

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
Faculdade de Biblioteconomia e Comunicação
Departamento de Ciências da Informação
Curso de Biblioteconomia

Andreza Lemke de Souza

**A presença dos periódicos da Ciência da Informação
no Google Scholar**

Porto Alegre
2018

Andreza Lemke de Souza

**A presença dos periódicos da Ciência da Informação
no Google Scholar**

Monografia apresentada como requisito para
obtenção do grau de bacharel em Biblioteconomia
pelo Departamento de Ciências da Informação da
Universidade Federal do Rio Grande do Sul.

Orientadora: Prof^a. Dr.^a Sônia Elisa Caregnato

Coorientadora: M.e Luciana Monteiro Krebs

Porto Alegre

2018

UNIVERSIDADE FEDERAL DO SUL

Reitor: Prof. Dr. Rui Vicente Oppermann

Vice-Reitora: Prof^a. Dr^a. Jane Fraga Tutikian

FACULDADE DE BIBLIOTECONOMIA E COMUNICAÇÃO

Diretora: Prof^a. Dr^a. Karla Maria Müller

Vice-Diretora: Prof^a. Dr^a. Ilza Maria Tourinho Girardi

DEPARTAMENTO DE CIÊNCIAS DA INFORMAÇÃO

Chefe: Prof^a. Dr^a. Jeniffer Alves Cuty

Chefe Substituta: Prof^a. Dr^a. Eliane Lourdes da Silva Moro

COMISSÃO DE GRADUAÇÃO DO CURSO DE BIBLIOTECONOMIA

Coordenadora: Prof^a. Dr^a. Rita do Carmo Ferreira Laipelt

Coordenador Substituto: Prof. Dr. Rene Faustino Gabriel Júnior

CIP - Catalogação na Publicação

Souza, Andreza Lemke de
A presença dos periódicos da Ciência da Informação
no Google Scholar / Andreza Lemke de Souza. -- 2018.
78 f.
Orientadora: Sônia Elisa Caregnato.

Coorientadora: Luciana Monteiro Krebs.

Trabalho de conclusão de curso (Graduação) --
Universidade Federal do Rio Grande do Sul, Faculdade
de Biblioteconomia e Comunicação, Curso de
Biblioteconomia, Porto Alegre, BR-RS, 2018.

1. Google Scholar. 2. Presença online. 3.
Periódicos eletrônicos. 4. Ciência da Informação. I.
Caregnato, Sônia Elisa, orient. II. Krebs, Luciana
Monteiro, coorient. III. Título.

Elaborada pelo Sistema de Geração Automática de Ficha Catalográfica da UFRGS com os dados fornecidos pelo(a) autor(a).

Universidade Federal do Rio Grande do Sul

Faculdade de Biblioteconomia e Comunicação

Departamento de Ciências da Informação

Rua Ramiro Barcelos, 2705 - Campus Saúde

CEP 90035-007 Porto Alegre/RS

Fone: (51) 3308 5067

Fax: (51) 3308 5435

E-mail: dci@ufrgs.br

Andreza Lemke de Souza

**A presença dos periódicos da Ciência da Informação
no Google Scholar**

Monografia apresentada como requisito para
obtenção do grau de bacharel em Biblioteconomia
pelo Departamento de Ciências da Informação da
Universidade Federal do Rio Grande do Sul.

Orientadora: Prof^a. Dr.^a Sônia Elisa Caregnato

Coorientadora: M.e Luciana Monteiro Krebs

Porto Alegre, 17 de dezembro de 2018.

Banca Examinadora

Prof^a. Dr.^a Sônia Elisa Caregnato – UFRGS (Orientadora)

M.e Luciana Monteiro Krebs – UFRGS (Coorientadora)

Prof. Dr. Rafael Port da Rocha – UFRGS (Examinador)

Prof^a. Dr.^a Samile Andrea De Souza Vanz – UFRGS (Examinadora)

Agradecimentos

No último dia de inscrição para o vestibular de 2013 fui convencida a me inscrever, agora que findo esta etapa e tenho a oportunidade de demonstrar a minha gratidão por aqueles que contribuíram com ela, achei justo começar por quem foi crucial ao seu início. Luís Eduardo Marques agradeço por ter acreditado em mim quando não acreditei, por ter contornado minha frustração prévia e insistido para que fizesse o vestibular. A tua insistência naquele dia tornou meu sonho possível e o transformei realidade graças ao teu amparo. Obrigada pelo apoio e incentivo constantes e por ser a minha ponte sobre as águas turbulentas da vida.

Agradeço às luzinhas da minha vida: Henrique de Souza Marques, serzinho que mais sofreu com essa empreitada, obrigada pela compreensão e ajuda nas épocas de final de semestre, 'prova, prova, trabalho!'. És, e sempre será, o meu leitor preferido! E à afilhada mais espetacular da galáxia, Joana de Souza Gallas, amo você de montão, agradeço e admiro tua empatia e compreensão. O sorriso e o amor de vocês me fortalecem para 'continuar a nadar'.

Agradeço aos professores: Glória Ferreira pela aconchegante recepção na matrícula e pelas aulas extremamente explicativas; Ivete Tazima pelas discussões acerca da profissão; Jackson Medeiros pelas aulas positivamente perturbadoras; Maria Lúcia Dias pela compreensão e empatia; Martha Bonotto, pela dedicação, delicadeza e paciência em ensinar; Rafael da Rocha, pela paciência de repetir incansavelmente quando a gente não entende o idioma da tecnologia e por aceitar compor a minha banca de avaliação na última hora; Rita Laipelt pela dedicação incessante aos alunos; Rodrigo Caxias por ser o professor-amigo a quem sempre podemos recorrer e Samile Vanz, pelas experiências proporcionadas em aula e por aceitar integrar a banca de avaliação.

Meu agradecimento especial para as professoras que me escolheram como monitora: Eliane Moro, tua visão cor de rosa do mundo mostrou-me as nuances das necessidades alheias; Helen Rozados, além da monitoria, proporcionou-me a primeira aproximação à pesquisa, agradeço as oportunidades, as aprendizagens, as conversas, as orientações e a dedicação a mim aplicadas; Sônia Caregnato que acolheu minhas indecisões e aceitou ser minha orientadora e que, junto à Luciana Krebs, conseguiu manter minha sanidade mental nos momentos conturbados, sou

grata a ambas pelo auxílio impecável na elaboração desta pesquisa e incentivo nos momentos difíceis.

Jeniffer Cuty, Ketlen Stueber e Marlise Giovanaz que não foram minhas professoras oficialmente, mas que nos breves encontros que tivemos cativaram minha admiração.

Minha gratidão a todos os professores pelos ensinamentos, em que muito extrapolaram o âmbito acadêmico, terão sempre meu respeito e admiração.

Aos funcionários da Biblioteca Central Irmão José Otão da PUCRS que me acolheram com carinho e dedicação, especialmente àqueles que receberam a tarefa de repassar seus conhecimentos à estagiária e o fizeram com maestria: Ana Paula Magnus, Anamaria Ferreira, Antônio Triches, Ariane Oliveira, Clarissa Selbach, Ednei Freitas, Ester Machado, Francione Rodrigues, Gabriel Hamester, Juliana Santos, Lígia de Freitas, Roberto Moura, Loiva Novak, Lucas Kern, Márcio, Sabrina Vicari, Salete Sartori, Tatiana Schuh, Tiago e Wagner Souza. Em particular, à Alessandra Fagundes, Aline Matte e Ísis Bitello pelo amparo. À Débora Kraemer e Michelangelo Viana pela oportunidade. À Isabel Crespo agradeço pela oportunidade e orientação do estágio e por aceitar participar da banca de avaliação, mesmo que não tenha se concretizado.

Meu agradecimento às monitoras mais incríveis e dedicadas Gabriela Nora e Marta de Oliveira. Aos colegas da turma de 2013/2 e aos amigos que tornaram o caminho mais agradável e as aprendizagens mais construtivas: Alberto Garcez, Anaida Batista, Ângelo Reginatto, Ediane Gheno, Estêvão Trindade, Fabíola Fagundes, Luiz Fernando Massoni e Samuel da Rosa. Em especial à Berenice Guilhermano pela amizade e parceria frequente nos trabalhos, à Luciana Milani e à Isadora Turcatel pela amizade e socorros nos momentos conturbados, sou grata por tê-las em minha vida.

Por fim, agradeço à minha família pelo apoio e amor incondicionais. Às luzes da minha vida: meus pais, João Ricardo e Helga e à minha *irmã* Greice, agradeço pelos exemplos de amor, empatia, comprometimento e retidão, sempre serão parte de mim. Vos amo eternamente do tamanho do mundo grande e espero o dia em que nos encontraremos novamente, porque somos uma só história...

Resumo

Apresenta pesquisa de presença online dos periódicos científicos correntes da área da Ciência da Informação no Google Scholar mediante a aplicação do comando *site*. O estudo possui tipologia básica, de cunho exploratório e de abordagem quantitativa, a amostra de periódicos analisados foi determinada de forma não-probabilística intencional. Utiliza a pesquisa documental como procedimento para coleta de dados nos *websites* dos periódicos analisados, no buscador Google Scholar com aplicação do comando *site* e na plataforma Sucupira. O formulário foi instrumento de coleta utilizado. O referencial teórico aborda o processo de comunicação científica, os periódicos científicos e contextualiza brevemente os subcampos dos estudos métricos e suas aplicações na *web*, finalizando com a apresentação do Google Scholar e dos estudos métricos aplicados a ele. No desenvolvimento da pesquisa, analisa individualmente as características: estratificação Qualis, existência de perfil no GS, formato de arquivo disponibilizados, padrão de preservação digital, padrão de metadados empregado no *site*, utilização de DOI da revista e versão do OJS utilizada, para posteriormente relacioná-las às taxas de presença dos periódicos no Google Scholar à procura de aspectos que evidenciem influenciar na taxa de presença dos periódicos no referido buscador. Conclui que as características relacionadas à taxa de presença não evidenciaram influência direta na presença dos periódicos no Google Scholar e que as métricas empregadas no ambiente virtual necessitam de exploração para seu desenvolvimento para a definição de estratégias eficazes.

Palavras-chave: Google Scholar. Presença online. Periódicos eletrônicos. Ciência da Informação.

Abstract

Presents research of online presence of current scientific journals in the area of Information Science in Google Scholar by applying the command site. The study has a basic typology, exploratory and quantitative approach, the sample of analyzed periodicals was determined in an intentional non-probabilistic way. It uses documentary research as a procedure for collecting data on the websites of analyzed journals, in the Google Scholar search engine with the site command and in the Sucupira platform. The form was a collection instrument used. The theoretical framework deals with the scientific communication process, the scientific journals and briefly contextualizes the subfields of the metric studies and their applications on the web, ending with the presentation of Google Scholar and the metric studies applied to it. In the development of the research, it analyzes individually the characteristics: Qualis classification, existence of profile in GS, available file format, archive standard, metadata standard used in the site, use of DOI of the magazine and version of OJS used, to the presence rates of the journals in Google Scholar in search of aspects that evidence to influence the presence rate of the journals in the referred search engine. It concludes that the characteristics related to the presence rate did not show a direct influence on the presence of journals in Google Scholar and that the metrics used in the virtual environment need exploration for its development in order to define effective strategies.

Keywords: Google Scholar. Online presence. Online journal. Information Science.

Lista de ilustrações

Figura 1 – Tela da caixa de busca por publicações no GSM (Página de h5-index do periódico).....	16
Figura 2 – Tela da caixa de busca por perfis (Perfil de pesquisador).....	16
Figura 3 – Tela de aplicação de comando <i>site</i> no GS.....	37
Figura 4 – Tela de busca da página de h5-index do periódico no GSM da Revista Em Questão	43
Figura 5 – Tela perfil de pesquisador da Revista Em Questão.....	44
Figura 6 – Tela de busca por perfil de pesquisador da Revista Em Questão	45
Figura 7 – Postagem em rede social do periódico Em Questão sobre perfil no GS .	45
Figura 8 – Tela de recuperação de busca com comando <i>site</i> inválido	49
Figura 9 – Destaque em registro que não é documento científico do periódico	51
Figura 10 – Busca por título de artigo da revista Em Questão	67
Gráfico 1 – Quantidade de periódicos por estratificação Qualis.....	54
Gráfico 2 – Existência de perfil dos periódicos no Google Scholar	56
Gráfico 3 - Padrão de metadados empregado no <i>site</i> do periódico	57
Gráfico 4 – Versões do OJS utilizadas pelos periódicos	58
Quadro 1 – Informações para cálculo da taxa de presença no GS	47
Quadro 2 – Características dos periódicos selecionados para o estudo	52
Tabela 1 – Relação Qualis x taxa de presença.....	59
Tabela 2 – Relação perfil no GS x taxa de presença.....	60
Tabela 3 – Formato de arquivo x taxa de presença.....	62
Tabela 4 – Padrão de metadados x taxa de presença.....	63
Tabela 5 – DOI x taxa de presença	64
Tabela 6 – Plataforma x taxa de presença	65

Sumário

1	Introdução	11
1.1	Identificação do Problema	12
1.2	Objetivos	12
1.2.1	<i>Objetivo Geral</i>	13
1.2.2	<i>Objetivos Específicos</i>	13
1.3	Justificativa	13
1.4	Questões Conceituais	14
1.5	Estrutura	17
2	Referencial Teórico	18
2.1	Comunicação Científica e Periódicos Acadêmicos	18
2.2	Estudos Métricos e Aplicações na <i>Web</i>	23
2.3	Google Scholar e os Estudos Métricos na <i>web</i>	27
3	Metodologia	33
3.1	Seleção do objeto de pesquisa	35
3.2	Procedimentos da Coleta de Dados	36
3.2.1	<i>Coleta de Dados no Google Scholar</i>	37
3.2.2	<i>Coleta de Dados nos Sites dos Periódicos</i>	38
3.3	Etapas do Processo de Coleta de Dados	40
3.4	Métodos para Tratamento dos Dados e Apresentação dos Resultados	41
3.5	Limites do Estudo	42
4	Apresentação e Análise dos Dados	46
4.1	Taxa de Presença dos Periódicos no Google Scholar	46
4.2	Descrição das Características de Análise dos Periódicos	51
4.3	Relações entre característica e taxa de presença	58
4.3.1	<i>Estratificação Qualis</i>	58
4.3.2	<i>Existência de perfil no GS</i>	60
4.3.3	<i>Formato de arquivo utilizado</i>	61
4.3.4	<i>Padrão de preservação digital</i>	63
4.3.5	<i>Padrão de metadados empregado no site</i>	63
4.3.6	<i>Utilização de DOI da revista</i>	64
4.3.7	<i>Versão do OJS</i>	65

4.4	Outras Proposições	66
5	Considerações finais.....	69
	Referências	71
	Apêndice – Formulário de coleta de dados.....	78

1 Introdução

O progresso científico está intrinsecamente ligado à comunicação científica, historicamente representada pelos periódicos científicos. Contudo, as transformações das tecnologias digitais e, posteriormente, o desenvolvimento do movimento de acesso aberto, remodelaram significativamente o processo, a disseminação e o acesso à comunicação científica.

Os avanços tecnológicos proporcionam o surgimento de redes de comunicação, amplificadas pela Internet. Neste contexto, emergiram as bases de dados digitais, com o propósito de reunir, organizar e medir a produção de cada área. Entretanto, ao mesmo tempo em que facilitam a busca e recuperação da informação, também ocasionam limitações, sejam elas por restrições temáticas, geográficas ou relativas ao idioma.

Tais limitações são reduzidas com a estratégia utilizada pelo buscador Google de reunir documentos relacionados e/ou que contenham as palavras mencionadas na busca, porém uma busca geral na *web* gera um número exorbitante de resultados e torna-se uma alternativa inviável para pesquisadores. Deste modo, em 2004, a empresa Google lançou o Google Scholar (Google Acadêmico), motor de busca voltado à pesquisa científica.

O Google Scholar (GS) funciona da mesma forma que seu gerador, entretanto restringe a pesquisa às páginas voltadas a atividades acadêmicas: universidades, institutos de pesquisa, repositórios acadêmicos, periódicos técnicos e científicos, bases de dados. Além disso, a lista de resultados contempla documentos com diferentes tipologias, idiomas ou formatos de acesso. O que inclui na resposta de busca livros, artigos, teses, dissertações, trabalhos de conclusão, trabalhos apresentados em eventos, patentes e outros, sem filtrar os resultados por idioma ou áreas temáticas.

Evidente que o modelo do buscador voltado à pesquisa também exhibe lacunas. Sabe-se que o idioma em que o usuário efetua a busca e seu perfil de busca interferem no ordenamento, inclusão e exclusão de resultados. Porém, não há divulgação, por parte da empresa Google, das diretrizes seguidas quanto aos critérios para definição da apresentação dos resultados da busca. Outros problemas são: a credibilidade frente aos pesquisadores, a possibilidade de manipulação da

visibilidade medida por estudos bibliométricos e lacunas na indexação, como as identificadas em análise preliminar e que incentivaram a realização dessa pesquisa.

O contínuo e ágil crescimento da produção informacional torna essencial o desenvolvimento de métodos que contribuam para a indexação e favoreçam a recuperação da informação, mas para obter este desenvolvimento é necessário conhecer seu funcionamento ou, mesmo, seu estado de arte. Este trabalho visa dar sua pequena parcela de contribuição para isso.

1.1 Identificação do Problema

Um dos grandes desafios de qualquer sistema de recuperação da informação é oferecer ao consulente o resultado mais completo e preciso possível. Conciliar revocação e precisão nos catálogos ainda desafia a Biblioteconomia e, aparentemente, a estratégia utilizada pela Google, no mínimo, reduziria a lacuna. Entretanto, apesar do buscador ser referência para busca de informações na atualidade e tendo seu nome como sinônimo de pesquisar, divergências de resultados são percebidas facilmente por qualquer usuário mais atento.

Ao se utilizar múltiplas estratégias para recuperar os artigos do periódico *Em Questão*, da Universidade Federal do Rio Grande do Sul (UFRGS), no referido buscador, por exemplo, foram constatadas divergências no número de resultados apresentados. O fato é preocupante, considerando que um número relevante de acessos provém do redirecionamento do buscador e que a exclusão de conteúdo interfere negativamente na visibilidade da publicação. Neste contexto, emerge o questionamento:

Quais características dos periódicos brasileiros da área da Ciência da Informação com publicações correntes em 2018 interferem na indexação e, conseqüentemente, em suas taxas de presença no Google Scholar?

1.2 Objetivos

Como estratégia para compreender e delimitar o problema de pesquisa proposto foram definidos o objetivo geral e os específicos, explicitados a seguir.

1.2.1 Objetivo Geral

Mapear a influência das características dos periódicos científicos brasileiros da área da Ciência da Informação nas suas taxas de presença no Google Scholar

1.2.2 Objetivos Específicos

São objetivos específicos:

- a) gerar a taxa de presença no Google Scholar dos periódicos de Ciência da Informação;
- b) caracterizar os periódicos selecionados quanto:
 - estratificação Qualis;
 - existência de perfil no GS;
 - formatos de arquivo disponibilizados;
 - padrão de preservação digital;
 - padrão de metadados empregado no *site*;
 - utilização de DOI da revista
 - versão do OJS utilizada;
- c) relacionar as características dos periódicos às taxas de presença no Google Scholar.

1.3 Justificativa

A mensuração das atividades de produção e uso do conhecimento relaciona-se com a Ciência da Informação e a Biblioteconomia desde o início. É natural medir acessos, busca, recuperação da informação e transcender para quantificar produção, citação e outros. Mais que natural, é essencial para atender as especificidades dos usuários. Com desenvolvimento da *World Wide Web (web)*, não é surpreendente que as técnicas de bibliometria, cientometria e informetria fossem adaptadas e reelaboradas para a análise métrica de ambientes *online*.

A *web* consiste no mais volumoso repositório e a maior fonte de informação já conhecida pela humanidade (VANTI, 2005). Grande parte deste repertório pode ser recuperado através dos serviços do Google, buscador tão utilizado que

coloquialmente seu nome é sinônimo de pesquisar. Devido a isso, pesquisas que identifiquem suas qualidades e limitações fazem-se necessárias, pois conhecer as fragilidades e lacunas é fundamental para montar estratégias para supri-las.

A análise da presença dos periódicos científicos no Google Scholar, segundo sua taxa de presença no GS, pretende evidenciar os estudos métricos nas bases de dados abertas, especialmente quanto à visibilidade, para identificar falhas e possíveis melhorias, e, assim, ampliar para o pesquisador as possibilidades de observação de um objeto de pesquisa. Desse modo, visa-se contribuir com os estudos métricos e de visibilidade em periódicos científicos digitais em motores de busca abertos e multidisciplinares.

Também se verificou que os estudos referentes à presença dos periódicos científicos no GS ainda são escassos, sendo um pouco mais comuns os estudos estrangeiros considerando a visibilidade, com base na indexação no buscador, em repositórios institucionais, como Arlitsch e O'Brien (2012) e Orduña-Malea e Delgado López-Cózar (2015). Por esta razão, entende-se ser esse um assunto de suma importância para a evolução e o melhoramento desse modelo de busca, que pretende atender a pesquisa científica de forma abrangente e multidisciplinar.

A motivação para elaboração desse estudo parte da premissa do acesso aberto, para propagação da ciência e invade o campo da visibilidade da comunicação científica na *web*. Acredita-se que o acesso aberto oportuniza o destaque à ciência e com isso favorece sua continuidade e aperfeiçoamento. Além de vislumbrar em motores de busca, como o GS, a multidisciplinaridade necessária para a compreensão dos diversos fatores que compõe cada questão levantada por pesquisadores, pois a estratégia de reunir documentos de assuntos similares, porém sem limitações disciplinares, como ocorre nas Bases de Dados comerciais, possibilita o contato com diferentes aspectos sobre o mesmo tema. Por fim, pressupõe-se a necessidade de maior inserção dos bibliotecários nos âmbitos das tecnologias da informação visando maximizar suas potencialidades.

1.4 Questões Conceituais

O objeto deste estudo são as publicações periódicas científicas, portanto não surpreende que o termo e seu conceito apareçam no texto diversas vezes,

porém, por economia linguística, dá-se prioridade ao termo periódico. O termo revista também é utilizado para evitar repetições subsequentes ou nos casos em que a diferenciação de outro termo se faz necessária, como exposto posteriormente. Por tanto, designamos os termos periódico(s) e revista(s) para representar o conceito de:

[...] publicações seriadas, que se apresenta sob a forma de revista, boletim, anuário etc., editada em fascículos com designação numérica e/ou cronológica, em intervalos pré-fixados (periodicidade), por tempo indeterminado, com a colaboração, em geral, de diversas pessoas, tratando de assuntos diversos, dentro de uma política editorial definida, e que é objeto de Número Internacional Normalizado (ISSN). (ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS, 2003)

O termo publicação não é utilizado para representar o conceito referido acima com o intuito de minimizar possíveis ambiguidades, pois pode também expressar o ato de publicar no sentido editorial, assim como, por vezes, faz-se sinônimo dos termos escolhidos. O termo revista foi adotado com vistas principalmente da necessidade de diferenciar a página de h5-index do periódico no GSM e a revista na qual se fala, como exemplificado no título da Figura 4. A expressão negritada no título da figura citada pretende chamar a atenção para a distinção entre os conceitos de página de h5-index do periódico no GSM e revista, pois foi necessário adequar os termos utilizados pelo Google Scholar (inglês) ou Google Acadêmico (português).

Quando a língua inglesa está habilitada na configuração de idioma, o termo utilizado na aba Metrics do Google Scholar para busca por periódico é *search publication*. Já com a opção de idioma configurada para português a tradução é 'pesquisar publicações' (Figura 1). Devido a possíveis ambiguidades causadas pelo termo publicações, adota-se a expressão **Página de h5-index do periódico no GS Metrics** para representar a página *web* de um periódico no GS Metrics (GSM). Enfatiza-se que a página de h5-index do periódico é gerada automaticamente pelos algoritmos do GS, através da indexação dos artigos publicados pelo buscador, sem intervenção direta de seus gestores.

Figura 1 – Tela da caixa de busca por publicações no GSM (Página de h5-index do periódico)



Fonte: Dados da pesquisa (2018)

Conforme ilustra Figura 2, o fenômeno ocasionado pela configuração de idioma também ocorre para a pesquisa de perfil. Como a palavra perfil isoladamente não carrega todo o conceito necessário para sua distinção, adota-se a expressão Perfil de Pesquisador para representar a página *web* de um pesquisador no GS, onde é possível administrar redes de coautoria, incluir e excluir documentos de autoria do pesquisador, verificar número de citação de cada publicação, entre outros além do h-index. Ressalta-se que diferentemente da página de h5-index do periódico no GSM, o perfil de pesquisador é criado e gerenciado pelo pesquisador, sendo essa sua função precípua. No caso da utilização da ferramenta por periódicos, o gerenciamento do perfil é feito pelos responsáveis pelo periódico.

Figura 2 – Tela da caixa de busca por perfis (Perfil de pesquisador)



Fonte: Dados da pesquisa (2018)

Quanto a adesão do termo taxa de presença, segundo as explicações de Arlitsch e O'Brien (2012) pode-se inferir que a contagem de páginas recuperadas representa a presença de cada periódico no buscador e a taxa de indexação é

relação percentual entre os itens indexados e os publicados na revista. Contudo, com o intuito de diminuir a possível comparação com o conceito de indexação já utilizado para os meios tradicionais de recuperação da informação na Biblioteconomia, aliado à recente estruturação dos estudos relacionados à *web* e a definição dos termos ainda em desenvolvimento, optou-se pela utilização do termo taxa de presença. Ressalta-se que o Google Scholar não possui política de indexação clara e disseminada.

1.5 Estrutura

Este estudo está estruturado em 5 seções primárias (capítulos). A seção 1 trata da introdução, dos objetivos, geral e específicos, e da justificativa do trabalho. Na seção 2 figura o referencial teórico utilizado para embasar o estudo. A seção 3 disserta sobre a metodologia empregada para execução da pesquisa, culminando na seção 4 na qual são apresentados e analisados os dados coletados. Esta seção subdivide-se na análise da taxa de presença dos periódicos selecionados para o estudo, atendendo ao objetivo proposto na alínea a) da seção 1.2.2; na análise individual das características selecionadas, descritas na alínea b) da mesma seção; seguida das análises que atendem à alínea c), com os resultados das relações entre as taxas de presença e as características pesquisadas e, ao final dessa seção, a descrição de achados identificados no decorrer da pesquisa, porém que não figuram relacionados a seus objetivos. A seção 5 explana as considerações finais. Por fim, são apontadas as referências que embasaram esta investigação.

2 Referencial Teórico

Esta seção destina-se a abordagem de assuntos fundamentais ao embasamento teórico e desenvolvimento do tema do presente estudo. Inicia-se com um breve histórico e conceituação de comunicação científica e dos periódicos científicos, assim como a contextualização de algumas características dos periódicos. Posteriormente, apresenta-se de forma concisa os estudos métricos e suas aplicações na Web a fim de contextualizar os estudos métricos aplicados ao Google Scholar.

2.1 Comunicação Científica e Periódicos Acadêmicos

A ciência origina-se da curiosidade inerente ao ser humano e da observação humana dos fenômenos a sua volta. Desenvolve-se a partir de um processo social, que abrange saberes provenientes da prática e de procedimentos de experimentação. Caracteriza-se por ser uma atividade coletiva e colaborativa, possui como objetivo alcançar um consenso de opinião racional de maior abrangência possível, sendo “[...] precisa, metódica, acadêmica, lógica e prática [...]” (ZIMAN, 1979, p. 17).

Apesar de ser essencial à evolução, indagar, testar e concluir torna-se uma estratégia inviável se utilizada isoladamente. Para solucionar uma questão, é necessário munir-se de conhecimento sobre os fatores que a compõe ou que nela incidem. Tal atividade é morosa e necessita ser exercida coletivamente como método para otimização do tempo e avanço. É no processo de construção do conhecimento que a comunicação exerce papel decisivo, pois a quantidade de informação prévia sobre o objeto interfere diretamente nos resultados.

Não é possível datar o início das atividades de comunicação científica, mas é possível detectar que a comunicação do conhecimento científico é responsável por sua avaliação e utilização (MEADOWS, 1999). É a partir do conhecimento comunicado que a avaliação das informações geradas se torna possível, sendo através do processo comunicação-colaboração-avanço que a ciência se estrutura. Meadows (1999) afirma que o processo de comunicação científica se constitui a partir da interação psicológica de influência contínua e recíproca entre os interesses

individuais ou de grupos. É natural que um pesquisador sinta-se desencorajado a seguir em frente com pesquisas que não foram bem recebidas por seus pares. O oposto também se faz verdadeiro, levando a pesquisas que suscitam novos campos de exploração.

A comunicação científica não compreende somente a publicação, ela abrange o ciclo completo da pesquisa, da identificação do problema ao uso da informação para novas pesquisas, passando pela publicação e aceitação pelos pares. Targino (1999) considera que a comunicação científica envolve as atividades de produção, disseminação e uso da informação desde o momento em que o pesquisador considera a realização da pesquisa até a validação dos resultados como parte integrante do conhecimento científico. Portanto, os canais de comunicação apresentam-se por meio de diferentes formatos, que podem ser considerados vias informais ou formais quanto à tipologia, cada uma delas pretende exercer funções definidas.

Os canais informais utilizam os meios de contato interpessoal para comunicar a informação que podem ser transmitidas de forma oral ou escrita. São exemplos de meio informal: cartas, reuniões, e-mail. Caracterizam-se por possuir uma via mais ágil de comunicação, mas dificultam a recuperação e o armazenamento da informação por atingir um número menor de pessoas. Enquanto isso, o meio formal torna-se mais lento devido à avaliação prévia pelos pares e se caracteriza pelo emprego da escrita como forma de comunicação. Porém, aumenta a facilidade de armazenamento e recuperação e amplia o acesso (MEADOWS, 1999).

O livro é um exemplo tradicional de meio de comunicação formal entre as comunidades científicas, pois utiliza a forma escrita de comunicação e facilita a recuperação e o armazenamento da informação. O periódico científico originou-se como uma alternativa para agilizar a comunicação entre as comunidades científicas e tornou-se o seu principal veículo. Transformou significativamente o modo de transmissão da ciência, além de formalizar o processo que antes era realizado principalmente através de livros e correspondências. A nova estratégia de divulgação foi amplamente aceita entre os pesquisadores e rapidamente houve aumento no número de novas publicações utilizando o modelo (MEADOWS, 1999). Weitzel (2006, p. 84) afirma que “[...] a revista científica tornou-se o principal marco da constituição da estrutura da comunicação científica, pois surgiu dessa

necessidade genuína de trocas de experiências científicas dos cientistas dos tempos modernos.”. Os periódicos realizam a importante função de manter atualizada a informação sobre as pesquisas nas diversas áreas do conhecimento e são instrumentos indispensáveis para a construção e desenvolvimento da ciência (SEVERINO, 2007).

Para esse estudo, os termos periódico, periódico científico, publicação, publicação científica, revista e revista científica foram considerados sinônimos, definidos como publicações seriadas, com um título comum, com corpo editorial formado por especialista da área, constituída pela reunião de contribuições relacionadas pela temática, mas sem interligação entre si, numerados em sequência e objeto de Número Internacional Normalizado (ISSN). Convergindo com a definição de Stumpf (1998, p. 2):

Assim, os periódicos se constituem em uma das categorias das publicações seriadas, que apresentam como características particulares serem feitas em partes ou fascículos, numeradas progressiva ou cronologicamente, reunidas sob um título comum, editadas em intervalos regulares, com a intenção de continuidade infinita, formadas por contribuições, na forma de artigos assinados, sob a direção de um editor, com um plano definido que indica a necessidade de um planejamento prévio.

Contudo, assim como transformou tantos outros aspectos da vida moderna, a vinda da Internet culminou na remodelação da comunicação científica e modificou a estrutura da informação e da comunicação do conhecimento quanto à interação entre as partes, à agilidade do ciclo de comunicação, à velocidade da interação, à estrutura da mensagem, que pode conter diferentes formatos em um mesmo documento, e à dimensão do espaço de comunicação (BARRETO, 1998, p.125). É possível elencar outras transformações do fluxo da comunicação científica para além da publicação eletrônica, como alterações nos padrões de comportamento dos pesquisadores e sua relação com a sociedade, assim como a velocidade de acesso e uso da informação (BARRETO, 1998; CASTRO, 2006). Targino (2000) acrescenta que comunicação eletrônica combina aspectos formais e informais da comunicação, por ampliar o acesso, característica dos canais formais, e adiciona agilidade ao processo de reposta dos pares, característica dos canais de comunicação informal, assim como diminui o tempo entre a divulgação e o uso da informação.

Com o novo modelo de comunicação científica eletrônica emerge a necessidade de estratégias de gestão das publicações eletrônicas. Neste contexto,

foi criado em 1998 na University of British Columbia o Public Knowledge Project (PKP). Hoje ancorada na Universidade de Stanford e na Biblioteca da Universidade Simon Fraser, a iniciativa desenvolve *softwares* de código aberto para melhorar a qualidade e alcance da publicação (PUBLIC KNOWLEDGE PROJECT, [s.d.]). Seu principal produto é o Open Journal Systems (OJS), um *software web* para gerenciamento completo do processo editorial e de publicação de periódicos em acesso aberto. Isso inclui todo o fluxo editorial, da criação da revista à manutenção dos fascículos publicados (BRITO *et al.*, 2018). Além do OJS, a PKP ainda conta com outras duas plataformas, o Open Monograph Press, para publicação de monografias em acesso aberto e o sistema para conferência acadêmica, Open Conference Systems.

O OJS permite a customização das configurações suprimindo as necessidades dos periódicos sem limitações geográficas ou idiomáticas. Algumas das configurações possíveis no OJS foram escolhidas para averiguar se causam influência na indexação dos artigos no GS, como a definição dos padrões de metadados utilizados, o tipo de preservação digital, a atribuição de DOI. Como qualquer sistema, o OJS necessita de determinados padrões e *plugins*¹ nativos para funcionar corretamente, entretanto, a utilização de *plugins* externos é possibilitada pela ferramenta para customização, por exemplo, dos padrões de metadados (BRITO *et al.*, 2018; PUBLIC KNOWLEDGE PROJECT, [s.d.]).

O OJS 3 oferece uma galeria com diversos *plugins* divididos por categorias, dentre elas, padrões de metadados, autenticação, formato de citação, de importação/exportação, metadados OAI, relatórios e visual. Alguns *plugins* vem pré-instalados no sistema, como o *plugin* para atribuição automática de DOI. O *Digital Object Identifier* (DOI) é uma identificação única e persistente composta por números e letras para documentos em redes digitais e pode ser atribuído a qualquer objeto digital, sejam eles livros, artigos, periódicos. (FERREIRA *et al.* 2015). A principal contribuição do DOI consiste em garantir o acesso, pois o *link* persistente garante que o objeto será encontrado mesmo após ser rearranjado na *web*, além do

¹ Plugins são módulos complementares do sistema que ampliam ou alteram funcionalidades, como por exemplo o idioma utilizado pelo periódico.

mais a identificação individual do documento unifica os diferentes formatos de arquivo (HTML, PDF) que o item possa ter.

Para o sistema de preservação digital, as opções comportadas pelo OJS são o LOCKSS e o CLOCKSS. O sistema de preservação digital LOCKSS (Lots of Copies Keep Stuff Safe, em português: Muitas cópias mantêm as coisas seguras), foi fundado em 1998 pela Stanford University Libraries. A instalação da LOCKSS *Box*, “Permite às bibliotecas participantes recolher, armazenar, preservar e fornecer o acesso a conteúdos que tenham subscrito ou adquirido.” (COSTA, 2015). O formato distribui réplicas dos documentos entre as bibliotecas participantes, deste modo, quando o acesso não for efetuado pelo *website* do editor por problemas de comunicação, a rede é ativada recuperando uma cópia do documento distribuída pelo sistema. No Brasil, o Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) aderiu ao sistema LOCKSS em 2013, através da Rede Brasileira de Serviços de Preservação Digital Cariniana (BRITO *et al.*, 2018). O CLOCKSS (Controlled LOCKSS) foi lançado em 2006 e baseia-se no sistema LOCKSS, porém um número limitado de bibliotecas possui a função de arquivar os documentos da rede. Assim, quando os arquivos não podem ser acessados diretamente dos editores são disponibilizados sob uma *licença Creative Commons* (COSTA, 2015).

Um dos aspectos comportados pela galeria de plugins é uso de padrões de metadados, item de importância para a indexação de documentos no GS segundo as orientações de indexação da empresa. Os metadados são comumente conceituados como dados sobre dados, mais claramente são modelos normativos padronizados para automatização da informação. Na biblioteconomia o padrão de metadados mais difundido para registros bibliográficos é o MARC21, contudo, existem diversos outros padrões como o Dublin Core (DC), XML, Highwire Press.

Outro aspecto importante em relação aos periódicos refere-se as formas de avaliação. No Brasil, a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES, 2018) “[...] desempenha papel fundamental na expansão e consolidação da pós-graduação *stricto sensu* (mestrado e doutorado) em todos os estados da Federação”. São funções atribuídas à instituição: avaliação dos programas de pós-graduação e apoio ao acesso e à divulgação da produção científica do País. Para isto, foi desenvolvido um modelo de estratificação das publicações produzidas, segundo o preenchimento de requisitos pré-determinados, denominado Qualis

Qualis-periódicos é o conjunto de **procedimentos** utilizados pela Capes para estratificação da qualidade da produção intelectual dos programas de pós-graduação. [...]. Como resultado, disponibiliza uma lista com a classificação dos veículos utilizados pelos programas de pós-graduação para a divulgação da sua produção.

A estratificação da qualidade dessa produção é realizada de forma indireta. Dessa forma, o Qualis afere a qualidade dos artigos e de outros tipos de produção, a partir da **análise da qualidade dos veículos de divulgação, ou seja, periódicos científicos**.

A classificação de periódicos é realizada pelas áreas de avaliação e passa por processo anual de atualização. Esses veículos são enquadrados em estratos indicativos da qualidade, de A1 a C, sendo A1 o estrato mais elevado (CAPES, 2017, n.p, grifo nosso).

Segundo Weitzel (2006), a Internet revolucionou o modo como os cientistas comunicam os resultados de suas pesquisas. E o Google e Google Scholar fazem parte desta revolução, pois contribuíram para a disseminação e ampliação do conteúdo virtual. Para a autora, tais buscadores passam a ter representatividade como as bases de dados convencionais, com a vantagem de indexarem um número maior de documentos e não possuem restrições como área do conhecimento, idioma e outras por ela mencionadas. Deste modo, os estudos relacionados à presença nesses buscadores demonstram sua contribuição para medir a divulgação científica.

No contexto *web* é percebido que os tradicionais indicadores bibliométricos não possuem representatividade fidedigna, o que torna indispensável a discussão sobre o uso das métricas no ambiente virtual.

2.2 Estudos Métricos e Aplicações na *Web*

Para Otlet (2018, p. 17) “[...] em qualquer área, a medida é uma forma superior assumida pelo conhecimento.” De modo amplo, os estudos métricos são técnicas de medição aplicadas à ciência como forma de compreender e retratar sua evolução (MUGNAINI; CARVALHO; CAMPANATTI-OSTIZ, 2006). Os estudos métricos da informação envolvem métodos e técnicas para estimar e avaliar de forma quantitativa a produção, a circulação e o uso da informação, os quais, possuem “[...] diversas abordagens teórico-metodológicas e diferentes denominações em função de seus objetivos e objetos de estudo.” (NORONHA; MARICATO, 2008, p. 117).

Devido ao recente crescimento dos subcampos dos estudos métricos promovido pelas tecnologias digitais é possível encontrar divergências terminológicas importantes, por exemplo, o uso do termo *webmetria* para organização e uso de *sites* (MACIAS-CHAPULA, 1998). Para melhor esclarecer e com base em Araújo (2015a;b), apresentam-se brevemente os subcampos dos estudos métricos, bibliometria, informetria, webmetria, cibermetria, altmetria e webometria, a fim de contextualizar e delimitar a localização do presente estudo.

O termo **bibliometria** surge da união do grego *biblion* e *metrikos*, que significam livro e mensuração, respectivamente. Usualmente seu conceito é definido por processos quantitativos de medidas relacionados a livros e documentos (BRUFEM; PRATES, 2005, p. 11). Tais medidas relacionam-se ao “[...] estudo dos aspectos quantitativos da produção, disseminação e uso da informação registrada.” (TAGUE-SUTCKIFFE, 1992, p. 1, tradução nossa) e possui na análise de citações um campo de estudo marcante e característico. Os meios digitais da informação propiciaram aumentos significativos e diversificação de objetos de análise para a área. Como no estudo bibliométrico de análise de citações da Em Questão de Bilhar (2013) em que os dados foram coletados exclusivamente do GS utilizando como interface o *software* Publish or Perish.

Informetria é a denominação dada ao estudo dos aspectos quantitativos da informação quando aplicados aos recursos de fluxo, busca, disseminação, recuperação e acesso à informação em qualquer tipo suporte, via canais formais e informais de comunicação. Têm por objetivo melhorar a eficiência da recuperação da informação, atualmente, incluindo sistemas de recuperação da informação (ARAÚJO, 2015; MACIAS-CHAPULA, 1998).

Já a análise informétrica que utiliza recursos e objetos da Internet recebe a denominação de **cibermetria** e, assim como a informetria, refere-se à análise quantitativa de canais formais e informais. A cibermetria ainda é recente e de baixa utilização (GOUVEIA, 2013), mas sua aplicação é bastante relevante na compreensão dos aspectos comunicacionais dos ambientes virtuais de interação. Desta forma, a cibermetria engloba métricas em “[...] bases de dados, páginas *web*, URLs, microblogs, *blogs*, salas de bate papo, *mailing lists*, comunidades virtuais, grupos de discussão, *muds*, ambientes virtuais de aprendizagem (AVA), *sites* de redes sociais.” (ARAÚJO 2015). A amplitude da cobertura da cibermetria implica sobreposições com a propagação de abordagens baseadas em ciência da

computação na análise de conteúdo da *web*, estruturas de *links* e uso da *web* e suas tecnologias.

Há ainda a **altmetria** que compreende as análises realizadas em canais formais de comunicação, análoga à bibliométrica, porém voltada ao uso de ambientes *online*. Nascimento e Oddone (2015) utilizaram um serviço altimétrico (Altmetrics.com) para aferição de impacto para avaliação de periódicos científicos brasileiros na área de Ciência da Informação e concluíram que poucos periódicos estão disponíveis para a geração de métricas alternativas de citação no serviço. Porém, observaram crescimento no uso das citações, principalmente pelo Twitter e Mendeley. Outros recursos utilizados para estudos altimétricos são plataformas como Facebook, Google+, Youtube, Figshare, CiteULike, Github, Slideshare, Research Gate, além de *sites* de notícias e postagens em blogs (NASCIMENTO; ODDONE, 2015).

A viabilidade de combinação de diferentes métricas permite a conferência e validação dos resultados obtidos isoladamente, Martín-Martín *et al.* (2016) detectaram uma rede científica de estudos métricos através do recurso de análise de citações do Google Scholar, composta por cientistas, documentos e fontes. Posteriormente, as informações coletadas foram comparadas às informações encontradas nos perfis sociais dos autores em redes sociais acadêmicas como ResearchGate e Mendeley e, também, em seus perfis no Twitter, exemplo de uso combinado de diferentes métricas e do GS como recurso métrico da informação. Em âmbito nacional, Araújo (2015b) coletou dados altimétricos dos artigos da revista Datagramazero e os relacionou com as citações recebidas, obtidas através do GS. Demonstrando, assim, como os dados conseguidos através das interações sociais na *web* dimensionam o impacto social dos artigos. Ratifica-se a possibilidade de uso em estudos com combinações entre os subcampos (ARAÚJO, 2014; GOUVEIA, 2012; VANTI, 2005).

A **webometria**, subcampo no qual se insere o problema de pesquisa já mencionado, trata das análises métricas de toda a *web*, incluindo as análises a motores de busca como o proposto anteriormente. Com maior escopo, a webometria abrange os estudos quantitativos de toda a Internet, da construção e do uso dos recursos da informação, estruturas e tecnologias na Web, a partir de abordagens informétricas e bibliométricas (BJÖRNEBORN; INGWERSEN, 2004). Tal “[...] medição aponta para o estudo ou análise comparativa da presença dos diversos

países na rede, das proporções de páginas pessoais, comerciais e institucionais.” (VANTI, 2002, p. 155).

As páginas podem ser analisadas por tipo (setor público, privado), classificações (pessoais, institucionais), categorias (páginas-documento, páginas-índice). As medições podem ser realizadas ao longo do tempo para fins de comparação de crescimento e evolução das informações existentes sobre determinado assunto ou páginas específicas. Para este estudo considerou-se a presença dos periódicos da Ciência da Informação (páginas específicas) no Google Scholar (recurso de informação) mediante aplicação do comando de busca descrito na seção Metodologia.

Os termos webometria e cibermetria são relacionados de forma geral e por vezes utilizados como sinônimos, mas possuem objetos distintos de estudo (BJÖRNEBORN; INGWERSEN, 2004), visto que o termo cibermetria é empregado para se referir quantitativamente à criação e utilização de conteúdos *web* de qualquer natureza usados no contexto da Internet, com abordagens informétricas, bibliométricas e cientométricas (ARAÚJO, 2015 a; b). Contudo é a webometria que comporta a análise de *links*, citações de *Uniform Resource Locator* (URL, em português localizador uniforme de recursos) e de mecanismos de busca. Em meio a vários subcampos existentes, alguns separados por limites tênues, a avaliação minuciosa para localizar e enquadrar uma pesquisa ao subcampo correspondente, a estratégia de identificar os recursos, assim como a natureza da informação e atores envolvidos são fatores facilitadores da tarefa.

A análise webométrica dos periódicos científicos é relevante para a compreensão dos fenômenos de comunicação científica na *web*, “[...] na medida em que estudos de webometria em periódicos científicos ainda se apresentam tímidos, muito provavelmente por questões voltados ao fator de impacto dominante nas avaliações.” (SHINTAKU, 2017, p. 268). Complementa-se com o fato de que para haver avaliação efetiva da produção científica na *web* é necessário conhecer suas falhas e com a identificação das falhas é possível corrigi-las.

Para a análise de tráfego de visitas em ambientes da Internet é utilizada a denominação **webmetria**, subcampo da webometria tratada a seguir. “Trata-se de um campo voltado para análises métricas de tráfego de vistas em *websites* (acesso e uso de informação na *web*) que auxilia no controle de qualidade de processos e recursos da *web*” (ARAÚJO, 2015, p. 46). Normalmente os dados para análise

webmétricas são coletados por meio de *logs* e *page taggings* (ARAÚJO, 2015). Gouveia e Lang (2013) fazem uma analogia com o levantamento de acessos a uma biblioteca para melhor compreensão da utilidade dos estudos webmétricos, como número de visitas, consultas ao acervo e retorno ou não do usuário à mesma biblioteca.

Devido a análise recente dos subcampos dos estudos métricos, devido a isso ainda se encontram em estruturação, portanto é possível observar divergências conceituais e terminológicas. Após uma breve inserção nos estudos métricos na Internet e suas respectivas aplicações, objetos e recursos, a seguir, são destacados estudos realizados utilizando o buscador Google Scholar.

2.3 Google Scholar e os Estudos Métricos na *web*

O Google (buscador geral) foi desenvolvido por Sergei Brin e Larry Page, lançado em 1998 e logo despontou na liderança entre os buscadores, tornando-se “a verdadeira porta de entrada da internet” e “seu caráter generalista” propiciou o uso acadêmico da ferramenta (TORRES-SALINAS; RUIZ-PÉREZ; DELGADO-LÓPEZ-CÓZAR, 2009, p. 501, tradução nossa). Os autores ainda afirmam que com sagacidade a empresa percebeu o uso diferenciado e de especificidades bem definidas do meio acadêmico, além do grande volume de negócios movimentados por ele e desenvolveu um buscador voltado a essas necessidades, e assim, em 2004, o Google Scholar foi lançado.

Não tardou para a comunidade científica perceber a potencialidade da ferramenta e iniciar a investigação para decifrar suas principais funcionalidades e usos como ferramenta quantitativa e avaliativa da produção científica. As conclusões vieram a favor (HARZING; VAN-DER-WAL, 2008) e contra, demonstrando que as vulnerabilidades do sistema também precisam ser mensuradas e conhecidas (JACSÓ, 2005). Comparações com *Web of Science* e *Scopus* para uso bibliométrico não ficaram de fora, igualmente as análises de cobertura documental, presença na *web* e análise e impacto de citações (ORDUÑA-MALEA; DELGADO LÓPEZ-CÓZAR, 2015; ORDUÑA-MALEA; SERRANO-COBOS; LLORET-ROMERO, 2009; ORTEGA *et al.*, 2008; TORRES-SALINAS; RUIZ-PÉREZ; DELGADO-LÓPEZ-CÓZAR, 2009).

Na página de informações sobre o GS² a empresa fornece algumas informações sobre a indexação de documentos a fim de auxiliar editores a terem seu conteúdo incluído nos resultados de pesquisa do GS. A recomendação para uma indexação eficaz no GS para repositórios acadêmicos é simples, usar a versão mais recente do software Eprints, Digital Commons ou Dspace para hospedagem do conteúdo (GOOGLE SCHOLAR 2018). Para editores de periódicos recomendam o uso dos serviços de hospedagem Atypon ou Highwire, também recomendam o uso do OJS para editores que dominam a plataforma de editoração. Também ressaltam que o uso de agregadores de periódicos em um único site, como o JSTOR e SciELO, também funciona bem (GOOGLE SCHOLAR 2018). Ainda orientam aos autores individuais a publicação do documento em *site* pessoal, assegurando que a extensão do arquivo seja PDF, terminada com ".pdf"; que o título do artigo esteja na parte superior da primeira folha seguida abaixo pelo nome dos autores em linha separada e ainda que contenham uma seção de referências.

Orduña-Malea e Delgado López-Cózar (2015) estimam que os mecanismos de busca Google e Google Scholar são o atual ponto de partida de qualquer acadêmico, por isso a relevância para os editores em garantir a presença de seus conteúdos recuperáveis pelo buscador. Dez anos antes o Centro de Bibliotecas de Computadores *Online* (OCLC) observou que 89% dos estudantes começaram suas buscas através de mecanismos de busca e apenas 2% o fizeram nos sites de bibliotecas (DEROSA e OCLC 2005) índice este que chegou a 0% na repetição da pesquisa cinco anos depois (DEROSA et al., 2010). Refletindo sobre a remodelação que a Internet proporcionou neste íterim, não é surpresa que mecanismos de busca sejam utilizados rotineiramente pelos estudantes (COOK et al. 2009; JAMALI; ASADI, 2010).

O Departamento de Física e Astronomia da Universidade College, em Londres, revelou que muitos dos artigos recuperados pelos cientistas da área da física foram encontrados na pesquisa geral do Google por 'acaso'. No estudo realizado através de entrevistas a cinquenta e seis pesquisadores, 46% relatam procurar diretamente nas bases de dados da área quando a intenção é recuperar artigos científicos, chegam a relatar surpresa por encontrar artigos científicos

² Fonte: <https://scholar.google.com/intl/en/scholar/about.html>

enquanto buscam informações gerais da área no Google (JAMALI; ASADI, 2010). Ainda em 2005, Nicholas *et al.* (2005) relataram resultado semelhante entre pesquisadores, indicando um alto percentual de artigos recuperados através do Google por estes investigadores. Ambas as pesquisas demonstram a crescente confiança, por parte dos cientistas, no Google como ferramenta de recuperação de informação por encontrarem literatura acadêmica com a especificidade necessária. Com o passar dos anos os resultados se ratificam. Arshad e Ameen (2017) sugerem uma diferença nos padrões de busca por informação entre doze disciplinas analisadas, porém novamente o Google, e agora o Google Scholar, passam a figurar no topo da lista dos recursos de busca da informação, este último, especificamente, envolvendo periódicos eletrônicos.

Haglund e Olsson (2008) observaram pesquisadores na Suécia de 2005 a 2006 e revelaram como o Google domina como um ponto de partida a busca de informações acadêmicas. Na mesma tendência, Schonfeld e Housewright (2010) pesquisaram 3.025 professores de faculdades e universidades dos Estados Unidos e concluíram que pesquisadores das áreas de ciências, ciências sociais e humanas, substituíram as visitas a bibliotecas e os recursos de pesquisa tradicionais pelo uso de ferramentas de busca *online*, no qual o Google e o Google Scholar figuram entre preferidas na localização de informações em periódicos científicos.

Herrera (2011) analisou os registros da biblioteca da Universidade do Mississippi e revela que a porcentagem de cliques originários do Google Scholar cresceu de 4% em 2006 para 27% em 2009. Devido ao amplo uso dos mecanismos de pesquisa, agregado ao fato de que muitos acessos são realizados sem passar pelo *site* inicial que hospeda o conteúdo, fica evidente a importância que o conteúdo acadêmico disponível na *web* seja indexado corretamente pelos motores de busca (ORDUÑA-MALEA; DELGADO LÓPEZ-CÓZAR 2015).

Málaga (2008) afirma que 62% dos usuários utilizam somente os resultados obtidos na primeira página do referido mecanismo de busca. Desta forma, caso os conteúdos não estejam presentes e bem posicionados nos resultados de busca há grandes possibilidades de serem ignorados pelos usuários. Assim todos que possuem o interesse em tornar seu conteúdo visível no buscador possuem em comum o desafio de ter seu conteúdo encontrado na imensidão da cobertura do GS e, portanto, precisam esmerar-se em incluí-los nos índices dos principais

mecanismos de busca, pois será onde a maioria procurará (ARLITSCH; O'BRIEN 2012).

A estratégia do GS pretende sanar o *gap* deixado pelas bibliotecas que falharam em centralizar conteúdos de diferentes fontes, potencializa a oportunidade de tornar o conteúdo visível, contribuindo para que possam ser localizados e, potencialmente, citados (ARLITSCH; O'BRIEN 2012), expectativa de qualquer pesquisador.

O Google e o Google Scholar expressam índices distintos. Enquanto o buscador Google recupera informações de todas as páginas indexadas por ele, o Google Scholar limita seus resultados às páginas de cunho acadêmico: universidades, institutos de pesquisa, revistas científicas, patentes. Martín-Martín *et al.* (2014) afirmam que o Google Scholar é capaz de percorrer quase toda a *web* acadêmica, incluindo editores, sociedades acadêmicas, bancos de dados disciplinares, repositórios institucionais e páginas pessoais da *web*. Tamanha abrangência e seletividade o tornam um recurso importante para avaliação dos *sites* institucionais e também da própria cobertura do buscador (COSTA, 2014; HARZING 2013 , 2014 ; BORNMANN *et al.* 2016 ; PRINS *et al.* 2016;).

O índice de presença dos conteúdos disponíveis nos repositórios digitais na Universidade de Utah aumentou 200% com aplicação de recomendações para melhorar a indexação no buscador, o que incluiu revisão dos padrões de metadados, com transformação de *tags* de metadados Dublin Core (DC) em *tags* bibliográficas Highwire Press (HP) de 0,1% para 62%, entre outras medidas que potencializaram a indexação dos conteúdos da universidade no GS (ARLITSCH; O'BRIEN 2012).

Arlitsch e O'Brian (2012) consideraram as taxas iniciais de indexação para os repositórios institucionais dos EUA muito baixas. Utilizaram a estratégia de busca com o comando de busca "*site:URL*" no GS. Os resultados através da aplicação do comando no buscador foram contados, multiplicados por cem e, posteriormente, divididos pelo número total de documentos hospedados pelo repositório. O resultado do cálculo é chamado pelos autores de taxa de indexação, pois representa a porcentagem de documentos que se fazem presentes no buscador através do referido comando de busca.

Estudos métricos de presença na *web* tiveram início em 2004 (PARK; KIM; BARNETT, 2004) e possuem diversos representantes com variados objetos de análise, incluindo produção acadêmica (BAR-ILAN, 2004), instituições (AFONSO;

MATIAS, 2016), conceitos (MARTÍN-MARTÍN *et al.*, 2016) e pesquisadores (ARAÚJO, 2014). Entretanto, há pouca menção à definição do conceito de presença na *web*, alguns estudos contêm a expressão inclusive no título, como em Araújo, (2014); Orduña-Malea; Serrano-Cobos; Lloret-Romero, (2009). Talvez deva-se ao fato de parecer singela no tocante que se refere “somente” aos itens recuperados mediante uma busca.

Presença *web* é um indicador de contagem que analisa a medida em que os repositórios (disciplinares, institucionais) são indexados nos motores de busca (ORDUÑA-MALEA; DELGADO LÓPEZ-CÓZAR, 2015). Na perspectiva do marketing digital o termo presença digital é utilizado em sentido semelhante, é a “[...] existência de algo (seres humanos, marcas, empresas, coisas, etc.) no ambiente digital.” (GABRIEL, 2010, p. 249). De forma simplificada, presença *web* refere-se ao sentido de estar visível, presente na *web*, passível de recuperação por meio de busca. Gulka e Lucas (2017) argumentam que a presença digital é a representação da existência real e sua ausência pode representar a sua inexistência.

Segundo Vitullo (2007), as análises webométricas munem-se de indicadores webométricos, tais como tamanho do sítio, visibilidade/popularidade, fator de impacto *web*, luminosidade, dentre outros, para mensuração quantitativa. Por este ângulo, pode parecer que a presença *web* é parcela exígua, coadjuvante, por ser um aspecto quantificado e analisado sob um contexto maior e mais complexo. Contudo, é essencial. O contexto realmente mostra-se mais complexo por apresentar a análise cruzada e outros aspectos. Apesar disto a presença *online* é a variável que fornece suporte para esse tipo de análise, seja para quantificar a presença de *links* externos e internos, quantidade de páginas, pesquisadores ou instituições na *web*.

Smith (1999), afirma que os motores de busca são ferramentas fundamentais para a realização de estudos webométricos, pois viabilizam a manipulação de volumes de informação. Arroyo (2004) complementa que são capazes de extrair informações sobre domínios ou subdomínios ou detectar a presença de certas palavras no espaço *web*. Smith (1999) destaca que os buscadores permitem quantificar o número total de páginas em um espaço *web*, entendendo o termo espaço *web* no sentido de domínio, que tanto pode ser um domínio de país quanto um domínio institucional.

A mesma estratégia de coleta de dados foi utilizada por Costa (2014), que emprega o termo taxa de visibilidade para o percentual de presença dos repositórios

brasileiros e portugueses no GS. Entretanto, apesar de a estratégia utilizada por Ortega et al., (2008) ser a mesma, o termo referido para o resultado é presença na web. Compreende-se que o campo está em desenvolvimento e variações terminológicas e conceituais são inerentes ao período, contudo deseja-se evitar possíveis equívocos na comparação entre indexação de catálogos tradicionais.

3 Metodologia

Segundo Richardson *et al.*, (2015, p. 22, grifos dos autores):

segundo a sua origem, **método** é o caminho ou a maneira para chegar a determinado fim ou objetivo, distinguindo-se assim, do conceito de metodologia, que deriva do grego *méthodos* (caminho para chegar a um objetivo) + *logos* (conhecimento). Assim, a **metodologia** são os procedimentos e regras utilizadas por determinado método. Por exemplo, o método científico é o caminho da ciência para chegar a um objetivo. A **metodologia** são as regras estabelecidas para o método científico, por exemplo: a necessidade de observar, a necessidade de formular hipóteses. a elaboração de instrumentos etc.

Portanto, esta seção descreve os procedimentos e métodos adotados com o intuito de atender os objetivos propostos na seção 1.2, seguindo o método científico. Inicia-se com a classificação metodológica da pesquisa, na sequência descrevem-se os critérios de seleção do objeto de pesquisa, o processo de coleta de dados, assim como a técnica de análise dos dados. Posteriormente, revela-se a convenção estabelecida para termos originários da língua inglesa em que a tradução não foi o suficiente para impedir ambiguidades e, por fim, os limites identificados no estudo.

Com tipologia de natureza básica, esta pesquisa visa “[...] gerar conhecimentos novos, úteis para o avanço da Ciência, sem aplicação prática prevista.” (GERHARDT; SILVEIRA, 2009, p.34). A identificação dos fatores que interferem ou não na indexação dos artigos de periódicos recuperados através da ferramenta de busca Google Scholar promove subsídios para balizar as decisões dos gestores de publicações científicas. Com isso objetiva-se gerar novos conhecimentos que colaborarem com estratégias que agreguem maior visibilidade aos periódicos científicos em geral, a seus autores e, conseqüentemente, à Ciência.

O desenvolvimento da pesquisa quanto aos objetivos estabelecidos possui cunho exploratório, pois propõe-se em identificar variáveis que influenciam a presença no GS através da análise dos padrões de presença dos periódicos da CI no buscador. Severino (2007, p. 123) aponta que a pesquisa com objetivo exploratório busca “[...] levantar informações sobre um determinado objeto, delimitando assim um campo de trabalho, mapeando as condições de manifestação desse objeto.” Neste estudo, coletam-se informações acerca da influência de algumas características dos periódicos na indexação pelo GS com o objetivo de delimitar e mapear a manifestação da presença destes no GS.

Gil (2002, p. 42) complementa que as pesquisas exploratórias caracterizam-se por propiciar o entendimento prático, a fim de “[...] torná-lo mais explícito [...]”, que possuem o intuito de descoberta na aplicação do estudo e que, a viabilidade de possuir um planejamento flexível, propicia o estudo de diversos aspectos do objeto. A análise da presença dos artigos dos periódicos da Ciência da Informação no GS, consiste no objeto explorado neste estudo para a compreensão dos fatos, recurso frequentemente utilizado em estudos exploratórios (GIL, 2002). Assim, essa investigação busca compreender os fatores que interferem na indexação e presença dos periódicos no buscador.

Quanto à abordagem, a pesquisa tem escopo quantitativo, pois “[...] recorre à linguagem matemática para descrever as causas de um fenômeno, as relações entre variáveis, etc.” (FONSECA, 2002, p. 20). Richardson *et al.* (2015, p. 70) complementam que esta é uma abordagem frequentemente aplicada nos estudos que “[...] investigam a relação de causalidade”. As relações investigadas são entre as características dos periódicos definidas em 1.2.2 e as taxas de presença dos periódicos correntes da CI. A amostra foi determinada de forma não-probabilística intencional, pois determinou-se a priori a análise dos periódicos correntes da CI, cujos critérios de seleção são descritos em 3.1. Entende-se que a amostra intencional representa um universo que corresponde às “[...] características típicas de todos os integrantes que pertencem a cada uma das partes da população.” (RICHARDSON *et al.*, 2015, p. 161), contudo alertam que dificilmente uma amostra intencional será a representação fidedigna do universo, pois para tanto é necessário conhecer a população detalhadamente.

A pesquisa documental foi o procedimento adotado para coleta dos dados. Este procedimento caracteriza-se por utilizar materiais que ainda não passaram por técnicas de análise, “[...] ou que ainda podem ser reelaborados de acordo com os objetivos da pesquisa.” (GIL, 2002, p. 45). As fontes da pesquisa documental que viabilizaram este estudo foram os *websites* dos periódicos, objeto desta pesquisa, de onde foram identificados o total de artigos publicados pela revista, entre outros descritos em 3.2.2, o buscador Google Scholar para identificação dos artigos localizados mediante argumento *site* e do uso dos perfis de periódico e de pesquisador descritos em 3.2.1.

A terceira fonte da pesquisa documental foi o website da plataforma Sucupira, de onde foram verificados os estratos Qualis dos periódicos selecionados.

A definição para a coleta dos dados de classificação Qualis foi limitada à área de pertinência do conteúdo veiculado, restrito à área Comunicação e Informação, sendo desconsideradas para este estudo a estratificação Qualis dos periódicos de outras áreas. A pesquisa foi realizada por título e ISSN do periódico. Foram consideradas somente a estratificação para a área de avaliação Comunicação e Informação no quadriênio 2013/2016.

3.1 Seleção do objeto de pesquisa

Inicialmente, pretendia-se dirigir o estudo aos periódicos da UFRGS, entretanto, a aplicação prévia do comando *site* em uma amostra experimental de 20 revistas mostrou o argumento inválido nos 20 itens da amostra, a saber, o portal de periódicos da UFRGS apresenta o total de 53 periódicos da instituição. Com isso, transferiu-se intencionalmente a aplicação do estudo aos periódicos da área da CI no Brasil, pois uma pesquisa somente com periódicos cujo argumento *site* é inválido não atingiria o objetivo de investigar a influência das características na taxa de presença.

A listagem dos periódicos com temáticas pertinentes à Ciência da Informação foi inicialmente baseada nos periódicos indexados pela Base de Dados Referenciais de Artigos de Periódicos em Ciência da Informação (Brapci), algo em torno de 58 periódicos. A eles foram acrescentados os periódicos Bibliomar e Convergências em Ciências da Informação que não constavam da lista de periódicos indexados pela Brapci, passando a um total de sessenta periódicos a serem analisados baseado nos critérios: foco temático na CI e publicação corrente. Vinte e um (21) periódicos foram excluídos do corpus da pesquisa, quinze (15) por não serem publicações correntes e seis (6) por possuírem temática interdisciplinar sem foco na Ciência da Informação.

Assim, os periódicos brasileiros da área da Ciência da Informação são o objeto de pesquisa, com exceção dos periódicos descontinuados, sem publicação corrente, ou cujo foco temático não é a Ciência da Informação. Contudo, na aplicação do comando *site* no GS com as *URL* das revistas, percebeu-se um grupo cujo comando foi inválido na pesquisa, não recuperando nenhum artigo do periódico. Orduña-Malea e Delgado López-Cózar (2015) alertam que o argumento *site* é

inválido em *sites* que não possuem domínio próprio, impossibilitando a recuperação de documentos indexados no GS através do método definido para a realização da pesquisa. Subdiretórios também apresentam impedimentos devido à limitação da precisão no mecanismo de busca.

3.2 Procedimentos da Coleta de Dados

Um formulário foi o instrumento utilizado como técnica de coleta de dados. Ele possibilita que o pesquisador formule as questões previamente e anote as respostas no momento da coleta dos dados (GIL, 2002). As questões elaboradas para o instrumento de coleta preocuparam-se em buscar informações necessárias que permitissem o cálculo da taxa de presença: número de artigos publicados e número de artigos recuperados no GS mediante argumento *site*. Também buscou levantar as informações para análise das características dos periódicos, definidas nos objetivos específicos, a estratificação Qualis; a existência de perfil de revista no GS; o formato de arquivo utilizado; o padrão de preservação digital; o padrão de metadados empregado no *site*; a utilização de DOI da revista; e versão do OJS utilizada.

A contagem do total de artigos publicados por revista e a de páginas recuperadas pelo GS mediante o comando "*site*", foram contabilizadas no período de 29 de junho à 7 de julho de 2018. Deste modo os dados coletados contemplaram o período de 6 meses de 2018. Para a coleta dos dados de contagem de registros presentes no GS foi utilizada a estratégia testada por Arlitsch e O'Brien (2012): comando (de busca) *site* [***site:URLdarevista***] no campo de busca inicial do GS. O comando *site* restringe os resultados da busca a um domínio ou parte dele, tornando possível a recuperação dos itens ali indexados.

A seguir, apresentam-se os procedimentos empregados para a coleta de dados subdividindo-os conforme a fonte de pesquisa documental utilizada, Google Scholar e *sites* dos periódicos.

3.2.1 Coleta de Dados no Google Scholar

O GS compõe a uma das três fontes documentais utilizadas para desenvolver este trabalho. O buscador mostrou, mediante a aplicação do comando *site* na caixa de busca, a quantidade de artigos indexados na ferramenta de busca. A Figura 3 ilustra a visualização da busca para os periódicos com comando *site* válido no GS. Em destaque encontra-se o número de documentos recuperados no buscador através da estratégia utilizada que constitui na aplicação da fórmula $TP = \text{artGS} \cdot 100 / \text{artpubl}$, explanada em 3.4. Deste modo, os dados coletados das duas fontes principais foram combinados para obtenção das taxas de indexação no GS, representando a presença dos periódicos no buscador.

Figura 3 – Tela de aplicação de comando *site* no GS

Fonte: Dados da pesquisa (2018)

A estratégia do comando *site* foi aplicada para todas as *URL* dos *sites* dos periódicos selecionados, o número de documentos recuperados por periódico no GS foi registrado no instrumento de coleta de dados, para posterior cálculo da taxa de presença. O formulário de coleta de dados foi gerado com o *software Microsoft Office Excel*®, cujo modelo encontra-se em Apêndice. A contagem de documentos levou em consideração o número daqueles recuperados pelo GS mediante comando definido (em destaque na Figura 3) e sem qualquer exclusão, mesmo nos casos em

que a publicação não era originalmente digital, como o caso da revista *Ágora*, que implantou a versão *online* em 2011, tendo completado o acervo retroativo (desde 1985) com digitalizações. Salienta-se que mais de 85% das publicações recuperadas com a estratégia utilizada correspondem ao período entre 1985 e 2011. Situação diferente foi identificada na coleta dos periódicos em que documentos como expediente, editorial e normas para autores não foram contabilizados, devido à análise preliminar em que foi identificada a ausência desses documentos, na maioria dos periódicos aferidos no GS mediante comando *site*.

Através do Google Scholar Metrics (Google Acadêmico Métricas), averiguou-se a existência da página de h5-index para todos os periódicos. Foram inseridos os nomes das revistas na caixa de busca e efetuada busca individuais. Após essa verificação, os nomes das revistas foram pesquisados na busca de perfil de pesquisadores, constando-se a presença de cinco revistas utilizando o perfil de pesquisador. A descoberta é discutida na seção 4.2 junto à apresentação dos resultados

3.2.2 Coleta de Dados nos Sites dos Periódicos

Seguindo os apontamentos de Gil (2002), utilizaram-se os *sites* dos periódicos para a coleta do número de artigos publicados pela revista, assim como informações das características a serem empregadas para posterior relacionamento com os níveis de taxa de presença. Os dados coletados são informações contidas no *site* disponibilizadas para informar os usuários sobre a plataforma e versão utilizados, tipo de preservação digital, além do número de formatos de arquivos disponibilizados identificados individualmente. Portanto, não possuem nenhuma análise técnica e foram “[...] reelaborados de acordo com os objetivos da pesquisa.”, (GIL, 2002, p. 45), caracterizando o procedimento como pesquisa documental.

Os dados coletados nos *sites* dos periódicos consistiram no número de artigos publicados pela revista, nos formatos em que os arquivos são disponibilizados, no tipo de preservação digital, no padrão de metadados, na utilização de DOI e na versão e plataforma de publicação utilizada pelo periódico. O número de artigos publicados em cada revista foi contabilizado individualmente em cada número publicado, diretamente do *site* do periódico. Uma amostragem

aleatória preliminar demonstrou distorções numéricas entre a quantidade de artigos publicados nas revistas e a indicada na base de dados Brapci, por isso optou-se pela contagem manual. Utilizou-se o *software Microsoft Office Excel*® para contabilizar os 9.549 documentos publicados nas 32 revistas em que o argumento *site* recuperou documentos no GS.

Documentos como resenhas, relatos, trabalhos de conclusão, entrevistas, entre outros também foram contabilizadas no número total de publicações, devido à estratégia de considerar a *URL* para recuperação da informação. Desde modo, cada item publicado indicado no sumário com um arquivo vinculado em qualquer formato, foi contado como uma unidade de presença na recuperação de busca no GS. Ressalta-se que os periódicos, mesmo digitalizados integralmente, foram contabilizados como 1 documento, quando estão disponibilizados mediante 1 arquivo e 1 uma *URL*, uma vez que o GS recupera através da *URL* no comando *site*.

Os formatos em que os arquivos são disponibilizados foram identificados visualmente na página de sumário de cada número. Para o estudo foram considerados formatos utilizados aqueles sempre disponibilizados pela revista, casos em que alguns números tiveram arquivos disponibilizados em outros formatos exceto o usual, foram ignorados. O padrão de preservação digital, plataforma de publicação e versão utilizados foram coletados através dos dados contidos na aba *Sobre*, que traz informações administrativas do periódico. Adverte-se que nem todas as revistas continham todas as informações.

Os padrões de metadados utilizados pelos *sites* dos periódicos foram detectados através do recurso exibir código fonte da página, acionada com o uso do botão direito do mouse na página de um artigo do último número publicado pelo periódico e aferido aleatoriamente em volumes mais antigos da revista. Os padrões Dublin Core (DC) e Highwire Press (HP) foram identificados pelo uso de *tags* DC. (por exemplo DC.Creator), e citation (por exemplo, citation_title), respectivamente, como mencionado nas diretrizes de inclusão para *webmasters* do recurso ajuda do GS. (GOOGLE, 2018). A adesão ao DOI foi averiguada na exibição e no código fonte para os periódicos que não continham a indicação expressa nas informações sobre o periódico. Assim como os outros formatos de arquivo disponibilizados, a adesão ao DOI foi considerada como positiva aos periódicos que o utilizaram durante todo o seu período da publicação.

3.3 Etapas do Processo de Coleta de Dados

Para o cumprimento do processo de coleta de dados foram seguidas as etapas e os procedimentos descritos a seguir:

Etapa 1: visita aos *sites* dos periódicos a fim de verificar se atendem os critérios definidos no estudo. As revistas que não atenderam esses critérios foram desconsideradas e as que atenderam seguiram para as etapas seguintes

Etapa 2: caracterizar os periódicos quanto à:

- formatos de arquivo disponibilizados;
- padrão de preservação digital;
- padrão de metadados empregado no *site*;
- utilização de DOI da revista
- versão do OJS utilizada

Etapa 3: verificar a existência de página de h5-index do periódico no GS Metrics

Etapa 4: averiguar a existência de perfil de pesquisador

Etapa 5: levantar a estratificação Qualis dos periódicos na plataforma Sucupira da Capes

Etapa 6: identificar a quantidade de artigos recuperados no GS para cada periódico selecionado, utilizando a estratégia de busca [*site:URLdarevista*]. Processo finalizado para os periódicos cujo comando mostrou-se inválido

Etapa 7: identificar a quantidade de artigos publicados em cada periódico do grupo a ser analisado

Etapa 8: determinar a presença no GS (através da taxa de indexação) de cada publicação no GS com a aplicação da fórmula Presença no GS (P)= n° de artigos recuperados pelo GS *100/ n° de artigos publicados na revista. TP= art.GS*100/art.publ.

Etapa 9: análise dos resultados.

Na sequência discorre-se sobre a metodologia adotada no tratamento dos dados obtidos.

3.4 Métodos para Tratamento dos Dados e Apresentação dos Resultados

Todos os dados coletados foram organizados em uma planilha utilizando-se o *software Microsoft Office Excel*®. A taxa de presença é a proporção de artigos recuperados no GS dentro do conjunto de artigos publicados no periódico, portanto o resultado esperado deve estar entre 0 e 1. Contudo, optou-se pela apresentação das taxas em percentual. Para, incluiu-se a multiplicação por 100, conforme a fórmula abaixo:

$$TP = \frac{\mathit{artGS} \cdot 100}{\mathit{artpubl}}$$

Em que **artGS** é o número de artigos recuperados no Google Scholar mediante aplicação do comando *site* e **artpubl** é o de artigos publicados no periódico.

Cabe notar que no Open Journal System, sistema utilizado pelos periódicos analisados, todos os recursos possuem URL próprias (BRITO *et al.*, 2018). Entretanto, cada publicação pode se distinguir utilizando marcações distintas que podem ser: URL da revista, marcação pela página inicial; URL da edição, marcado pela página de sumário e URL do artigo, marcado pela página de identificação do artigo. Desta forma, utilizou-se o domínio do *site* do periódico para recuperar as publicações sem exclusão de diferentes marcações. Constatou-se que grande parte das publicações possuem marcações por artigo, com exceções mínimas de números que foram digitalizados integralmente e que receberam uma marcação, contando, assim como uma *URL* na recuperação, mas na realidade são compostas pelos vários artigos do número. Este é o caso da Revista Brasileira de Biblioteconomia e Documentação (RBBD) que digitalizou os artigos de 1970 a 2000. Foram 48 números marcados pela edição, contabilizados como 48 unidades, mesmo contendo vários artigos cada.

Para o resultado do percentual das taxas de presença foram consideradas duas casas decimais. A estratégia originou 33 resultados distintos, dificultando a análise dos grupos. Na perspectiva de facilitar a descrição e a análise, as taxas de presença foram agrupadas nas seguintes faixas: igual a zero, maior que zero e

menor ou igual a 20% (0-20); maior que 20% e menor ou igual a 40% (20-40); maior que 40% e menor ou igual a (40-60); maior que 60% e menor ou igual a 80% (60-80); maior que 80% e menor ou igual a 100% (80-100) e maior que 100% (*100).

Todos os dados coletados foram organizados em uma planilha utilizando o *software Microsoft Office Excel*. O *software* também foi adotado no tratamento dos dados. Os recursos de filtro possibilitaram a visualização por agrupamento de características para análise e o de tabela dinâmica demonstrou como a taxa de presença se relaciona com as características. Os resultados foram apresentados em formato de quadros, gráficos e tabelas. Telas de eventos ilustrativos foram adicionados como figuras.

3.5 Limites do Estudo

O uso do operador “*site:URL*” deve ser utilizado com ponderação, visto que os resultados obtidos através da estratégia exibem somente versões primárias dos documentos. Deste modo, versões adicionais do mesmo não serão contabilizadas e somente são reveladas após o uso do link “versões” (ARLITSCH; O’BRIEN, 2012; ORDUÑA-MALEA; DELGADO LÓPEZ-CÓZAR, 2015) O alerta dos autores identifica uma falha possível na análise de repositórios utilizando o comando de busca “*site:URL*”, devido ao fato de que a pesquisa foi aplicada a repositórios acadêmicos. Consideram, prudentemente, que num mesmo repositório pode haver registros de diferentes versões de um único documento, como um artigo publicado em anais de um evento e em um periódico, por exemplo.

Shintaku (2017) sinaliza um alerta semelhante em sua pesquisa conceitual sobre estudos métricos na *web*, entretanto, tal equívoco não é prenunciado na análise de periódicos científicos, no qual cada revista mantém uma única versão de registro do documento, inculcando precisão à estratégia utilizada no estudo, pois é exatamente a análise das versões primárias que anunciam os resultados da presença *online* dos periódicos, segundo a indexação no referido buscador. Contudo, a ponderação indica uma das limitações do estudo, porém ressalta-se que a intenção é medir a presença *online* dos documentos publicados em periódicos e recuperados pelo Google Scholar vinculados à *URL* da revista e não a sua presença geral na *web*.

O periódico Em Questão evidenciou o uso do perfil de pesquisador por periódicos, o que abriu a questão às demais revistas. Apesar de ter a página de h5-index do periódico no GSM, a revista Em Questão possui poucos artigos recuperados através dela no GS, conforme exhibe a Figura 4. Para contornar a situação e aumentar a visibilidade do periódico no GS, foi criado um perfil de pesquisador para a revista ao qual foram adicionados todos os artigos publicados, como mostra a Figura 5.

Figura 4 – Tela de busca da página de h5-index do periódico no GSM da Revista Em Questão

Em Questão			
h5-index: 0		h5-median: 0	
Title / Author	Cited by	Year	
Mídias sociais e comunicação científica: análise alométrica em artigos de periódicos da ciência da informação R Ferreira Araujo Em Questão 21 (1)	10	2015	
Produção científica em áreas multidisciplinares: educação a distância no Brasil F Schweitzer, RS Rodrigues Em Questão 19 (1), 156-172	7	2013	
Análise das influências intelectuais na produção científica da área de Ciência da Informação: um estudo sobre os bolsistas de produtividade em pesquisa (PQ-CNPq) L Santiago Bufrem, F Mascarenhas e Silva, N Vitor Sobral Em Questão 23	6	2017	
Indicadores dos grupos de pesquisa da área de Gestão da Informação na Região Nordeste: um enfoque para a colaboração em artigos de periódicos GA Santana, FM Silva, NV Sobral, MH Ferreira Em Questão 20 (3), 229-252	6	2014	
A interdisciplinaridade na Ciência da Informação pela perspectiva da análise de citações M Moraes, AE Carelli Em Questão 22 (1)	6	2016	
A Perspectiva de estudos sobre os sujeitos na Arquivologia, na Biblioteconomia e na Museologia CAÁ Araújo Em Questão 19 (1), 213-238	6	2013	

Data and citation counts are estimated and are determined automatically by a computer program.

Fonte: Dados da pesquisa (2018)

Figura 5 – Tela perfil de pesquisador da Revista Em Questão



Fonte: Dados da pesquisa (2018)

O problema supostamente foi solucionado buscando os nomes de todos os periódicos selecionados para a pesquisa nas duas ferramentas, busca por página de h5-index do periódico no GSM e por perfil de pesquisador. Diz-se supostamente solucionado, pois o próprio **perfil de pesquisador** da revista Em Questão não é recuperado na busca por perfil de pesquisador (Figura 6), tal perfil foi identificado pela autora em uma postagem da página da Revista Em Questão em rede social, como demonstrado na Figura 7, e a verificação com os responsáveis pelo periódico foi propiciada pela proximidade entre a autora e os responsáveis pelo periódico por estarem vinculados à mesma instituição. O limite do estudo gerado pela situação é a possibilidade de o fenômeno ocorrer em outros periódicos e seus perfis existentes não serem recuperados normalmente através da busca por perfis.

Figura 6 – Tela de busca por perfil de pesquisador da Revista Em Questão



Fonte: Dados da pesquisa (2018)

Figura 7 – Postagem em rede social do periódico Em Questão sobre perfil no GS



Fonte: Em Questão, https://www.facebook.com/pg/EmQuestao/posts/?ref=page_internal, 12 abr. 2018

Situações como estas ocorrem devido à dinamicidade da web, que provoca alterações constantes nos resultados das buscas no GS, bem como modificações nos metadados, que podem demorar de 6 a 9 meses para serem detectadas pelos *web crawlers* do GS (GOOGLE SCHOLAR, 2018).

4 Apresentação e Análise dos Dados

Nesta seção são expressos os dados obtidos na pesquisa, sendo posteriormente analisados, conforme os objetivos específicos definidos na seção 1.2.2. Inicia-se com os resultados levantados através da utilização do comando *site:URL* no GS, que forneceu o número de artigos recuperados pelo buscador, em relação ao número de artigos publicados no periódico, resultando na taxa de presença no GS. Logo após, são descritos individualmente os resultados das características determinadas para observação: estratificação Qualis, página de h5-index do periódico no GSM, padrão de preservação digital e de metadados, assim como utilização de DOI e versão do OJS. Na sequência, relaciona-se tais variáveis às taxas de presença identificadas, na procura por padrões que evidenciem uma maior presença no GS dos artigos de periódicos.

4.1 Taxa de Presença dos Periódicos no Google Scholar

O objetivo específico descrito na alínea a) da seção 1.2.2 propõe a criação de um ranking dos periódicos da Ciência da Informação no Google Scholar, através de suas taxas de presença no buscador. Este objetivo foi atingido com a contagem da presença dos artigos de cada periódico no GS e a utilização do comando *site* na caixa de pesquisa do GS que recuperou os artigos existentes no buscador vinculados à *URL*. Mediante o número de artigos recuperados pelo GS e o total de artigos publicados no periódico foi possível calcular a taxa de presença no GS aplicando a fórmula: $TP = \frac{\text{artGS} \cdot 100}{\text{artpubl}}$. Contudo, por motivos ainda não desvendados completamente, o comando *site* não funciona para a busca de todos os *sites* (ARLITSCH; O'BRIEN, 2012; ORDUÑA-MALEA; DELGADO LÓPEZ-CÓZAR, 2015). Neste caso, foram indicados com o termo (inválido) após a adoção do argumento *site:URL* utilizado, conforme pode ser observado no Quadro 1.

Quadro 1 – Informações para cálculo da taxa de presença no GS

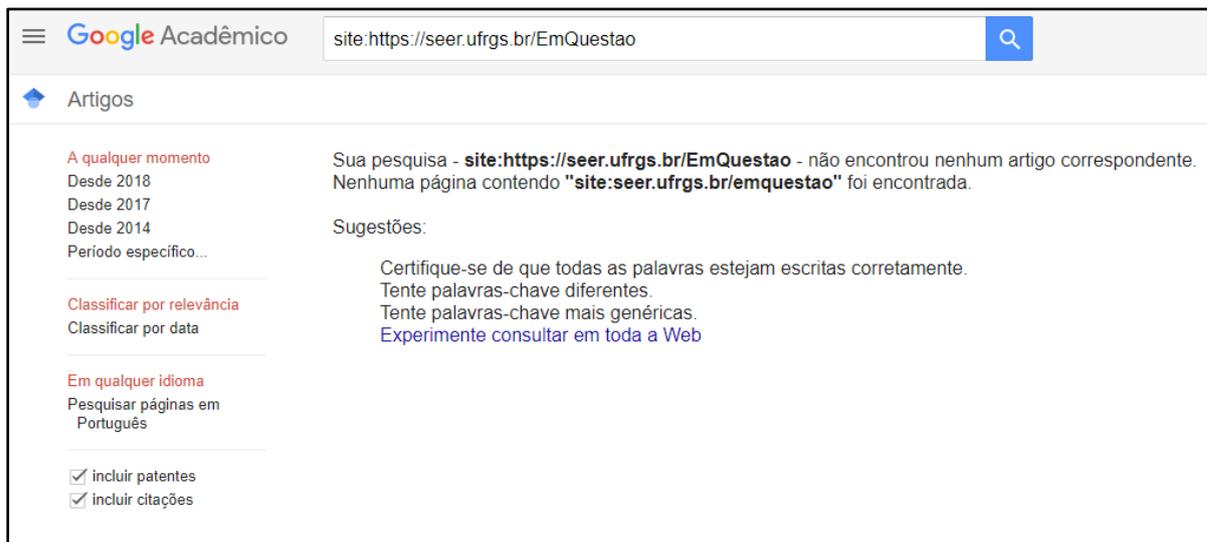
Revista	Comando <i>site:URL</i>	Art. Publ.	Art. Recup. GS	Presença no GS	
Bibliomar	<i>site:http://www.periodicoeletronicos.ufma.br/index.php/bibliomar</i>	30	38	126,66%	1,27
Ciência da Inf. em Revista	<i>site:www.seer.ufal.br/index.php/cir</i>	84	98	116,67%	1,17
Informação@Profissões	<i>site:http://www.uel.br/revistas/uel/index.php/infoprof</i>	89	95	106,74%	1,07
Conhecimento em ação	<i>site:https://revistas.ufrj.br/index.php/rca</i>	47	50	106,38%	1,06
AtoZ	<i>site:https://revistas.ufpr.br/atoz</i>	71	75	105,60%	1,06
Convergências em CI	<i>site:https://seer.ufs.br/index.php/conci</i>	5	5	100%	1
Liinc em revista	<i>site:http://revista.ibict.br/liinc</i>	373	372	99,73%	0,99
Múltiplos Olhares em CI	<i>site:http://portaldeperiodicos.eci.ufmg.br/index.php/moci</i>	214	213	99,53%	0,99
Encontros Bibli	<i>site:https://periodicos.ufsc.br/index.php/eb</i>	596	591	99,16%	0,99
Logeion	<i>site:http://revista.ibict.br/fiinf</i>	53	52	98,11%	0,98
Informação & Informação	<i>site:http://www.uel.br/revistas/uel/index.php/informacao</i>	537	500	93,11%	0,93
Rev. Analisando em CI	<i>site:http://racin.arquivologiauepb.com.br</i>	54	48	88,89%	0,89
Folha de Rosto	<i>site:https://periodicos.ufca.edu.br/ojs/index.php/folhaderosto</i>	69	58	84,06%	0,84
Bibliotecas Universitárias	<i>site:https://seer.ufmg.br/index.php/revistarbu</i>	41	34	82,86%	0,83
Comunicação & Informação	<i>site:https://www.revistas.ufg.br/ci</i>	420	330	78,57%	0,78
Ciência da Informação	<i>site:http://revista.ibict.br/ciinf</i>	1314	874	66,51%	0,66
Inf. na Sociedade Contemporânea	<i>site:https://periodicos.ufrn.br/informacao</i>	45	29	64,44%	0,64
Revista Ibero-Americana de CI	<i>site:ehhttp://periodicos.unb.br/index.php/RICI</i>	287	182	63,41%	0,63
Perspectivas em CI	<i>site:http://portaldeperiodicos.eci.ufmg.br/index.php/pci</i>	1252	786	62,78%	0,63
CAJUR	<i>site:http://www.cajur.com.br/index.php/cajur</i>	75	43	57,33%	0,57
Rev. Comum. Inf. & Inov. Saúde	<i>site:https://www.reciis.icict.fiocruz.br/index.php/reciis</i>	643	359	55,83%	0,56
Rev. Bras. de Bibl. e Documentação	<i>site:https://rbbd.febab.org.br/rbbd</i>	655	358	54,65%	0,55
BIBLOS	<i>site:https://periodicos.furg.br/biblos</i>	479	244	50,94%	0,51

Revista	Comando <i>site:URL</i>	Art. Publ.	Art. Recup. GS	Presença no GS	GS
Àgora	<i>site:https://agora.emnuvens.com.br/ra</i>	446	210	47,08%	0,47
Ponto de Acesso	<i>site:https://portalseer.ufba.br/index.php/revistaici</i>	291	126	43,30%	0,43
IRIS	<i>site:https://periodicos.ufpe.br/revistas/IRIS</i>	15	6	40%	0,4
InCID	<i>site:https://www.revistas.usp.br/incid</i>	208	78	37,50%	0,37
Rev Brasileira de Ed. em CI	<i>site:http://abecin.org.br/portalderevistas/index.php/rebecin</i>	63	21	33,33%	0,33
Revista ACB	<i>Site:https://revista.acbsc.org.br/racb</i>	510	149	29,21%	0,29
Informação em Pauta	<i>site:http://www.periodicos.ufc.br/informacaoempauta</i>	40	11	27,50%	0,27
Brazilian J. of Information Science	<i>site:http://www2.marilia.unesp.br/revistas/index.php/bjis</i>	180	32	17,78%	0,18
Rev. Digital de Bibl. & CI	<i>site:https://periodicos.sbu.unicamp.br/ojs/index.php/rdbci</i>	363	1	0,27%	0,027
Archeion online	<i>site:http://www.periodicos.ufpb.br/ojs/index.php/archeion (inválido)</i>	-	0	0	0
Informação & Sociedade	<i>site:http://periodicos.ufpb.br/ojs2/index.php/ies (inválido)</i>	-	0	0	0
Persp. em Gestão & Conhecimento	<i>site:http://periodicos.ufpb.br/ojs/index.php/pgc (inválido)</i>	-	0	0	0
Biblionline	<i>site:http://periodicos.ufpb.br/ojs2/index.php/biblio (inválido)</i>	-	0	0	0
Em Questão	<i>site:https://seer.ufrgs.br/EmQuestao (inválido)</i>	-	0	0	0
Biblioteca Escolar em Revista	<i>site:http://www.revistas.usp.br/berev/issue/view/10847 (inválido)</i>	-	0	0	0
Transinformação	<i>site:http://www.scielo.br/scielo.php?script=sci_serial&pid=0103-3786&lng=en&nrm=isso (inválido)</i>	-	0	0	0

Fonte: Dados da pesquisa (2018)

A seguir, apresentam-se os resultados referentes à taxa de presença no GS dos periódicos selecionados para o estudo expostos no Quadro 1. Iniciam-se pelos sete (7) periódicos que não retornaram nenhum artigo na busca no GS utilizando o comando *site*, representando um total de 17,95% perante os 39 periódicos analisados. Nestes casos, sem retorno de artigos presentes no GS utilizando o comando *site*, a taxa de presença no GS é inexistente (0). Deste modo, a quantidade de artigos publicados por estes periódicos não foi contabilizada, devido à clara inexistência de presença no buscador. A Figura 8, exposta a seguir, demonstra a visualização exibida nos casos dos sete periódicos nos quais não foram recuperados resultados com o comando, incluindo a Revista Em Questão, periódico vinculado ao Programa de Pós-Graduação em Comunicação e Informação da Universidade do Rio Grande do Sul, ponto de partida inicial para a execução da presente investigação.

Figura 8 – Tela de recuperação de busca com comando *site* inválido



Fonte: Dados da pesquisa (2018)

No resultado do percentual das taxas de presença foram consideradas duas casas decimais, o que originou 33 resultados distintos, que para facilitar a descrição, foram agrupados nas seguintes faixas: igual a zero, maior que zero e menor ou igual a 20% (0-20); maior que 20% e menor ou igual a 40% (20-40); maior que 40% e menor ou igual a (40-60); maior que 60% e menor ou igual a 80% (60-80); maior que 80% e menor ou igual a 100% (80-100) e maior que 100% (*100). Dos 39 periódicos analisados, apenas dois possuem taxa de presença entre 0 e 20% (5,13% do total),

e há valores discrepantes entre eles. A Revista Digital de Biblioteconomia e Ciência da Informação com apenas 1 artigo recuperado, resulta em uma taxa de presença de 0,03% em relação aos 363 artigos publicados até o momento da coleta de dados, e o periódico Brazilian Journal of Information Science apresenta-se com uma taxa de presença de 17,78%.

Cinco periódicos manifestam taxas de presença entre 27,50% e 40%, agrupando-se na faixa entre 20-40% representando 12,82% do total de periódicos. Seis periódicos (15,38%) ficaram na faixa entre 40-60%, com taxas de presença entre 43% e 55%. Cinco periódicos (12,82%) figuraram na faixa de 60-80% de taxa de presença e 23,08% dos periódicos (9) possuem entre 80-100%, no qual somente o novo periódico Convergências em Ciência da Informação atingiu a taxa de 100% de presença no GS. Todavia, ratifica-se que o periódico é recente, possuindo o total de 5 artigos publicados na data de coleta dos dados sendo que análises mais elaboradas carecem de novas coletas para confirmar ou refutar a presença total dos artigos publicados pelo periódico no GS.

As investigações de Arlitsch e O'brien, (2012) e de Orduña-Malea e Delgado López-Cózar (2015) sugerem que as taxas de presença que ultrapassam 100% podem decorrer da duplicação de registros, pois não são raros os repositórios institucionais que possuem mais de um registro do mesmo documento. Tal distorção aparentemente poderia ser evitada na análise de repositórios com um único registro de cada documento, caso dos periódicos mencionados.

Entretanto, detectou-se que em cinco (5) periódicos (12,82%) do *corpus* de 39, o número de resultados na busca no GS ultrapassou o número de artigos publicados efetivamente nos periódicos. Esta distorção pode ter causa na recuperação de documentos no GS como editoriais, expedientes e diretrizes para autores publicados no periódico que equivocadamente foram indexados no GS, como evidencia a seta indicativa na Figura 9, em que a busca no GS utilizando o comando *site* para a revista Informação@Profissões exhibe um documento Expediente em PDF. Situação idêntica foi encontrada no periódico AtoZ, que recuperou a sessão Bem-vindos a AtoZ de 2011, porém neste caso a revista disponibiliza o arquivo em formato PDF e HTML.

Figura 9 – Destaque em registro que não é documento científico do periódico

The image shows a Google Scholar search results page. The search query is 'site http://www.uel.br/revistas/uel/index.php/infoprof'. The results are displayed in a list format. The first result is 'Mediação da informação no campo da moda' by M da Silva Finamor. The second result is 'Biblioteca escolar: discussões constantes, mudanças frágeis' by S Bortolin. The third result, 'Expediente' by MLC de Miranda, is highlighted with a red arrow. The search filters on the left include 'A qualquer momento', 'Classificar por relevância', and 'Em qualquer idioma'.

Fonte: Dados da pesquisa (2018), grifo nosso

O fenômeno não foi previsto e necessita de comparação e aferição individual entre cada artigo no periódico e na recuperação de busca do GS para que a hipótese seja confirmada ou refutada. Todavia, a conferência extrapola os objetivos deste estudo, não tendo sido realizada, mas sua execução mostra-se importante na identificação dos aspectos que interferem na indexação do GS, sendo sugerida para estudos futuros.

Ainda é válido salientar que os periódicos que possuem mais documentos indexados no GS do que publicados na revista têm em comum o fato de apresentarem menos de cem artigos publicados. Entretanto não há evidências que indiquem que este aspecto interfere na taxa de presença devido ao fato de haver outras revistas com menos de cem artigos publicados e taxas maiores. Contudo, esta é a única categoria que agrupa o fenômeno e chama a atenção na planilha de coleta de dados, ver conforme observa-se no Quadro 1.

4.2 Descrição das Características de Análise dos Periódicos

O objetivo específico b) da seção 1.2.2, que trata de caracterizar os periódicos quanto: à estratificação Qualis; à existência de perfil de revista no GS; ao formato de arquivo utilizado; ao padrão de preservação digital ; ao padrão de metadados empregado no *site*; à utilização de DOI da revista; à versão do OJS utilizada, visa

propiciar subsídios para relacionar essas características com a taxa de presença de cada periódico no GS, em busca de evidências de sua influência na promoção de índices mais elevados de presença no buscador. O Quadro 2 expõe os dados coletados. A apresentação segue a ordem decrescente das taxas de presença.

Quadro 2 – Características dos periódicos selecionados para o estudo

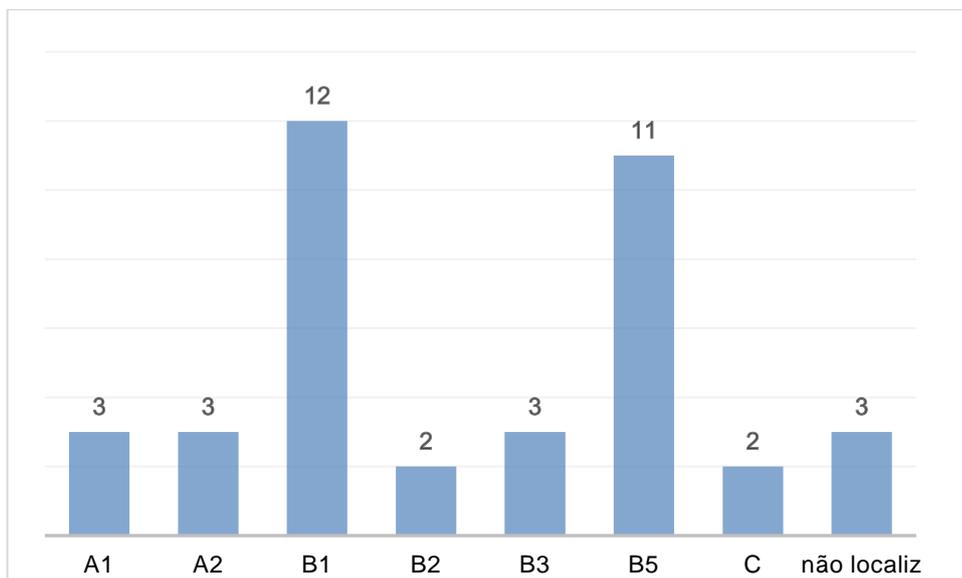
Revista	% presença	QUALIS	Perfil GS	Formato de arquivo	Arquiv.	Pad. de metadados	DOI Periód	Versão OJS	Categorias
Bibliomar	126,66%	não loc	não	PDF	LOCKSS	DC/HP	não	2.4.8.0	*100
Ciência da Informação em Revista	116,67%	B5	não	PDF	LOCKSS	DC/HP	não	2.4.8.1	*100
Informação@Profissões	106,74%	B5	não	PDF	LOCKSS	DC/HP	sim	2.4.8.1	*100
Conhecimento em ação	106,38%	B5	não	PDF	NI	DC/HP	não	2.4.8.0	*100
AtoZ	105,60%	B2	perfil pesq.	PDF/HTML/EPUB	NI	DC/HP	sim	2.4.8.1	*100
Convergências em CI	100%	não loc	não	PDF	LOCKSS	DC/HP	não	2.4.8.0	81-100
Liinc em revista	99,73%	B1	perfil revista	PDF	LOCKSS	DC/HP	sim	2.4.8.3	81-100
Múltiplos Olhares em CI	99,53%	B5	perfil revista	PDF	NI	DC/HP	não	2.4.8.1	81-100
Encontros Bibli	99,16%	A2	perfil revista	PDF	LOCKSS	DC/HP	sim	2.4.8.3	81-100
Logeion	98,11%	B5	não	PDF	LOCKSS	DC/HP	sim	2.4.8.3	81-100
Informação & Informação	93,11%	A2	perfil pesq./periódico	PDF	LOCKSS	DC/HP	não	2.4.8.1	81-100
Rev. Analisando em CI	88,89%	B5	não	URL.PDF	NI	NI	não	NI	81-100
Folha de Rosto	84,06%	B5	não	PDF	NI	DC/HP	não	2.4.6.0	81-100
Bibliotecas Universitárias	82,86%	C	não	PDF	NI	DC/HP	não	2.4.8.3	81-100
Comunicação & Informação	78,57%	B2	não	PDF	LOCKSS	DC/HP	sim	2.4.8.2	61-80
Ciência da Informação	66,51%	B1	perfil revista	PDF	LOCKSS	DC/HP	não	2.4.8.3	61-80
Inf. na Sociedade Contemporânea	64,44%	não loc	não	PDF	NI	NI	não	NI	61-80

Revista	% presença	QUALIS	Perfil GS	Formato de arquivo	Arquiv.	Pad. de metadados	DOI Periód	Versão OJS	Categoria
Revista Ibero-Americana de CI	63,41%	B1	não	PDF	LOCKSS	DC/HP	não	2.4.8.0	61-80
Perspectivas em CI	62,78%	A1	perfil revista	PDF	LOCKSS	DC/HP	não	2.4.8.1	61-80
CAJUR	57,33%	B5	perfil pesq.	PDF	NI	DC/HP	não	2.4.8.1	41-60
Rev. Comum. Inf. & Inov. Saúde	55,83%	B1	perfil revista	PDF	NI	DC/HP	não	2.4.8.1	41-60
Rev. Bras. de Bibl. e Documentação	54,65%	B1	perfil revista	PDF	NI	DC/HP	não	2.4.8.3	41-60
BIBLOS	50,94%	B3	perfil pesq.	PDF	LOCKSS	DC/HP	não	2.4.8.3	41-60
Àgora	47,08%	B1	perfil revista	PDF	LOCKSS	DC/HP	não	2.4.8.3	41-60
Ponto de Acesso	43,30%	B1	não	PDF	LOCKSS	DC/HP	não	2.4.6.0	41-60
IRIS	40%	B3	não	PDF	NI	DC/HP	não	2.4.8.2	20-40
InCID	37,50%	B1	perfil revista	PDF	NI	HP	não	NI	20-40
Rev Brasileira de Ed. em CI	33,33%	B5	não	PDF	LOCKSS	DC/HP	não	2.4.8.1	20-40
Revista ACB	29,21%	B1	perfil revista	PDF	NI	DC/HP	não	2.4.8.3	20-40
Informação em Pauta	27,50%	B5	perfil pesq.	PDF	LOCKSS	DC/HP	não	2.4.8.0	20-40
Brazilian J. of Information Science	17,78%	B1	perfil revista	PDF	NI	DC/HP	sim	NI	0-20
Rev. Digital de Bibl. & CI	0,27%	B1	perfil revista	PDF	LOCKSS	DC/HP	sim	2.4.8.3	0-20
Archeion <i>online</i>	0	C	não	PDF	LOCKSS	DC/HP	não	2.4.8.2	0
Biblionline	0	B5	perfil revista	PDF	NI	DC/HP	não	NI	0
Biblioteca Escolar em Revista	0	B3	não	PDF	NI	HP	sim	NI	0
Em Questão	0	A2	perfil pesq.	PDF	LOCKSS	DC/HP	não	2.4.8.0	0
Informação & Sociedade: Estudos	0	A1	perfil revista	PDF	LOCKSS	DC/HP	não	2.4.8.2	0
Persp. em Gestão & Conhecimento	0	B1	perfil revista	PDF	LOCKSS	DC/HP	não	2.4.8.2	0
Transinformação	0	A1	perfil revista	PDF	NI	HP	sim	Scielo	0

Fonte: Dados da pesquisa (2018)

Quanto à estratificação Qualis, os trinta e nove (39) periódicos analisados distribuem-se da seguinte forma: as classificações Qualis A1 e A2 possuem três (3) periódicos cada e representam individualmente 7,69%. A estratificação B1 é a que possui maior número de periódicos, totalizam doze (12) (30,77%). Dois (2) periódicos apresentam Qualis B2 (5,13%) enquanto B3 é representado por três (3) revistas (7,69%). O Qualis B5 é o segundo colocado em número de periódicos, sendo onze (11) revistas com essa classificação que representam 28,11%. Na estratificação C incluem-se dois (2) periódicos (5,13%) e três (3) periódicos não tiveram suas classificações Qualis localizadas na plataforma Sucupira, são eles: Bibliomar, criado em 2015, Convergências em Ciência da Informação, criado em 2018 e Informação na Sociedade Contemporânea, criado em 2014. As classificações B1 e B5 apresentam juntas a maior representatividade dentro do quesito Qualis, como demonstra o Gráfico 1, e totalizam 58,98% dos periódicos analisados.

Gráfico 1 – Quantidade de periódicos por estratificação Qualis



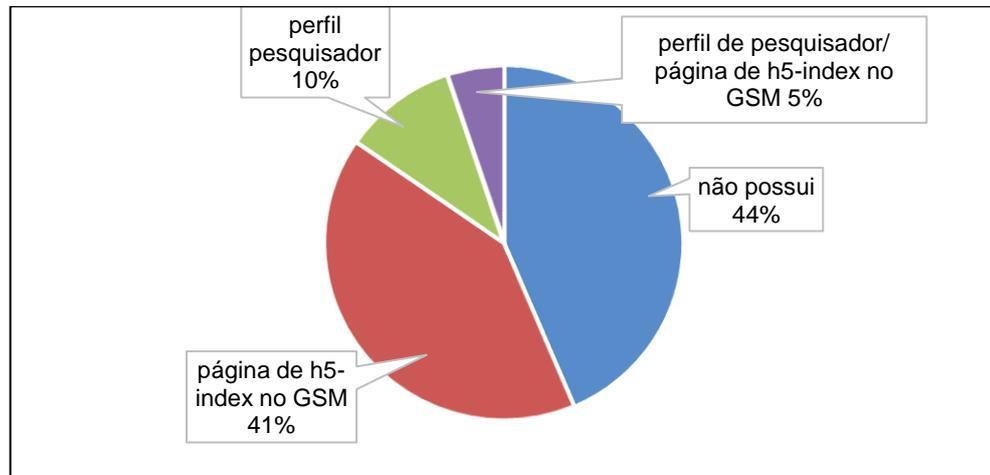
Fonte: Dados da pesquisa (2018)

Para a existência de página de h5-index do periódico no GSM, inicialmente esperou-se resposta binária: sim, para os periódicos que possuíssem perfil no GS, ou não, para os periódicos que não possuíssem perfil no GS. Entretanto, percebeu-se através do periódico Em Questão, que há alguns que criam perfis de pesquisador para a divulgação do mesmo, visto que não são recuperados na pesquisa por

periódicos no GS Metrics. Devido a essa identificação, a categoria 'sim' foi subdividida em: página de h5-index do periódico no GSM, para os periódicos recuperados a partir da busca por periódico no GS Métricas e perfil de pesquisador, para periódicos recuperados através da busca por pesquisadores no ícone My profiles no GS. Como consequência, houve duplicação de categorias para o uso de perfil, de sim ou não inicialmente engendrados, emergem perfil de pesquisador, página de h5-index do periódico no GSM, perfil de pesquisador e página de h5-index do periódico no GSM e não possui nenhum perfil.

Ressalta-se que a utilização do termo página de h5-index do periódico no GSM foi definida com o intuito de excluir ambiguidades, pois o GS, com configuração de idioma inglês, utiliza o termo *search publications*, cuja tradução é pesquisar publicação, contudo, julga-se que o termo publicação é utilizado para explicitar diferentes conceitos, como um documento, um periódico ou uma obra impressa, por exemplo. Devido a isso, optou-se pela adesão do termo página de h5-index do periódico no GSM, que não deixa dúvidas quanto ao objeto a que se refere, questão esclarecida em 1.4.

O Gráfico 2 exibe a quantificação dos periódicos segundo as categorias geradas quanto a existência de perfil no GS. Não possui nenhum tipo de perfil identificado no GS o total de dezessete (17) revistas, que compõe 43,59% dos periódicos analisados. Na sequência, com dezesseis (16) periódicos (41,03%), está o grupo que possui a página de h5-index do periódico no GSM. Apenas quatro (4) periódicos (10,26%) têm perfil de pesquisador, enquanto duas (2) revistas (5,13%), Em Questão e Informação & Informação, criaram perfil de pesquisador concomitante à página de h5-index do periódico no GSM.

Gráfico 2 – Existência de perfil dos periódicos no Google Scholar

Fonte: Dados da pesquisa (2018)

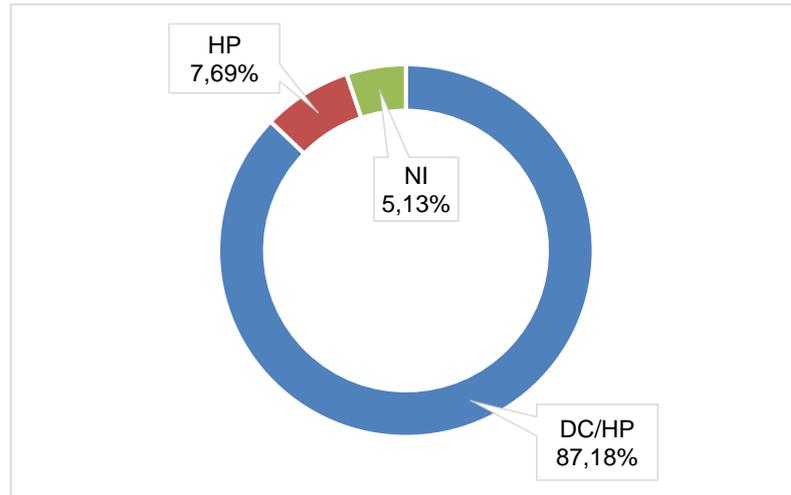
Todos os trinta e nove (39) periódicos analisados disponibilizam os artigos em formato PDF. O periódico AtoZ é o único que disponibiliza os arquivos dos artigos em outros formatos, especificamente. Quanto ao padrão de preservação digital adotado pelos periódicos, vinte e dois (22) informam utilizar o sistema LOCKSS para preservação digital, equivalendo à 56,4%, enquanto dezessete (17) periódicos (43,5%) não cientificam qual o padrão de preservação digital utilizado. Brito et al. (2018) afirmam que, no Brasil, a preservação digital realizada com o sistema LOCKSS utiliza a rede Cariniana, distribuída por todo o País, mas para que efetivamente funcione é necessário que o periódico participe da rede e que o gestor habilite corretamente a coleta de conteúdo para sua distribuição e preservação.

Há a indicação do Google Scholar (2018) que disponibilizar os *links* para os artigos individualmente e com formato *URL.pdf* melhora indexação pelo buscador, contudo, o único periódico analisado que disponibiliza os arquivos dos artigos neste formato é o Analisando em Ciência da Informação e revela uma taxa de presença de 88,89%. Por não haver comparativos não é possível afirmar que este seja o motivo de uma taxa de presença elevada.

A não adesão ao DOI predomina entre os periódicos analisados. Vinte e nove (29) revistas não usaram o recurso de identificação na maior parte do seu período de existência, representando 74,36% do total pesquisado. O restante, 10 periódicos (25,64%), favoreceu-se do uso do identificador. O Gráfico 3 exibe o uso dos padrões de metadados. O uso concomitante dos padrões Dublin Core e

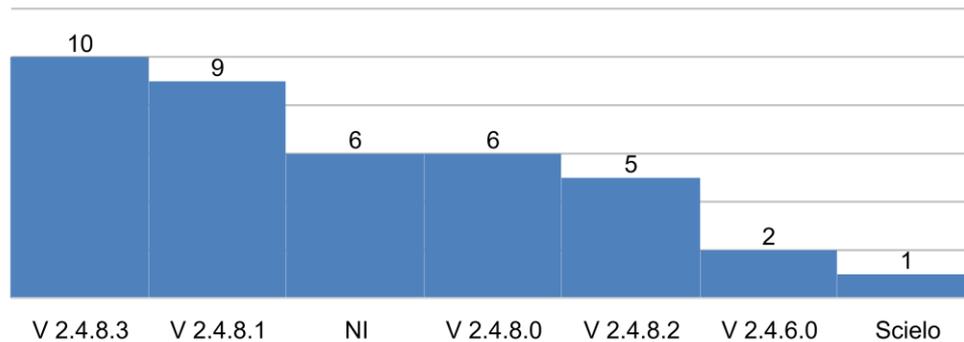
Highwire Press foi identificado em trinta e quatro (34) periódicos, equivalendo a 87,18% das revistas investigadas. O padrão também é utilizado por mais 3 periódicos (7,69%) que não adotam o Dublin Core (DC). Dois (2) periódicos (5,13%) utilizam padrões de metadados não identificados.

Gráfico 3 - Padrão de metadados empregado no *site* do periódico



Fonte: Dados da Pesquisa (2018)

O Gráfico 4 expõe o uso, em ordem decrescente, das versões OJS pelos periódicos. A versão OJS 2.4.8.3 é a mais utilizada e compõe 25,64% do total, sendo empregada por dez (10) periódicos, seguida da versão OJS 2.4.8.1, adotada por nove (9) periódicos. Seis periódicos não informam em seus *sites* qual a versão do OJS que empregam, correspondendo a mesma quantidade que utiliza a versão OJS 2.4.8.0, constituindo, cada um, 15,38% do total estudado. Em seguida, aparece a versão OJS 2.4.8.2, adotada por cinco (5) periódicos (12,82%). A versão mais antiga do grupo, a OJS 2.4.6.0, foi encontrada em apenas duas (2) duas revistas (5,13%). Somente a revista *Transformação pública* seus artigos através da plataforma de publicação Scielo e representa os 2,56% restantes. A seguir, explicitam-se as análises das relações entre as características anteriormente descritas com as taxas de presença no GS.

Gráfico 4 – Versões do OJS utilizadas pelos periódicos

Fonte: Dados da pesquisa (2018)

4.3 Relações entre característica e taxa de presença

Esta subseção discorre sobre as relações entre os cruzamentos dos dados obtidos para as características selecionadas e sua taxa de presença no GS. A análise das relações entre características e taxa de presença inicia-se com a estratificação Qualis. Posteriormente são analisadas a existência de página de h5-index do periódico no GSM, o formato de arquivo utilizado, o padrão de preservação digital, o padrão de metadados empregado no *site* do periódico, a adoção do DOI pela revista, finalizando com a versão do OJS utilizada.

4.3.1 Estratificação Qualis

Como constatado anteriormente, as classificações Qualis B1 e B5 são as que mais possuem representantes, com 12 e 11 periódicos respectivamente. Dos doze (12) periódicos com estratificação Qualis B1, quatro (4) apresentam taxas de presença entre 40% a 60%. No entanto, os periódicos com Qualis B5, com qualificação inferior, portanto, mostram maior representatividade nas taxas de presença entre 80% a 100%, no qual aparecem 4 dos 11 periódicos B5. O resultado fornece indícios de que a estratificação Qualis mais elevada não propicia a maior presença no GS. Contudo, este resultado foi baseado na contagem unitária de periódicos relacionados a cada categoria de taxa de presença. Entende-se, por lógica, que como os Qualis B1 e B5 possuem maior volume de periódicos apareceriam mais nas divisões das categorias de presença.

A Tabela 1 exibe os dados reorganizados para que as classificações Qualis possam ser equiparadas segundo as subdivisões das categorias de taxa de presença. Assim é possível detectar as taxas de presença mais frequentes para cada estratificação Qualis individualmente e determinar a interferência do Qualis na presença no GS. Cada linha comporta os dados de uma estratificação Qualis individualmente (100% dos periódicos que possuem o referido Qualis) e as colunas integram as categorias de taxa de presença, tornando possível a comparação entre o percentual de presença e cada estratificação Qualis. Para melhor visualização os dados com resultado zero foram excluídos.

É notável observar que o Qualis A1 aponta que 66,67% dos periódicos nele classificados não possuem presença no GS, diferentemente do Qualis A2, em que a mesma porcentagem situa-se na margem entre 80% a 100% de presença, enquanto B1 e B5 demonstram maior dispersão dentre as categorias de presença. Ratifica-se que elas também são as mais numerosas em volume de periódicos A1 (3), A2 (3), B2 (2), B3 (3), C (2) e Qualis não localizados (3) e concentram-se em duas ou três categorias de presença no GS.

Tabela 1 – Relação Qualis x taxa de presença

Qualis (quant. De periódicos)	0%	0-20%	20-40%	40-60%	60-80%	80-100%	+100%	Total Geral
A1 (3)	66,67%				33,33%			100,00%
A2 (3)	33,33%					66,67%		100,00%
B1 (12)	8,33%	16,67%	16,67%	33,33%	16,67%	8,33%		100,00%
B2 (2)					50,00%		50,00%	100,00%
B3 (3)	33,33%		33,33%	33,33%				100,00%
B5 (11)	9,09%		18,18%	9,09%		36,36%	27,27%	100,00%
C (2)	50,00%					50,00%		100,00%
não local. (3)					33,33%	33,33%	33,33%	100,00%
Total Geral	17,95%	5,13%	12,82%	15,38%	12,82%	23,08%	12,82%	100,00%

Fonte: Dados da pesquisa

Com as duas análises realizadas, considerando o Qualis em relação à taxa de presença no GS e culminando na convergência de resultados, é possível afirmar, com certa convicção, que a atribuição de Qualis com qualificações superiores não influencia a presença no GS. A seguir, pondera-se sobre a existência de perfil no GS relacionado à taxa de presença no GS.

4.3.2 Existência de perfil no GS

Como exposto na seção 4.2, os resultados identificados para perfil no GS mostraram que: dezessete (17) periódicos não possuem nenhum perfil no GS (43,59%), dezesseis (16) possuem página de h5-index do periódico no GSM (41,03%), quatro (4) possuem perfil de pesquisador (10,26%) e dois (2) possuem ambos perfis (5,13%), de página de h5-index do periódico e perfil de pesquisador. Estes são a revista *Em Questão*, com taxa de presença inexistente no GS, e o periódico *Informação & Informação*, com 93,11% de taxa de presença. A seguir, relacionam-se esses resultados às taxas de presença identificadas no GS, através das categorias de presença criadas para a análise, como pode ser consultado na Tabela 2.

Tabela 2 – Relação perfil no GS x taxa de presença

Perfil no Google Scholar	0	0-20%	20-40%	40-60%	60-80%	80-100%	*100	Total Geral
Não possui	11,76%		11,76%	5,88%	17,65%	29,41%	23,53%	100,00%
Página de h5-index do periódico no GSM	25,00%	12,50%	12,50%	18,75%	12,50%	18,75%		100,00%
Perfil pesquisador			25,00%	50,00%			25,00%	100,00%
Perfil pesq./periódico	50,00%					50,00%	0,00%	100,00%
Total Geral	17,95%	5,13%	12,82%	15,38%	12,82%	23,08%	12,82%	100,00%

Fonte: Dados da pesquisa

Dentre os periódicos que não possuem perfil no GS, há uma variação significativa dos percentuais de presença no GS, de 126,66% do periódico *Bibliomar* a zero por cento (0%) com as revistas em que o argumento *site* foi inválido. Destacam-se duas categorias, a faixa que abarca os periódicos que apresentam taxas entre 80% e 100%, com 29,41% dos periódicos analisados que não possuem perfil no GS, e a categoria com taxas maiores de 100% que representa 23,53% dos periódicos estudados. Não houve nenhum periódico na faixa entre 0% e 20% de presença que não possuísse perfil no GS. Um periódico do *corpus* encontra-se na faixa de presença entre 40% e 60% (5,88%). Para taxas inexistentes e faixa entre 20% a 40% figuram 2 periódicos cada (11,76%).

Para aqueles que dispõem de página de h5-index do periódico no GSM, a dispersão é semelhante. Porém, com esta característica não há nenhum periódico

que ultrapasse o percentual de presença de 100% e a maior concentração de periódicos (25%) localiza-se naqueles que não foram recuperados no buscador por meio do argumento '*site*', com taxa de presença zero. As categorias entre taxas de presença de 0% a 20%, 20% a 40% e 60% a 80% concorrem com dois periódicos cada (12,5%) e entre 40%-60% e 80%-100%, três periódicos cada (18,75%). Portanto, a característica página de h5-index do periódico no GSM, em ordem decrescente de taxa de presença, apresenta taxas que vão de 99,73%, com o periódico Liinc, a zero por cento (0%), com os periódicos Biblionline, Informação e Sociedade, Perspectivas em Gestão e Conhecimento e Transinformação, não evidenciando que a página de h5-index do periódico no Google Scholar Metrics proporcione taxas mais elevadas.

Dos quatro (4) periódicos que utilizam o perfil de pesquisador, dois apresentaram taxas entre 40% a 60%, representando 50% do total. Um possui taxa superior a 100% e o outro tem taxa entre 20% a 40%, assim, cada um representa 25% do grupo de periódicos que utiliza o perfil de pesquisador.

Dois periódicos são encontrados no Google Scholar Metrics com página de h5-index, além disto recorreram ao uso de perfil de pesquisador. No entanto, perante os resultados obtidos, observa-se uma diferença discrepante entre eles. Enquanto a revista Informação & Informação exibe uma taxa de presença de 93,11%, a revista Em Questão não recupera nenhum artigo com o comando *site* e aparece entre os periódicos com taxa de presença zero no buscador. Nota-se que a utilização dos dois perfis, pesquisador e periódico, não geram evidências de elevação na taxa de presença. Comparações efetivas não puderam ser efetuadas pelo fato de haver poucos periódicos nos quais se percebe esta situação. Todavia, também não refuta a hipótese de a revista Em Questão manifestar outros aspectos que dificultem a indexação dos artigos no buscador.

4.3.3 Formato de arquivo utilizado

A disponibilização dos arquivos de repositórios em formato PDF é uma recomendação do Google Scholar (2018), seguida por todos os trinta e nove (39) periódicos analisados. No entanto, foram identificados dois periódicos que disponibilizam seus arquivos em outros formatos além do PDF. Na revista Atoz é

possível visualizar os artigos no formato HTML, na tela do navegador, além de disponibilizá-los também em e-Pub. A taxa de presença dos artigos da revista no GS ultrapassa o total de artigos publicados em 105,60%. Conjectura-se que a ocorrência se deve ao fato de haver indexação equivocada de documentos, como editorial e expediente, já expostos na seção 4.2.

A disponibilidade de diferentes formatos foi considerada somente para a utilização em todos os volumes do periódico. Na coleta de dados identificou-se que o periódico Encontros Bibli disponibilizou os artigos no formato HTML em alguns números, também apresentando elevada taxa de presença no GS. Contudo, não é possível afirmar que tal característica influenciou a indexação no GS, sem a averiguação de todos os artigos individualmente e a comparação com sua presença ou ausência na recuperação de busca no GS.

A Tabela 3, a seguir exposta, anuncia os achados relativos ao cruzamento das informações de formato de arquivo disponíveis nos *sites* dos periódicos e as faixas de taxa de presença. Diferentemente da Tabela 1 e da Tabela 2, em que o total de cada linha representa uma divisão da característica avaliada, a Tabela 3 reflete o uso do formato PDF por 100% dos periódicos do *corpus*, porém compara a utilização exclusiva do PDF ao uso concomitante de outros formatos.

Tabela 3 – Formato de arquivo x taxa de presença

Formato de arquivo	0	0-20	20-40	40-60	60-80	80-100	+100	Total Geral
URL.PDF						2,56%		2,56%
PDF	17,95%	5,13%	12,82%	15,38%	12,82%	20,51%	10,26%	94,87%
PDF/HTML/EPUB							2,56%	2,56%
Total Geral	17,95%	5,13%	12,82%	15,38%	12,82%	23,08%	12,82%	100,00%

Fonte: Dados da pesquisa (2018)

O uso do formato *URL.pdf* é utilizada pelo periódico Analisando em Ciência da Informação que exibe na busca com comando *site* no GS 88,89% de seus artigos publicados. Os periódicos que utilizam predominantemente o formato PDF dispersam-se por todas as faixas de presença no GS. A revista AtoZ, que disponibiliza os arquivos em PDF, HTML e e-PUB, apresenta presença superior a 100%, conforme suposições já mencionadas.

4.3.4 Padrão de preservação digital

O uso do padrão de preservação digital LOCKSS foi identificado em vinte e dois (22) dos periódicos analisado, representando 56,41% do total. Contudo, dezessete (17) periódicos não informam, na documentação do *site*, qual o padrão de preservação digital adotado. Brito et al. (2018, p.60) afirmam que no Brasil “[...] a preservação é realizada com o sistema LOCKSS, por meio da rede Cariniana [...]”. Mas não há menção de uso exclusivo do padrão LOCKSS no País, o que não deixa claro a possibilidade ou não do uso do padrão CLOCKSS para o restante das revistas.

4.3.5 Padrão de metadados empregado no site

A Tabela 4 demonstra a comparação do uso dos padrões identificados na exibição do código fonte dos *sites* dos periódicos. Nas diretrizes de inclusão para *webmasters* o Google Scholar aconselha o uso do padrão de metadados Highwire Press, Eprints e BE Press. E recomenda:

Use as *tags* do Dublin Core (por exemplo, DC.title) como último recurso - elas funcionam mal para os artigos de periódicos porque o Dublin Core não possui campos não ambíguos para o título, volume, edição e números de página da revista. (GOOGLE, 2018, n.p.)

Tabela 4 – Padrão de metadados x taxa de presença

Padrão de metadados	0	0-20%	20-40%	40-60%	60-80%	80-100%	+100%	Total Geral
DC/HP	14,71%	5,88%	11,76%	17,65%	11,76	23,53%	14,71%	100,00%
HP	66,67%			33,33%				100,00%
Não identificado						50,00%	50,00%	100,00%
Total Geral	17,95%	12,82%	5,13%	12,82%	15,38%	12,82%	23,08%	100,00%

Fonte: Dados da pesquisa (2018)

Divergindo da recomendação, trinta e quatro (34) periódicos fazem uso do padrão Dublin Core, porém concomitante ao Highwire Press. A maior concentração de periódicos encontra-se na faixa de 80% a 100% de presença no GS, sendo 8 periódicos que representam 23,53% do total deles que emprega DC e HP. Mas há

ampla variação, de zero (0%) a mais de 100% de presença. Em cada extremo, no 0% e no 100% situam-se 5 periódicos em cada faixa.

Ao contrário da recomendação do GS, 2 dos 3 periódicos que usam o padrão de metadados Highwire Press não recuperam nenhum artigo. Apesar do pequeno número de periódicos com essa característica, é possível afirmar que a recomendação do GS do uso do padrão Highwire Press como forma de melhorar a indexação pelo buscador não se mostrou verdadeira.

4.3.6 Utilização de DOI da revista

Referente à utilização de DOI, 74,36% das revistas analisadas não fazem uso do recurso, contra as 25,64% que o utilizam. A média de presença desses periódicos é de 56,17%, mas na Tabela 5 percebe-se que esta média não traduz uma representação útil, frente à variação entre as taxas de presença. Para comparação, a média de presença entre os que possuem DOI é de 60,60%. A diferença é relativamente pequena, mas se mostra significativa em confronto com a análise individual das duas respostas possíveis e posterior comparação entre elas.

Tabela 5 – DOI x taxa de presença

Possui DOI	0	0-20	20-40	40-60	60-80	80-100	+100	Total Geral
não	17,24%	0,00%	17,24%	20,69%	13,79%	20,69%	10,34%	100,00%
sim	20,00%	20,00%	0,00%	0,00%	10,00%	30,00%	20,00%	100,00%
Total Geral	17,95%	5,13%	12,82%	15,38%	12,82%	23,08%	12,82%	100,00%

Fonte: Dados da pesquisa (2018)

Desta forma constata-se que as categorias sim e não se organizam distintamente dentre as faixas de presença. Ambas parecem não ser aspectos cruciais para a definição da indexação, caso contrário haveria destaque evidente para alguma faixa. A categoria sim mostra o mesmo percentual para periódicos com taxa zero e com taxa maior que 100%, além de que a faixa entre 80% a 100% não se difere muito da anterior. A categoria não, apesar de mais volumosa, também apresenta uniformidade, reforçando a não interferência na indexação.

4.3.7 Versão do OJS

A Tabela 6 apresenta os dados de plataformas utilizadas pelos periódicos em relação às taxas de presença. A não identificação das versões do OJS adotadas por seis (6) periódicos, assim como a existência de apenas um (1) periódico que utiliza a plataforma Scielo, inviabiliza a comparação destas categorias com as demais. As versões 2.4.6.0 e 2.4.8.2 do OJS são utilizadas por 2 e 5 periódicos respectivamente, a versão 2.4.6.0 foi adotada por dois periódicos, um pertence à faixa de 40% a 60% de presença e o outro à faixa de 80% a 100% e a versão OJS 2.4.8.2 é utilizada por 5 periódicos, porém 3 deles não apresentam presença sob o comando *site*. As demais versões apresentam dispersão entre as faixas de presença, sendo o resultado mais alto de 40% para a faixa de 80% a 100% da versão 2.4.8.3. Contudo, a dispersão do restante dos resultados evidencia que a versão utilizada não interfere na indexação dos artigos de periódicos no GS e que o indício sinalizado pela localização de 4 periódicos, na faixa de 80% a 100% da versão 2.4.8.3, dá-se pelo fato de ser a versão com maior representação numérica.

Tabela 6 – Plataforma x taxa de presença

Plataforma	0	0-20	20-40	40-60	60-80	80-100	+100	Total Geral
Não identificada	33,33%	16,67%	16,67%		16,67%	16,67%		100,00%
OJS 2.4.6.0				50,00%		50,00%		100,00%
OJS 2.4.8.0	16,67%		16,67%		16,67%	16,67%	33,33%	100,00%
OJS 2.4.8.1			11,11%	22,22%	11,11%	22,22%	33,33%	100,00%
OJS 2.4.8.2	60,00%		20,00%		20,00%			100,00%
OJS 2.4.8.3		10,00%	10,00%	30,00%	10,00%	40,00%		100,00%
Scielo	100,00%							100,00%
Total Geral	17,95%	5,13%	12,82%	15,38%	12,82%	23,08%	12,82%	100,00%

Fonte: Dados da pesquisa (2018)

A versão 2.4.8.0 é utilizada por seis (6) periódicos, 2 apresentam taxas de presença maiores que 100%, os outros se dispersam nas demais faixas. Observa-se que esta versão é a utilizada pela revista *Em Questão*, cujo comando *site* foi inválido. Brito *et al.* (2018, p. 21) afirmam que:

Nos casos dos portais, cada periódico tem seu próprio endereço (URL), podendo ter sua própria interface visual e funcionalidades. Quando uma revista compartilha uma mesma instalação OJS, alguns recursos ficam no portal, como, por exemplo, os dados cadastrais dos usuários e alguns plugins que afetam o portal como um todo.

Sabe-se que a revista *Em Questão* compõe o portal de periódicos da UFRGS, juntamente com outros 52 periódicos. Desde modo, surge a indagação da possibilidade de o compartilhamento do domínio interferir na indexação dos artigos do referido periódico no GS. Outra possibilidade, também para os demais periódicos, é o uso de plugins de metadados externos ao sistema OJS interferirem na indexação. O uso de plugins externos é um cenário aprovado pelo sistema de editoração, mas há também os plugins nativos, necessários ao sistema, cuja desativação não é permitida, existindo um padrão de plugins utilizados por todos os periódicos que adotam a plataforma de editoração OJS (PUBLIC KNOWLEDGE PROJECT, [s.d.]).

Observa-se que recomendação do Google Scholar (2018) de uso da plataforma *web* OJS para publicação de periódicos eletrônicos não demonstra eficácia no quesito promoção de presença no buscador.

4.4 Outras Proposições

Após as avaliações descritas, estima-se que os aspectos que interferem na indexação no Google Scholar são mais dinâmicos do que a suposição inicial, o que gerou a abordagem desenvolvida nesta subseção, englobando as questões identificadas durante o processo de coleta dos dados que extrapolaram a proposta deste estudo, mas que merecem destaque.

Visto que a coleta de dados foi realizada no período de 29 de junho a 7 de julho de 2018, esperava-se que a realização de buscas com os mesmos critérios, porém em datas posteriores, trariam um número maior de artigos devido a publicação de novos fascículos pelos periódicos. Não obstante, observando atentamente a Figura 9, é possível verificar que o *print* da tela de recuperação de busca no GS, realizado em 26 de outubro de 2018, apresenta um montante de 83 resultados recuperados, contudo, verificando o número de artigos recuperados no período na coleta de dados, constata-se que foram 95 resultados recuperados para o mesmo periódico, no caso Informação@Profissões.

O extremo oposto ocorreu com o periódico Revista Digital de Biblioteconomia e Ciência da Informação, que no momento da coleta de dados apresentou 1 artigo na busca no GS com o uso do comando *site* e, dois meses após, exibiu 247 artigos no buscador, dos 363 publicados contabilizados no período da coleta. Os editores da revista foram contatados e questionados sobre alguma alteração ocorrida, especialmente relacionada aos metadados, que pudessem culminar em tal mudança, porém a resposta foi que nada havia sido alterado no periódico, no período em questão.

As revistas Atoz e Analisando em Ciência da Informação mostraram-se casos isolados na característica formato de arquivo disponível, porém a predominância de artigos presentes no GS, com taxas de 105,60% e 88,89%, respectivamente, são ocorrências que merecem posterior atenção, pois podem trazer indícios de aspectos que culminam na elevada presença no GS, considerando que há recomendações do Google Scholar (2018) referentes à disponibilização de formatos de arquivos.

Durante a coleta de dados observou-se que os artigos da revista Em Questão, cujo comando *site* foi inválido, quando pesquisados por título no GS não retornavam com os arquivos vindos da revista como versão principal, mas sim de outros *sites*, bases de dados, catálogos ou repositórios, conforme está destacado na Figura 10.

Figura 10 – Busca por título de artigo da revista Em Questão

The image shows a Google Scholar search interface. The search bar contains the text "bibliometria evolução histórica e questões atuais". Below the search bar, it indicates "Articles" and "About 1,070 results (0.05 sec)". On the left side, there are filters for "Any time" (with sub-options: Since 2018, Since 2017, Since 2014, Custom range...), "Sort by relevance" (selected), "Sort by date", "include patents" (checked), "include citations" (checked), and "Create alert". The main search results area shows two entries for the article "Bibliometria: evolução histórica e questões atuais" by CAA Araújo, 2006. The first entry is from "unirioja.es" and is labeled "[PDF] unirioja.es". The second entry is from "brapci.inf.br" and is labeled "[HTML] brapci.inf.br". A red rectangular box highlights these two entries.

Fonte: Dados da pesquisa (2018)

A partir deste fato supôs-se que situação similar aconteceria às outras revistas em que o argumento *site*, no GS, foi inválido. Entretanto, a suposição não

se mostrou verdadeira com os demais periódicos nesta situação, com exceção da revista *Perspectivas em Gestão & Conhecimento*, que demonstrou o mesmo comportamento da *Em Questão*. Outra observação a ser feita é que nem mesmo ao solicitar o *link* 'outras versões' a versão contida na própria revista foi recuperada, evidenciando que o GS não é capaz de recuperar informações vinculadas à *URL* dos periódicos mencionados. A revista *Transinformação* não foi reavaliada por utilizar a plataforma Scielo.

5 Considerações finais

Shintaku (2017) salienta que não há consenso sobre qual a melhor métrica para utilização em ambiente virtual e considera que os estudos métricos são essenciais para avaliação da plataforma. Alia-se a convicção de Veríssimo (2012), que propõe ao bibliotecário mobilizar-se para responder à altura aos desafios resultantes das tecnologias da informação, pois não é somente um desafio, é uma oportunidade que deve ser aproveitada para que ele possa dominar saberes que auxiliam o usuário, contribuem para a visibilidade da produção científica da sua instituição, fomentam a pesquisa e o compartilhamento de informação, cernes da profissão. Tal medida não se limita ao GS, entende-se que o bibliotecário deve dominar os recursos de acesso à informação em qualquer suporte, principalmente nos suportes e meios que são mais utilizados pelos usuários. Neste sentido, este é um estudo que mostra os vestígios no uso cotidiano da web para acesso à informação e sua importância para a disseminação e o uso da informação.

Em um contexto de ciberespaço, oferecer produtos e serviços de informação para os usuários, de qualidade e fora do ambiente da biblioteca, seja para uso educativo, informativo ou lúdico, é de suma importância para a própria sobrevivência da biblioteca. Para que estes serviços tenham qualidade, o uso de métricas que o avaliem são fundamentais e indispensáveis. A presença da internet no cotidiano das pessoas faz com que a biblioteca precise estar atenta e atualizada com esta ferramenta e suas infinitas possibilidades.

Neste contexto o GS é uma das mais recentes ferramentas de recuperação da informação, possibilitando um *rol* de exploração mais abrangente que os mecanismos de busca comerciais por não impor limitações disciplinares, geográficas, idiomáticas e financeiras para recuperação da informação. Porém, como demonstrado em estudos já realizados, pesquisas baseadas no GS, são escassas na área da Ciência da Informação, possivelmente por ser um recurso novo e pelo reduzido grupo de pesquisadores da área que se identifica com as tecnologias virtuais. Seu uso é mais evidente em análises bibliométricas, especialmente nas análises de citações, índice-h e assuntos relacionados, ou estudos já estabelecidos em outros meios de comunicação científica que foram adaptados, remodelados e empregados no buscador.

Contudo, a limitada exploração do buscador, aliado aos nebulosos critérios de indexação utilizados pela empresa, é um campo fértil e promissor para investigações, além de serem extremamente necessárias, visto que essa ferramenta de pesquisa é amplamente utilizada pelos usuários da informação. Como forma de colaborar em investigações futuras, indica-se a investigação comparativa entre periódicos com taxas de presença elevadas e periódicos com baixas taxas de presença, a fim de identificar quais os fatores interferem na indexação do GS.

Avalia-se que os objetivos propostos foram alcançados visto que o objetivo geral foi atingido ao relacionar as taxas de presença calculadas às características selecionadas. Contudo, não foi possível determinar os aspectos que interferem na elevação na taxa de presença no GS, mas compreendeu-se que nem mesmo as recomendações do GS para indexação são eficazes isoladamente.

Referências

AFONSO, Raffaella Dayane; MATIAS, Márcio. Análise webométrica das menções web entre os arquivos públicos estaduais e o Arquivo Nacional. **ÁGORA: Revista do Curso de Arquivologia da UFSC**, [s. l.], v. 26, n. 53, p. 347–372, 2016. Disponível em: <https://agora.emnuvens.com.br/ra/article/view/634>.

AGUILLO, Isidro; ORTEGA, José-Luis; FERNÁNDEZ, Mario. Webometric ranking of world universities: introduction, methodology and future developments. **Higher education in Europe**, [s.l.], v. 33, n. 2/3, p. 234-244, 2008. https://www.researchgate.net/publication/263449934_Webometric_Ranking_of_World_Universities_Introduction_Methodology_and_Future_Developments. Acesso em: 25 abr. 2018.

ARAÚJO, Ronaldo Ferreira De. Ciência 2.0 e a presença online de pesquisadores: visibilidade e impacto. **Ciência da Informação em Revista**, [s. l.], v. 1, n. 3, p. 32–40, 2014. Disponível em: <http://www.seer.ufal.br/index.php/cir/article/view/1608>. Acesso em: 18 jun. 2018.

ARAÚJO, Ronaldo Ferreira De. Da altmetria à análise de citações: uma análise da revista Datagramazero. **Datagramazero**, Paraíba, v. 16, n. 1, 2015. a. Disponível em: <https://www.brapci.inf.br/v/a/18494>. Acesso em: 15 maio. 2018.

ARAÚJO, Ronaldo Ferreira De. Estudos métricos da informação na web e o papel dos profissionais da informação. **Revista Bibliotecas Universitárias**, Belo Horizonte, v. 2, n. esp., p. 42–64, 2015. b. Disponível em: <https://www.bu.ufmg.br/rbu/index.php/localhost/article/viewFile/110/70>. Acesso em: 18 jun. 2018.

ARLITSCH, Kenning; O'BRIEN, Patrick S. Invisible institutional repositories. **Library Hi Tech**, [s. l.], v. 30, n. 1, p. 60–81, 2012. Disponível em: <http://www.emeraldinsight.com/doi/10.1108/07378831211213210>. Acesso em: 1 abr. 2018.

ARROYO, Natalia; PAREJA, Víctor Manuel. Metodología para la obtención de datos con fines cibernéticos. *In: TALLER DE INDICADORES BIBLIOMÉTRICOS*, 3., 2003. Madri. **Proceedings** [...]. Madrid: 2003, documento não paginado. Disponível em: <http://eprints.rclis.org/6372/>. Acesso em: 30 set. 2018.

ARSHAD, Alia; AMEEN, Kanwal. Scholarly communication in the age of Google: Exploring academics' use patterns of e-journals at the University of the Punjab. **Electronic Library**, [s. l.], v. 35, n. 1, p. 167–184, 2017. Disponível em: <https://www.emeraldinsight.com/doi/abs/10.1108/EL-09-2015-0171>. Acesso em: 24 ago. 2018.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 6021**: informação e documentação - publicação periódica científica impressa - apresentação. Rio de Janeiro, ABNT, 2003.

BAR-ILAN, Judit. A microscopic link analysis of academic institutions within a country—the case of Israel. **Scientometrics**, Budapeste, v. 59, n. 3, p. 391–403, 2004. Disponível em: <https://link.springer.com/article/10.1023/B%3ASCIE.0000018540.33706.c1>. Acesso em: 13 jul. 2018.

BARRETO, Aldo de Albuquerque. Mudança estrutural no fluxo do conhecimento: a comunicação eletrônica. **Ciência da Informação**, Brasília, v. 27, n. 2, p. 122–127, 1998. Disponível em: <http://www.scielo.br/pdf/%0D/ci/v27n2/2729803.pdf>. Acesso em: 18 jun. 2018.

BILHAR, Suelen Spíndola. **Revista em questão**: uma análise das citações recebidas. 99 f. 2013. Trabalho de conclusão de curso (Graduação em Biblioteconomia) – Faculdade de Biblioteconomia e Comunicação, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2013. Disponível em: <https://lume.ufrgs.br/handle/10183/78372>. Acesso em: 10 ago. 2018.

BJÖRNEBORN, Lennart; INGWERSEN; Peter. Towards a basic framework for webometrics. **Journal of American Society for Information Science and Technology**, v. 55, n. 14, p. 1216-1227, 2004. Disponível em: https://www.researchgate.net/publication/264289068_Toward_a_basic_framework_for_Webometrics. Acesso em 12 ago. 2018.

BRITO, Ronnie Fagundes De *et al.* **Guia do usuário OJS 3**. Brasília, DF: IBICT, 2018. Disponível em: http://bibjuventude.ibict.br/jspui/bitstream/192/252/1/IBICT_OJS_3_2018.pdf.

BRUFEM, Leilah; PRATES, Yara. O saber científico registrado e as práticas de mensuração da informação. **Ciência da Informação**, Brasília, DF, v. 34, n. 2, p. 9-25, maio/ago. 2005. Disponível em: <http://revista.ibict.br/ciinf/article/view/1086>. Acesso em: 17 ago. 2018.

CAPES (Brasil). História e Missão. In: CAPES (Brasil). **Missão**. [Brasília, DF]: CAPES, 2018. Disponível em: <http://www.capes.gov.br/historia-e-missao>. Acesso em: 18 jun. 2018.

CASTRO, Regina C. Figueiredo. Impacto da Internet no fluxo da comunicação científica em saúde. **Revista de Saúde Pública**, São Paulo, v. 40, n. esp., p. 57–63, 2006. Disponível em: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0034-89102006000400009&lng=pt&tlng=pt. Acesso em: 18 jun. 2018.

COSTA, Michelli. A visibilidade no Google Scholar dos repositórios digitais de acesso aberto brasileiros e portugueses. **Cadernos BAD**, [s. l.], v. jul/dez, n. 2, p. 41–53, 2014. Disponível em: <http://eprints.rclis.org/28967/>. Acesso em: 12 ago. 2018.

COSTA, Teresa. A b-on e a preservação digital de periódicos eletrônicos : O projeto Portico. In: 12º CONGRESSO NACIONAL DE BIBLIOTECÁRIOS, ARQUIVISTAS E DOCUMENTALISTAS, 12., 2015, Évora. **Anais** [...]. Évora: Universidade de Évora, 2015. documento não paginado. Disponível em:

<http://www.bad.pt/publicacoes/index.php/congressosbad>. Acesso em: 15 set. 2018.

FERREIRA, Elisabete *et al.* Digital object identifier (doi): o que é, para que serve, como se usa?. **AtoZ: Novas Práticas em Informação e Conhecimento**, Curitiba, v. 4, n. 1, p. 5-9, 2015. Disponível em:

<https://revistas.ufpr.br/atoz/article/view/42369/26039>. Acesso em: 10 out. 2018.

FONSECA, João José Saraiva. **Metodologia da pesquisa científica**. Fortaleza: Universidade Federal do Ceará, 2002. Apostila. Disponível em:

http://leg.ufpi.br/subsiteFiles/lapnex/arquivos/files/Apostila_-_METODOLOGIA_DA_PESQUISA%281%29.pdf.

GABRIEL, Martha. O impacto da Internet nos negócios. *Site*. 2015. Disponível em: <http://www.adflip.com.br/2015/06/08/o-impacto-da-internet-nos-negocios>. Acesso em: 29 set. 2018.

GERHARDT, Tatiana Engel; SILVEIRA, Denise Tolfo. **Métodos de pesquisa**. Porto Alegre: Editora da UFRGS, 2009.

GIL, Antônio Carlos. **Como elaborar projetos de pesquisa**. 4. ed. São Paulo: Atlas, 2002.

GOOGLE SCHOLAR. **About**. *Site*. 2018. Disponível em:

<https://scholar.google.com/intl/en/scholar/about.html>. Acesso em: 20 ago. 2018.

GOUVEIA, Fábio Castro. Novos caminhos e alternativas para a Webometria. **Em Questão**, Porto Alegre, v. 18, n. 3, p. 249–261, 2012. Disponível em:

<http://seer.ufrgs.br/EmQuestao/article/view/33270>.

GOUVEIA, Fábio Castro; LANG, Pamela. Da webometria à altmetria: uma jornada por uma ciência emergente. *In*: ALBAGLI, Sarita (org). **Fronteiras da ciência da informação**. Brasília, DF: IBICT, 2013. p. 172-195. Disponível em:

<http://repositorio.ibict.br/bitstream/123456789/453/1/Fronteiras%20da%20Ci%C3%Aancia%20da%20Informa%C3%A7%C3%A3o.pdf>. Acesso em: 16 ago. 2018.

GULKA, Juliana Aparecida; LUCAS, Elaine Rosangela de Oliveira. Presença digital em portais de periódicos : proposta de análise. **Em Questão**, Porto Alegre, v. 23, n. Edição especial 5 EBBC, p. 159–179, 2017. Disponível em:

<http://seer.ufrgs.br/EmQuestao/article/view/68085/40209>. Acesso em: 15 maio. 2018.

HARTMAN, Karen; MULLEN, Laura-Bowering. Google Scholar and academic libraries: an update. **New library world**, [s.l.], v. 109, n. 5-6, p. 211-222, 2008.

Disponível em: <http://eprints.rclis.org/13820/1/GSfinalupdate.pdf>. Acesso em: 11 set. 2018.

HERRERA, Gail. Google Scholar Users and User Behaviors: An Exploratory Study.

College & Research Libraries, Chicago, v. 72 n. 4 p. 316-330, jul 2011. Disponível em: <https://eric.ed.gov/?id=EJ1064846>. Acesso em: 12 maio 2018.

JACSÓ, Péter. As we may search - Comparison of major features of the Web of Science, Scopus, and Google Scholar citation-based and citation-enhanced databases. **Current science**, [s.l.], v. 89, n. 9, p. 1537-1547, 2005. Disponível em: https://www.researchgate.net/publication/234144770_As_we_may_search-Comparison_of_major_features_of_the_Web_of_Science_Scopus_and_Google_Scholar_citation-based_and_citation-enhanced_databases. Acesso em: 11 set.2018.

JAMALI, Hamid R.; ASADI, Saeid. Google and the scholar: the role of Google in scientists' information-seeking behaviour. **Online Information Review**, Bingley, v. 34, n. 2, p. 282–294, 2010. Disponível em: https://www.researchgate.net/publication/220207532_Google_and_the_scholar_The_role_of_Google_in_scientists%27_information-seeking_behaviour. Acesso em: 24 ago. 2018.

KOUSHA, Kayvan; THELWALL, Mike. Sources of Google Scholar citations outside the Science Citation Index: a comparison between four science disciplines. **Scientometrics**, Budapeste, v. 74, n. 2, p. 273-294, 2008. Disponível em: <https://link.springer.com/article/10.1007/s11192-008-0217-x>. Acesso em: 11 jun. 2018.

MACIAS-CHAPULA, Cesar. A. O papel da informetria e da cienciometria e sua perspectiva nacional e internacional. **Ciência da Informação**, Brasília, DF, v.27, n.2, p. 134-140, maio/ago. 1998. Disponível em: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-19651998000200005&lng=en&nrm=iso&tlng=pt. Acesso em: 7 set. 2018.

MALAGA, Ross A. Worst practices in search engine optimization. **Communications of the ACM**, Nova Iorque, v. 51 n. 12, p. 147-150. 2008. Disponível em: <https://dl.acm.org/citation.cfm?doid=1409360.1409388>. Acesso em: 15 maio 2018.

MARTÍN-MARTÍN, Alberto *et al.* **The counting house, measuring those who count**: Presence of Bibliometrics, Scientometrics, Informetrics, Webometrics and Altmetrics in Google Scholar Citations, ResearcherID, ResearchGate, Mendeley & Twitter. *In*: ECR Working Papers, 21., 2015, Granada. **Proceedings** [...]. Granada: 2015. Disponível em: <https://arxiv.org/abs/1602.02412>. Acesso em: 2 abr. 2018.

MEADOWS, Arthur Jack. **A comunicação científica**. Brasília, DF: Briquet de Lemos, 1999.

MUGNAINI, Rogério; CARVALHO, Telma; CAMPANATTI-OSTIZ, Heliane. Indicadores de produção científica: uma discussão conceitual. *In*: POBLACION, Dinah Aguiar; WITTER, Geraldina Porto; SILVA, José Fernando Modesto da. **Comunicação & produção científica**: contexto, indicadores e avaliação. São Paulo: Angellara, 2006. p.313-340.

NASCIMENTO, Andrea Gonçalves do; ODDONE, Nanci. Uso de altmetrics para avaliação de periódicos científicos brasileiros em Ciência da Informação. **Ciência da Informação em Revista**, Maceió, v. 2, n. 1, p. 3-12, jan./abr. 2015. Disponível em: <http://www.seer.ufal.br/index.php/cir/article/view/1745>. Acesso em: 23 mai. 2018.

NORONHA, Daisy Pires; MARICATO, João de Melo. Estudos métricos da informação: primeiras aproximações. **Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação**, Florianópolis, v. 13, n. esp., p. 116–128, 2008. Disponível em: <http://www.brapci.inf.br/index.php/article/view/0000005005/2d364e9cfdb8d8e923ed352f0733af72>. Acesso em: 18 jun. 2018.

ORDUÑA-MALEA, Enrique; DELGADO LÓPEZ-CÓZAR, Emilio. The dark side of open access in google and google scholar: The case of latin-american repositories. **Scientometrics**, Budapeste, v. 102, n. 1, p. 829–846, 2015. Disponível em: <https://doi.org/10.1007/s11192-014-1369-5>. Acesso em: 15 mar. 2018.

ORDUÑA-MALEA, Enrique; SERRANO-COBOS, Jorge; LLORET-ROMERO, Nuria. Las universidades públicas españolas en Google Scholar: presencia y evolución de su publicación académica web. **El Profesional de la Información**, [s. l.], v. 18, n. 5, p. 493–500, 2009. Disponível em: <https://recyt.fecyt.es/index.php/EPI/article/view/epi.2009.sep.02>. Acesso em: 15 set. 2018.

ORDUÑA-MALEA, Enrique; SERRANO-COBOS, Jorge; LLORET-ROMERO, Nuria. Spanish public universities in Google Scholar: presence, evolution and coverage of their scientific output. **El Profesional de la Información**, [s.l.], v. 18 n. 5, p. 493-500, 2009. Disponível em: <http://dx.doi.org/10.3145/epi.2009.sep.02>. Acesso em: 15 jun. 2018.

ORTEGA, Jose Luis *et al.* Maps of the academic web in the European Higher Education Area - An exploration of visual web indicators. **Scientometrics**, Budapeste, v. 74, n. 2, p. 295–308, 2008. Disponível em: <https://link.springer.com/content/pdf/10.1007%2Fs11192-008-0218-9.pdf>. Acesso em: 20 jun. 2018.

OTLET, Paul. Divisão e modos de expressão. In: OTLET, Paul. **Tratado de documentação: o livro sobre o livro teoria e prática**. Tradução de Taiguara Villela Aldabalde, *et al.* Brasília, DF: Briquet de Lemos / Livros, 2018. p. 15-27. Disponível em: https://archive.org/download/Otlet_tratado_de_documentao.pronto/otlet_tratado_de_documenta%C3%A7%C3%A3o.pronto.pdf. Acesso em: 9 set. 2018.

PARK, Han Woo; KIM, Chun-Sik; BARNETT, George. Socio-communicational structure among political actors on the web in South Korea: The dynamics of digital presence in cyberspace, **New Media & Society**, [s.l.], v. 6, n. 3, p. 403-423, 2004. Disponível em: <http://journals.sagepub.com/doi/abs/10.1177/1461444804042522?legid=spnms%3B6%2F3%2F403&patientinform-links=yes>. Acesso em: 18 set. 2018.

PRIEM, Jason; GROTH, Paul; TARABORELLI, Dario. The altmetrics collection. **PLOS One**, [s.l.], v. 7, n. 11, p. e48753, nov. 2012. Disponível em: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0048753>. Acesso em: 27 maio 2018.

PUBLIC KNOWLEDGE PROJECT. In: PUBLIC KNOWLEDGE PROJECT. **OJS 3.1 Learning**: A visual guide for opening journal systems. [s.l.]: PUBLIC KNOWLEDGE PROJEC, [s.d.]. Disponível em: <https://docs.pkp.sfu.ca/learning-ojs/en/>. Acesso em: 19 nov. 2018.

RICHARDSON, Roberto Jarry *et al.* **Pesquisa social**: métodos e técnicas. 3 ed. São Paulo: Atlas, 2015. Disponível em: <https://acervodigital.ssp.go.gov.br/pmgo/bitstream/123456789/355/8/Livro%20-%20Pesquisa%20Social%20-%20Métodos%20Quantitativos%20e%20Qualitativos%20-%20Capítulo%205.pdf>.

SEVERINO, Antônio Joaquim. **Metodologia do trabalho científico**. 23 ed. ed. São Paulo: Cortez, 2007.

SHINTAKU, Milton. Webometria e os periódicos científicos eletrônicos. **Informacao e Sociedade**: Estudos, João Pessoa, v. 27, n. 3, p. 267–275, 2017. Disponível em: <http://www.periodicos.ufpb.br/ojs/index.php/ies/article/view/34796/pdf>. Acesso em: 10 maio. 2018.

SMITH, Alastair. A tale of two web spaces: comparing *sites* using web impact factors. **Journal of Documentation**, [s.l.], v. 55, n. 5, p. 577-592, 1999.

STUMPF, Ida Regina Chitto. Reflexões sobre as revistas brasileiras. **Intexto**, Porto Alegre, v. 1, n. 3, p. 1–10, 1998. Disponível em: <http://www.seer.ufrgs.br/intexto/article/viewFile/3369/3953>. Acesso em: 17 jun. 2018.

TAGUE-SUTCKIFFE, J. An introduction to informetrics. **Information Processing & Management**, [s.l.], v. 28, n. 1, p. 1-3, 1992. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/030645739290087G>. Acesso em: 11 set. 2018.

TARGINO, Maria das Graças. Comunicação científica na sociedade tecnológica: periódicos eletrônicos em discussão. **Comunicação & Sociedade**, São Paulo, v. 3, n. 31, p. 71–98, 1999. Disponível em: <https://www.metodista.br/revistas/revistas-ims/index.php/CSO/article/view/7890>. Acesso em: 17 maio. 2018.

TARGINO, Maria das Graças. Comunicação científica: uma revisão de seus elementos básicos. **Informação & Sociedade**: Estudos, João Pessoa, v. 10, n. 2, p. 37–85, 2000. Disponível em: <http://www.brapci.inf.br/index.php/article/view/0000001182/6f78994cb4aa4e65990e1d764d794d5b>. Acesso em: 17 maio. 2018.

TORRES-SALINAS, Daniel; RUIZ-PÉREZ, Rafael; DELGADO-LÓPEZ-CÓZAR, Emilio. Google Scholar como herramienta para la evaluación científica. **El profesional de la información**, [s. l.], v. 18, n. 5, p. 501–5010, 2009. Disponível em: <https://doi.org/10.3145/epi.2009.sep.03%09>. Acesso em: 18 set. 2018.

VANTI, Nadia Aurora Peres. Os links e os estudos webométricos. **Ciência da Informação**, Brasília, DF, v. 34, n. 1996, p. 78–88, 2005. Disponível em: <http://revista.ibict.br/ciinf/article/view/1104>. Acesso em: 14 abr. 2018.

VERÍSSIMO, Joaquim Manuel Domingues. **As bibliotecas universitárias face ao desafio do Google Scholar: ameaça ou oportunidade?**. 2012. 104f. Dissertação (Mestrado em Informação, Comunicação e Novos Media) - Faculdade de Letras, Universidade de Coimbra, Coimbra, 2012. Disponível em: <http://hdl.handle.net/10316/18222>. Acesso em: 20 ago. 2018.

VITULLO, Nadia Aurora Vanti. **Links Hipertextuais na Comunicação Científica: análise webométrica dos sítios acadêmicos latino-americanos em Ciências Sociais**. 2007. 299 f. Tese (Doutorado em Comunicação e Informação) - Faculdade de Biblioteconomia e Comunicação, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2007. Disponível em: <https://lume.ufrgs.br/handle/10183/11175>. Acesso em: 18 ago. 2018.

WEITZEL, Simone da Rocha. Fluxo da informação científica. *In*: POBLACION, Dinah Aguiar; WITTER, Geraldina Porto; SILVA, José Fernando Modesto da (Org.). **Comunicação e produção científica: contexto, indicadores, avaliação**. São Paulo: Angellara Editora, 2006. p. 85-114.

ZIMAN, John Michael. O que é ciência? *In*: ZIMAN, John Michael. **Conhecimento público**. Belo Horizonte: Itatiaia, 1979. p. 17–25.

Apêndice – Formulário de coleta de dados

Revista	
Comando <i>site</i> : <i>URL</i>	
Número de artigos publicados	
Número de artigos recuperados no GS	
Taxa de presença no GS	
Qualis	
Perfil no GS	
Formatos de arquivos disponibilizados	
Preservação digital	
Padrão de metadados do <i>site</i>	
DOI do periódico	
Versão do OJS utilizada	