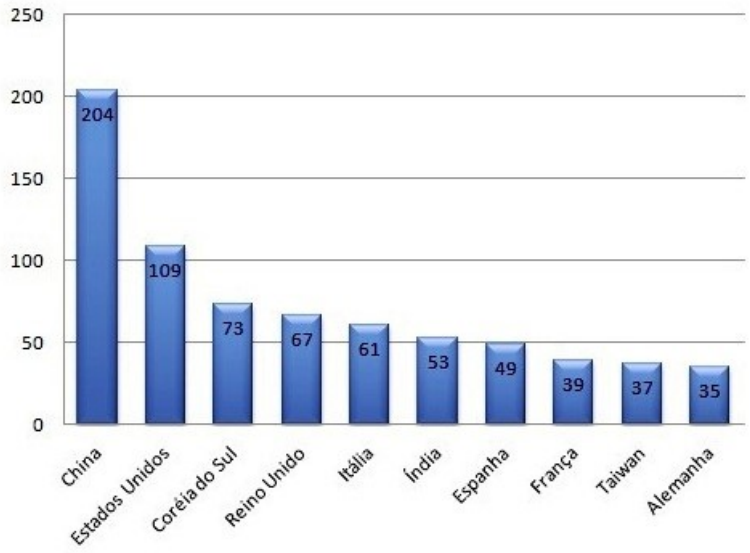


Figura 7. Publicações conforme o país de origem – Base Scopus (2001-junho/2016).
Fonte: os autores.

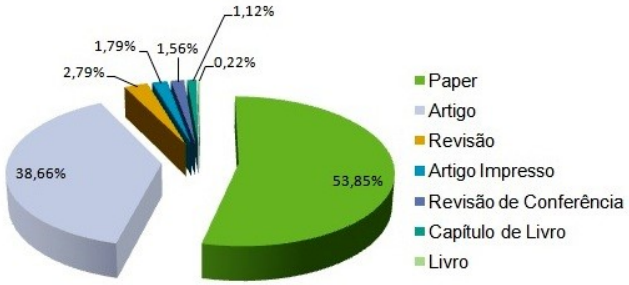


Figura 8. Publicações segundo o tipo – Base Scopus (2001-junho/2016).
Fonte: os autores.

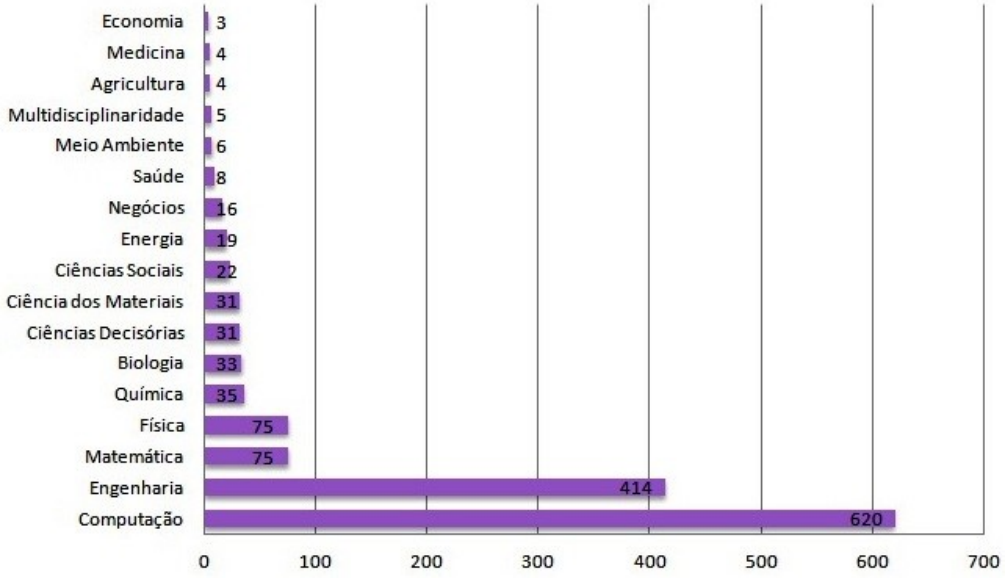


Figura 9. Publicações conforme a área de conhecimento – Base Scopus (2001-junho/2016).
Fonte: os autores.

os autores comentam sobre o desenvolvimento de fontes conversoras de voltagem, com eficiência energética na ordem de 62%, a serem utilizadas em dispositivos eletrônicos.

O artigo *Recommender systems survey*, de Bobadilla, Ortega, Hernando, e Gutiérrez (2013), recebeu 259 citações, ficando em primeiro lugar neste quesito. O foco do documento concentra-se em sistemas de recomendação e

sua afinidade com a Internet das Coisas, não fazendo menção à "eficiência energética".

O trabalho de Sheng et al. (2013), denominado *A survey on the ietf protocol suite for the internet of things: standards, challenges, and opportunities*, apresenta aspectos relacionados à Internet das Coisas, enfatizando sua estrutura, principalmente os protocolos e tecnologias afins, como o IPv6, 6LoWPAN, 802.15.4 e ZigBee. Pontua-se também, sobre a necessidade da adoção de dispositivos que manifestem um reduzido consumo energético. O artigo apresenta uma quantidade de 101 citações a ele associadas. Nesta mesma linha de raciocínio, com considerações consonantes, está o artigo de Zeng, Guo, e Cheng (2011), intitulado *The Web of Things: A survey*, com 84 citações.

Aziz, Sekercioglu, Fitzpatrick, e Ivanovich (2013), no artigo *A survey on distributed topology control techniques for extending the lifetime of battery powered wireless sensor networks*, explanam sobre a utilização de energia em dispositivos associados à Internet das Coisas, principalmente em "redes de sensores sem fio". Segundo os autores, o uso eficiente deste recurso é fundamental para a sua operação e está condicionado à topologia da rede. Desta maneira, o artigo relata características, vantagens e desvantagens, demonstrando algoritmos que visam obter o máximo de eficiência da rede, subtraindo a menor quantidade possível de energia das baterias dos sensores. O documento foi citado 87 vezes.

O artigo *How low energy is Bluetooth low energy? Comparative measurements with ZigBee/802.15.4*, de Siekkinen, Hiienkari, Nurminen, e Nieminen (2012) é citado 47 vezes, faz um comparativo entre as tecnologias Bluetooth e ZigBee em relação ao consumo de bateria, enfatizando a tecnologia denominada BLE (*Bluetooth Low Energy*), considerada pelos autores, extremamente eficaz neste quesito. Como ponto negativo, existem restrições quanto à pilha de protocolos adotada pela BLE.

O trabalho de Hancke, de Carvalho e Silva, e Hancke Junior (2012), citado 46 vezes e denominado *The role of advanced sensing in smart cities*, apresenta aspectos de como a Internet das Coisas, juntamente com as redes de sensores sem fio podem contribuir para a sustentabilidade e o uso eficiente de recursos energéticos nas cidades, em áreas como: saúde, transportes, distribuição de água e energia, monitoramento ambiental e serviços públicos. Na percepção dos autores, a infraestrutura necessária para a concepção de um projeto deste porte, caracteriza-se por custos financeiros elevados e extrema complexidade, indicando a existência de consideráveis desafios a serem superados.

No artigo intitulado *Internet of Things and BOM-based life cycle assessment of energy-saving and emission-reduction of products*, escrito por Tao, Zuo, Xu, LV, e Zhang (2014), especificam-se procedimentos para reduzir o consumo de energia e minimizar a emissão de poluentes em processos industriais, utilizando a Internet das Coisas, propondo uma arquitetura de quatro níveis: acesso à percepção, dados, serviços e aplicação. Contudo, na ótica dos autores a pesquisa para produtos desta natureza está apenas começando, sendo que ainda existe um longo caminho a ser percorrido. Este artigo foi citado 15 vezes.

Huang, Meng, Gong, Liu, e Duan (2014), sugerem através do artigo *A novel deployment scheme for Green Internet of Things*, um ambiente voltado para a chamada Internet das Coisas Verde (*Green IoT*). Este artigo recebeu 14 citações e conforme os autores, o modelo proposto demonstra-se mais eficaz no consumo de energia quando comparado às tradicionais redes de sensores sem fio.

Em se tratando de documentos oriundos do Brasil, o artigo de Kamienski et al. (2015), denominado *Context-aware energy efficiency management for smart buildings*, discorre sobre a necessidade de aplicar conceitos relacionados à "eficiência energética" em prédios públicos, devido à estas edificações frequentemente serem desprovidas de mecanismos desta natureza. Para tal, os autores preconizam o uso da Internet das Coisas. O texto apresenta detalhes de como implementar esta filosofia e demonstra resultados obtidos através do desenvolvimento de um protótipo, baseado no microcontrolador Arduino, concebido para esta finalidade e contextualizado na sala de aula de uma universidade.

CONCLUSÕES

Neste artigo objetivou-se investigar a evolução das pesquisas relativas a associação da Internet das Coisas à Eficiência Energética, apresentando definições, aplicações e características inerentes a estes dois conceitos, através da adoção dos métodos de revisão sistemática da literatura e análise quantitativo-qualitativa.

A respeito do paradigma de Internet das Coisas (IoT), um ponto importante atenta-se justamente à definição e abrangência do termo. Contemplam-se incertezas quanto à maneira, intensidade e o contexto de sua concepção. Os autores pesquisados discordam em aspectos técnicos, e frequentemente defendem sua aplicação em circunstâncias pontuais, voltadas a cenários peculiares, demonstrando divergências em relação à essência do conceito.

Quanto aos aspectos inerentes à "eficiência energética", evidencia-se a dependência do ser humano pela eletricidade, sendo esta, fundamental para o desenvolvimento econômico, social e para o bem-estar. Demonstra-se premente, a adoção de métodos de geração limpos, menos poluentes e agressivos ao meio ambiente, baseados

em fontes renováveis. Tal qual estas medidas, a modificação dos hábitos de consumo da população e o incremento de novas tecnologias, se mostra relevante para alcançar resultados sustentáveis na utilização da energia.

Os dados obtidos pela revisão sistemática da literatura e pela análise bibliométrica, através da junção dos dois conceitos, sinalizam um extenso volume de publicações, indicando a relevância do tema. Em contrapartida, constata-se que os documentos discorrem acerca do uso eficiente de energia nos dispositivos, principalmente os sensores, associados à Internet das Coisas, apresentando pouca ênfase, em comparação à quantidade total de publicações existentes, na aplicação efetiva da tecnologia ao conceito de "eficiência energética", vislumbrando-se desta maneira, um considerável universo investigativo a ser perscrutado.

REFERÊNCIAS

Atzori, L., Iera, A., & Morabito, G. (2010, Oct.). The Internet of Things: A survey. *Computer Networks*, *54*(15), 2787–2805. doi: 10.1016/j.comnet.2010.05.010

Aziz, A. A., Sekercioglu, Y. A., Fitzpatrick, P., & Ivanovich, M. (2013). A survey on distributed topology control techniques for extending the lifetime of battery powered wireless sensor networks. *IEEE Communications Surveys & Tutorials*, *15*(1), 121–144. doi: 10.1109/SURV.2012.031612.00124

Barros, T. H. C. (2011). *Tecnologias da informação e comunicação (TICs) na educação: professores - mediadores - mentores* (Dissertaçao de Mestrado em Meios e Processos Audiovisuais, Universidade de São Paulo, Sao Paulo). Recuperado de http://www.teses.usp.br/teses/disponiveis/27/27161/tde-12032013-161611/pt-br.php

Bobadilla, J., Ortega, F., Hernando, A., & Gutiérrez, A. (2013, July). Recommender systems survey. *Knowledge-Based Systems*, *46*, 109–132. doi: 10.1016/j.knosys.2013.03.012

Bohlen, H. P., & Wright, E. L. (2004). An improved concept for a higher-order mode IOT. In *Fifth IEEE International Vacuum Electronics Conference* (p. 311–). doi: 10.1109/IVELEC.2004.1316335

Burattini, M. P. T. C. (2008). *Energia: uma abordagem multidisciplinar*. São Paulo: Livraria da Física.

Empresa de Pesquisa Energética. (2015). *Balanço Energético Nacional 2015 – Ano base 2014: Relatório Síntese*. Rio de Janeiro: EPE. Recuperado de https://ben.epe.gov.br/downloads/S%C3%ADntese%20do%20Relat%C3%B3rio%20Final_2015_Web.pdf

Ferreira, H. G. C. (2014). *Arquitetura de Middleware para Internet das Coisas* (Dissertação de Mestrado em Engenharia Elétrica, Universidade de Brasília, Brasília). Recuperado de http://repositorio.unb.br/handle/10482/17251

Freire, P. S. (2013). *Aumente a qualidade e a quantidade de suas publicações científicas: manual para elaboração de projetos e artigos científicos*. Curitiba: Crv.

Gogliano Sobrinho, O. (2013). *Serviço de resolução e descoberta de informações sobre objetos em sistemas baseados em RFID* (Tese de Doutorado em Engenharia, Universidade de São Paulo, Sao Paulo). Recuperado de http://www.teses.usp.br/teses/disponiveis/3/3141/tde-16102013-162918/pt-br.php

Goldemberg, J., & Lucon, O. (2012). *Energia, meio ambiente e desenvolvimento* (3a. ed.). São Paulo: EDUSP.

Hancke, G. P., de Carvalho e Silva, B., & Hancke Junior, G. P. (2012). The role of advanced sensing in smart cities. *Sensors*, *13*(1), 393–425. doi: 10.3390/s130100393

Hinrichs, R. A., Kleinbach, M., & Reis, L. B. (2015). *Energia e meio ambiente* (5a. ed.). São Paulo: Cengage Learning.

Huang, J., Meng, Y., Gong, X., Liu, Y., & Duan, Q. (2014). A novel deployment scheme for Green Internet of Things. *IEEE Internet of Things Journal*, *1*(2), 196–205. doi: 10.1109/jiot.2014.2301819

Kamienski, C., Borelli, F., Biondi, G., Rosa, W., Pinheiro, I., Zyrianoff, I., … Pramudianto, F. (2015). Context-aware energy efficiency management for smart buildings. In *2015 IEEE 2nd World Forum on Internet of Things (WF-IoT)* (p. 699–704).

Kurose, J. F., & Ross, K. W. (2013). *Redes de computadores e a internet: uma abordagem top-down* (6a. ed.). São Paulo: Pearson.

Lacerda, F. (2015). *Arquitetura da informação pervasiva: projetos de ecossistemas de informação na internet das coisas* (Tese de Doutorado em Ciência da Informação, Universidade de Brasília, Brasília). Recuperado de http://repositorio.unb.br/handle/10482/19646

Lacerda, F., & Lima-Marques, M. (2015). Da necessidade de princípios de arquitetura da informação para a internet das coisas. *Perspectivas em Ciência da Informação*, *20*(2), 158–171. doi: 10.1590/1981-5344/2356

Lamberts, R., Ghisi, E., Pereira, C. D., & Batista, J. O. (2010). *Casa eficiente: consumo e geração de energia*. Florianópolis: UFSC LabEEE.

Morandi, M. I. W. M., & Camargo, L. F. R. (2015). Revisão sistemática da literatura. In A. Dresch, D. P. Lacerda, & J. A. V. Antunes Junior (Eds.), *Design science research: método de pesquisa para avanço da ciência e tecnologia* (p. 141–172). Porto Alegre: Bookman.

Nascimento, M. H. R. (2011). *As tecnologias da informação e comunicação - TICs influenciam o desempenho acadêmico?* (Dissertação de Mestrado em Economia, Universidade Católica de Brasília, Brasília). Recuperado de https://bdtd.ucb.br:8443/jspui/bitstream/123456789/454/1/Manoel%20Henrique%20Reis%20Nascimento.pdf

Oliveira, I. P. (2015). *Eficiência energética em sistemas de energia elétrica: um estudo de caso em uma indústria de lubrificantes* (Dissertaçao de Mestrado em Eficiência Energética e Sustentabilidade, Universidade Federal de Mato Grosso do Sul, Campo Grande). Recuperado de http://repositorio.cbc.ufms.br:8080/jspui/bitstream/123456789/2602/1/IGORPETRIDEOLIVEIRA.pdf

Palz, W. (2002). *Energia solar e fontes alternativas*. Curitiba: Hemus.

Pereira, C. D. (2010). Introdução. In R. Lamberts, E. Ghisi, C. D. Pereira, & J. O. Batista (Eds.), *Casa eficiente:consumo e geração de energia* (p. 17–21). Florianópolis: UFSC LabEEE.

Pinho, J. A. G. (2011). Sociedade da informação, capitalismo e sociedade civil: reflexões sobre política, internet e democracia na realidade brasileira. *RAE: Revista de Administração de Empresas*, *51*(1), 98–106. doi: 10.1590/S0034-75902011000100009

Roméro, M. A., & Reis, L. B. (2012). *Eficiência energética em edifícios*. Barueri: Manole.

Sampieri, R. H., Collado, C. F., & Lucio, M. P. B. (2013). *Metodologia de pesquisa* (5a. ed.). Porto Alegre: Penso.

Sato, S. K. (2015). *Mobilidade, comunicação e consumo: expressões da telefonia celular em Angola, Brasil e Portugal* (Tese de Doutorado em Comunicação, Universidade de São Paulo, São Paulo). Recuperado de http://www.teses.usp.br/teses/disponiveis/27/27153/tde-01062015-163806/pt-br.php

Serafim, E. (2014). *Uma estrutura de rede baseada em tecnologia IoT para atendimento médico a pacientes remotos* (Dissertação de Mestrado em Ciência da Computação, Faculdade Campo Limpo Paulista, Campo Limpo Paulista). Recuperado de http://www.cc.faccamp.br/Dissertacoes/Edivaldo_2014.pdf

Sheng, Z., Yang, S., Yu, Y., Vasilakos, A. V., Mccann, J. A., & Leung, K. K. (2013). A survey on the ietf protocol suite for the internet of things: standards, challenges, and opportunities. *IEEE Wireless Communications*, *20*(6), 91–98. doi: 10.1109/mwc.2013.6704479

Siekkinen, M., Hiienkari, M., Nurminen, J. K., & Nieminen, J. (2012). How low energy is bluetooth low energy? comparative

measurements with ZigBee/802.15.4. In *2012 IEEE Wireless Communications and Networking Conference Workshops* (p. 232–237). doi: 10.1109/wcncw.2012.6215496

Tao, F., Zuo, Y., Xu, L., LV, L., & Zhang, L. (2014). Internet of Things and BOM-based life cycle assessment of energy-saving and emission-reduction of products. *IEEE Transactions on Industrial Informatics*, *10*(2), 1252–1261. doi: 10.1109/tii.2014.2306771

Trsic, M., & Fresqui, M. C. (2012). *Curso de química para engenharia*. Barueri: Manole.

Wu, G., Talwar, S., K., J., Himayat, N., & Johnson, K. D. (2011). M2M: from mobile to embedded internet. *IEEE Industrial Electronics Magazine*, *49*(4), 36–43. doi: 10.1109/mcom.2011.5741144

Yu, X., Cecati, C., Dillon, T., & Simões, M. G. (2011). The new frontier of smart grids. *IEEE Industrial Electronics Magazine*, *5*(3), 49–63. doi: 10.1109/mie.2011.942176

Zeng, D., Guo, S., & Cheng, Z. (2011). The Web of Things: A survey. *Journal of Communications*, *6*, 424–438. doi: 10.4304/jcm.6.6.424-438

Como citar este artigo (APA):
Sônego, A. A., Marcelino, R. & Gruber, V. (2016). A Internet das Coisas aplicada ao conceito de eficiência energética: uma análise quantitativo-qualitativa do estado da arte da literatura. *AtoZ: novas práticas em informação e conhecimento, 5*(2), 80 – 90. Recuperado de: http://dx.doi.org/10.5380/atoz.v5i2.47860

AtoZ
novas práticas em informação e conhecimento

ISSN: 2237-826X
ARTIGO | PAPER

Rede de coautoria da Universidade Federal de Goiás: o impacto de publicações internacionais

Co-authorship network at Federal University of Goiás: the impact of international publications

Eduardo Alves Silva[1], Dalton Lopes Martins[1]

[1] *Universidade Federal de Goiás - UFG, Goiânia, GO, Brasil*

Autor para correspondência/*Mail to:* Eduardo Alves Silva (easilva91@gmail.com)

Recebido/*Submitted*: 28 Jul. 2016; **Aceito/*Approved*:** 20 Out. 2016

Copyright © 2016 Silva & Martins. Todo o conteúdo da Revista (incluindo-se instruções, política editorial e modelos) está sob uma licença Creative Commons Atribuição-NãoComercial-CompartilhaIgual 3.0 Não Adaptada. Ao serem publicados por esta Revista, os artigos são de livre uso em ambientes educacionais, de pesquisa e não comerciais, com atribuição de autoria obrigatória. Mais informações em http://revistas.ufpr.br/atoz/about/submissions#copyrightNotice.

Resumo

Introdução: Propõe-se demonstrar a construção da coautoria dentro da Universidade Federal de Goiás (UFG) e o possível impacto da internacionalização dessa produção.

Método: foi realizada a extração de dados referentes a publicações da UFG na base de dados Web of Science (WoS) no período histórico de dezesseis anos (1999-2015), efetivando-se o tratamento do nome dos coautores das publicações de forma a construir a rede de coautoria com o uso de metodologia de análise de redes sociais (ARS).

Resultados: obteve-se a rede de coautoria de maior impacto internacional da UFG respondendo à questão sobre a existência da rede e suas características, assim como percepções sobre a internacionalização e desenvolvimento da instituição de acordo com a rede.

Conclusão: A criação da rede de coautoria demonstra a divisão de grupos e fatores que a compõem, introduzindo dessa forma contextos que vão além da coautoria, como a internacionalização e o uso das informações obtidas na gestão da pesquisa cientifica em uma instituição.

Palavras-chave: Análise de Redes Sociais; Internacionalização acadêmica; Comunicação científica

Abstract

***Introduction**: proposes to demonstrate the construction of co-authoring within the Federal University of Goiás (UFG) and the possible impact of this internationalization production*

***Method: a**: data extraction (concerning the temporal arch of 1999 to 2015) was performed at Web of Science (WoS) database in order to retrieve UFG publications. The items retrieved were treated – meanly the co-authors names - to build the network of co-authorship using social network analysis methodology (SNA).*

***Results**: the co-authorship network showed the international impact of UFG which answered the question about the existence of such network and its characteristics, as well as perceptions and characteristics of internationalization and the development of the network .*

***Conclusions**: The creation of co-authorship network demonstrates cluster divisions and related features that compose it, thus introducing contexts beyond the co-author realm, such as internationalization and the use of information obtained in the management of scientific research in an institution.*

***Keywords:** Social Network Analysis; Academic internationalization; Scientific communication.*

INTRODUÇÃO

A utilização da análise de redes sociais (ARS) como método para ampliar a visão a respeito de diferentes informações e fatores que se encontram nos elementos da rede, traz possibilidades para a e análise de redes de coautoria, uma vez que esse gênero de rede se caracteriza pela colaboração entre dois ou mais autores que fazem uma publicação em conjunto. Neste tipo de rede as relações são representadas por um conjunto de nodos que estão interconectados por setas, e cuja de sua análise demonstram, univocamente, áreas e contextos de seus integrantes.

Atualmente a análise de redes de coautoria, assim como o estudo da internacionalização de produção científica servem como indicadores de desenvolvimento de um país. Como mencionado por Fiorin (2007, p. 264) "é a inserção global que indica a participação de um país na produção científica mundial", ou no caso aqui relatado, o desenvolvimento de uma instituição. Por vezes, as redes de coautoria dizem respeito a pós-graduação ou áreas isoladas, sendo que o estudo de redes de coautoria em uma instituição permite agregar distintas áreas de conhecimento e mapear a internacionalização de sua produção como um todo.

A análise de redes sociais é definida por Marteleto (2001, p. 72), como um "sistema de nodos e elos; uma estrutura sem fronteiras, uma comunidade não geográfica; um sistema de apoio ou um sistema físico que se pareça com uma árvore ou uma rede". Este sistema traz uma abordagem expressiva que representa o todo em um meio social, e "[...]estuda as ligações relacionais (da expressão em inglês *relational tie*) entre atores sociais" (Matheus & Silva, 2006, p. 1).

A construção dessas ligações relacionais, tendo como foco a coautoria de uma instituição, pode demonstrar o perfil de produção científica institucional, que se relaciona com a área do conhecimento de maior impacto na

rede, eventualmente revelando quem são os autores que se destacam (centralidade na rede), os que possuem mais conexões, aqueles que mais se articulam com outros autores, e aqueles que mais se aproximam dos demais, criando caminhos de circulação de informação e conhecimento. Além disso, também é possível identificar comunidades ou agrupamentos de autores em torno de articulações sociais que podem revelar formas organizacionais de produzir pesquisa e publicar seus resultados podendo nos trazer novas maneiras de entender o modo como a ciência se organiza por meio da leitura de suas relações sociais.

Vale destacar que também se pode utilizar a análise de redes para avaliar o impacto real de uma área do conhecimento interna e externamente a instituição, auxiliando na visibilidade da internacionalização da produção científica. É possível, a partir da análise da rede, compreender seu desenvolvimento ao passar dos anos e perceber fatores externos que levaram determinada área do conhecimento a se destacar na rede de coautoria.

Embora a coautoria internacional possa refletir interesses, motivações e características pessoais da pesquisa de um pesquisador (Glänzel, 2001), a mesma auxilia na internacionalização de produção cientifica de um país e, nomeadamente, da área de estudo e instituição à qual se agrega o pesquisador.

MOTIVAÇÃO

A análise de redes sociais pode auxiliar em fatores estratégicos, oferecendo uma visão mais ampla de problemas ou relações que por vezes podem ser ricas ou problemáticas e não serem notadas por métodos tradicionais de análise de contexto social. Ao se aplicar a ARS à coautoria no contexto de uma instituição é possível que essa análise venha a servir de auxílio para os gestores de uma universidade que deseja produzir uma política de incentivo a qualificação de sua produção científica, bem como de estímulo a internacionalização dessa produção. Tem-se a possibilidade, igualmente, de usar a ARS como ferramenta analítica de apoio aos gestores nos processos de diagnóstico, planejamento e qualificação das redes e articulações já existentes nessa instituição. Ou seja, uma vez que se tem conhecimento das redes de coautoria internacional que se conectam a instituição, é possível produzir políticas de incentivo a essas redes, mobilizando recursos que por elas fluem para empoderar diferentes nodos, outras pessoas, e outros grupos para que estes criem caminhos de internacionalização, dado que os percursores já foram identificados, facilitando o processo. Durante a construção dessa pesquisa foi possível avaliar que isso pode beneficiar a gestão da própria instituição, alavancando ainda mais a internacionalização de produção e o amadurecimento do conhecimento em pesquisa e colaboração dos pesquisadores.

Como o estudo de análise de redes sociais aplicada a redes de coautoria apresenta peculiaridades devido ao impacto internacional representado na rede, pretende-se, nesse artigo, discutir a construção de relações de coautoria de todas as áreas do conhecimento de uma instituição, no caso a Universidade Federal de Goiás, e como essa produção relacional tem impactado na internacionalização do conhecimento.

ANALISE DE REDES SOCIAIS E COAUTORIA

O surgimento do conceito ou palavra rede está datada no seu aparecimento na língua francesa (*réseau*) no século XII, vindo do latim *retilous*, a rede designa, redes de caça ou peças de tecidos, uma malhagem têxtil que envolve o corpo (Musso, 2004, p. 18). Já em outra instância a palavra *network*, segundo da Silva (2014), "surge em 1908, na Conferência Internacional de Bibliografia e de Documentação, em Bruxelas, primando pelos resultados da cooperação universal disponíveis para todos, num projeto coletivo".

A conceituação e a origem das variações da palavra rede são notáveis, sendo que suas primeiras menções foram fatores que vieram a determinar a atual conceituação que se tem a respeito de redes e análise de redes sociais. Uma colocação que engloba de forma tanto o conceito de redes quanto o de ARS é a de Musso (2004), para o qual "a rede é um veículo que nos transmuda em "passantes", sempre mergulhados nos fluxos (de informações, de imagens, de sons, de dados…) ". Em termos gerais a possibilidade de se enxergar o todo torna a ARS um ambiente de compreensão social, informacional e causal, que por sua vez traz para o estudo de coautoria possibilidades valiosas de análise.

Os estudos sobre análise de redes sociais com temas voltados para colaboração e coautoria tem aumentado, como mencionado por de Fátima S. Maia e Caregnato (2008, p. 20)):

> Associado aos trabalhos sobre co-autoria, atualmente também se observa um fortalecimento do método de análise de redes sociais nos estudos sobre colaboração científica, o que possibilita uma visão abrangente das interações entre as partes, diferentemente das análises que enfocam características individuais.

Assim, é possível se constatar o interesse relacionado à colaboração cientifica e um dos seus muitos aspectos a coautoria, embora o estudo da coautoria não demonstre todas as tramas da colaboração cientifica, tal como explicitado por de Souza Vanz e Stumpf (2010, p. 45), no trecho a seguir:

> A colaboração científica aparece muitas vezes na literatura relacionada à co-autoria. Frequentemente, os dois termos são considerados sinônimos pelos pesquisadores, mas convém afirmar que a co-autoria é apenas uma faceta da colaboração científica, pois ela não mede a colaboração na sua totalidade e complexidade.

Ao ser associada à análise de redes sociais, uma rede de coautoria desperta interesse pelas possibilidades de compreensão e visualização de forma ampla de interações, harmonizando dessa forma a compreensão de questões relacionadas ao amadurecimento de uma instituição no desenvolvimento científico (e nomeadamente na pesquisa) dado o envolvimento e impacto de colaborações e estudos de seus pesquisadores. Uma vez que a coautoria ocorre por vezes entre áreas e instituições distintas, há, potencialmente, um maior impacto em publicações.

A escolha do estudo da rede de coautoria, em suma, se dá por facilitadores, tais como a possibilidade de coleta de dados na instituição, e as oportunidades da ARS como mencionadas por de Oliveira e Silva, Matheus, Parreiras, e Parreiras (2006, p. 77) no trecho a seguir:

> O uso da ARS vem crescendo significativamente nos últimos 20 anos. Tal crescimento vem ocorrendo em função do aumento da quantidade de dados disponíveis para análise, do desenvolvimento nas áreas de informática e processamento de dados – com o consequente aumento do poder computacional à disposição dos pesquisadores -, e da ampliação dos assuntos de interesse e das áreas de conhecimento que utilizam a ARS.

Além de indicadores de rede que disponibilizam a análises de fatores que circulam a rede, é possível utilizar uma forma de visualização de um ambiente único.

Para o estudo e análise da rede podem ser utilizados como métricas os indicadores de rede destacados por Wasserman e Faust (1994), além de Cherven (2015), em um livro que descreve de forma detalhada o uso da ferramenta de análise de redes complexas Gephi, que é a ferramenta utilizada no desenvolvimento dessa pesquisa. A Quadro 1 lista os indicadores utilizados e seus significados.

Indicadores	Significado
Nós	Representação de pessoas que fazem parte da rede.
Relações (links)	Número de relações existentes entre as pessoas.
Centralidade de Grau	Ator da rede com maior atividade ou relações, em relação aos demais.
Centralidade de Proximidade	Um segundo ponto de vista de centralidade baseado em proximidade ou distância. A medida incide sobre o quão perto um ator é para todos os outros intervenientes no conjunto de atores.
Centralidade de Intermediação	Demonstram nodos da rede que gera uma interação entre dois atores não adjacentes, servindo como ponte para se chegar a um determinado ator ou comunidade.
Modularidade	Abordagem estatística que mede o agrupamento numa rede, definindo comunidades na rede.
Coeficiente de Clusterização	O coeficiente de agrupamento possibilita medir em que nível os nós são agrupados em conjunto, demonstrando como os nós estão inseridos em sua vizinhança.
Componentes Conectados	Número de componentes distintos na rede. Normalmente componentes que estão separados ou isolados da rede.

Quadro 1. Indicadores de análise de redes sociais.
Fonte: Wasserman e Faust (1994); Cherven (2015).

METODOLOGIA

O processo que envolve o desenvolvimento da rede de coautoria de maior impacto internacional da Universidade Federal de Goiás dividiu-se em três etapas distintas, sendo elas a coleta e tratamento dos dados, passando pela sua estruturação e a análise dos resultados.

De forma a se estruturar toda a pesquisa, foram utilizadas as seguintes ferramentas/plataformas: a base de dados da Web of Science (WoS); o *kit Office* como ferramenta de tratamento de dados; a construção de um script na linguagem web PHP (que auxiliou no tratamento dos dados necessários para se gerar a rede de coautoria); e, o software de análise de redes complexas Gephi, que tem como papel auxiliar na visualização e análise da rede de coautoria.

Coleta e Tratamento de Dados

A utilização da Web of Science como base de dados para a coleta de informação se deu pela vasta gama de publicações armazenadas, pois "...inclui mais de 10.000 periódicos e é composto por sete bases de dados diferentes de citações, incluindo diferentes informações coletadas de revistas, conferências, relatórios, livros e séries de livros" (Chadegani et al., 2013, p. 19, tradução nossa), assim como sua facilidade de busca em relação aos dados que se mostraram necessários durante o desenvolvimento da pesquisa.

Para obtenção dos dados necessários, coletaram-se todas as publicações da Universidade Federal de Goiás dentro da WOS com o uso de algoritmo de busca demonstrado no Quadro 2, efetuando-se a coleta de todas publicações da UFG - entre o período 1999 e 2015 - resultando em 3.007 publicações (em sua maioria artigos).

Prefixo de Busca	Chave de busca
AD = Address	Universidade Federal de Goiás

Quadro 2. Algoritmo de busca para obtenção de publicações da Universidade Federal de Goiás.
Nota: AD= Nome da instituição a qual deseja que as publicações sejam pesquisadas.
Fonte: Thomsom Reuters – Web Of Science.

Uma vez que a WoS possibilita que a exportação dos resultados da da busca em planilha de apenas 500 registros por vez, foram necessárias várias exportações até se conseguir todos os registros. Ao final da exportação os dados foram reunidos em uma única planilha.

O tratamento dos dados incluiu a separação dos coautores das publicações, uma vez que as planilhas exportadas da WoS trazem todos coautores de uma publicação em uma mesma célula, separados apenas por ponto e vírgula (Quadro 3).

Publicações	**Coautores**
Publicação 1	Coautor 1; Coautor 2; Coautor 3

Quadro 3. Representação de dados da planilha exportada a partir da WoS.
Fonte: os autores.

Ao se detalhar da divisão por ponto e vírgula pretende-se justificar a etapa de separação dos coautores, a qual influencia na criação de relações para o desenvolvimento da rede de coautoria. Os coautores foram, portanto, extraídos dos dados gerais e separados entre colunas em um arquivo de formato CSV (dados separados por vírgulas).

Ao se obter o arquivo com todos os nomes devidamente separados por colunas e não mais em uma única coluna com ponto e vírgula, fez-se necessária a criação de relações entre esses coautores (Quadro 4).

Publicações	**Relações**
Publicação 1	Coautor 1 - Coautor 2 Coautor 1 - Coautor 3 Coautor 2 - Coautor 3

Quadro 4. Representação de dados da planilha exportada a partir da WoS.
Fonte: os autores.

O Quadro 4 representa as relações entre os coautores de uma publicação. Essa representação é crítica ao presente estudo devido aos fatores de análise de redes sociais que necessitam da construção de ligações para gerar a rede de coautoria. Para que a construção de laços entre coautores se desse da forma apresentada na tabela 3, de acordo com os dados já tratados, criou-se um pequeno programa computacional em PHP (um *script*) (Apêndice A), que fez a leitura do arquivo .csv (onde os nomes se encontravam), separando-os em colunas. É importante ressaltar que um dos resultados obtidos a partir do script é a possibilidade de coleta dos dados necessários para a representação da lista de arestas ou ligações exigidas no Gephi, e é similar ao demonstrado no Quadro 4.

Um fator que deve ser levado em consideração na criação da rede é o que indica que, em termos de ligações, o número é absoluto, pois a variação de nomes não interfere nas mesmas. No entanto, o número de coautores pode vir a sofrer alguma alteração decorrente da repetição de nomes (e não em termos de ligações) (Quadro 5).

Publicações	**Nomes**
Publicação 1	Eduardo, A. S. Silva, Eduardo A. A, Eduardo Silva

Quadro 5. Representação da repetição/variação de nomes de um coautor.
Fonte: os autores.

Essa variabilidade de representações de nomes pode ser um fator prejudicial para a análise da rede, uma vez que irá produzir conexões diferentes do mesmo coautor, não demonstrando de forma totalmente coerente sua importância na rede. Tal situação gera uma quantidade significativa de nomes a serem desambiguados. Efetuar tais correções seria um esforço manual intenso e, dado o tempo disponível para a pesquisa, impeditivo. Entre as diferentes formas de se resolver tal problema, a sugerida por Newman (2001) explora o seguinte processo:

> Em primeiro, nós identificamos cada autor por seu sobrenome e sua primeira inicial apenas. Este método é claramente propenso a confundir duas pessoas para uma, mas raramente irá falhar para identificar dois nomes que genuinamente se referem a mesma pessoa. [...], nós identificamos autores pelo sobrenome e suas iniciais. Este método pode distinguir com mais fiabilidade os autores um dos outros, mas também identifica uma pessoa como diferente se suas iniciais forem diferentes em diferentes trabalhos. (Newman, 2001, p. 3, tradução nossa)

Em termos de correção dos nomes, o método utilizado segue o mesmo princípio: caso um nome tenha suas iniciais similares a um nome já identificado - e não se obtenha dentro da base qualquer outro nome com essas

iniciais - se dá como definido que se trata do mesmo nome, fazendo assim sua alteração. Em casos onde existe a ocorrência de iniciais e dois ou mais sobrenomes, a busca por similaridade de nomes se dá por exclusão de sobrenomes.

No entanto, por conta da quantidade de nomes recuperados na base propôs-se, para este estudo, um outro método. Desambiguaram-se apenas os nodos que possuíam um número mínimo de conexões em rede. Ou seja, o critério adotado foi de relevância pelo impacto que esse nodo produz na rede. Aqueles nodos que tiveram um número significativo de conexões (a ser estabelecida quando da análise dos dados), foram identificados em maior detalhe, inclusive para busca de atributos qualitativos (unidade de origem, área de atuação etc.). Dessa maneira, assumiu-se como pressuposto que esse formato reduziria o tempo de trabalho e entregaria um resultado que permita interpretar as estratégias de relação entre os nodos sem perda significativa de nosso potencial de entendimento da rede.

Para se chegar aos nomes que deveriam ser tratados de acordo com o critério de relevância, propôs-se uma análise pelo corte de grau. Como todos os coautores têm um número de aparições na base de dados, esse número na rede se converte em grau de centralidade, demonstrando os coautores com mais ligações e, portanto, que podem ser considerados mais relevantes na rede. Dessa forma, a partir da contagem de nomes obteve-se o grau (número de vezes que aparecem na base) para cada autor. Verificou-se, então, quantas vezes o grau de cada coautor se repete nos dados. Após essa verificação, alguns graus podem aparecer mais de uma vez (como é o caso de coautores que se repetem apenas uma vez nos dados) o que representa a quantidade de vezes que esse nome apareceu na rede. Assim, passou-se ao corte de grau, a partir de intervalos de aproximadamente nove ou dez nomes que, somando-se o grau (consoante a sua repetição) revelou uma maior ocorrência.

Torna-se importante atentar a esta etapa, pois, como mencionado anteriormente, ao se procurar os nomes com maior relevância, ou seja, que podem aparecer mais vezes nos dados, isto indica que o seu valor de grau é muito maior que o dos outros. Sendo assim, um coautor que aparece mais na rede e que exige o tratamento de seu nome, terá maior impacto de forma geral.

Um outro tratamento exigido foi o da classificação dos periódicos que publicam os artigos coletados. Efetivou-se uma contagem dos periódicos sem que houvesse duplicação dos títulos de tais periódicos. Após a contagem fez-se a busca manual e a identificação de proveniência, sendo que, posteriormente, realizou-se a classificação entre periódicos nacionais e internacionais. Esse processo mostrou-se trabalhoso, uma vez que muitos periódicos não apresentavam identificação de país de procedência, agregavam diferentes idiomas ou não possuíam site com mais informações. Dessa forma, para contornar essas dificuldades, periódicos brasileiros com foco internacional foram classificados como internacionais e alguns periódicos foram identificados como internacionais não identificados.

Estruturação de dados e criação da rede

Após o tratamento dos nomes dos autores, passou-se para a estruturação dos dados para a criação a rede de coautoria. Como já mencionado, foram necessários dois arquivos predispostos de maneiras diferentes, uma planilha com os nodos (*nodes*) e outra com as arestas (*edges*).

Para estruturar o primeiro arquivo, a planilha de nodos, foram utilizados os nomes sem repetição, estando ordenados ou não. A primeira coluna elencava tais nomes e outra coluna, nomeada de ID, se atribuiu uma identificação numérica para esses nomes (Quadro 6).

Label	ID
Coautor 1	1
Coautor 2	2
Coautor 3	3

Quadro 6. Identificação numérica dos nomes dos coautores - exemplo.
Fonte: os autores.

O resultado no Quadro 6 é uma classificação numérica dos nomes cujos ids são únicos. Os números associados aos nomes irão estruturar as relações no arquivo de arestas, como é mostrado no Quadro 7.

Source	*Target*	*Weight*
1	2	1
1	3	1
2	3	1

Quadro 7. Planilha de arestas (*edges*).
Fonte: os autores.

A Planilha de arestas traz algumas diferenças se comparada a de nodos em decorrência das relações em si - que estão classificadas não pelos nomes, mas sim pelo ID de cada nome - além de trazer alguns classificadores

diferentes. A estrutura ocorre a partir do link entre os coautores (*source*) e para onde esse link está direcionado (*target*), gerando, assim, as arestas ou ligações entre autores. Adiciona-se a esta estrutura o indicador que atribui o valor do peso das arestas (*weight*), cujo cálculo é feito pelo próprio Gephi, demonstrando a importância ou peso de cada conexão.

Com a estruturação dos dados de base foi possível criar a rede de coautoria, importando estes arquivos para o Gephi.

RESULTADOS

Tratamento de nomes

A obtenção de resultados deu-se a partir do tratamento inicial dos dados explicitado em tópicos anteriores. Um dos primeiros resultados diz respeito ao número de coautores e ligações obtidos (Tabela 1).

Relações	Número de Coautores
48.857	8.482

Tabela 1. Relações/ligações obtidas e o número de Coautores.
Fonte: os autores.

Em seguida obteve-se novos resultados a partir da classificação de nomes dos coautores com base em seus graus, dispostos do maior para o menor grau. Os intervalos, por sua vez, foram classificados no sentido contrário, ou seja, do menor para o maior grau. A Tabela 2 representa os intervalos e a soma do grau entre esses intervalos.

Intervalo	Soma de ocorrências de grau por intervalo
0-10	6177
11-20	1782
21-30	438
31-40	163
41-50	84
51-60	42
61-70	31
71-80	28
81-90	20
91-100	18
101-110	12
111-120	11
121-125	5

Tabela 2. Classificação de ocorrência por intervalo.
Fonte: os autores.

Percebe-se que o intervalo inicial 0-10 tem uma soma maior de ocorrências, sendo que são nomes da base que ocorreram várias vezes, mas têm um grau menor, e seu tratamento seria dispendioso em termos de tempo e esforço, não resultando em impacto significativo na visualização posterior dos dados. O tratamento se deu a partir do corte de grau de 41-50, onde se encontra o declínio de ocorrência até 121-125, ou seja: todo o intervalo anterior a 41-50 tem uma repetição significativa de grau - onde, por exemplo, cerca de 6177 autores aparecem apenas uma vez na rede. Com isso houve uma redução de 329 nomes dos 8.811 iniciais, em um total final de 8.482. Embora, em termos gerais, esta redução possa parecer inócua, o real impacto está nas relações entre os nomes da rede.

De forma a exemplificar este impacto, a Figura 1 permite visualizar os vinte primeiros nomes mais expressivos derivados dos dados, omitidas as repetições/duplicações. A representatividade dos nomes e o número de aparições de cada um tem impacto direto nos indicadores de rede.

Internacionalização da produção cientifica

Quanto à internacionalização dos periódicos, verificou-se que a publicação em periódicos nacionais representa 60% do total de títulos, enquanto outros 33% são publicações em periódicos internacionais.

Como mencionado por Katz (2000) a partir do seu estudo sobre indicadores e avaliação de pesquisa, o tamanho de uma instituição ou nação, pode influenciar em relação ao nível de colaboração e coautoria internacional.

> Demonstrou-se que o reconhecimento e impacto aumentaram em uma relação de lei de potência com o tamanho de uma instituição, comunidade ou nação. Colaboração expõe uma relação de lei de potência com o tamanho de uma instituição. (Katz, 2000, p. 29, tradução nossa)

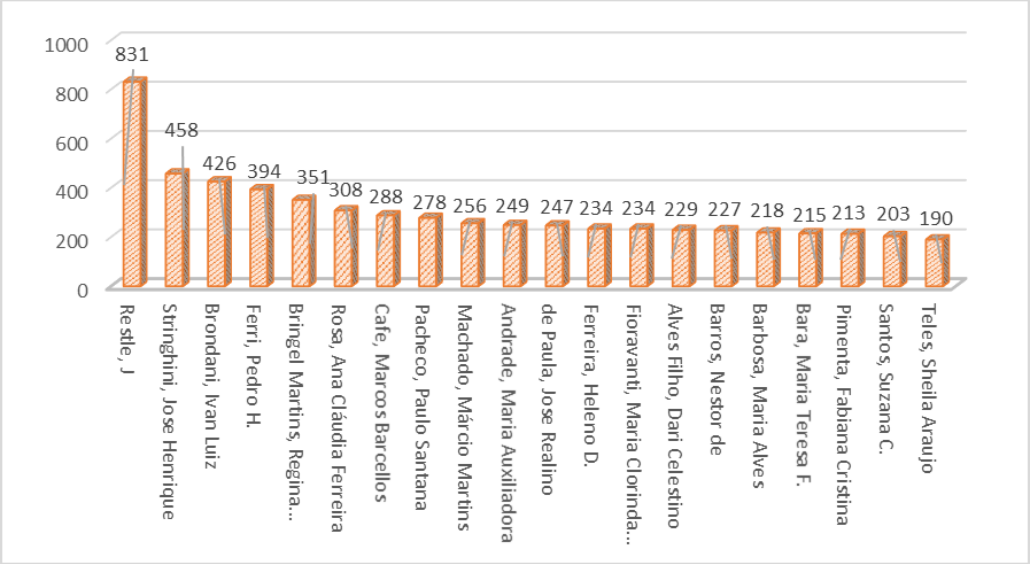


Figura 1. Frequência - nomes dos coautores – Universidade Federal de Goiás (1999-2015).
Fonte: os autores.

Dessa forma ao se relacionar esta afirmativa com os dados obtidos, algumas considerações podem ser levantadas, nomeadamente:

a) uma vez que o tamanho da instituição exerce impacto sobre o nível de colaboração/coautoria internacional, pode-se dizer que a UFG necessita avanças na internacionalização.

b) embora a afirmação anterior possa ser verdadeira, um outro fator pode ser considerado. Como nação, a extensão do território brasileiro pode levar a colaboração científica em território nacional, com instituições com melhor posicionamento nacional e mundial, como é o caso da Universidade de São Paulo.

c) outra possibilidade seria a da dificuldade de pesquisadores em publicarem com colaboração ou coautoria internacional devido a fatores tais como: interesse pessoal ; possível falta de estrutura (que pode influenciar na qualidade e coerência do trabalho, o que pode prejudicar a colaboração científica no escopo nacional e internacional).

De forma a elucidar os resultados apresentados, a Tabela 3 apresenta a variação geral em relação a internacionalização, ressaltando-se uma frequência significativa de publicações em títulos publicados no território nacional, comparativamente à soma dos demais países.

Países	Soma de ocorrências de grau por intervalo
Brasil	237
Brasil (internacional)	44
Holanda	14
USA	11
Chile	10
Internacional (n/identificado)	10
Inglaterra	7
Colômbia	6
Portugal	5
Argentina	3
Espanha	3
Alemanha	1
Canadá	1
Costa Rica	1
México	1
Peru	1
Uruguai	1

Tabela 3. Internacionalização na Universidade Federal de Goiás - quantidade publicações por títulos periódicos – ordenados por País (1999-2015).
Fonte: os autores.

Ainda que a Holanda apareça como um dos principais países que publicam os periódicos nos quais os pesquisadores da UFG divulgam suas contribuições, isso pode ser derivado da área de publicação ou mesmo a centralização desse repositório entre aqueles em que o pesquisador pode optar por publicar. Tais dados parecem indicar que a UFG - ao se levar em conta que o "peso" da produção científica de cada país, "é aquela publicada em periódicos reconhecidos mundialmente, lidos, citados" (Fiorin, 2007, p. 267) – necessita aumentar os esforços em direção

a publicações internacionais.

Entretanto, não existem ainda estudos voltados para fatores de internacionalização dentro de uma instituição especifica. Porém, em números absolutos a UFG apresenta avanços em busca da visibilidade a nível internacional, orientando-se para os processos de cooperação e internacionalização como forças motoras do desenvolvimento científico e tecnológico (Balbachevsky, 1998, p. 200).

Análise descritiva dos dados

A partir da análise feita sobre os dados obtidos, foram geradas algumas informações adicionais, tais como a média de coautoria anual da instituição, entre outros fatores. A Figura 2 representa a quantidade total de artigos publicados por ano, de acordo com os dados coletados.

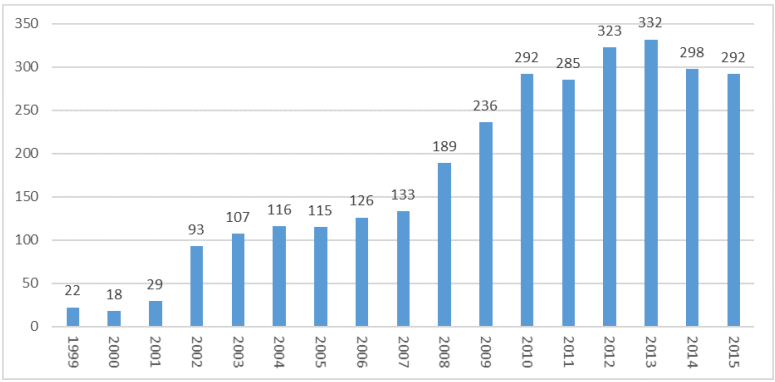


Figura 2. Quantidade de artigos publicados por ano - Universidade Federal de Goiás (1999-2015).
Fonte: os autores.

É possível perceber que, a partir dos anos de 2002, 2003, inicia-se um crescimento em relação a quantidade de artigos publicados por ano. Tal crescimento vai até o ano de 2013 havendo uma queda subsequente nos dois anos seguintes. No período entre 2003 e 2013, o crescimento pode estar relacionado ao Programa de Reestruturação e Expansão das Universidades Federais (Reuni) do Governo Federal - que esteve em atividade entre os anos de 2003 e 2012. Tal Programa se propôs a alavancar o crescimento e desenvolvimento das Universidades Federais em pontos estratégicos relacionados expansão física, acadêmica e pedagógica. Porém, tal relação necessitaria um estudo mais aprofundado. A Figura 3 representa a média de coautoria anual, da instituição durante o período estudado.

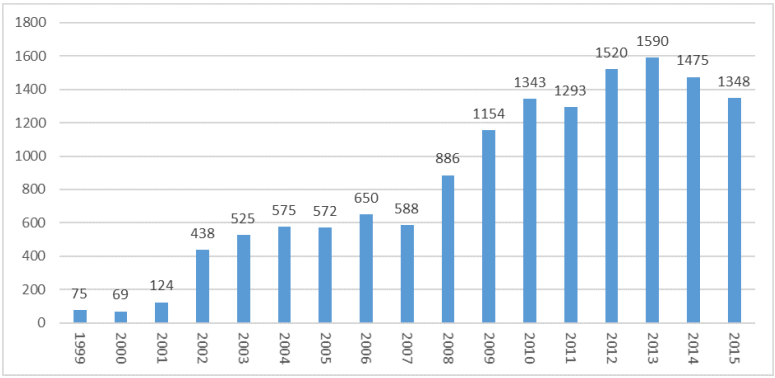


Figura 3. Número de coautores por ano - Universidade Federal de Goiás (1999-2015).
Fonte: os autores.

A similaridade da Figura 3 com a Figura 02 permite verificar que o crescimento no número de artigos normalmente está diretamente relacionado a mais autores e por consequente, mais coautores. No ano de 2013, por exemplo, há um total de 332 artigos publicados com 1.590 coautores. A Figura 4, por sua vez, representa a média de coautoria por artigo e por ano.

É possível perceber que a média de coautoria segue padrões bem definidos durante cada período, tendo um leve aumento entre 2000 e 2002, se mantendo entre 2002 e 2010, onde volta a baixar, ainda que não de forma abrupta. Embora essa média seja constante, não se pode dizer que ela realmente representa o que ocorre nos dados, pois há publicações com dois coautores (mínimo encontrado) e, publicações com o máximo de dezessete coautores. Assim, a média aqui apresentada tem o objetivo de ilustrar uma discrepância em relação aos dados.

A relação da quantidade de periódicos, na WoS, onde autores da UFG publicam a cada ano demonstra-se variada ainda que em evolução constante (Figura 5). Constata-se que alguns periódicos são comuns aos coautores da

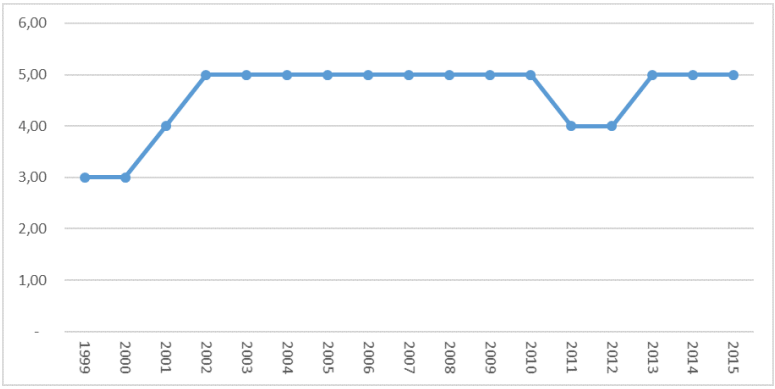


Figura 4. Média de coautoria por artigo - Universidade Federal de Goiás (1999-2015).
Fonte: os autores.

UFG, sendo que a média de publicação por periódico seja em torno de 2,12 publicações nos 382 periódicos da WoS nos quais os pesquisadores da UFG publicaram nestes 16 anos.

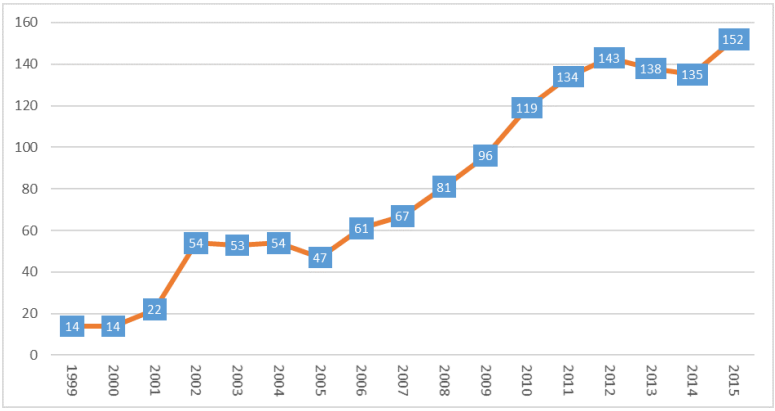


Figura 5. Média de Periódicos por Ano - Universidade Federal de Goiás (1999-2015).
Fonte: os autores.

Dentre os periódicos listados, alguns que aparecem com mais frequência na base de dados estão provavelmente ligados à área de atuação dos pesquisadores mais expressivos da rede (Tabela 4).

Periódicos	Quantidade de aparições na base
Revista Brasileira de Zootecnia	120
Ciência Rural	90
Revista da Sociedade Brasileira de Medicina Tropical	83
Memorias do Instituto Oswaldo Cruz	71
Radiologia Brasileira	63
Arquivo Brasileiro de Medicina Veterinária e Zootecnia	62
Journal Of The Brazilian Chemical Society	55
Pesquisa Agropecuária Brasileira	55
Arquivos Brasileiros de Oftalmologia	44
Ciência e Agro Tecnologia	43

Tabela 4. Periódicos mais expressivos na base de dados WoS - - Universidade Federal de Goiás (1999-2015).
Fonte: os autores.

Os periódicos mais expressivos estão voltados em sua maioria para as áreas das Ciências Agrarias, isso é um forte reflexo da localização geográfica a instituição, uma região que tem fortes características agropecuaristas.

Análise da rede de coautoria da Universidade Federal de Goiás

A Figura 6a permite visualizar - por cores e alguns *clusters* ou comunidades circuladas - a rede de coautoria da UFG. As comunidades circuladas são comunidades com ligações mais fortes entre si gerando assim atores da rede que estão mais próximos. Em alguns casos é possível reparar por conta das cores que existem outras comunidades que não se mostram tão densas, mas sim bastante dispersas na rede, isso pode ocorrer por conta do quão forte são suas ligações.

Embora a modularidade aplicada a uma rede de tal densidade possa demonstrar algumas divergências se comparada a atributos que dividem a mesma – tais como as áreas de atuação ou ano de ingresso nos dados

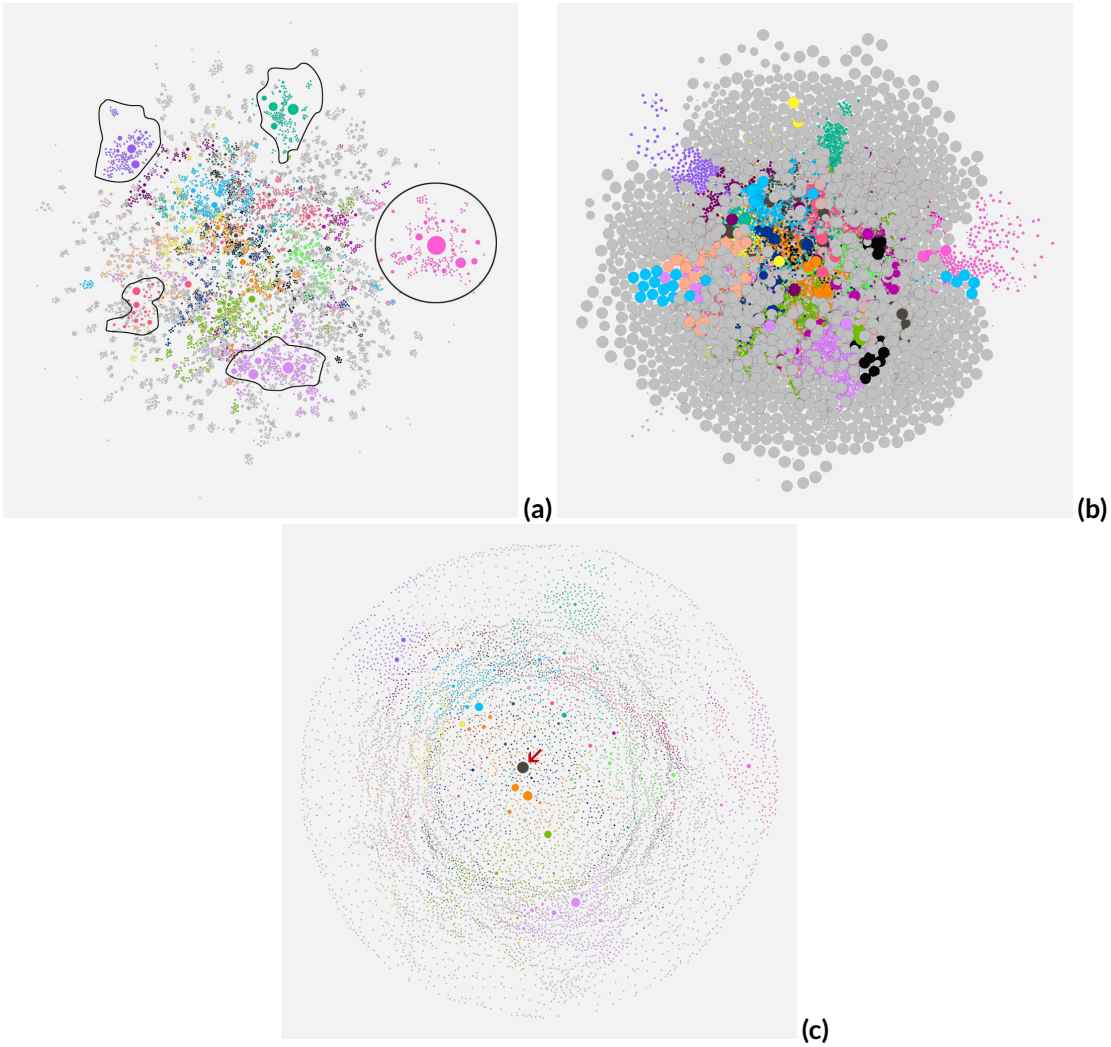


Figura 6. **(a)** Rede de Coautoria – Universidade Federal de Goiás - (1999-2015). **(b)** Rede de coautoria - centralidade de proximidade (Universidade Federal de Goiás - 1999-2015). **(c)** Rede de Coautoria - centralidade de intermediação (Universidade Federal de Goiás - 1999-2015).
Fonte: os autores.

- o número de comunidades identificadas é expressivo (total de 548 comunidades). Além disso, um segundo indicador diz respeito a centralidade de grau. Pode-se perceber que a comunidade rosa mais clara, a direita da figura, tem o nodo com maior centralidade de grau de toda a rede.

A Figura 6b demonstra que todos os nodos maiores são aqueles que estão próximos de todos os outros dentro da rede. O que torna essa visualização interessante é o fato do nodo que anteriormente tinha a maior representação enquanto a centralidade de grau, nessa visualização não demonstra tanta relevância. Ou seja, caso um ator externo queira adentrar mais a rede, talvez o nodo com maior centralidade de grau não seja a melhor opção, uma vez que - embora o mesmo tenha muitas ligações e grande importância na rede - quando se diz respeito a se conectar ou estar próximo de outros nodos da rede, ele não se demonstra tão relevante.

O último parâmetro analisado é o indicador de centralidade de intermediação, que demonstra os nodos que servem como ponte dentro da rede. Ou seja, para se chegar a uma determinada comunidade ou nodo dentro da rede, esse nodo em específico será o intermediário. Como demonstrado na Figura 6c (seta vermelha), o nodo/ator com maior grau de centralidade de intermediação dentro da rede interconecta diferentes outros atores de comunidades distintas. Em alguns casos ele se conecta a atores importantes de determinadas comunidades. Sendo assim, este é o ator da rede a ser procurado caso se queira conhecer novas áreas da rede.

Um dos indicadores que deve ser considerado durante a análise de redes é o de componentes conectados. Este indicador demonstra a quantidade de componentes da rede que estão fortemente conectados, assim como os que têm ligações mais fracas com o todo da rede. Neste último caso as ligações fracas são aquela que, normalmente, provêm de atores que estão mais afastados da rede. No caso da UFG foram encontrados 514 nodos com ligações fracas. Em termos gerais não é um bom fator ter nodos afastados; no entanto, muitos desses nodos estão conectados a redes externas à estudada, podendo assim trazer uma maior expansão a rede atual, ou servir de intermediação para um maior nível de colaboração entre redes distintas, tornando-se assim um fator importante a ser observado na análise de redes.

CONCLUSÃO

O estudo de análise de redes sociais traz inúmeras informações a respeito de uma determinada comunidade ou grupo que interage entre si, neste estudo foi demonstrado o potencial desse tipo de análise aplicado a algo bastante comum em termos acadêmicos, a coautoria.

A visão macro de toda a rede e a possibilidade e aplicabilidade do uso de atributos que tornam a rede única, fazem com que o estudo e obtenção de possíveis fatores que expliquem os "porquês" das relações, sejam um ativo informacional bastante relevante para uma instituição ou mesmo governo (caso se pretenda a discussão em nível nacional). Uma visão macro e a apresentação de indicadores tende a facilitar a busca de informações adicionais, tais como o porquê de determinado ator ser o nodo da rede com mais ligações, e o que faz com que ele seja o mais importante? O investimento do governo ou da própria instituição está atrelado a esse fato? A formação, networking ou mesmo o posicionamento geográfico da instituição, influenciam para que este ator seja o mais relevante em termos de conexão? Os fatores são diversos e a visualização da rede em si não oferece respostas completas ainda que auxilie na construção de pressupostos e hipóteses a serem discutidas posteriormente.

Contudo, os fatores de internacionalização da Universidade Federal de Goiás merecem estudos mais aprofundados quanto ao impacto das políticas públicas nacionais e da própria instituição neste sentido.

Em um contexto geral, a análise de redes sociais, traz *insights* como demonstrados neste estudo, possibilitando recortes mais exploratórias e imersivas que por sua vez irão agregar informações em diferentes níveis para aqueles que fazem uso desse método. Os gestores de uma instituição, com o uso desses métodos de análise, podem ter um fator a mais de tomada de decisão voltada aos investimentos em diferentes segmentos/áreas do conhecimento com impactos mensuráveis no desenvolvimento cientifico da instituição.

REFERÊNCIAS

Balbachevsky, E. (1998, Jun.). Uma janela de oportunidades para a sociologia da ciência. *Revista Brasileira de Ciências Sociais*, *13*(37), 197–201. doi: 10.1590/S0102-69091998000200012

Chadegani, A. A., Salehi, H., Yunus, M. M., Farhadi, H., Fooladi, M., Farhadi, M., & Ebrahim, N. A. (2013). A Comparison between Two Main Academic Literature Collections: Web of Science and Scopus Databases. *Asian Social Science*, *9*(5), 18–26. doi: 10.5539/ass.v9n5p18

Cherven, K. (2015). *Mastering Gephi network visualization*. Birmingham: Packt Publishing.

da Silva, A. K. A. (2014). A dinâmica das redes sociais e as redes de coautoria. *Perspectivas em Gestão & Conhecimento*, *4*(Número Especial), 27–47. Recuperado de http://periodicos.ufpb.br/ojs2/index.php/pgc/article/view/21275

de Fátima S. Maia, M., & Caregnato, S. E. (2008, Maio/Ago.). Co-autoria como indicador de redes de colaboração científica. *Perspectivas em Ciência da Informação*, *13*(2), 18–31. doi: 10.1590/S1413-99362008000200003

de Oliveira e Silva, A. B., Matheus, R. F., Parreiras, F. S., & Parreiras, T. A. S. (2006, Jan./Abr.). Análise de redes sociais como metodologia de apoio para a discussão da interdisciplinaridade na ciência da informação. *Ciência da Informação*, *35*(1), 72–93. doi: 10.1590/S0100-19652006000100009

de Souza Vanz, S. A., & Stumpf, I. R. C. (2010, Maio/Ago.). Colaboração científica: revisão teórico-conceitual. *Perspectivas em Ciência da Informação*, *15*(2), 42–55. doi: 10.1590/S1413-99362010000200004

Fiorin, J. L. (2007). Internacionalização da produção científica: a publicação de trabalhos de ciências humanas e sociais em periódicos internacionais. *Revita Brasileira de Pós-Graduação*, *4*(8), 263–281. doi: 10.21713/2358-2332.2007.v4.133

Glänzel, W. (2001, April). National characteristics in international scientific co-authorship relations. *Scientometrics*, *51*(1), 69–115. doi: 10.1023/A:1010512628145

Katz, J. S. (2000). Scale-independent indicators and research evaluation. *Science and Public Policy*, *27*(1), 23–36. doi: 10.3152/147154300781782156

Marteleto, R. M. (2001, Jan./Abr.). Análise de redes sociais: aplicação nos estudos de transferência da informação. *Ciência da Informação*, *30*(1), 71–81. doi: 10.1590/S0100-19652001000100009

Matheus, R. F., & Silva, A. B. O. (2006). Análise de redes sociais como método para a Ciência da Informação. *DataGramaZero*, *7*(2), 1–17.

Musso, P. (2004). A filosofia da rede. In A. Parente (Ed.), *Tramas da rede: Novas dimensões filosóficas, estéticas e políticas da cognição* (p. 17–38). Porto Alegre: Sulinas.

Newman, M. E. J. (2001). Scientific collaboration networks: Network construction and fundamental results. *Physical Review E*, *64*(1). doi: 10.1103/PhysRevE.64.016131

Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications (structural analysis in the social sciences)*. Cambridge: Cambridge University.

Como citar este artigo (APA):
Silva, E. A. & Martins, D. L. (2016). Rede de coautoria da Universidade Federal de Goiás: o impacto de publicações internacionais. *AtoZ: novas práticas em informação e conhecimento, 5*(2), 91 – 103. Recuperado de: http://dx.doi.org/10.5380/atoz.v5i2.47854

APÊNDICE A - CÓDIGO EM PHP PARA GERAR RELAÇÕES ENTRE COAUTORES

```
<?php
$lines = file($nome_final, FILE_IGNORE_NEW_LINES); //importa o arquivo que será usado
foreach ($lines as $key) //linhas são transformadas em chaves
{
    $vetor = explode (";",$key); // divide as linhas tendo em consideração o ponto e vírgual(;)
    $tamanho = count ($vetor);// contador
    echo $key."<br />";
    echo $tamanho."<br />";
    $titulorevista = $vetor[0];
    //$titulorevista = $vetor[1];
    //$ano = $vetor[2];
    for ($i=1; $i < $tamanho-1 ; $i++){//começo da matriz

        for ($j=2; $j < $tamanho; $j++){
         if (strcmp($vetor[$i],$vetor[$j]) != 0) //faz comparação dos nomes
             echo $titulorevista./*"--".$vetor[$i]."-".$vetor[$j].*/"<br />";
        }
    }
//echo "<br /><br />";
}
//"|".$vetor[$j]. "--".$vetor[$j].
?>
```

ISSN: 2237-826X
ARTÍCULO | PAPER

El efecto del flipped classroom en la motivación por el aprendizaje del inglés como lengua extranjera en estudiantes de nivel pre-intermedio

The effect of the flipped classroom method on pre-intermediate level students' motivation to learn English as a foreign language

José Luis Ruiz Robles[1]
[1] *Universidad del Atlántico - UA, Colombia*

Correo a/*Mail to:* José Luis Ruiz Robles (joselruiz@mail.uniatlantico.edu.co)

Recibido/*Submitted*: 20 Oct. 2016; **Aceptado/*Approved*:** 13 Dic. 2016

Copyright © 2016 Ruiz Robles, J. L.. Todo el contenido de la revista (incluyendo instrucciones, política editorial y modelos) está bajo una licencia Creative Commons Atribución-NoComercial-CompartirIgual 3.0 No Adaptada. Siendo publicados por esta revista, artículos son de libre uso en ambientes educativos, de investigación y no comerciales, con atribución de autoría obligatoria. Más información en http://revistas.ufpr.br/atoz/about/submissions#copyrightNotice.

Resumen

Introdución: el presente estudio tuvo como objetivo principal determinar la efectividad del *flipped classroom* para incrementar la motivación de los estudiantes en un curso de nivel pre-intermedio de inglés como lengua extranjera.
Método: a fin de alcanzar el objetivo trazado, un estudio cuasi-experimental de corte cuantitativo fue llevado a cabo. La población objeto de estudio para esta investigación estuvo conformada por 48 estudiantes de una escuela pública en la ciudad de Barranquilla, Colombia. Por medio de un muestreo no probabilístico, 25 de ellos compusieron el grupo control del estudio y 23 el grupo experimental. La variable independiente fue el uso del *flipped classroom* como estrategia instruccional, mientras que la dependiente fue el nivel de motivación por el aprendizaje del inglés como lengua extranjera, la cual fue medida a través de un escalamiento tipo Likert adaptado y previamente validado. Este instrumento fue administrado antes y después de la intervención.
Resultados: la información obtenida fue analizada con el soporte de un software de análisis cuantitativo y la definición de significancias estadísticas mediante el cálculo *t* de Student.
Conclusión: estos procesos permitieron concluir que el uso del *flipped classroom* aumentó la motivación por el aprendizaje del inglés como lengua extranjeras en la población objeto de estudio.

Palabras clave: Tecnologia; Inovación en educación; Procesos de enseñanza y aprendizaje; Idioma inglés; Inovación en el aprendizaje

Abstract

***Introduction**: the investigation arises from a research project that had the objective to determine the effectiveness of the flipped classroom as an instructional strategy to increase the pre-intermediate level students' motivation to learn English as a foreign language.*
***Method**: in order to achieve this aim, a quasi-experimental study with quantitative techniques was conducted. The target population is composed of 48 students from a public school in the city of Barranquilla, Colombia. By means of a non-probabilistic sampling, 25 of them composed the control group and 23 of them the experimental group. The independent variable was the use of the flipped classroom as an instructional strategy while the dependent variable was the motivation level to learn English as a foreign language, which was measured via a previously adapted and validated Likert scale. This instrument was administrated before and after the intervention.*
***Results**: the gathered information was analyzed through a quantitative analysis software and the definition of statistical significances with the t-Student calculus.*
***Conclusions**: consequently, it is concluded that the use of the flipped classroom increased the motivation to learn English as a foreign language of the target population of the study.*

***Keywords:** Technology; Innovation in education; Teaching and learning processes; English language; Innovation in learning*

INTRODUCCIÓN

El mundo actual se caracteriza por la comunicación intercultural, por el creciente ritmo de los avances científicos-tecnológicos y por los procesos de internacionalización (Powell-Davies & Otero, 2011). Estas circunstancias plantean la necesidad de un idioma común que le permita a la sociedad internacional acceder a este nuevo mundo globalizado (Crystal, 2003). En Colombia, la importancia del inglés se evidencia en la implementación de un prominente número de acuerdos multilaterales. "En el plano comercial, por ejemplo, se tienen acuerdos vigentes, suscritos o negociaciones en curso con una buena cantidad países o zonas económicas, tales como Estados Unidos" (Sánchez, 2013, p. 4).

Conscientes de lo anterior, en Colombia, por parte de instituciones gubernamentales y no gubernamentales, se han aunado esfuerzos significativos para fortalecer la enseñanza y aprendizaje del inglés como lengua extranjera en las instituciones del sistema educativo colombiano. No obstante, a pesar de estas reglamentaciones, Colombia es un país con un bajo nivel de bilingüismo entre la población estudiantil que egresa del sistema educativo, específicamente, con un bajo nivel de inglés. "La proporción de estudiantes bilingües o de aquellos que por lo menos superan el nivel pre-intermedio de inglés es muy reducida, mientras que la cantidad de estudiantes que

no alcanza el nivel de principiante es significativamente alta" (Ruiz-Robles, 2014, p. 3). Esto se ratifica con los más recientes resultados de las pruebas de estado, en los cuales se pone de relieve que, en promedio, el 5% de estudiantes que egresan de las escuelas del país tienen un nivel de proficiencia intermedio (Ministerio de Educación Nacional [MEN], 2015).

Así las cosas, se hace necesario la gestación de propuestas pedagógicas, iniciativas, inclusión de nuevos métodos, enfoques, entre otras estrategias a fin de fomentar la promoción de la enseñanza y el aprendizaje del inglés como lengua extranjera. Particularmente, el presente estudio abarca el diseño de una secuencia instruccional que emplea el *flipped classroom*, así como los recursos tecnológicos necesarios para la aplicación del mismo, analizando su impacto sobre la motivación de los participantes. Esta secuencia fue implementada durante dos bimestres lectivos en las sesiones de clases de inglés de nivel pre-intermedio con estudiantes próximos a egresar del sistema escolar. A continuación, se pone de relieve la pregunta problema del estudio y, consecuentemente, el objetivo del mismo:

Pregunta problema: ¿Qué efectividad tiene el *flipped classroom* para incrementar la motivación de los estudiantes de un curso de inglés pre-intermedio como lengua extranjera?

Objetivo de la investigación: Determinar la efectividad del *flipped classroom* para incrementar la motivación de los estudiantes de un curso de inglés pre-intermedio como lengua extranjera.

FUNDAMENTACIÓN TEÓRICA

Motivación

El término motivación deriva del verbo latino *movere* que significa moverse. Por consecuencia, la idea del movimiento está implícita en las ideas de sentido común sobre la motivación (Pintrich, Meece, & Schunk, 2014). No obstante, actualmente, existe un vasto espectro de literatura especializada sobre la motivación en múltiples campos del conocimiento: psicología, educación, lingüística, marketing, etcétera. (Kleinginna & Kleinginna, citados en Mayer, Roberts, & Barsade, 2008), sin llegar a un consenso entre las múltiples ópticas de estudio, compilaron cientos de definiciones de motivación, concebidas desde las diferentes perspectivas o corrientes teóricas y ópticas de estudio.

Se han abordado, a lo largo del siglo pasado y del presente, diversas aproximaciones teóricas sobre la motivación. Desde la fisiológica, se afirma que la motivación se usa para "describir las fuerzas que actúan sobre un organismo o en su interior para que inicie y dirija la conducta" (Petri & Govern, 2013, p. 16). Por su parte, desde la psicología, se preceptúa que "la motivación es un constructo teórico-hipotético que designa un proceso complejo que causa la conducta. En la motivación intervienen múltiples variables (biológicas y adquiridas) que influyen en la activación, direccionalidad, intensidad y coordinación del comportamiento encaminado a lograr determinadas metas" (Bisquerra, 2008, p. 165). En cuanto a la educación, se anota que la motivación es un estado interno que activa, dirige y mantiene el comportamiento de los estudiantes hacia la consecución de logros académicos y de aprendizaje (Woolfolk, 2010). Así pues, se puede apreciar que desde los diferentes enfoques, la motivación presenta elementos en común: activación, dirección y persistencia.

Dado la naturaleza del presente estudio, se profundiza en la teoría de motivación integradora de Gardner y Lambert (citados en Dörnyei & Ushioda, 2009). En esta, se distingue una dicotomía motivacional: orientación instrumental y orientación integradora. Para Gardner y Lambert, el concepto de orientación instrumental hace alusión "al conjunto de motivos de índole práctica por los que un aprendiente estudia una nueva lengua" (Pérez-Ruiz, 2015, p. 3-4). En consecuencia, se orienta hacia objetivos pragmáticos concretos de aprendizaje (v.g. aprobar una prueba, obtener un título universitario o empleo, ascensos, alcanzar determinado nivel socio-económico, entre muchos otros). En contraste, la orientación integradora hace referencia a

> una disposición positiva hacia la comunidad hablante con el deseo de interactuar, o el anhelo de asemejarse e identificarse con el modo de vida y los valores de la nueva cultura. Por lo tanto, los dos objetivos principales serían el deseo de comunicación o de integración con el grupo social de la cultura de la lengua meta. Esta orientación implica una empatía y un conjunto de sentimientos positivos hacia la nueva lengua y cultura (Pérez-Ruiz, 2015, p. 4).

Flipped Classroom

El término *flipped classroom* —expresión del idioma inglés que, literalmente, puede ser entendida como "dar la vuelta a la clase" o "una clase al revés"— fue acuñada por primera vez en el año 2000. No obstante, no fue sino hasta el 2007 que esta estrategia de instrucción tuvo implicaciones prácticas y el término fue consolidado en la comunidad americana (Bergmann & Sams, 2012). Dos profesores del Instituto de Colorado en los Estados Unidos empezaron a grabar presentaciones en PowerPoint de los contenidos de sus clases con narraciones explicativas capturadas en vídeo con el objetivo de facilitarlas a los alumnos con dificultades para asistir a las sesiones de clase. "Para sorpresa de los docentes, estas grabaciones, en principio creadas para estudiantes no presenciales de sus clases, eran también seguidas por otros estudiantes" (Paz, Serna, Ramírez, Valencia, & Reinoso, 2014, p. 397). Desde ese entonces, hasta la actualidad, casi una década después, estas prácticas siguen

replicándose a lo largo de las escuelas del mundo, especialmente las de Estados Unidos, Australia y el centro de Europa (Herreid & Schiller, 2013).

Hoy día, el *flipped classroom* se ha consolidado como una estrategia instruccional y, además, una modalidad de aprendizaje semi-presencial. El objetivo principal de la misma se centra en invertir la forma en que los contenidos son entregados para dar mayor tiempo a la práctica, ejercitación, transferencia, construcción y aplicación de teorías y conceptos. Los contenidos son revisados en casa, mediante la reproducción de objetos virtuales de aprendizaje por los estudiantes y la práctica es realizada en clase. Esta perspectiva, dentro de los procesos de enseñanza y aprendizaje, es considerada por Bergmann y Sams (2012), como una metodología y/o perspectiva pedagógica que propone un modelo de aprendizaje caracterizado por el rol activo del aprendiz.

> El potencial de esta perspectiva radica en que el tiempo invertido en explicar la materia de manera magistral, queda relegado al trabajo que el estudiante realiza tranquilamente en casa a través de la explicación estratégica en grabaciones de vídeo. Dichos materiales son estudiados por los estudiantes en su casa, con la ventaja de que pueden hacerlo cuantas veces considere necesario. Las tareas tradicionales que el docente explica en clase y que luego son realizadas por el estudiante en casa, pueden ser realizadas en el salón de clase con el beneficio que esto tiene para el estudiante, debido a que las dudas, opiniones, y resoluciones de las mismas se pueden llevar a cabo mediante la interacción entre compañeros y docente, posibilidad que no se presenta al realizar este tipo de actividades en casa (Paz et al., 2014, p. 397).

De otra parte, cabe anotar que múltiples estudios (Bergmann & Sams, 2012; Berrett, 2012; Brunsell & Horejsi, 2011; Fulton, 2012; Ronchetti, 2010) sugieren que métodos instruccionales como *flipped classroom*, los cuales soportan el aprendizaje activo, beneficia a los estudiantes más que el modelo de clase tradicional basada en la lectura y compromisos escolares en casa. Entre estos beneficios, se encuentra documentando el impacto positivo en la motivación por el aprendizaje. Esto se debe, en ilación con Touchton (2015), a que los estudiantes, en el marco del aula invertida, ejecutan las tareas de menos exigencia cognitiva afuera del aula de clases, mientras que las asignaciones con más carga cognitiva —atendiendo a la taxonomía de Bloom (citado en Touchton, 2015)— se llevan a cabo dentro de esta, con el soporte del docente e intervenciones significativas.

Entre las más notables características del *flipped classroom* —en consonancia con Brunsell y Horejsi (2011)— se anotan las siguientes:

a) Los objetos virtuales de aprendizaje (v.g. vídeos, esquemas, libros electrónicos, pistas de audio, software educativo, etcétera) sustituyen al docente en la instrucción directa.

b) Fomenta la interacción y el tiempo de contacto personalizado entre alumnos y profesores.

c) El contenido curricular queda archivado en los repositorios o plataformas de enseñanza institucionales que, eventualmente, puede ser revisado, supervisado o reparado.

d) Los estudiantes son responsables de su propio aprendizaje.

e) Se lleva a cabo una combinación entre la instrucción directa y el aprendizaje constructivista.

f) Los alumnos están involucrados plenamente en su propio aprendizaje.

A fin de construir o seleccionar los objetos virtuales de aprendizaje necesarios para soportar las secuencias de enseñanza-aprendizaje, los docentes y diseñadores curriculares, actualmente, cuentan con amplio catálogo de herramientas para la elaboración de recursos educativos digitales abiertos. Como se plasma en la Cuadro 1, los formatos más comunes de estos recursos son presentaciones interactivas, vídeos, pistas de audio, documentos, documentos con texto enriquecido, gráficos y esquemas conceptuales.

Luego de la selección o diseño de objetos virtuales de aprendizaje para el soporte de las secuencias didácticas, el docente debe seleccionar una plataforma para su difusión. Así, los estudiantes tienen acceso al contenido dispuesto e inician con la ejecución o reproducción del mismo desde sus hogares para su posterior reflexión, aplicación y trasferencia en el aula de clase, a través de actividades significativas. Como se pone de relieve en la Cuadro 2, las principales herramientas para la difusión de estos recursos son Moodle, Edmodo, blogs, páginas web, wikis y redes sociales como Twitter y Facebook.

METODOLOGÍA

Diseño

Dada la naturaleza del presente estudio, este se enmarca en el enfoque cuantitativo de la investigación científica, de alcance correlacional y con un diseño de tipo experimental, específicamente, cuasiexperimental (Hernández, Fernández, & Baptista, 2014).

En cuanto al enfoque cuantitativo, el proceso de investigación de este estudio —como se esquematiza en la Figura 1— inicia con la formulación de una serie hipótesis, previo a la inmersión inicial y profunda en el campo objeto de estudio. Posteriormente, estas hipótesis son probadas para su verificación a través de la experimentación estructurada, en este caso, a través de la aplicación del *flipped classroom* durante las sesiones de inglés en un grupo experimental, mientras en un grupo control se llevan a cabo sesiones de clase según el currículo

Formato	Herramientas	Características
Presentaciones	Microsoft PowerPoint	Software para la elaboración de presentaciones en diapositivas con gran reconocimiento y múltiples herramientas de edición, nativo del sistema operativo Windows.
	Keynote	Software nativo y de uso exclusivo en el sistema operativo Macintosh para la elaboración de presentaciones en diapositivas. Cuenta con una variedad de plantillas y herramientas de enriquecimiento estético y gráfico. Se debe adquirir una licencia paga para su uso.
	Prezi	Software y aplicativo web para el diseño de presentaciones interactivas que pueden ser visualizadas en internet, o bien, mediante un archivo en formato flash. Cuenta con una versión básica para uso gratuito.
	Slideshare, Scribd	Plataformas web para la creación de presentaciones, publicación de documentos y difusión de los mismos empleando código HTML que puede ser embebido en páginas web, wikis y blogs.
Vídeos y pistas de audio	Windows Movie Maker	Editor de vídeos nativo y de uso exclusivo para el sistema operativo Windows. Cuenta con herramientas básicas para la creación de vídeos basados en secuencias de imágenes.
	Powtoon, Go-animate!, Animoto	Plataformas web para la creación de vídeos basados en secuencias sincrónicas de imagen, texto, transiciones, animaciones y sonidos. Cuentan con versiones gratuitas y de uso limitado.
	YouTube	Plataforma para la creación, edición y difusión de vídeos de diversa índole.
	Jing	Software freeware para la realización de objetos de aprendizaje basados en screencast o grabación de pantalla, en conjunto con otros periféricos (v.g. micrófono, cámara).
	Adobe Presenter	Software con licencia paga para el soporte de instrucción a distancia basado en screencast o grabación de pantalla con una integración de diferentes soluciones ofimáticas (v.g. Microsoft PowerPoint). Disponible para Windows y Macintosh.
	Soundcloud, Ivoox	Plataforma para la creación y difusión de podcasts.
Documentos, documentos con texto enriquecido, gráficos y esquemas conceptuales	Microsoft Publisher	Software con licencia de pago para la elaboración de una amplia gama de documentos impresos y digitales que hace hincapié en el diseño y la maquetación. Disponible para Windows y Macintosh.
	Pages	Software con licencia de pago, nativo y de uso exclusivo de Macintosh, para la elaboración de una amplia gama de documentos impresos y digitales.
	Microsoft Word	Software editor de textos con licencia de pago. Disponible para los sistemas operativos Windows y Macintosh.
	Mindomo	Plataforma web para la creación de mapas de pensamiento interactivos. Cuenta con versiones básicas de uso gratuito.
	Cmaps	Software de licencia abierta para el diseño de mapas de pensamiento disponible para Windows.

Cuadro 1. Principales herramientas para la creación o selección de objetos virtuales de aprendizaje.
Fuente: los autores.

Herramienta	Características	Acceso a información adicional
Moodle	Software para la creación y gestión de cursos en línea. Contempla el desarrollo de un entorno virtual de aprendizaje que engloba actividades de comunicación, comprensión y evaluación. Cuenta con una vasta gama de herramientas tanto para docentes como para estudiantes que facilitan los procesos de enseñanza-aprendizaje.	https://docs.moodle.org/all/
Edmodo	Plataforma educativa gratuita con características similares a la de una red social que permite la creación de grupos cerrados entre alumnos y profesores con el objetivo de compartir mensajes, enlaces, documentos, eventos, pruebas, entre otros.	https://edmodo.com/about/
Blogs y páginas web	Espacio web que permite la difusión de material multimedia de forma cronológica y permite a los alumnos comentar las entradas y los comentarios de sus iguales.	https://wordpress.com/about/ https://blogger.com/about/ https://tumblr.com/about/ https://edublogs.org/features/
Wikis	A diferencia del blog, los wikis son colecciones concatenadas de páginas de internet que pueden ser creadas, editadas o eliminadas por todos los usuarios.	http://wikispaces.com/content/classroom/about/
Twitter y Facebook	Redes sociales abiertas que posibilitan la creación de grupos privados y públicos, así como listas de difusión para la publicación de contenido multimedia. Asimismo, permite a los alumnos comentar las entradas y los comentarios de sus iguales.	https://facebook.com/ https://twitter.com/

Cuadro 2. Principales herramientas para la difusión de los objetos virtuales de aprendizaje.
Fuente: los autores.

institucional. Este proceso de prueba sistemática envuelve un mecanismo estadístico a fin de determinar relaciones entre las variables. Con el objetivo de medir estas variables, se adapta un escalamiento tipo Likert que mide los niveles motivacionales de los estudiantes antes y después de la intervención experimental. En cuanto al diseño cuasiexperimental, se anota que se manipula deliberadamente una variable independiente (el *flipped classroom* como método instruccional) para observar su efecto y relación con la variable dependiente

(motivación por el aprendizaje del inglés).

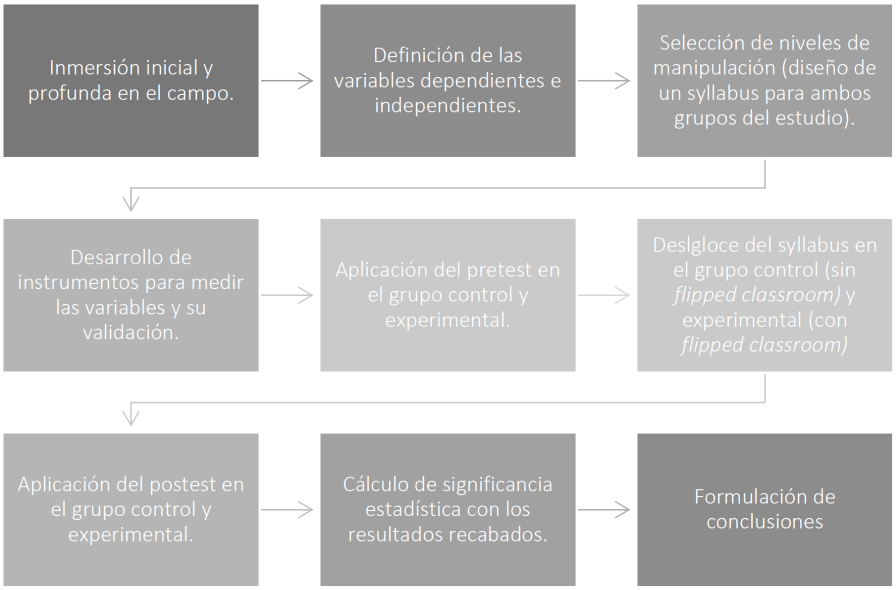


Figura 1. Procesos secuenciales de la investigación.
Fuente: los autores.

Variables

Variable dependiente: motivación por el aprendizaje del inglés como lengua extranjera. En esta variable se distinguen dos constructos: la orientación integradora y la orientación instrumental.

Variable independiente: uso del *flipped classroom* como estrategia instruccional durante las sesiones de clase de inglés. La manipulación de la variable dependiente, mediante el uso del *flipped classroom*, se llevó a cabo mediante el diseño y aplicación de una propuesta de intervención (véase Cuadro 3).

Componentes instruccionales para el grupo control y experimental	
Curso formativo	Inglés como lengua extranjera (EFL).
Nivel de proficiencia	Pre-intermedio atendiendo al *Marco común europeo de referencias para las lenguas* (MCERL).
Intensidad horaria/duración de la intervención	Cuatro meses con una intensidad horaria de cuatro horas semanales.
Competencias	Competencia comunicativa: lingüística, sociolingüística y pragmática.
Habilidades de comprensión y producción	Lectura, escritura, escucha y habla.
Contenidos y logros instruccionales.	Alineados a los *Estándares básicos de competencia* del Ministerio de Educación Nacional [MEN] (2006).
Estrategias de evaluación	Portafolios, rúbricas, matrices, evaluación por pares, autoevaluación, evaluación sumativa mediante test estandarizados de proficiencia.
Recursos educativos	Computadores, conexión de banda ancha, diademas, altavoces, proyectores, recursos basados en software (procesadores de texto, presentaciones en diapositivas, navegador web, reproductor de multimedia, etc.).
Metodología	En el grupo control se emplea una método instruccional de corte tradicional, en consonancia a los preceptos institucionales. En el aula, se hace la revisión teórica, de forma magistral de los patrones gramaticales. En casa, se le asignan a los estudiantes tareas de investigación, creación y aplicación.
	En el grupo experimental se realiza una concatenación de las actividades de aprendizaje atendiendo la taxonomía de Bloom (citado en Marzano & Kendall, 2006). Así pues, en casa, se llevan a cabo procesos de comprensión y retención de la información. En contraste, en la clase, se dan procesos de creación, evaluación, análisis y aplicación de los conceptos y constructos comprendidos en casa, mediante la mediación docente.

Cuadro 3. Propuesta de intervención.
Fuente: los autores.

Hipótesis

H0. No hay una diferencia estadísticamente significativa entre el nivel de motivación por el aprendizaje del inglés como lengua extranjera antes y después de la implementación del *flipped classroom*.

H1. Hay una diferencia estadísticamente significativa (positiva) entre el nivel de motivación por el aprendizaje del inglés como lengua extranjera antes y después de la implementación del *flipped classroom*.

H2. Hay una diferencia estadísticamente significativa (negativa) entre el nivel de motivación por el aprendizaje del inglés como lengua extranjera antes y después de la implementación del *flipped classroom*.

Población y muestra

La población objeto de estudio consta de 68 estudiantes, distribuidos en tres aulas de clase, que cursan undécimo grado en una institución educativa pública en Barranquilla —ciudad ubicada en la costa norte de Colombia. Así pues, los participantes, durante el estudio de experimentación, se encuentran próximos a egresar del sistema escolar. Estos tienen edades que oscilan entre los 15 a 17 años y cuentan con condiciones socio-económicas modestas. Cabe resaltar que los involucrados en este estudio asisten a sesiones de clases de inglés integradas en el currículo durante cuatro horas a la semana y su nivel de proficiencia es pre-intermedio.

Con respecto a la muestra, teniendo en cuenta la naturaleza de este estudio, se lleva a cabo un proceso de muestreo no probabilístico —en consonancia con los presupuestos de Hernández et al. (2014). Una porción (25 estudiantes de undécimo grado de un aula de clases específica) de la población total compone el grupo control del estudio y otra porción (23 estudiantes de undécimo grado de otra aula de clases) componen el grupo experimental, para un total de 48 participantes. De acuerdo con Hernández et al. (2014) y considerando los diseños cuasiexperimentales, este tipo de estudio deben ser llevados a cabo con un mínimo de 15 individuos. Así pues, el presente estudio cumple con este criterio de validez para el diseño de investigación acogido.

Instrumentos de recolección de la información

A fin de medir la variable dependiente del estudio, se adaptó y se tradujo al español el cuestionario *Attitude/Motivation Test Battery* de Gardner (2004) para el International AMTB Research Project, incluyendo algunos enriquecimientos lingüísticos (véase Apéndice A). Gardner (2004) pone de relieve una serie de instrumentos de recolección de la información para medir la motivación y los componentes asociados a esta: la orientación integradora y la integración instrumental. Entre estos instrumentos, Gardner diseña escalamientos tipo Likert y diferenciales semánticos fácilmente adaptables a un vasto número de contextos y poblaciones objeto de estudio. Dado los objetivos del presente estudio, se seleccionó el escalamiento tipo Likert (Gardner, 2004) que consta de 25 afirmaciones relacionadas con la motivación por el aprendizaje del inglés. Cada afirmación es evaluada por los participantes con cinco opciones de respuesta: *totalmente de acuerdo, de acuerdo, ni en acuerdo ni en desacuerdo, en desacuerdo y muy en desacuerdo*. El cuestionario tiene un puntaje final mínimo de 25 y máximo de 125.

En cuanto a la validación, cabe resaltar que la *Attitude/Motivation Test Battery* ha sido aplicada y validada en un vasto número de investigaciones sobre la motivación de los estudiantes de sexto a undécimo grado, en cursos de lengua nativa, segunda lengua y lengua extranjera (Dörnyei, 2001; Gardner, 2004). Los coeficientes de Cronbach α para los ítems que conforman el cuestionario fueron superiores a un valor de 0.70; la consistencia interna del instrumento es sustancial, luego de su administración en la prueba piloto. Cabe anotar que Dörnyei (2007) plantea que un valor de Cronbach α entre .60 y .70 son suficientes en el campo de lingüística aplicada. De otra parte, Gardner (1985) presenta información considerablemente relevante de la validez convergente y discriminante de las escalas e ítems.

Instrumentos de análisis de la información

Durante el proceso de análisis de los datos, se utiliza el software Statistical Package for the Social Sciences —SPSS— de IBM versión 23 para Macintosh, a fin de tabular los resultados proporcionados por el escalamiento tipo Likert y emplear el cálculo t de Student. De este modo, se determina la significancia estadística entre los resultados de los test que tuvieron lugar antes y después de la intervención en los grupos control y experimental, dando lugar a la aceptación o rechazo de las hipótesis del estudio. Asimismo, por medio de este software estadístico, se ahonda en los promedios, desviaciones estándar y porcentajes de cada una de las categorías de análisis. Además de esto, se emplea Microsoft Office Visio y Microsoft Office Excel para representar los datos en gráficas de distribución y tendencias, así como también Microsoft Office Word para las tablas de información correspondientes.

RESULTADOS Y ANÁLISIS DE DATOS

A continuación, se ponen de relieve los resultados recabados y su análisis estadístico correspondiente. En el Apéndice B se encuentran los puntajes de los participantes en el *Attitude/Motivation Test Battery* antes y después de la intervención.

De otra parte, se anota que las puntuaciones de los escalamientos tipo Likert antes de la intervención eran de media similares y aproximadamente normales. La prueba ANOVA, ejecutada en el software SPSS, no mostró una diferencia estadísticamente significativa. Asimismo, la prueba de Levene para la homogeneidad de varianzas

confirmó que se podía asumir la igualdad de varianzas para los resultados de los instrumentos antes de la intervención. Finalmente, las desviaciones de la normalidad no afectan la validez del análisis.

Ahondando en el análisis de los datos, en el grupo control —como puede apreciarse en la Tabla 1 y Figura 2—, se obtuvo un promedio de 45.64 de 125 puntos posibles antes de la experimentación y, posterior a esta, 46.40, en el instrumento de recolección. En consecuencia, y considerando el análisis de significancia estadística (véase Tabla 2), en el grupo control no hubo una diferencia estadísticamente significativa en los promedios del cuestionario de motivación por el aprendizaje del inglés. Esto era de esperarse, puesto que, en el grupo control, no hubo ningún tipo de intervención; las sesiones de clase fueron llevadas a cabo de forma usual, siguiendo los preceptos del currículo institucional.

	Grupo control		Grupo experimental	
	Pretest	**Postest**	**Pretest**	**Postest**
Promedio	45.64	46.40	43.00	66.00
Desviación estándar	17.00	17.02	13.61	17.95
Muestra	25	25	23	23

Tabla 1. Contraste de los promedios y desviaciones estándar de los puntajes obtenidos en el pre y pos test en el grupo control y experimental.
Fuente: los autores.

En cuanto al grupo experimental —como se esboza en la Tabla 1 y Figura 2—, se señala que este tuvo un promedio de 43 de 125 puntos posibles, frente a 45.6 del grupo control durante el pretest. En contraste, luego de la intervención mediante la aplicación del *flipped classroom* los resultados del postest alcanzaron, en promedio, 66 puntos en el grupo experimental. Por consiguiente, como se devela en el análisis de significancia estadística —plasmado en la Tabla 2—, en el grupo experimental, la diferencia entre los puntajes del *Attitude/Motivation Test Battery* antes y después de la intervención son altamente significativos, ascendiendo el valor p a -0.0001.

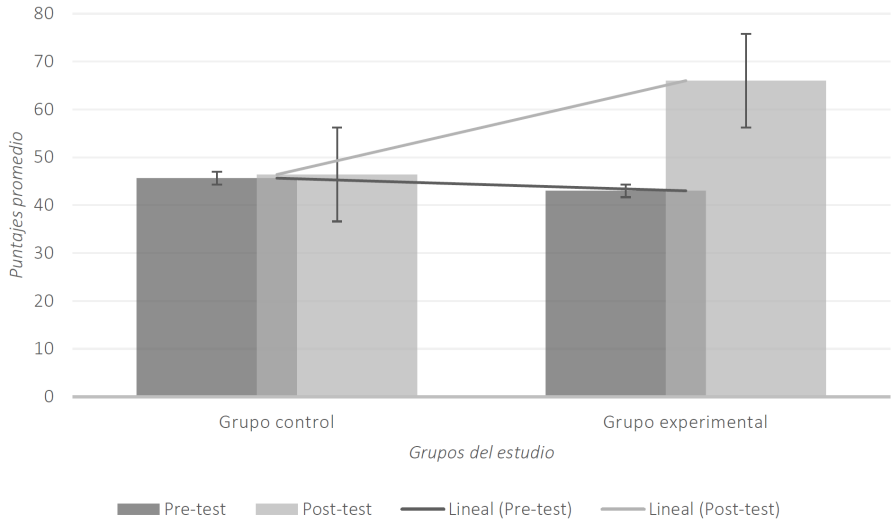


Figura 2. Tendencias en los resultados del *Attitude/Motivation Test Battery* en los grupos control y experimental.
Fuente: los autores.

	Grupo control	Grupo experimental
Valor p	-0.76	-0.0001
Valor t	0.1580	4.8966
Grados de libertad df	48	44
Error estándar de diferencia	4.811	4.697
¿Hay una diferencia estadísticamente significativa entre el pretest y postest?	No	Sí
Muestra	25	23

Tabla 2. Análisis de significancia estadística.
Fuente: los autores.

Así las cosas, la primera hipótesis de la investigación —hay una diferencia estadísticamente significativa (positiva) entre el nivel de motivación por el aprendizaje del inglés como lengua extranjera antes y después de la implementación del *flipped classroom*— es aceptada. A su vez, la hipótesis nula es rechazada. A tenor de lo anterior, cabe remarcar el papel del uso del *flipped classroom* como estrategia instruccional en el incremento de la motivación de los estudiantes durante las sesiones de clase de inglés.

CONCLUSIONES

Luego de la experimentación mediante el uso del *flipped classroom* durante las sesiones de clase de inglés como lengua extranjera, la correspondiente administración de instrumentos de recolección de la información, recabado de datos y su volcamiento en el software de análisis estadístico, se concluye lo siguiente:

a) El uso del *flipped classroom* —aula invertida— como estrategia instruccional previamente diseñada, teniendo en cuenta los múltiples factores condicionantes del contexto, permite aumentar la motivación por el aprendizaje del inglés de forma significativa y sus componentes; esto es, orientación instrumental y orientación integradora de forma significativa. Estos hallazgos constatan investigaciones anteriores sobre las implicaciones pedagógicas del *flipped classroom* (Bergmann & Sams, 2012; Berrett, 2012; Brunsell & Horejsi, 2011; Fulton, 2012; Ronchetti, 2010).

b) Además de las implicaciones en los niveles motivacionales de los estudiantes, se pudo determinar que el logro académico de los mismos aumentó considerablemente.

c) Se alcanzó todos los objetivos instruccionales trazados para el curso, alcanzando el nivel de proficiencia de la lengua contemplado por el currículo institucional. En contraste, en el grupo control esto no sucedió. Así pues, el uso del *flipped classroom*, no solo tiene efectos positivos en la motivación, sino también en el incremento del logro académico y niveles de proficiencia.

d) La creación de objetos de aprendizaje, considerando las necesidades educativas de los estudiantes, en el marco de los recursos educativos digitales abiertos, se configura como una apuesta de gran preponderancia, debido a la exponencial acogida de este tipo de recursos en la comunidad educativa internacional.

e) Cabe anotar que, para lograr la implementación un programa formativo, empleando la estrategia de instrucción del *flipped classroom*, se hace indispensable que los estudiantes y docentes cuenten con las competencias y habilidades necesarias para el uso apropiado de las herramientas tecnológicas. Asimismo, es de vital importancia contar con los recursos tecnológicos adecuados en el aula de clase para la mediación de las actividades, así como también los dispositivos y periféricos necesarios en las casas de los estudiantes. Así pues, se deben reunir ciertas características de acceso a este tipo de recursos. En poblaciones educativas donde existe una determinada brecha digital, este tipo de estrategias constituye un gran reto a la hora de su implementación. No obstante, con la conjugación de esfuerzos por parte de los diferentes estamentos educativos, podrían llevarse a cabo pruebas piloto para la posterior aplicación sistemática del *flipped classroom*.

f) En cuanto las líneas de continuidad para futuras investigaciones, se propone el estudio de los efectos del *flipped classroom* como estrategia instruccional en otras variables de estudio, tales como ambiente escolar, logro académico (nivel de proficiencia), motivación de logro y teorías motivacionales de corte metacognitivo.

REFERENCIAS

Bergmann, J., & Sams, A. (2012). *Flip your classroom reach every student in every class every day*. Alexandria: ASCD-ISTE.

Berrett, D. (2012). How 'Flipping' the classroom can improve the traditional lecture. *Chronicle of Higher Education*, *12*, 1–14. Descargado de http://www.chronicle.com/article/How-Flipping-the-Classroom/130857/

Bisquerra, A. (2008). *Educación emocional y bienestar*. Barcelona: Wolters Klwwer.

Brunsell, E., & Horejsi, M. (2011). Science 2.0: "Flipping" your classroom. *The Science Teacher*, *78*(2), 10. Descargado de http://www.jstor.org/stable/24148320

Crystal, D. (2003). *English as a global language* (2a ed.). New York: Cambridge University.

Dörnyei, Z. (2001). New themes and approaches in second language motivation research. *Annual Review of Applied Linguistics*, *21*, 43–59. doi: 10.1017/S0267190501000034

Dörnyei, Z. (2007). *Research methods in applied linguistics: quantitative, qualitative, and mixed methodologies*. Oxford: Oxford University.

Dörnyei, Z., & Ushioda, E. (2009). Motivation, language identities and the L2 self: A theoretical overview. En Z. Dörnyei & E. Ushioda (Eds.), *Motivation, language identity and the L2 self.* Bristol/New York/Canadá: Multilingual Matters.

Fulton, K. (2012, June/July). Upside down and inside out: Flip your classroom to improve student learning. *Learning & Leading with Technology*, 12–17. Descargado de http://www.learningandleading-digital.com/learning_leading/20120607?pg=14#pg14

Gardner, R. C. (1985). *Social psychology and second language learning: the role of attitudes and motivation*. London: Edward Arnold.

Gardner, R. C. (2004). *Attitude motivation test battery: International AMTB research project*. Ontario: The University of Western Ontario.

Hernández, R., Fernández, C., & Baptista, M. (2014). *Metodología de la investigación* (6a ed.). México: McGraw-Hill.

Herreid, C. F., & Schiller, N. (2013). Case study: Case studies and the flipped classroom. *Journal of College Science Teaching*, *42*(5). Descargado de http://www.nsta.org/store/product_detail.aspx?id=10.2505/4/jcst13_042_05_62

Marzano, R. J., & Kendall, J. S. (2006). *The new taxonomy of educational objectives*. Thousand Oaks: Corwin Press.

Mayer, J. D., Roberts, R. D., & Barsade, S. G. (2008). Human abilities: Emotional intelligence. *Annual Review of Psychology*, *59*, 507–536. doi: 10.1146/annurev.psych.59.103006.093646

Ministerio de Educación Nacional. (2006). *Estándares básicos de competencias en lenguas extranjeras: Inglés.* Bogotá.

Ministerio de Educación Nacional. (2015). *Programa nacional de inglés 2015-2025: Documento de socialización.* Bogotá.

Paz, A. P., Serna, A., Ramírez, M. I., Valencia, T., & Reinoso, J. (2014). Hacia la perspectiva de aula invertida (Flipped Classroom) en la Pontificia Universidad Javeriana desde una tipología de uso educativo del Sistema Lecture Capture (S.L.C). En *Novena conferencia latinoamericana de objetos y tecnologías de aprendizaje* (pp. 395–403). Manizales, Colombia. Descargado de http://www.laclo.org/papers/index.php/laclo/article/view/265/247

Petri, H., & Govern, J. (2013). *Motivation: Theory, research, and application* (6a ed.). New York: Thomson-Wadsworth.

Pintrich, P., Meece, J., & Schunk, D. (2014). *Motivation in education: Theory, research, and applications* (4a ed.). New York: Pearson.

Powell-Davies, P., & Otero, J. (2011). *Word for word: The social, economic and political impact of spanish and english*. Madrid: British Council; Instituto Cervantes.

Pérez-Ruiz, J. (2015). Actitudes sociolingüísticas en estudiantes universitarios taiwaneses de ELE. *Marco ELE*, *21*, 1-18. Descargado de http://marcoele.com/actitudes-sociolinguisticas-en-estudiantes-taiwaneses-de-ele/

Ronchetti, M. (2010). Using video lectures to make teaching more interactive. *International Journal of Emerging Technologies in Learning*, *5*(2), 45–48. doi: 10.3991/ijet.v5i2.1156

Ruiz-Robles, J. (2014). *Fortalecimiento motivacional para el aprendizaje del inglés a través del uso de recursos educativos digitales abiertos en estudiantes de quinto grado.* (Proyecto de grado publicado. Escuela Normal Superior del Distrito de Barranquilla.)

Sánchez, A. (2013). *El bilingüismo en los bachilleres colombianos*. Cartagena de Indias: Banco de la República.

Touchton, M. (2015). Flipping the classroom and student performance in advanced statistics: Evidence from a quasi-experiment. *Journal of Political Science Education*, *11*, 28–44. doi: 10.1080/15512169.2014.985105

Woolfolk, A. (2010). *Educational psychology* (13a ed.). New York: Prentice Hall-Pearson.

Cómo citar este artículo (APA):
Ruiz-Robles, J. L. (2016). El efecto del flipped classroom en la motivación por el aprendizaje del inglés como lengua extranjera en estudiantes de nivel pre-intermedio. *AtoZ: novas práticas em informação e conhecimento, 5*(2), 104 – 114. Descargado de: http://dx.doi.org/10.5380/atoz.v5i2.48940

APÉNDICE A

Versión adaptada del *Attitude/Motivation Test Battery* de Gardner (2004).

Instrucciones. Nos gustaría conocer tu opinión acerca de cada uno de los siguientes enunciados. Tacha la alternativa que más se acerque a tu opinión. Por favor, no dejes preguntas sin contestar. Recuerda que este cuestionario es anónimo; tus respuestas no afectarán tus calificaciones.

Opciones de respuesta:

Totalmente de acuerdo—5

De acuerdo—4

Ni en acuerdo ni en desacuerdo—3

En desacuerdo—2

Muy en desacuerdo—1

	Pregunta	Respuesta				
		1	2	3	4	5
1	Estudiar inglés es fabuloso.					
2	Estudiar inglés es importante porque me permitirá sentirme cómodo/a cuando me relacione con personas que hablan ese idioma.					
3	Siento un fuerte deseo por conocer todos los aspectos del inglés.					
4	Pongo mucha atención para tratar de entender todo el inglés que veo y escucho.					
5	Estudiar inglés es importante porque lo necesitaré en mi desempeño profesional.					
6	Preferiría pasar más tiempo en mi clase de inglés que en otras clases.					
7	Desearía poder leer periódicos, libros y revistas en muchas otras lenguas.					
8	Realmente disfruto aprender inglés.					
9	Siento que estudiar inglés es importante porque me permitiría comunicarme con un número mayor de personas de diferentes culturas en el mundo.					
10	Si por mi fuera, me pasaría todo el tiempo aprendiendo inglés.					
11	Me mantengo al corriente con mi clase de inglés trabajando a diario en ella.					
12	Desearía tener a muchos hablantes nativos del inglés como amigos.					
13	Disfruto las actividades de mi clase de inglés más que las de las otras clases.					
14	Estudiar inglés es importante porque me permitirá entender y apreciar mejor la cultura anglófona.					
15	Los hablantes nativos del inglés tienen mucho de qué sentirse orgullosos porque han hecho contribuciones valiosas para el mundo.					
16	Cuando tengo problemas para entender algo en mi clase de inglés, siempre le pido ayuda a mi maestra/o.					
17	Estudiar inglés es importante porque me será útil para obtener un buen empleo.					
18	Me gusta tanto la clase de inglés que quiero seguir estudiando esta lengua en el futuro.					
19	Quiero aprender tanto inglés como sea posible.					
20	Es importante estudiar inglés porque tendré la oportunidad de interactuar más fácilmente con gente que habla esa lengua.					
21	Yo realmente pongo mucho empeño en aprender inglés.					
22	Estudiar inglés es importante porque la gente me respetará más si domino esta lengua.					
23	Me gustaría dominar el inglés a nivel nativo.					
24	Cuando estoy estudiando inglés ignoro las distracciones y le pongo atención a la actividad que estoy realizando.					
25	La clase de inglés es una de mis favoritas.					
Puntaje total						
Observaciones						

APÉNDICE B - PUNTAJES DEL ATTITUDE/MOTIVATION TEST BATTERY DE GARDNER (2004) ANTES Y DESPUÉS DE LA INTERVENCIÓN.

Grupo control			Grupo experimental		
Estudiante	Puntaje pretest	Puntaje postest	Estudiante	Puntaje pretest	Puntaje postest
1	35	36	1	36	70
2	30	33	2	33	65
3	25	25	3	25	40
4	80	81	4	76	90
5	45	43	5	47	64
6	50	51	6	38	66
7	52	52	7	48	50
8	35	37	8	49	85
9	40	45	9	33	50
10	100	101	10	25	25
11	45	44	11	40	70
12	50	50	12	42	61
13	37	38	13	36	58
14	41	42	14	33	52
15	75	76	15	50	80
16	45	47	16	81	115
17	38	38	17	43	69
18	46	48	18	51	73
19	50	49	19	52	65
20	32	33	20	37	59
21	25	25	21	45	78
22	39	40	22	36	74
23	43	42	23	33	59
24	36	36			
25	47	48			

ISSN: 2237-826X
SHORTPAPER

Producción de cursos en-línea basados en objetos de aprendizaje: una propuesta metodológica orientada a estudiantes de carreras técnicas

Production of on-line courses based on learning objects: a methodological proposal aimed at students of technical careers

Margarita Elizabeth Zambrano[1], Walter Marcelo Fuertes-Diaz[1], Diego David Pérez-Herrera[1], César Javier Villacís[1], Ernesto Pérez-Estévez[2]

[1]*Universidad de las Fuerzas Armadas - ESPE, Sangolquí, Ecuador*
[2]*Consorcio Ecuatoriano para el Desarrollo de Internet Avanzado - CEDIA, Cuenca, Ecuador*

Correo a/Mail to: *Walter Marcelo Fuertes-Diaz (wmfuertes@espe.edu.ec)*

Agradecimientos/Acknowledgments: *PhD. Jenny Gabriela Torres Olmedo, Escuela Politécnica Nacional, Quito, Ecuador; PhD. Luis Enrique Sánchez Crespo, Universidad Castilla La Mancha, España; y Mag. Mario Bernabé Ron, Universidad de las Fuerzas Armadas - ESPE, Sangolquí, Ecuador por su aporte invaluable como miembros del Proyecto de Investigación referido.*

Financiamiento/Funding: *Este trabajo ha sido parcialmente financiado en el marco del proyecto de investigación titulado "Plataformas Computacionales de Entrenamiento, Experimentación, Gestión y Mitigación de Ataques a la Ciberseguridad", Código: ESPE-2015-PIC-019, Fase 1 con el cofinanciamiento del Consorcio Ecuatoriano para el Desarrollo de Internet Avanzado (CEDIA).*

Recibido/*Submitted*: 21 Dez. 2016; Aceptado/*Approved*: 21 Dez. 2016

Copyright © 2016 Zambrano et al.. Todo el contenido de la revista (incluyendo instrucciones, política editorial y modelos) está bajo una licencia Creative Commons Atribución-NoComercial-CompartirIgual 3.0 No Adaptada. Siendo publicados por esta revista, artículos son de libre uso en ambientes educativos, de investigación y no comerciales, con atribución de autoría obligatoria. Más información en http://ojs.c3sl.ufpr.br/ojs2/index.php/atoz/about/submissions#copyrightNotice.

Resumen

Introducción: Este estudio tiene como objetivo estimular la comprensión conceptual y procedural de estudiantes de carreras técnicas a través de una propuesta metodológica que se aplica durante el diseño e implementación de Cursos en-línea basados en objetos de aprendizaje (OA).
Método: Se han aplicado de manera armonizada el método de diseño de hipermedia orientado a objetos (OOHDM) y la metodología para la creación de OA (CROA) para la producción de un curso de Seguridad de la Información, como prueba de concepto. El curso fue desarrollado bajo licencia académica*Articulate Studio* para contenidos y evaluaciones, y con *Ardora* para la elaboración de actividades de aprendizaje. Ambos fueron implantados dentro de *Moodle* como entorno virtual de E/A (EVEA). Para las instancias de auto evaluación y reflexión, se aplicó el aprendizaje basado en problemas (ABP), que además promueve experiencias colaborativas entre estudiantes y aprendientes. Para la evaluación de la usabilidad del curso, se aplicó la matriz para la evaluación de OA *LORI*.
Resultados: Los resultados muestran, por un lado el procedimiento metodológico teórico y práctico validado, y por otro, muestran la satisfacción de los estudiantes con respecto a la organización, contenido y funcionalidad del curso.
Conclusión: La aplicación conjunta de las metodologías OOHDM, CROA y ABP, permiten estructurar sistemáticamente y pedagógicamente los materiales educativos en base a OA de cursos en línea, para mejorar la comprensión conceptual y procedural de los estudiantes de carreras técnicas.

Palabras clave: Objetos de aprendizaje (OA); Metodología de criação de objetos de aprendizagem (CROA); Diseño Instruccional; Entorno Virtual de Enseñanza-Aprendizaje (EVA); Aprendizaje Basado en Problemas (ABP); método de diseño de hipermedia orientado a objetos (OOHDM)

Abstract

Introduction*: This research main aim has been to stimulate the conceptual and procedural understanding of students of technical careers through a methodology that is applied during the design and implementation of online courses, which are based on Learning Objects.*
Method*: We applied a harmonized Object-Oriented Hypermedia Design Method (OOHDM) and the methodology for the Creation of Learning Objects (CROA) for the production of a course in Information Security as proof of the concept. The course has been developed with an academic license of Articulate Studio for the content and assessments as well as with Ardora for developing learning activities. Both of them were introduced within Moodle as virtual teaching and learning environment (VLE). For cases of self-assessment and reflection, the Learning Based on Problems (ABP) has been applied, which also promotes collaborative experiences between students and learners. For the evaluation of the usability of the course, the LORI matrix for the evaluation of learning objects we have used.*
Results*: The results demonstrate firstly the theoretical and practical validated methodological procedure, and secondly, they indicate the student's satisfaction regarding the organization, content and functionality of the course.*
Conclusio*: n: The blend application of the OOHDM, CROA and LBP methodologies allow to structure systematically and pedagogically educational materials, which are based on Learning Objects for online courses to improve conceptual and procedural comprehension in students of technical careers.*

Keywords: *Learning objects (OA); Methodology of creation of learning objects (CROA); Instructional design; Object Oriented Hypermidia Design Model (OOHDM); Virtual Learning Environment (AVA); Problem Based Learning (PBL)*

INTRODUCCIÓN

En los últimos años la comunidad académica y científica se ha preocupado por innovar los entornos virtuales de enseñanza-aprendizaje (EVEA) en la Educación Superior incorporando la aplicación de Objetos de Aprendizaje

(OA). Algunas investigaciones como García Vera, Roig Vila, Marti Ciriquian, y García Ferrández (2012) describen el uso de OA para cursos de asignaturas técnicas. Maldonado, Carvallo, y Siguencia (2015), presentan un estudio comparativo de propuestas metodológicas para el diseño de OA a nivel Iberoamericano, donde se destaca CROA. Soto de Giorgis, Palma Muñoz, y Roncagliolo de la Horra (2002), explican los modelos de diseño conceptual, de contexto y el modelo navegacional de cursos en línea para EVEAs. De Giusti et al. (2015), explican los resultados preliminares de la utilización de OA para la enseñanza universitaria, aplicando el diseño instruccional alineado con la metodología CROA. Alonso, Castillo, Martínez, y Muñoz (2013), presentan la Metodología MEDOA para el desarrollo de OA que es una metodología similar a CROA. Mestre, Torres, Díaz, y Gómez (2014), exponen los lineamientos pedagógicos, comunicativos y tecnológicos para la producción de cursos mediados por Tics en la Educación a Distancia.

A pesar de estos esfuerzos, en estos trabajos no se distingue los lineamientos metodológicos para el diseño de cursos en línea en base a principios de Ingeniería de Software para garantizar un material digital didáctico reutilizable. Además no se especifican los procedimientos de diseño, desarrollo, implantación y almacenamiento de los OA, ni tampoco contempla una licencia de autor para el uso comercial o académico de dicha metodología.

Ante este panorama, este estudio intenta aportar con una propuesta metodológica que coadyuve a estimular la capacidad de comprender o aprehender (i. e. traducción, interpretación y extrapolación, según la Taxonomía de Bloom). En consecuencia, nuestra premisa de investigación a resolver sería: ¿En qué medida el diseño y desarrollo de cursos on-line con OA combinados con aprendizaje basado en problemas (ABP), mejoraría los niveles de comprensión que poseen los estudiantes del 5to nivel de una carrera de ingeniería?

Este estudio tiene como objetivo estimular la comprensión conceptual y procedural de los estudiantes a través de una propuesta metodológica que se aplica sistemáticamente durante el diseño e implementación de Cursos On-Line basados en OA. Para llevarlo a cabo, se han aplicado de manera armonizada el Método de Diseño de Hipermedia Orientado a Objetos (OOHDM) (Pressman, 2010) para el diseño navegacional del sitio Web del curso y el diseño conceptual de la aplicación; y la metodología para la creación de OA (CROA) (Sanz, Moralejo, & Barranquero, 2014). Esto mantiene una apropiada organización de los contenidos del curso y permite reducir contenidos extensos en pequeñas unidades pedagógicas compactas e interactivas.

Como prueba de concepto, se implementaron los OA del curso de Fundamentos de Seguridad de la Información, que fueron desarrollados bajo licencia académica *Articulate Studio* para contenidos y evaluaciones, donde también se utilizó *Ardora* para la elaboración de actividades de aprendizaje. Estos fueron integrados mediante *Reload* para formar un paquete SCORM, que puede ser implantado dentro de la plataforma Moodle, que es un entorno virtual de distribución gratuita en donde se pueden llevar a cabo varios procesos pedagógicos sobre la Web. Para las instancias de auto evaluación y reflexión, se aplicó el ABP que promueve experiencias colaborativas entre estudiantes y aprendientes. Para la evaluación de la usabilidad del curso, se aplicó la matriz de evaluación de OA LORI. Los resultados muestran que esta integración permite estructurar los materiales educativos en base a OA de los cursos en línea, que mejoran la comprensión conceptual y procedural de los estudiantes de quinto nivel de carreras de ingeniería, población objeto de este estudio.

PROCEDIMIENTOS METODOLÓGICOS Y RESULTADOS

Desde el punto de vista académico, esta propuesta está fundamentada en tres pilares: i) El Constructivismo, como modelo educativo que implica generar estrategias de aprendizaje para desarrollar las competencias de un estudiante, no solo en lo referente a conocimientos sino también a destrezas, habilidades, actitudes y valores; ii) La adquisición de mayores conocimientos estimulando la comprensión. Es decir desarrollar en los educandos el pensamiento crítico, el análisis de texto, y el uso de las herramientas informáticas para facilitar el proceso de adquisición de conocimientos; y iii) El desarrollo de entornos virtuales de enseñanza-aprendizaje, también conocidos como sistemas de gestión de conocimiento (LMS), que posibilitan la interactividad entre profesores y estudiantes fomentando el aprendizaje colaborativo. La articulación de estos tres pilares permite lograr en los estudiantes efectos satisfactorios en su aprendizaje (figura 1).

Tal como se puede apreciar en la figura 1, esta propuesta está dividida en dos grandes bloques. El primero que tiene que ver con el proceso ingenieril de desarrollo e implantación del software a través de OOHDM. El segundo concerniente a la creación de OA con la Metodología CROA, con cada una de sus etapas. Esto ha sido articulado con la técnica didáctica de aprendizaje basado en problemas, para desarrollar en los estudiantes las habilidades de comunicación, trabajo colaborativo, investigación, selección de información y auto-aprendizaje. Finalmente, muestra cómo se vincula la etapa de Diseño con la Taxonomía de Bloom, para evaluar dos de sus dominios: a) el conocimiento, que es la información, los principios y las ideas sobre un tema específico; b) la comprensión, que es la capacidad que desarrollan los estudiantes para entender, explicar, describir, comparar y distinguir objetos del mundo real. De esta propuesta se han derivado una guía metodológica y un curso en-línea de Fundamentos de Seguridad de la Información. A continuación se describe brevemente las fases de la propuesta:

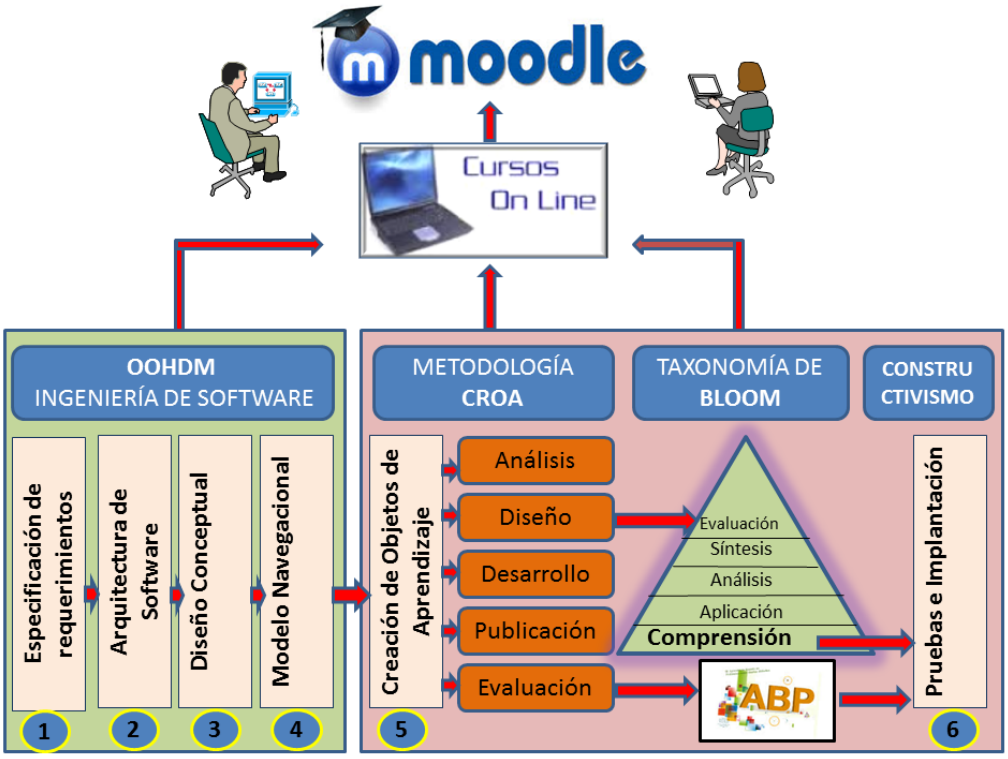


Figura 1. Elementos de la propuesta metodológica y sus interrelaciones.
Fuente: los autores.

Fase 1: Especificación de Requerimientos

Como requisitos esenciales (conforme al estándar IEEE-830, 1998), el curso en-línea de Fundamentos de Seguridad de la Información debía tener al menos tres unidades didácticas: (1) Programas indeseables o maliciosos; (2) Tipos de ataques de Ciberseguridad; (3) Análisis de vulnerabilidades. Debía ser subido a un entorno virtual de E/A, por lo que fue subido en Moodle, en un aula virtual de CEDIA (Red Nacional de Educación e Investigación del Ecuador). Debía incluir actividades de aprendizaje individuales y colaborativas aplicando el aprendizaje basado en problemas. Debía vincular las interrelaciones de sus dos actores principales, el estudiante autenticado y el tutor. El primero, que cumpliría las siguientes tareas y roles: (i) Acceder al aula virtual; (ii) Acceder a los materiales del curso; (iii) Acceder a las lecciones del curso; (iv) Acceder a los OA del curso; (v) Interactuar con herramientas de colaboración; (vi) Revisar los problemas resueltos (i.e., revisar los objetivos, acceder a los contenidos, resolver las actividades de aprendizaje y el test de auto-evaluación); (vii) Resolver los problemas propuestos. El segundo actor, por su parte realizaría el seguimiento académico, revisaría los foros y verificaría la resolución de los problemas propuestos.

Fase 2: Arquitectura de Software

El curso en-línea ha sido diseñado con una arquitectura Cliente/Servidor en n-capas las cuales son: a) Capa del Usuario, que le permite navegar por el sitio Web; b) Capa del LMS, que permite acceder a los cursos de formación y capacitación que ofrece CEDIA, a la bases de datos de Moodle y a los OA que tiene cada uno de los tres cursos desarrollados.

Fase 3: Diseño Conceptual

En esta fase se generó el modelo conceptual de la aplicación, donde las clases, relaciones y cardinalidades se definen de acuerdo a las reglas que se aplican sobre los Diagramas de Interacción de Usuario. Este modelo ha considerado las entidades que se relacionan dentro de Moodle con los OA del curso, donde se distinguen metadatos de tipo educacional.

Fase 4: Modelo Navegacional

Una vez realizado el diseño conceptual, hay que diseñar el Modelo Navegacional, que es una vista del modelo conceptual que refleja la información accesible a un usuario, y los caminos y estructuras de acceso para llegar a ella. De modo que, en esta propuesta se deben diseñar los aspectos navegacionales, teniendo en cuenta qué objetos del modelo básico van a ser navegables, qué tipo de relaciones, estructuras de composición y conexiones

existen entre los objetos. Esto permite a la aplicación Web ejecutar todas las tareas requeridas por los actores. En este estudio, se han considerado el uso de formularios Web de Moodle y los OA que han sido generados con Articulate Studio y Ardora.

Fase 5: Creación de OA

De acuerdo con Sanz et al. (2014), CROA es una metodología empleada para la creación de OA basada en el Diseño Instruccional, compuesta por cinco fases: Análisis, Diseño, Desarrollo, Publicación, y Evaluación. La figura 2 muestra que en cada fase existen varios insumos, técnicas, herramientas y grupos de documentos que cumplen con los principios del Diseño Instruccional y la Taxonomía de Bloom. Los OA del presente proyecto de investigación fueron diseñados y desarrollados en base a esta metodología. En esta fase de nuestra propuesta se describirán las cinco etapas:

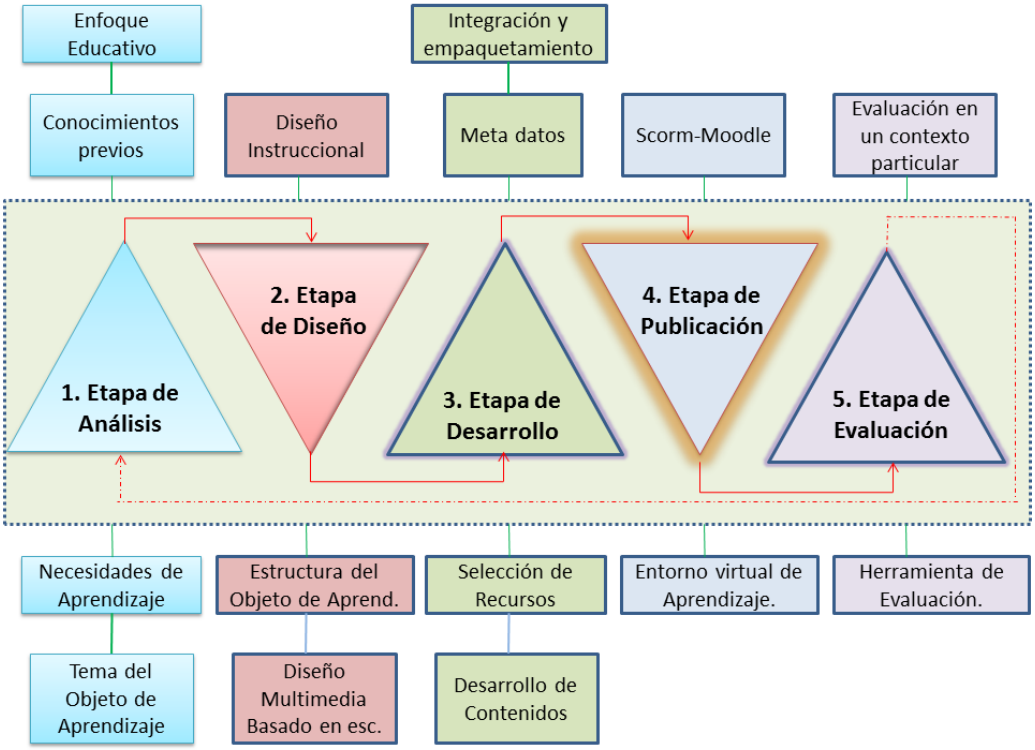


Figura 2. Fases de la Metodología CROA.
Fuente: ilustración adaptada de Sanz et al. (2014).

a) En la etapa de Análisis se identifican las necesidades educativas que originan la creación del OA y se propone responder a una serie de preguntas que orientan este análisis.

b) En la etapa de Diseño, se aborda los aspectos de diseño instruccional y de la estructura que tendrá el OA. Se analiza el diseño multimedia, a partir de una serie de escenarios que se proponen por defecto, a manera de guía para los autores.

c) En la etapa de Desarrollo, se debe abordar la selección de recursos (imágenes, videos, documentos de texto, presentaciones) para incorporar en el OA. Luego, se deben utilizar las plantillas elegidas, desarrollarlas en la/s herramienta/s de autor correspondientes, y finalmente integrarlas para conformar el Objeto citado, y empaquetarlo.

d) En la etapa de Publicación se propone publicar el Objeto de Aprendizaje en un repositorio de un Entorno Virtual de E/A. En el caso de nuestro proyecto, tal como se ha explicado, se levantaron sobre la plataforma Moodle de CEDIA.

e) En la etapa de Evaluación se analizan las posibles barreras encontradas en el uso del objeto de aprendizaje en contextos específicos. Hasta la realización de este proyecto, se han encontrado diferentes métodos de evaluación, como la Herramienta de Evaluación de Calidad de OA basada en CodA (Fernández, Domínguez, & de Armas, 2012).

Fase 6: Pruebas, Implantación y evaluación de resultados

El curso en-línea de Fundamentos de Seguridad de la Información, está implantado en Moodle de CEDIA, cuya dirección IP es: 190.15.141.97. Este curso está organizado por lecciones o módulos (temas) y por OA (subtemas),

donde cada objeto se compone por un objetivo de aprendizaje, contenido educativo, actividades de aprendizaje, autoevaluación y metadatos. El primer curso denominado *Programas Indeseables o Maliciosos* contiene nueve OA y se compone por tres lecciones que son: a) Conceptos Generales del Malware; b) Amenaza de Denegación de Servicios; c) Ingeniería Social. El segundo curso denominado *Tipos de Ataques de Ciberseguridad* contiene nueve OA y se compone por dos lecciones que son: a) Conceptos Generales de los ataques informáticos; b) Tipos de Ataques. El tercer curso denominado *Análisis de Vulnerabilidad* contiene ocho OA y se compone por una lección que es: a) Análisis de Vulnerabilidades. Cada curso contiene un grupo de problemas propuestos y resueltos utilizando la técnica didáctica de ABP.

Para corroborar los beneficios del diseño didáctico de cursos en-línea con OA, se aplicó una evaluación en base a la Matriz COdA (Fernández et al., 2012), a dos grupos de 22 y 28 estudiantes respectivamente, del cuarto nivel de la carrera de Ingeniería de Sistemas e Informática que requieren este conocimiento. Este instrumento evaluó los siguientes criterios: i) objetivos y coherencia didáctica; ii) calidad de contenidos; iii) capacidad de generar reflexión, crítica e innovación; iv) interactividad y adaptabilidad del objeto de aprendizaje; v) motivación; vi) formato y diseño; y vii) usabilidad. Luego del procesamiento estadístico, se pudo determinar que se obtuvieron resultados entre 4 y 5 (de una escala ente 1 y 5), para los siete parámetros de evaluación seleccionados, según el siguiente detalle:

a) Los resultados muestran un equilibrio de calidad en los parámetros de evaluación.

b) El parámetro con mayor puntaje es el de Objetivos y coherencia didáctica, esto quiere decir que lo que se aprende con el curso, está de muy acuerdo a los objetivos planteados, los destinatarios sienten que las habilidades adquiridas son las necesarias y las instrucciones de usos están bien definidas.

c) Los parámetros evaluados tanto de carácter didáctico y tecnológico (OA y EVEA) se pudieron valorar de forma conjunta, esto quiero decir que el desarrollo de cursos online implementando ambas metodologías, proporcionan buenos resultados.

Como segundo aspecto de evaluación, se compararon los beneficios del diseño didáctico de cursos en-línea con OA, *versus* los cursos en-línea tradicionales. Para este propósito, se aplicó una adaptación de la matriz de evaluación de diseño didáctico de cursos virtuales de la Universidad Pontificia Bolivariana. Esta matriz fue aplicada a los mismos grupos de estudiantes que se les aplicó la herramienta COdA. Este instrumento evaluó lo siguiente: i) presentación y contextualización; ii) planteamiento de objetivos; iii) coherencia de contenidos; iv) actividades de aprendizaje; v) evaluaciones, vi) usabilidad e interactividad. Luego del procesamiento estadístico, se pudo determinar que en todos los parámetros evaluados, el diseño de cursos online con OA supera significativamente al diseño de cursos online Tradicionales, lo cual contrasta nuestra hipótesis (figura 3).

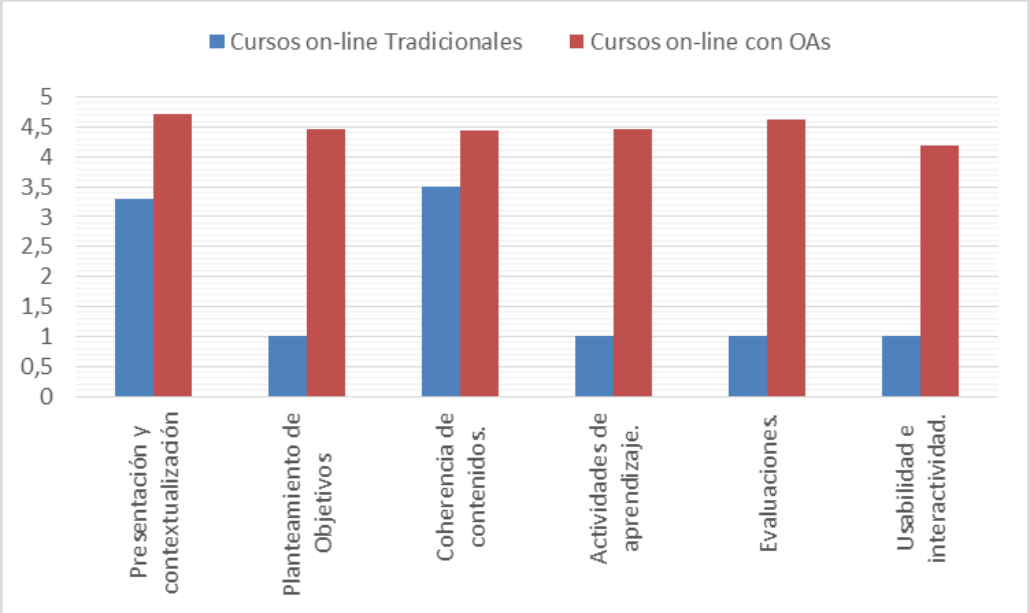


Figura 3. Resultados de la evaluación del diseño didáctico de cursos virtuales on-line.
Fuente: los autores.

Finalmente, contrastando con la premisa de investigación de este estudio, se ha demostrado que se cumple los dos primeros niveles de complejidad creciente: conocimiento y comprensión de la taxonomía de Bloom, pues se ha integrado OA, EVEAS con ABP, lo que ha permitido articular los tres ámbitos fundamentales en la educación: enseñanza, aprendizaje (i.e. ámbito cognitivo, afectivo y psicomotor) y evaluación (Bloom, 1956), utilizando tecnología de la era digital.

CONCLUSIONES

Este estudio tuvo como objetivo estimular la comprensión conceptual y procedural de estudiantes de carreras técnicas, a través de una propuesta metodológica para el diseño e implementación de cursos en-línea basados en OA. Para ello, se aplicaron OOHDM para el diseño navegacional del entorno virtual de aprendizaje de cursos en línea y el diseño conceptual de la aplicación y la metodología CROA para el diseño de los AO, donde la principal ventaja de su aplicación era lograr una adecuada estructuración y organización de los OA desde un punto de vista pedagógico y tecnológico, sin importar las herramientas de autor que se utilicen para armar los mismos, ni tampoco dependería de los EVEAS donde se suben los OA de un curso. En esta investigación se diseñó, se implementó y se evaluó un curso en-línea de fundamentos de Seguridad de la Información utilizando OA y Moodle. Los resultados muestran la satisfacción de los estudiantes en cuanto a la organización del curso, contenidos, su utilidad, empleo de casos prácticos, uso de medios audiovisuales, utilización de dinámicas de grupo, duración, material empleado, todo con una percepción general favorable desde el punto de vista del diseño instruccional que plantea la metodología CROA. Por lo expuesto, este curso en línea permite estimular la comprensión conceptual y procedural de los estudiantes de carreras de ingeniería. Como trabajo futuro se planea desarrollar una herramienta de autor para crear ambientes virtuales y que agregue automáticamente las bondades de objetos SLOODLE (Simulation Linked Object Oriented Dynamic Learning Environment) *open source*.

REFERENCIAS

Alonso, M., Castillo, I., Martínez, V., & Muñoz, Y. (2013). MEDOA: Metodología para el Desarrollo de Objetos de Aprendizaje. En *12a Conferencia Iberoamericana en Sistemas, Cibernética e informática: CISCI.*

Bloom, B. S. (1956). *Taxonomy of educational objectives: the classification of educational goals*. Longmans Green and Company.

De Giusti, A., Zangara, A., Sanz, C., Moralejo, L., Barranquero, F., & Naiouf, M. (2015). Producción de Objetos de Aprendizaje para la enseñanza universitaria. En *21o Congreso Argentino de Ciencias de la Computación.*

Fernández, A., Domínguez, E., & de Armas, I. (2012). *Herramienta de evaluación de la calidad de Objetos de Aprendizaje (Herramienta COdA): Guía para la producción y evaluación de materiales didácticos digitales* (1.1 ed.). Universidad Complutense de Madrid. (Licencia de Creative Commons: Reconocimiento – No comercial)

García Vera, V. E., Roig Vila, R., Marti Ciriquian, P., & García Ferrández, P. (2012). Uso de recursos digitales en el ámbito de ingeniería de edificación. En *10a Jornadas de Redes de Investigación en Docencia Universitaria.* Universidad de Alicante.

Maldonado, J., Carvallo, J., & Siguencia, J. (2015). *Consorcio Ecuatoriano para el Desarrollo de Internet Avanzado (CEDIA)* [Repositorio Digital].

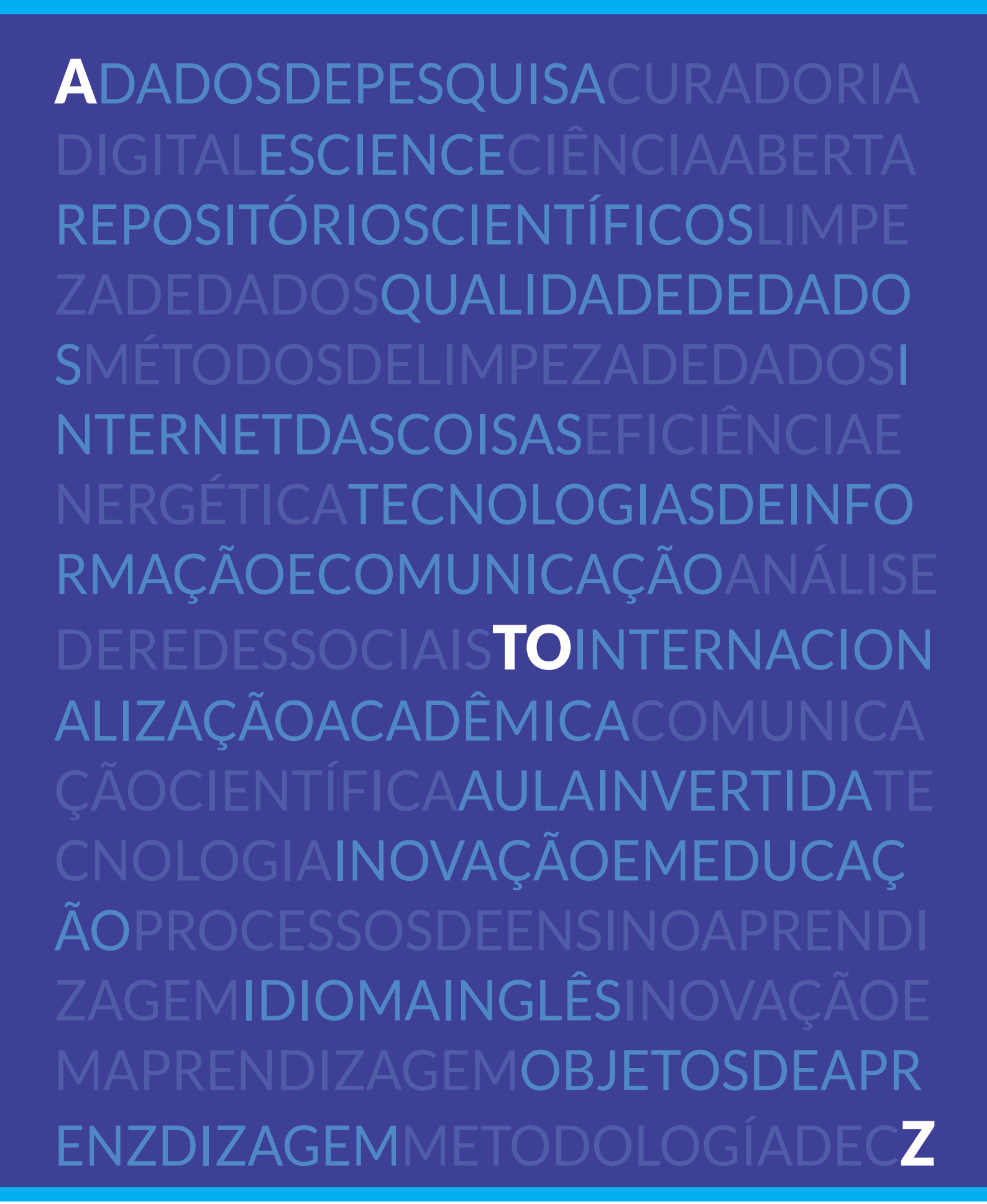
Mestre, G., Torres, E., Díaz, D., & Gómez, J. (2014). *Lineamientos pedagógicos, comunicativos y tecnológicos para la producción de cursos mediados por TICs en la educación a distancia* (1a ed.). Cartagena de Indias, Colombia: Universidad Tecnológica de Bolívar.

Pressman, R. S. (2010). *Ingeniería del software: un enfoque práctico* (6a ed.). México DF, México: Editorial McGraw-Hill.

Sanz, C., Moralejo, L., & Barranquero, F. (2014). *Metodología CROA* [Materiales del Curso de Doctorado]. Universidad Nacional de la Plata, Argentina. Descargado de http://croa.info.unlp.edu.ar

Soto de Giorgis, R., Palma Muñoz, W., & Roncagliolo de la Horra, S. (2002). *Propuesta de un modelo navegacional para el desarrollo de aplicaciones basadas en OOHDM.*

Cómo citar este shortpaper (APA):
Zambrano, M. E., Fuertes-Diaz, W. M., Pérez-Herrera, D. D., Villacís, C. J. & Pérez-Estévez, E. (2016). Producción de cursos en-línea basados en objetos de aprendizaje: una propuesta metodológica orientada a estudiantes de carreras técnicas. *AtoZ: novas práticas em informação e conhecimento, 5*(2), 115 – 121. Descargado de: http://dx.doi.org/10.5380/atoz.v5i2.49884



Universidade Federal do Paraná
Setor de Ciências Sociais Aplicadas
Programa de Pós-Graduação em Ciência, Gestão e Tecnologia da Informação
Av. Prefeito Lothário Meissner, 632 - Campus III - Jardim Botânico
80210-170 - Curitiba, PR - Brasil
homepage: http://www.revistas.ufpr.br/atoz | e-mail: revistaatoz@ufpr.br