

UNIVERSIDADE FEDERAL DE SÃO CARLOS
CENTRO DE EDUCAÇÃO E CIÊNCIAS HUMANAS
DEPARTAMENTO DE CIÊNCIA DA INFORMAÇÃO

CESAR AUGUSTO LOPES

**MODELOS COLABORATIVOS DE INDEXAÇÃO SOCIAL
PARA REPOSITÓRIOS DIGITAIS**

SÃO CARLOS - SP

2021

CESAR AUGUSTO LOPES

**MODELOS COLABORATIVOS DE INDEXAÇÃO SOCIAL
PARA REPOSITÓRIOS DIGITAIS**

Trabalho de Conclusão de Curso apresentado ao Departamento de Ciência da Informação como requisito parcial para a obtenção do título de Bacharel em Biblioteconomia e Ciência da Informação pela Universidade Federal de São Carlos.

Orientadora: Profa. Dra. Paula Regina Dal'Evedove.

SÃO CARLOS - SP

2021

Lopes, Cesar Augusto

Modelos colaborativos de indexação social para repositórios digitais / Cesar Augusto Lopes -- 2021. 70f.

TCC (Graduação) - Universidade Federal de São Carlos, campus São Carlos, São Carlos

Orientador (a): Paula Regina Dal'Evedove

Banca Examinadora: Paula Regina Dal'Evedove, Zaira Regina Zafalon, Joyce Mirella dos Anjos Viana

Bibliografia

1. Modelos Colaborativos. 2. Indexação social. 3. Repositórios digitais. I. Lopes, Cesar Augusto. II. Título.

Ficha catalográfica desenvolvida pela Secretaria Geral de Informática (SIn)

DADOS FORNECIDOS PELO AUTOR

Bibliotecário responsável: Ronildo Santos Prado - CRB/8 7325

CESAR AUGUSTO LOPES

Trabalho de Conclusão de Curso apresentado ao Departamento de Ciência da Informação como requisito parcial para a obtenção do título de Bacharel em Biblioteconomia e Ciência da Informação pela Universidade Federal de São Carlos.

Aprovado em: 03 de dezembro de 2021.

BANCA EXAMINADORA

Profa. Dra. Paula Regina Dal'Evedove
Universidade Federal de São Carlos (UFSCar)
Departamento de Ciência da Informação (DCI)

Profa. Dra. Zaira Regina Zafalon
Universidade Federal de São Carlos - UFSCar
Departamento de Ciência da Informação - DCI

Ma. Joyce Mirella dos Anjos Viana
Mestre em Ciência da Informação (PPGCI/UFSCar)
Pesquisadora do Instituto Brasileiro de Informação em Ciência e
Tecnologia (IBICT)

AGRADECIMENTOS

Às pessoas que mesmo com gestos despretensiosos no decorrer dessa jornada fizeram algo decisivo para que seu desenlace fosse possível.

À **minha avó Ilza**, que mesmo sem compreender muito a lógica das minhas necessidades ou escolhas, nunca me deixou desagasalhado. Me deu condições para a permanência nesta cidade, a “grande” “capital da tecnologia”.

Ao meu colega André, que mesmo sem questionar muito a complexidade da minha pessoa, me resgatou um dia de uma potencial desistência. Sem ele eu não sei se teria retornado ou como eu teria resolvido isso sozinho.

Aos colegas de sala Evaldo e José, que algumas vezes me ajudaram a transigir com a barreira da inabilidade social, do pânico de me sentir excluído e perdido. Realmente não foi uma fase fácil.

Aos meus pais, que mesmo de longe me fizeram sentir protegido. E ainda tentam esclarecer comigo o problema que me ocorre.

Por fim, à minha professora Dra. Paula, que me orientou e com maestria/objetividade ensinou, no fim, como se deixa um relatório, na medida do possível, dentro dos conformes. Para mim, parecia inatingível.

A todos, obrigado por abrandarem, de alguma forma, a aflitiva realidade daquela frase do Facebook “tudo passa, nem que seja por cima de você, mas passa” (autor desconhecido).

RESUMO

Modelos colaborativos de indexação social visam uma indexação democrática que segue as premissas da inteligência coletiva. No contexto da Ciência Aberta e do movimento de acesso aberto, os repositórios digitais se enquadram como importantes sistemas que podem incorporar esses modelos. Contudo, questiona-se: Quais modelos colaborativos de indexação social para repositórios digitais foram delineados e estão divulgados na literatura especializada? Nesta perspectiva, a presente pesquisa possui como objetivo geral investigar as estratégias de inovação e tecnologia para a indexação social em sistemas híbridos de informação, a partir da investigação dos modelos colaborativos de indexação social para repositórios digitais, tendo como objetivos específicos: a) contextualizar Ciência Aberta, repositórios digitais e representação colaborativa da informação em ambientes digitais pelo viés da Ciência da Informação; b) delinear as principais perspectivas de estudo e aplicação de modelos colaborativos de indexação social como recursos estratégicos e de inovação em repositórios digitais; e c) apresentar os principais avanços e perspectivas futuras das pesquisas dedicadas aos modelos colaborativos de indexação social para repositórios digitais. Para tanto, conduziu-se estudo exploratório e descritivo conduzido a partir de pesquisa bibliográfica interdisciplinar, direcionada à literatura científica da Ciência da Computação e Ciência da Informação. Os resultados mostram que diferentes estudos defendem a hibridização dos vocabulários controlados e da folksonomia para a representação da informação em repositórios digitais e apresentam propostas de modelos colaborativos para indexação. Frente aos resultados, conclui-se sobre a necessidade de aperfeiçoamento das metodologias e ferramentas que consideram a integração da participação dos usuários no processo de indexação da informação.

Palavras-chave: Modelos Colaborativos. Indexação social. Repositórios digitais. Ciência da Informação.

ABSTRACT

Collaborative social indexing models aim at democratic indexing that follows the premises of collective intelligence. In the context of Open Science and the open access movement, digital repositories stand as important systems that can incorporate these models. However, the question is: What collaborative social indexing models for digital repositories have been outlined and are published in the specialized literature? In this perspective, this research has as a general objective to investigate the innovation and technology strategies for social indexing in hybrid information systems, from the investigation of collaborative models of social indexing for digital repositories, having as specific objectives: a) contextualize Science Open, digital repositories and collaborative representation of information in digital environments through the Information Science bias; b) outline the main perspectives for the study and application of collaborative social indexing models as strategic and innovative resources in digital repositories; and c) present the main advances and future perspectives of research dedicated to collaborative social indexing models for digital repositories. Therefore, an exploratory and descriptive study was conducted based on an interdisciplinary bibliographical research, directed to the scientific literature of Computer Science and Information Science. The results show that different studies defend the hybridization of controlled vocabularies and folksonomy for the representation of information in digital repositories and present proposals for collaborative models for indexing. Based on the results, it is concluded that there is a need to improve methodologies and tools that consider the integration of user participation in the information indexing process.

Keywords: Collaborative Models. Social indexing. Digital repositories. Information Science.

LISTA DE ILUSTRAÇÕES

Figura 1 – Modelo colaborativo de indexação social: conceituação.

Figura 2 – Níveis de controle de vocabulário

Figura 3 – Representação iterativa: dinâmica de funcionamento

Figura 4 – Esboço de categorização facetada: do tipo *menu dropdown*

Figura 5 – Modelo colaborativo baseada em *tags* categorizadas: dinâmica de estruturas e funcionamento

LISTA DE QUADROS

Quadro 1 – Pesquisadores e estudos sobre modelos colaborativos de indexação social.

Quadro 2 - Protótipos que propõem uma etiquetagem facetada.

Quadro 3 - Protótipos com aperfeiçoamento da interface.

Quadro 4 – Elementos identificados nos modelos colaborativos de indexação social analisados.

LISTA DE ABREVIATURAS E SIGLAS

BDTD	Biblioteca Digital Brasileira de Teses e Dissertações
BRAPCI	Base de Dados Referenciais de Artigos de Periódicos em Ciência da Informação
CDD	Classificação Decimal de Dewey
CDU	Classificação Decimal Universal
CNPq	Conselho Nacional de Desenvolvimento Científico e Tecnológico
CSF	Collaboratively built Structured Folksonomy
E-LIS	E-prints in Library and Information Science
IBICT	Instituto Brasileiro de Informação em Ciência e Tecnologia
NDLTD	Networked Digital Library of Theses and Dissertations
NISO	National Information Standards Organization
OAI	Open Archives Initiative
OAI-PMH	Open Archives Initiative Protocol for Metadata Harvesting
PIBITI	Programa Institucional de Bolsas de Iniciação em Desenvolvimento Tecnológico e Inovação
PMEST	Personality, Matter, Energy, Space and Time
RUP	Rational Unified Process
SciELO	Scientific Electronic Library Online
UFMG	Universidade Federal de Minas Gerais
UML	Unified Modeling Language
W3C	World Wide Web Consortium
XML	Extensible Markup Language

SUMÁRIO

1 INTRODUÇÃO	12
2 FUNDAMENTAÇÃO TEÓRICA	16
2.1 Ciência Aberta e o movimento de acesso aberto.....	16
2.2 Repositórios digitais.....	19
2.3 Representação colaborativa.....	22
3 MODELOS COLABORATIVOS DE INDEXAÇÃO SOCIAL PARA REPOSITÓRIOS DIGITAIS.....	29
4 AVANÇOS DA TEMÁTICA NA CIÊNCIA DA INFORMAÇÃO BRASILEIRA.....	51
5 CONSIDERAÇÕES FINAIS	56
REFERÊNCIAS	58
APÊNDICE A - LISTA DE REFERÊNCIAS DOS ESTUDOS IDENTIFICADOS	

1 INTRODUÇÃO

Com o avanço tecnológico digital, a informação se configurou como ferramenta de fácil acesso e essencial para o desenvolvimento pessoal e coletivo. A quantidade nunca antes vista de informação produzida, em diversos formatos e suportes, clama por mecanismos capazes de representá-la e recuperá-la de forma mais eficiente. Com esse objetivo, novas discussões, produtos e serviços que auxiliam os múltiplos usuários/atores no processo de representação de recursos informacionais avançam na Ciência da Informação.

Dentre as novas formas de pensar a representação da informação, os modelos colaborativos de indexação social visam uma indexação democrática que segue as premissas da inteligência coletiva. Considerando que a indexação é parte importante de um sistema de informação, por condicionar os resultados de uma estratégia de busca, esses modelos figuram como propostas metodológicas que consideram a folksonomia para o preenchimento dos metadados na estrutura de sistemas de recuperação da informação de caráter híbrido. Por sistemas híbridos, podemos considerar os repositórios digitais e a prática do autoarquivamento, em que bibliotecários e autores atuam no preenchimento dos metadados.

Diante das intempéries emergidas com as novas perspectivas de indexação social dos recursos informacionais, como o descontrole terminológico advindos da autonomia dos usuários no referido processo, esta pesquisa se fundamenta na seguinte pergunta de investigação: Quais modelos colaborativos de indexação social para repositórios digitais foram delineados e estão divulgados na literatura especializada? De modo complementar, essa questão central de pesquisa é amparada por outras duas questões, em particular, a saber: Qual a aplicabilidade desses modelos colaborativos de indexação social na resolução dos atuais desafios da representação da informação? Em quais teorias e perspectivas esses modelos se apoiam?

Face ao problema apresentado, parte-se da hipótese de que os modelos colaborativos de indexação social no contexto dos repositórios digitais contribuem com o cenário de grande produção, compartilhamento e reuso de dados e informações.

A fim de oferecer contribuições acerca das tendências investigativas sobre essas questões, a presente pesquisa possui como objetivo geral investigar as estratégias de inovação e tecnologia para a indexação social em sistemas híbridos de informação, a partir da investigação dos modelos colaborativos de indexação social para repositórios digitais.

Neste sentido, foram estabelecidos os seguintes objetivos específicos:

1. Contextualizar Ciência Aberta, repositórios digitais e representação colaborativa da informação em ambientes digitais pelo viés da Ciência da Informação;
2. Delinear as principais perspectivas de estudo e aplicação de modelos colaborativos de indexação social como recursos estratégicos e de inovação em repositórios digitais; e
3. Apresentar os principais avanços e perspectivas futuras das pesquisas dedicadas aos modelos colaborativos de indexação social para repositórios digitais.

Considera-se que esta temática colabora para o avanço das discussões na Ciência da Informação brasileira, campo que nos últimos anos se direciona para as problemáticas decorrentes do ambiente digital e suas especificidades, sobretudo, no que tange ao tratamento, acesso e disponibilidade de dados e informações para a sociedade atual – complexa e atenta aos avanços e usos tecnológicos diversos. Além disso, o seu desenvolvimento está amparado nas pesquisas e preocupações teóricas e metodológicas conduzidas no âmbito do Grupo de Pesquisa “Organização do Conhecimento e Humanidades Digitais”¹, o qual congrega estudantes e

¹ Disponível em: <http://dgp.cnpq.br/dgp/espelhogrupo/340090>

pesquisadores preocupados com a representação de recursos informacionais em tempos de humanidades e tecnologias digitais.

Para tanto, o trabalho foi desenvolvido mediante estudo exploratório e descritivo conduzido a partir de pesquisa bibliográfica interdisciplinar, direcionada à literatura científica de Ciência da Informação e Ciência da Computação. As buscas por materiais digitais foram realizadas no período de janeiro a abril de 2021, durante a condução da iniciação científica pelo pesquisador no período de setembro de 2020 a agosto de 2021, com bolsa concedida pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), a partir do Programa Institucional de Bolsas de Iniciação em Desenvolvimento Tecnológico e Inovação (PIBITI).

A pesquisa bibliográfica viabilizou a fundamentação teórica desta pesquisa, a qual favoreceu uma contextualização geral dos principais tópicos contemplados; e amparou a busca por estudos dedicados aos modelos colaborativos de indexação social.

A descrição detalhada as etapas conduzidas para o desenvolvimento desta pesquisa, são descritas a seguir:

1ª Etapa: Levantamento bibliográfico e seleção do material obtido. Recorreu-se às informações pertinentes ao tema proposto a fim de criar embasamento teórico ao estudo. O levantamento bibliográfico foi realizado em nível nacional e internacional em fontes bibliográficas, levando em consideração materiais impressos e digitais;

2ª Etapa: Leitura, interpretação, análise e sistematização das informações. Buscou-se oportunizar o desenvolvimento da base teórica para a discussão dos diferentes pontos de vista identificados na literatura sobre o tema proposto, possibilitando assim, criar subsídios para a elaboração da pesquisa;

3ª Etapa: Análise e delineamento das características fundamentais extraídas da literatura. Identificam-se os fundamentos teóricos e

metodológicos que servem de aporte à representação colaborativa da informação em repositórios digitais, a partir dos modelos existentes;

4ª Etapa: Sistematização, análise e discussão dos dados identificados; e

6ª Etapa: Elaboração da redação final da pesquisa para divulgação à comunidade científica e realização da Banca de Defesa.

Após esta seção introdutória, a pesquisa se divide em outras quatro seções. Primeiramente são apresentadas questões relacionadas à Ciência Aberta e ao movimento de acesso aberto, aos repositórios digitais como um sistema híbrido de informação, além da representação colaborativa da informação. Todos esses tópicos são explorados na seção 2 intitulada "FUNDAMENTAÇÃO TEÓRICA", com especial atenção às pesquisas publicadas na Ciência da Informação. Na sequência, a seção 3 intitulada "MODELOS COLABORATIVOS DE INDEXAÇÃO SOCIAL PARA REPOSITÓRIOS DIGITAIS" explora os modelos colaborativos de indexação social para repositórios digitais identificados na literatura científica dos campos de Ciência da Computação e Ciência da Informação. Em seguida, a seção 4 intitulada "AVANÇOS DA TEMÁTICA NA CIÊNCIA DA INFORMAÇÃO BRASILEIRA" apresenta as perspectivas de avanço da temática na Ciência da Informação brasileira. Ao final, a seção 5 "CONSIDERAÇÕES FINAIS" expõe as conclusões advindas com a pesquisa, tendo como compromisso principal apontar sugestões para o avanço dos estudos informacionais dedicados aos modelos colaborativos de indexação social para o ambiente de repositórios digitais.

2 FUNDAMENTAÇÃO TEÓRICA

Esta seção contempla as principais temáticas trabalhadas pela pesquisa. Inicialmente, realiza-se uma breve exposição da Ciência Aberta e o movimento de acesso aberto nos últimos tempos. Na sequência, dedica-se a atenção para os repositórios digitais como os principais representantes do referido movimento, sendo esta exposição seguida por questões próprias da representação colaborativa em ambientes digitais.

2.1 Ciência Aberta e o movimento de acesso aberto

Altos custos para publicação e restrições para o acesso a materiais científicos publicados em periódicos pagos têm levado a comunidade científica a repensar sobre o fluxo da informação. Em um mundo globalizado e capitalista neoliberal alguns movimentos reacionários refletem uma nova forma de encará-la, especialmente, de maneira democrática, acessível, transparente e reprocessável.

A Iniciativa dos Arquivos Abertos (“Open Archives Initiative” - OAI) é considerada um movimento pioneiro no sentido de Ciência Aberta. Originada da Convenção de Santa Fé (1999), resultou na formalização do “Open Archives Initiative Protocol for Metadata Harvesting” (OAI-PMH). Segundo Weitzel (2006), a OAI se caracteriza pelo possível autodepósito² (*upload*) da produção científica pelos próprios autores em múltiplos formatos de texto e a descrição padronizada dos metadados³; pelo acesso livre (pesquisa e *download*) e pelo emprego do protocolo OAI-PMH. Dessa

² Usualmente conhecido por autoarquivamento. Corresponde ao “[...] arquivamento efetuado pelos próprios pesquisadores, de suas respectivas produções científicas, em repositórios digitais de acesso aberto” (FERRARI; PIRES, 2014, p. 24).

³ Conforme a Organização Nacional de Padrões de Informação (NISO), “metadados é a informação estruturada que descreve, explica, localiza, ou possibilita que um recurso informacional seja fácil de recuperar, usar ou gerenciar” (NISO, 2004, p.1).

forma, o OAI-PMH promove normas de interoperabilidade⁴ entre diferentes arquivos aperfeiçoando a recuperação e o compartilhamento de conteúdo científico.

Com perspectiva similar, a Declaração de Budapeste (2002), visou a provocação de mudanças nos modelos vigentes de armazenamento, disseminação, visibilidade e acesso às pesquisas científicas, se voltando à situação de oligopólio⁵ das editoras. Na reunião que deu origem à referida declaração foram ainda definidas duas estratégias para viabilizar o acesso à produção científica: a instauração da Via Verde (*Green Road*), que visou a criação de repositórios temáticos/disciplinares, repositórios institucionais e repositórios de dados científicos para o depósito, organização e disseminação de publicações científicas e outros recursos informacionais, aderindo ao autoarquivamento e a Via Dourada (*Golden Road*), que visou a criação de revistas de acesso aberto na *web*. Na mesma perspectiva, surgiram ainda as Declarações de Bethesda⁶, de 2003 e a Declaração de Berlim⁷, do mesmo ano, advindas das preocupações de áreas do conhecimento específicas.

Segundo Silva e Silveira (2019, p. 2), "a Ciência Aberta (*Open Science*) é um movimento que incentiva a transparência da pesquisa científica desde a concepção da investigação até o uso de *softwares* abertos", que amplificam a divulgação. Dessa forma, o movimento pela Ciência Aberta também coincide com um conjunto de inovações relativamente recentes, associadas às novas formas de colaboração e culturas de uso em ambientes digitais, mais especificamente à disseminação da chamada cultura livre digital (ALBAGLI; CLINIO; RAYCHTOCK, 2014). Destaca-se que, do final do século XX aos dias de

⁴ Interoperabilidade é definida como "[...] a capacidade de múltiplos sistemas com diferentes plataformas de *hardware* e *software*, estruturas de dados e interfaces de trocar dados com perda mínima de conteúdo e funcionalidade." (NISO, 2004, p. 2).

⁵ Situação de mercado em que poucas empresas detêm o controle da maior parcela do mercado.

⁶ Área biomédica.

⁷ Área de Humanidades.

hoje, a magnitude da produção informacional tomou proporções tão gigantescas que o gerenciamento e o tratamento da informação muitas vezes só se tornam possíveis com o auxílio de ferramentas, métodos e *softwares* especializados.

Esse cenário de quebra de paradigma atual demanda expertise do cientista da informação, visto que, como pontuado por Hey (2010), a incapacidade de aproveitar as ferramentas sofisticadas de computadores que manipulem dados poderá impedir até mesmo o perito mais altamente treinado a descobrir os *insights* que agora estão começando a entrar em foco.

A Ciência Aberta, conforme Albagli (2014), pode ainda ser entendida como um termo guarda-chuva que permite múltiplas interpretações, por englobar práticas e abordagens distintas, em que seu exercício representa um novo *modus operandi* para se projetar, executar e comunicar pesquisas, em grande parte do processo fazendo uso de ferramentas para gestão, armazenamento e acesso à informação.

Albagli (2015) elucida que no desenvolvimento da Ciência Aberta devem ser consideradas, além dos aspectos técnicos e tecnológicos, as questões de cunho ético, legal, cultural, político e institucional, que podem interferir no caráter aberto ou proprietário dessas práticas. Esse ponto reforça e converge com o fato de que a ciência moderna já experimenta o “quarto paradigma científico”, voltado para a valorização do compartilhamento dos dados de pesquisa (HEY; TANSLEY; TOLLE, 2009), também implicando em um cenário de recorrente busca por confluência e aprimoramento de teorias, métodos, modelos, práticas e funcionalidades diferentes e a serem revalidadas.

O emprego dos princípios e das ferramentas voltadas para a Ciência Aberta dá permissão a vantagens que se relacionam em cadeia: ocorre o aumento da visibilidade da pesquisa; do nível de reprodutibilidade, conferindo transparência às metodologias e é concebido o acesso aos resultados de forma mais completa, ocasionando a eficiência do investimento público e privado em atividades científicas que aceleram o

processo de inovação. Essa nova forma de comunicar ciência visa trazer maior produtividade ao empreendimento científico, recorrendo ao uso intensivo de recursos tecnológicos que permitem esta colaboração *online*, à distância, em tempo real e cada vez mais aberto à ampla colaboração.

Nesse panorama, as Humanidades Digitais representam a pertinente ligação entre as tecnologias digitais e a investigação em Humanidades. Destarte, floresce um grande eixo de atuação para os profissionais da Computação e da Ciência da Informação, conjuntamente. As Humanidades Digitais permitem reflexões e práticas inovadoras que contribuem com a busca por produtos e serviços que suprem as necessidades de sujeitos e instituições informacionais. O papel da Ciência da Informação, em específico, tem sido prover meios para o acesso às informações relevantes, desenvolvendo ainda estudos que agregam aperfeiçoamento à perspectiva supracitada.

Na concepção da Ciência Aberta, os repositórios digitais têm sido apontados como os importantes sistemas de informação, que tornam a informação mais gerenciável e que possibilitam a livre transferência da informação científica e tecnológica nas respectivas áreas do conhecimento.

2.2 Repositórios digitais

Segundo o Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) (2012), definem-se “repositórios digitais” como bases de dados *online* que reúnem de maneira organizada a produção científica de uma instituição ou área temática. No geral, repositórios digitais surgiram em prol de uma ciência democrática, se tornando ferramentas que armazenam, organizam, recuperam e disseminam informação técnico-científica.

Repositórios digitais, tradicionalmente, são divididos entre [1] “institucionais”, quando lidam com a produção científica de determinada

instituição e [2] “temáticos” ou “disciplinares”, quando lidam com a produção científica de determinada área do conhecimento, não necessariamente considerando limites institucionais. De forma geral, podem disponibilizar três modalidades de acesso, sendo: [1] “acesso aberto”, quando ocorre acesso ao recurso digital de maneira integral, direta e imediata; [2] “acesso restrito”, quando não permite acesso imediato ao recurso digital de forma direta, imediata e integral; [3] “acesso fechado”, quando não permite acesso ao recurso digital, tampouco aos seus metadados (IBICT, 2012). Os dois últimos não se configuram dentro da Ciência Aberta.

No Brasil, o IBICT foi o órgão governamental que liderou as articulações para a implantação do Movimento de Acesso Aberto, realizando em 2009 duas iniciativas importantes: a providência de um projeto-piloto que originou alguns repositórios institucionais nacionais e o “edital IBICT - FINEP/PCAL/XDBD Nº 002 2009”, que contemplou algumas instituições com *kits* tecnológicos para implementação do Dspace⁸ e de seus respectivos repositórios, incluindo treinamento de pessoal e suporte técnico e informacional. Essa estratégia conferiu a essas instituições maior visibilidade (VEIGA; MACENA, 2015).

Como reflexo desse incentivo, segundo estudos de Xavier (2019), desde a instituição da Via Verde no Brasil, a partir de 2011 o número de repositórios digitais duplicou; no final de 2015, seu crescimento foi de quase 218% e somente nos anos seguintes o aumento nos números se deu de forma moderada.

Na perspectiva de avanço da ciência, a disponibilização dos dados levantados e produzidos pelas pesquisas é uma das condições para a realização de uma ciência efetivamente aberta. Nos últimos anos, a criação de repositórios denominados repositórios científicos ou de dados de pesquisa⁹ vem sendo um dos caminhos para a preservação de dados

⁸ Pacote de *software* de código aberto, normalmente usado para criar repositórios de acesso aberto.

⁹ Denominados também “repositórios de dados científicos” ou “repositórios de dados de investigação abertos”.

gerados pela pesquisa científica. A partir dos anos 2000, as políticas que advogam pelo acesso aberto da comunicação da pesquisa científica, financiada com recursos públicos, têm impulsionado a difusão e o compartilhamento dos dados produzidos nos processos de investigação (CREASER, 2011).

Segundo Galvino, Rosa e Oliveira (2020, p. 36), nos dias atuais,

[...] duas questões importantes podem impulsionar universidades na criação de seus repositórios de dados, sendo: a crescente exigência das agências de fomento para que os pesquisadores adotem o planejamento da gestão dos dados de suas pesquisas e a valorização dos princípios da Ciência Aberta pela sociedade.

Dessa maneira, os repositórios de dados de pesquisa facilitam o fluxo informacional, levando em consideração as necessidades de informação de pesquisadores, de instituições e de sociedades científicas, proporcionando maior visibilidade aos resultados de pesquisas e possibilitando a preservação da memória científica institucional. Esses sistemas de informação armazenam, preservam e divulgam materiais em formato digital, que podem ser acessados por diversos provedores de serviços nacionais e internacionais.

Entretanto, grande parte dos repositórios brasileiros não possui política de preservação digital, ou quando possui, esta não está devidamente formalizada. Sayão (2007) sinaliza que é imprescindível que todo repositório tenha uma política de preservação digital formalizada, políticas estas, que abarquem os processos de curadoria, preservação, arquivamento e compartilhamento de coleções de dados de pesquisa e que garantam sustentabilidade. Na mesma direção, Weitzel e Mesqyusa (2015, p. 189) afirmam que,

[...] a política de preservação digital, quando formalizada e pública, ajuda a sedimentar um posicionamento transparente do repositório em relação à preservação digital, orientando os autores (depositantes) a respeito dos procedimentos adotados e reforçando a garantia de perpetuação e acesso.

No contexto brasileiro, conforme estudos de Viana (2020), no que tange à construção de políticas de preservação voltadas aos repositórios de dados científicos, parte das instituições estão desenvolvendo suas políticas, definindo suas estruturas de *hardware* e *software*, seus ambientes digitais e aprimorando suas redes. No entanto, outras ainda não desenvolveram consciência sobre a importância dessas políticas, ou estão lidando com elas apenas como exigência administrativa.

Ressalta-se que as políticas de preservação, no geral, devem ser de interesse dos pesquisadores, curadores de dados, comunidades científicas, órgãos de fomento à pesquisa, editores de periódicos científicos, gestores e executores de programas e projetos de dados de pesquisa, profissionais da computação e da Ciência da Informação, dentro das suas delimitações.

2.3 Representação colaborativa

Em definição ampla, a NISO (1997) afirma que indexação cobre qualquer sistema ou procedimento em que a seleção e a organização de palavras-chave exigem decisões humanas, podendo corresponder ao ato de criação de índices (de autor, de título, de assunto ou de publicações) ou ainda ao ato de criação de catálogos ou banco de dados, em bibliotecas ou centros de informação.

De forma mais restrita, outros pesquisadores, como Robredo e Cunha (1994), Lancaster (2010) e Langridge (2006), concordam que a indexação consiste em uma atividade integrante do tratamento temático da informação, parte de um conjunto de procedimentos que objetivam expressar/representar o conteúdo temático de itens informacionais através de linguagens de indexação ou documentárias¹⁰, visando a posterior recuperação. Essas linguagens têm a finalidade de padronização

¹⁰ Sinônimos para "linguagens de indexação" e "linguagens documentárias": "linguagem documental", "linguagem formal", "linguagem de informação", "linguagem controlada" (GUEDES, 2010), "linguagem artificial" (BORGES, 2008).

terminológica de determinada área do conhecimento que confere coesão e consistência ao processo de representação da informação.

“Vocabulário controlado” pode ser definido como um conjunto ou lista de termos autorizados, expressos a partir de algum instrumento específico, de forma estruturada ou não-estruturada, podendo ou não dispor de hierarquias ou polihierarquias entre si. Ontologias¹¹, tesauros¹², taxonomias¹³, listas de cabeçalho de assunto¹⁴ e listas de autoridades¹⁵ são exemplos de vocabulários controlados. Em um sistema de recuperação de informação, os termos supracitados são pontos de acesso para os documentos, haja vista que é possível que se recupere esses objetos informacionais com a submissão no campo de busca das palavras-chave que foram empregadas no ato da sua representação ou mesmo por meio da combinação delas.

Os metadados produzidos podem ser empregados para representar um recurso em sua totalidade, em uma de suas partes ou, ainda, em suas diversas formas de apresentação/formatos. Nessa perspectiva, “os objetos informacionais devem passar por um conjunto de tarefas de

¹¹ Conforme Jacob (2003, p. 19) “ontologias são categorias de coisas que existem ou podem existir em um determinado domínio particular, produzindo um catálogo que apresenta as relações entre os tipos e até os subtipos do domínio, provendo um entendimento comum e compartilhado do conhecimento de um domínio que pode ser comunicado entre pessoas e programas de aplicação”.

¹² Conforme Cavalcanti (1978, p. 27), tesouro é definido como “uma lista estruturada de termos associados empregada por analistas de informação e indexadores, para descrever um documento com a desejada especificidade, a nível de entrada, para permitir aos pesquisadores a recuperação da informação que procuram”.

¹³ Conforme Edols (2001), taxonomias são considerados esquemas que classificam coisas – organismos vivos, produtos, livros – em uma série de grupos hierárquicos para serem mais fáceis de identificar, estudar e localizar.

¹⁴ Conforme Carlan (2010, p. 36), cabeçalho de assunto é definido como um “conjunto de termos controlados que representam o assunto de um item de uma coleção. Pode ser extensivo e cobrir vários assuntos, embora sua estrutura seja, geralmente, muito superficial e limitada”.

¹⁵ Conforme Carlan (2010, p. 35), lista de autoridades é definida como “lista de termos usada para controlar a variedade de nomes para entidades. [...] geralmente não inclui uma organização ou estrutura complexa. A apresentação pode ser alfabética ou organizada por um esquema de classificação superficial. Por exemplo: nomes de países, indivíduos ou instituições”.

armazenamento, organização e tratamento para que seja possível a sua recuperação, acesso e uso em sistemas de recuperação da informação” (SANTOS; CORRÊA, 2017, p. 4).

Desde a “web 2.0” – já se consideram uma “web. 3.0” e uma “web 4.0” - o massivo fluxo de informação clama por mecanismos capazes de organizá-la e promover a recuperação de forma rápida e precisa. Esse fato acarreta novas alternativas de organização e representação da informação, consideradas mais adequadas ao contexto, apoiadas na filosofia colaborativa e que não se baseiam em regras de indexação com controle de vocabulário na descrição dos recursos.

Brooks (2001) menciona que o modelo clássico de representação da informação pode não ser mais aplicável e econômico no ambiente digital. Driblando a complexidade informacional desses ambientes, o emprego de etiquetas (*tags*) que condicionam melhor representação dos recursos, de forma colaborativa e interativa entre usuários, contribui com o cenário da grande produção, compartilhamento e reuso de dados.

Dessa maneira, a descrição dos objetos informacionais acaba por não ficar apenas a cargo dos gestores de unidades de informação, mas também de outros atores, com a realização do processo de etiquetagem¹⁶. Conforme Guedes e Dias (2010, p. 48), “antes mesmo da popularização da modalidade colaborativa de representação da informação, autores que trabalham com formas de indexação alternativas já estudavam o assunto”. Brown et al. (1996) chamaram a atenção para a necessidade de um tratamento democrático da indexação, em que os usuários acrescentariam aos registros termos de sua própria escolha, quando isso fosse necessário e apropriado.

A indexação de maneira colaborativa contribui para a popularização de novas perspectivas em classificação de recursos digitais, e significam

¹⁶ Em inglês, destacam-se outros possíveis termos para se nomear a etiquetagem na Web: “*Tagging*”; “*Tagging Systems*”; “*Social Tagging*”; “*Social Tagging Systems*”; “*Collaborative Tagging*”; “*Collaborative Tagging Systems*”; “*Social Classification*”; “*Bookmarking*”; “*Social Bookmarking*”; “*Social Bookmarks Manager*”; “*Social Ontologies*”; “*Dinamic Taxonomy*” (CATARINO; BAPTISTA, 2009).

uma alternativa inovadora para a organização da informação na medida em que propõem a “[...] integração da folksonomia em metadados de sistemas de recuperação da informação” (SANTOS; CORRÊA, 2017).

Assim sendo, a folksonomia é utilizada na categorização de conteúdos disponíveis na rede. Segundo Santos (2016), além de conferirem autonomia aos usuários do sistema e favorecerem o desenvolvimento de metadados, modelos existentes de colaboração na indexação contribuem com a qualidade da descrição dos recursos informacionais em ambientes digitais, bem como para o aperfeiçoamento dos processos, produtos e instrumentos (SANTOS, 2016). Além disso, a folksonomia ainda tem o poder de agregar a inteligência dos indivíduos, que podem colaborar na ativa produção de metadados semânticos em sistemas que utilizam modelos colaborativos.

Segundo Catarino (2009, p. 59),

[...] pressupõe-se que a folksonomia permite uma nova forma de organização de recursos da *web* que, naturalmente, poderá também ser adotada pelos repositórios institucionais, para que seus utilizadores tenham uma forma de organizar os recursos conforme suas necessidades.

De forma conceitual, *folksonomia* é a tradução do termo criado por Thomas Vander Wal (1966-), a partir da junção das palavras *folk* (povo) com *taxonomy* (taxonomia). Wal (2005) define folksonomia como “resultado de atribuição livre e pessoal de *tags* (etiquetas) a informações ou objetos (recursos na *web*), visando a sua recuperação”. Segundo Santos e Corrêa (2019), existem outras designações para se referir a folksonomia, tais como: “marcação social”, “etiquetagem colaborativa”, “representação colaborativa”, “classificação distribuída” e “indexação social”, no entanto, “folksonomia” consiste no termo amplamente aceito e mais popular na literatura científica.

Análises de Santos e Corrêa (2017) indicam que não ocorre na literatura um consenso conceitual e terminológico acerca da folksonomia, tendo em vista que em situações distintas ela é considerada como fenômeno; como inovação; como sistema; como classificação; como

vocabulário; como taxonomia; como método ou como resultado de um processo. “Folksonomia” pode ainda ser considerada uma antítese ao termo taxonomia, na medida em que, ao invés de utilizar uma forma hierárquica e centralizada de categorização de alguma coisa, o usuário escolhe palavras-chave (“tags”) para classificar a informação, de forma colaborativa. Esse processo diverge da taxonomia, onde coisas precisam ser encaixadas em classificações pré-estabelecidas.

Na perspectiva de arquitetura da informação de *websites*, Reis (2007) afirma que a folksonomia pode se manifestar de formas distintas em um sistema de informação no momento da indexação colaborativa, a saber: a) apresentando ao usuário instruções para escolha das *tags*; b) sugerindo descritores pré-definidos, sinônimos ou semelhantes já utilizados por outros usuários; c) aplicando corretores ortográficos; d) indicando o assunto ao qual a *tag* se refere. Essas formas podem operar juntas ou isoladamente.

No entanto, a utilização da folksonomia de forma isolada nos sistemas informacionais digitais, ainda que conveniente, pode conferir inconsistências no processo de descrição de recursos, como a falta de padronização de termos, considerando ausência de regras para o emprego de etiquetas. Essa possível circunstância pode impactar negativamente na qualidade da recuperação.

Brandt e Medeiros (2010, p. 120) evidenciam que:

[...] o vocabulário [na folksonomia] é totalmente livre, gerando os conhecidos problemas que costumam ser resolvidos por um vocabulário controlado. Entende-se que a folksonomia, comparada aos outros esquemas de representação do conhecimento, é gerada de forma inversa: primeiro se classificam os objetos informacionais, e, posteriormente, surge uma folksonomia. Já nas outras ferramentas como as taxonomias e os tesouros, os objetos informacionais são classificados somente quando elas já existem: um tesouro, por exemplo, é usado para a escolha dos termos que irão descrever o conteúdo de um documento.

Fazendo alusão aos vocabulários controlados, Aquino (2007) considera a folksonomia como uma espécie de vocabulário

“descontrolado” na medida em que trata de um mecanismo de representação, organização e recuperação de informações que não é feito por especialistas e que refletem um modo em que os próprios indivíduos que buscam informação na rede ficam livres para representá-la, organizá-la e recuperá-la, realizando ações com base no senso comum e na linguagem natural. Pondera-se ainda, que no caso do autodepósito, ao preencher os metadados de um recurso digital, o autor pode não conhecer as técnicas documentárias de tratamento da informação.

Ainda assim, se apoiando na inteligência coletiva, o “taggeamento colaborativo” vem ganhando cada vez mais espaço por ser uma forma de classificação que não utiliza taxonomias ou vocabulários pré-estabelecidos, em grande parte dos casos não há restrição hierárquica ou outras para definir *tags*.

Mecanismos têm sido sistematicamente empregados aos sistemas, como forma de assegurar maior precisão no processo de categorização e recuperação da informação, diminuindo a incidência de instabilidades operacionais comuns na representação colaborativa, como as sinalizadas nos estudos de Amstel (2007), Gouvêa; Loh (2007) e Catarino; Baptista (2009), a exemplo: a presença de descritores homógrafos; de erros ortográficos; de expressões regionais; de etiquetas em diferentes idiomas e com letras variavelmente maiúsculas e minúsculas e de etiquetas assimétricas (*tags* que não possuem relações entre si). Outras inconsistências também são apontadas nos estudos de Noruzi (2007), a saber: a discrepância entre de singulares e plurais, a existência de polissemia, emprego de sinonímia e desajuste na profundidade/especificidade da descrição.

Conforme Lancaster (2004), métodos colaborativos de indexação são viáveis em ambientes de biblioteca digital por garantirem que usuários ofereçam novos termos de indexação aos itens que consultam e esses termos fornecidos poderem ser armazenados em um novo campo de registro. Para este autor, certos tipos de materiais, como artigos de

periódicos, livros, ou relatórios técnicos, a indexação orientada pelo usuário pode impactar positivamente na recuperação.

Contudo, autores como Cintra (2002), enfatizam que as linguagens documentárias mais consistentes dispõem de um vocabulário que integra, de um lado, a linguagem de especialidade e, de outro, a linguagem natural e ainda, a linguagem dos autores dos documentos.

Por sua vez, Santarém Segundo (2011) considera este tipo de indexação um processo altamente importante para ser aplicado ao contexto dos repositórios digitais, visto que permite a construção de inteligência coletiva.

A partir do exposto, a próxima seção apresenta iniciativas sobre modelos colaborativos de indexação social para repositórios digitais, compreendidos pela literatura científica nacional e internacional, indo ao encontro do segundo objetivo específico desta pesquisa.

3 MODELOS COLABORATIVOS DE INDEXAÇÃO SOCIAL PARA REPOSITÓRIOS DIGITAIS

Estudos sobre a representação da informação e do conhecimento nunca foram tão significativos e pertinentes, uma vez que a sociedade da informação tem produzido quantidades nunca antes vistas de informação e de conhecimento, em diversos formatos e suportes. Nos ambientes digitais, como em repositórios digitais, o processo de análise de assunto deve primar pela a qualidade da recuperação. Espera-se que quanto mais assertiva for a conduta do indexador, melhor será a qualidade da indexação feita por ele e, como resultado, melhor será a qualidade da recuperação da informação.

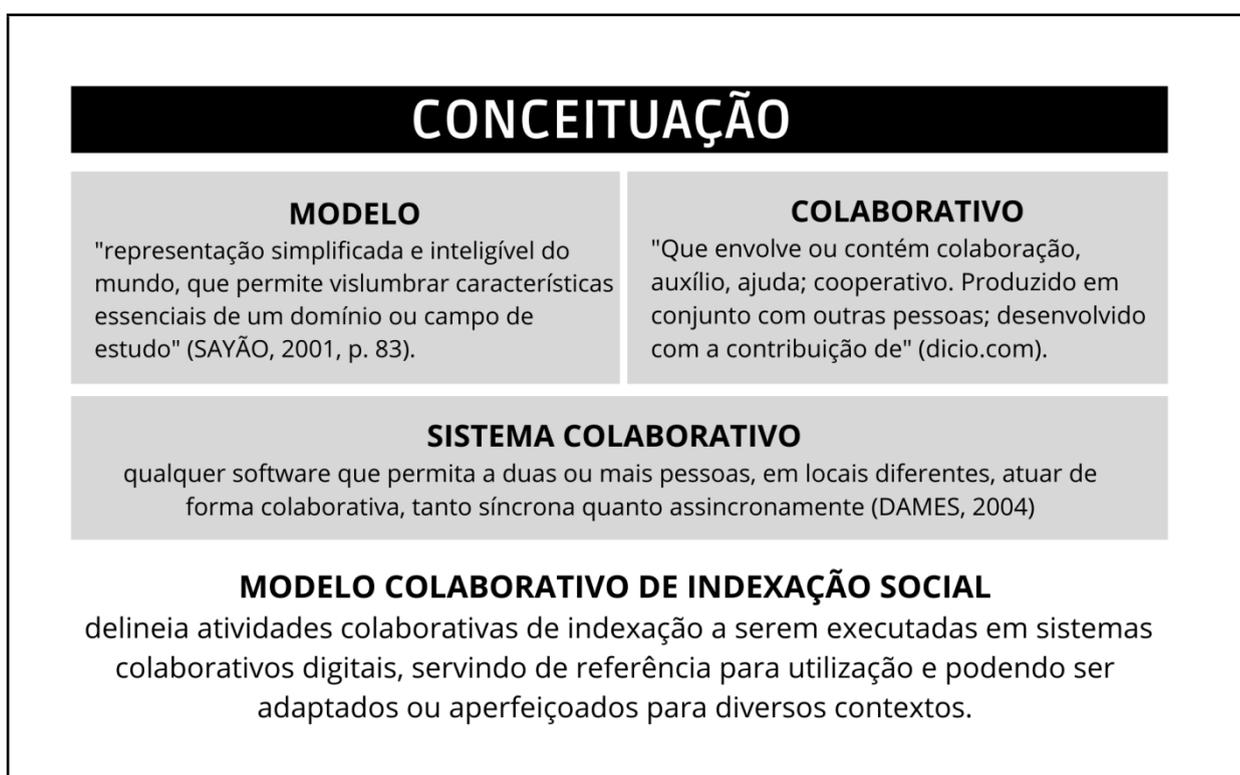
Segundo Fujita (1999), a indexação é reconhecida como parte mais importante de um sistema de informação por condicionar os resultados de uma estratégia de busca. Dessa forma, a indexação pode ser dividida em momentos distintos dentro do sistema, sendo: a "entrada", que corresponde ao tratamento temático da informação e a "saída", que corresponde à recuperação da informação, em que o usuário acessa uma parcela dos documentos disponíveis mediante uma estratégia de busca.

Gracioso (2010) documenta a necessidade de uma proposta metodológica para a criação de um sistema de recuperação que considere a folksonomia em sua estrutura e de um sistema eficiente que utilize algum modelo colaborativo no processo de indexação. Na mesma direção, Santos e Corrêa (2015) concordam sobre a escassez de metodologias que considerem o bom proveito dos metadados gerados pelos próprios usuários do sistema no momento da busca de conteúdos e que podem impactar positivamente na representação e na recuperação da informação em bibliotecas digitais.

De forma conceitual, Sayão (2001, p. 83) define "modelo" como "uma representação simplificada e inteligível do mundo, que permite vislumbrar características essenciais de um domínio ou campo de estudo". Por sua vez, "sistema colaborativo", de acordo com Dames (2004)

corresponde a qualquer *software* que permita a duas ou mais pessoas, em locais diferentes, atuar de forma colaborativa, tanto síncrona quanto assincronamente. Dessa forma, os modelos colaborativos de indexação social delineiam atividades colaborativas de indexação a serem executadas em sistemas colaborativos digitais, servindo de referência para utilização e podendo ser adaptados ou aperfeiçoados para diversos contextos (Figura 1).

Figura 1 – Modelo colaborativo de indexação social: conceituação.

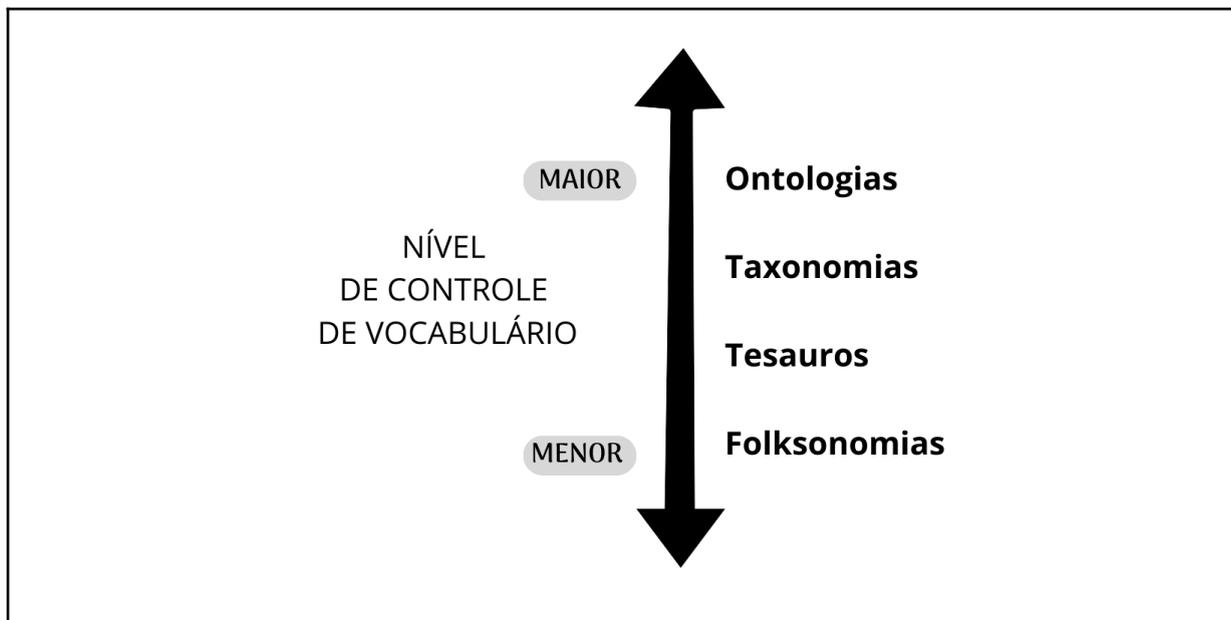


Fonte: Elaborado pelo autor.

A integração da folksonomia em metadados na representação temática de recursos digitais é indicada, muitas vezes, combinada com estruturas controladas de representação do conhecimento, como ontologias, tesouros e taxonomias, que também concebem diferentes níveis de controle de vocabulário entre si (Figura 2). Dessa forma, os modelos colaborativos de indexação em ambientes digitais, adiante apresentados, seguem as premissas da inteligência coletiva, se configurando como alternativas a serem incorporadas às estruturas de

representação do conhecimento utilizadas em ampla escala, objetivando maior adaptabilidade e coerência para a etapa de recuperação.

Figura 2 – Níveis de controle de vocabulário



Fonte: Elaborado pelo autor.

No ambiente digital, Levy (2003) defende que o aproveitamento da inteligência coletiva - um tipo de inteligência compartilhada que surge da colaboração de mais de um indivíduo em suas diversidades - é um dos princípios básicos da web 2.0. Dessa maneira, a inteligência coletiva é consolidada em um ambiente de participação ativa de pessoas, que resulta no "reconhecimento e enriquecimento mútuo". Para esse autor, o conceito de inteligência coletiva parte da ideia de que "uma inteligência está distribuída por toda parte [...], o saber está na humanidade, já que ninguém sabe tudo, porém todos sabem alguma coisa" (LEVY, 2003, p. 23).

No intuito de apresentar um panorama sobre a emergência de modelos colaborativos de indexação social em ambientes digitais (e que consideram a inteligência coletiva) dirigiu-se à literatura científica das áreas pertinentes. Assim, o Quadro 1 sistematiza cronologicamente as principais pesquisas voltadas à abordagem destes modelos.

A coleta dos dados apresentados foi executada mediante o exame de fontes impressas como livros e a buscas em periódicos científicos eletrônicos e em repositórios digitais. Foram considerados materiais publicados em língua portuguesa do Brasil e em língua inglesa e coletados no período de janeiro a abril de 2021, desprezando qualquer delimitação de data de publicação para a consideração de materiais.

Para pesquisa à nível nacional optou-se por utilizar a “Base de Dados Referenciais de Artigos de Periódicos em Ciência da Informação” (BRAPCI)¹⁷; a “Biblioteca Digital Brasileira de Teses e Dissertações” (BDTD)¹⁸, do IBICT; a Biblioteca Digital de Teses e Dissertações Universidade Federal de Minas Gerais (UFMG)¹⁹ e a biblioteca digital de periódicos “*Scientific Electronic Library Online*” (SciELO)²⁰. Para pesquisa à nível internacional foram utilizados o repositório de acesso livre “*E-prints in Library and Information Science*” (E-LIS)²¹, com documentos da área da Biblioteconomia e Ciência da Informação; a “*Networked Digital Library of Theses and Dissertations*” (NDLTD)²²; a ferramenta de busca especializada para documentos científicos “Scholar”²³, do Google e a “ACM Digital Library”²⁴, com publicações na área de Computação e Tecnologia da Informação.

No processo de busca nas referidas ferramentas, bases de dados e repositórios digitais foram submetidos nos campos assunto, título, resumo e palavras-chave os termos de busca: “representação colaborativa”, “folksonomia”, “indexação social”, “classificação social”, “etiquetagem colaborativa”, “indexação colaborativa” e “representação colaborativa da informação”.

A triagem dos materiais se deu mediante a leitura dos resumos, da introdução, dos métodos e das considerações finais das pesquisas

¹⁷ Disponível em: <https://www.brapci.inf.br/>

¹⁸ Disponível em: <http://bdtd.ibict.br>

¹⁹ Disponível em: <http://www.bibliotecadigital.ufmg.br>

²⁰ Disponível em: <http://www.scielo.org>

²¹ Disponível em: <http://eprints.rclis.or>

²² Disponível em: <http://www.ndltd.org>

²³ Disponível em: <http://scholar.google.com.br>

²⁴ Disponível em: <http://dl.acm.org>

recuperadas, de modo a identificar e considerar os materiais em que os modelos colaborativos de indexação social tenham sido objeto de estudo principal. Dispondo-se da análise qualitativa dos dados coletados, realizou-se a leitura integral do material e a elaboração de síntese teórica em torno dos principais aspectos abordados, dirigindo-se à sistematização dos dados e informações, como abaixo apresentada. O “Apêndice A” apresenta de forma listada as referências dos respectivos estudos.

Quadro 1 – Pesquisadores e estudos sobre modelos colaborativos de indexação social.

	Autores e Ano de publicação	Título	Produto
1	(DEVADASON, 2002)	<i>Faceted Indexing Based System for Organizing and Accessing Internet.</i> (Sistema baseado em indexação facetada para organização e acesso na internet).	Artigo científico (Ciência da Computação e Gestão da Informação)
2	(QUINTARELLI; RESMINI; ROSATI, 2006)	<i>FaceTag: Integrating Bottom-up and Top-down Classification</i> (FaceTag: Integração de classificação de baixo para cima e de cima para baixo)	Artigo científico (Ciência da Computação)
3	(DEVADASON, 2009)	<i>Faceted Indexing Based System for Organizing and Accessing Internet Resources</i> (Sistema baseado em indexação facetada para organizar e acessar recursos da Internet)	Artigo científico (Ciência da Computação e Gestão da Informação)
4	(LAI, P., 2009)	<i>Facette : using facets to improve tag-based bookmarking</i> (Facette: usando facetas para melhorar o <i>bookmarking</i> baseado em <i>tags</i>)	Dissertação de Mestrado (Engenharia Elétrica e Ciência da Computação)
5	(MALY, K; WU, H. ZUBAIR. M., 2009)	<i>Automated support for a collaborative system to organize a collection using facets</i> (Suporte automatizado para um sistema colaborativo para organizar uma coleção usando facetas)	Artigo científico (Ciência da Computação)
6	(SIGURBJÖRNSSON; ZWOL, 2010)	<i>Tagexplorer: faceted browsing of flickr photos</i> (Tagexplorer: navegação facetada de fotos do flickr)	Artigo científico (Ciência da Computação)
7	(SANTARÉM SEGUNDO, J. E., 2010)	Representação Iterativa: um modelo para Repositórios Digitais	Tese de Doutorado (Ciência da Informação)

8	(MALY, K; WU, H.; ZUBAIR, M., 2010)	<i>A collaborative faceted categorization system - user interactions</i> (Um sistema colaborativo de categorização facetada por meio de interações do usuário)	Anais (Ciência da Computação)
9	(SANTARÉM SEGUNDO, J. E.; VIDOTTI, S. A. B. G., 2011)	Rede de tags para recuperação da informação no contexto da Representação Iterativa	Artigo científico (Ciência da Informação)
10	(SANTARÉM SEGUNDO, J. E., 2011)	Representação Iterativa e folksonomia assistida para Repositórios Digitais	Artigo derivado da Tese de Doutorado [7] (Ciência da Informação)
11	(SILVA, M. F., 2013)	Proposta de modelo de colaboração para catálogo <i>web</i> facetado.	Tese de Doutorado (Ciência da Informação)
12	(YOO, D. et al., 2013)	<i>Building and evaluating a collaboratively built structured folksonomy</i> (Construção e avaliação de uma folksonomia estruturada construída de forma colaborativa)	Artigo Científico (Ciência da Informação)
13	(SANTOS, R. F.; CORRÊA, R. F., 2015)	A folksonomia e a representação colaborativa da informação em ambientes digitais	Artigo científico (Ciência da Informação)
14	(SANTOS, R. F.; CORRÊA, R. F., 2015)	Modelos colaborativos de indexação social e sua aplicabilidade em bibliotecas digitais	Artigo científico (Ciência da Informação)
15	(SANTOS, R. F., 2016)	Modelos colaborativos de indexação social e a sua aplicabilidade na Base de Dados Referencial de Artigos de Periódicos em Ciência da Informação (BRAPCI)	Dissertação de Mestrado (Ciência da Informação)
16	(SANTOS, R. F. D.; CORRÊA, R. F., 2019)	Modelos colaborativos de indexação: aspectos conceituais e elementos constituintes	Anais (Ciência da Informação)
17	(SILVA, B. F. M., 2019)	A folksonomia assistida aplicada na construção de Corpus de Referência para avaliação de sistemas de indexação automática	Dissertação de Mestrado (Ciência da Informação)
18	(SILVA, B. F. M.; CORRÊA R. F., 2020)	Aplicação da folksonomia assistida na construção de corpus de referência em ciência da informação	Artigo derivado da Dissertação de Mestrado [17] (Ciência da Informação)

Legenda:

Azul: Representação Iterativa.

Amarelo: Modelo colaborativo para catálogo *web* facetado.Rosa: Modelo Colaborativo baseado em *tags* categorizadas.

Verde: Análise ou apresentação de modelos, englobando dois ou três.

Fonte: Elaborado pelo autor.

O conteúdo das pesquisas apresentadas mostrou que a abordagem tem seu cerne na área da Ciência da Computação, de onde decorre o

desenvolvimento de protótipos pioneiros que consideram a folksonomia em metadados. As pesquisas elencadas apresentam três propostas de modelos colaborativos de indexação social tratados pela literatura de Ciência da Computação e Ciência da Informação, sendo eles: modelo colaborativo de representação iterativa; modelo colaborativo para catálogo *web* facetado; e modelo colaborativo baseado em *tags* categorizadas.

O levantamento revela que as teses e as dissertações (literatura cinzenta média) defendidas no Brasil, dentro da delimitação temporal adotada na presente pesquisa, representam boa parte das pesquisas sobre a abordagem e possuem profundo significado ao apresentarem efetivamente propostas de metodologias que visam à hibridização/coexistência dos vocabulários controlados e da folksonomia para a representação da informação em ambientes digitais.

De forma exploratória, a maior parte dos artigos científicos coletados discorre o tema sob os espectros de análise, correlação, aplicação e avaliação de desempenho dos modelos em determinados ambientes diversificados. Nota-se que a produção científica com direcionamento restrito aos modelos colaborativos de indexação social é regular e persistente desde o ano de 2002, ainda que moderada em quantidade. Há ainda a recorrência de pesquisadores brasileiros na abordagem da temática, voltada à Ciência da Informação.

Abaixo são explanados os modelos colaborativos de indexação social identificados, com destaque para as principais características e elementos constituintes de cada um. Além disso, são abordadas algumas das teorias e perspectivas em que esses modelos se apoiam.

3.1 Modelo colaborativo de representação iterativa

Santarém Segundo (2010), em sua tese de doutorado intitulada “Representação Iterativa: um modelo para Repositórios Digitais” propõe um modelo estrutural que objetiva a recuperação semântica da

informação em repositórios digitais, fundamentalmente considerando os recursos da web 2.0 (colaborativa) e da web 3.0 (semântica). Por meio de sua pesquisa, o autor prima adequar os repositórios digitais para que haja o aproveitamento da folksonomia amparada por alguma linguagem de especialidade na representação de recursos informacionais.

Partindo da visão da “web dos dados conectados”, a web semântica confere aos usuários a possibilidade de criarem repositórios de dados na web, construírem vocábulos e definirem regras para interoperarem com esses dados (W3C, 2011). Essa web consiste em um conjunto de padrões destinados a fazer com que o material publicado na web possa ser recuperado de “forma semântica”, congregando informações com o mesmo significado, independentemente de sua estrutura sintática.

Considerando as premissas da “web semântica”, o modelo de representação interativa resulta em uma camada de informação construída de forma social e cíclica, em que a estrutura ontológica é construída inicialmente a partir de um conjunto controlado de termos, porém sujeita a reciclagem de acordo com a ambiência e o conhecimento dos usuários utilizadores do sistema (SANTARÉM SEGUNDO, 2010). À vista disso, o modelo se reestrutura em uma rede que organiza, reorganiza e associa ideias e termos à medida que termos são agregados, permitindo ao usuário navegar pelos termos registrados através de uma matriz de adjacência feita da construção de grafos – estrutura matemática e computacional escolhida para representá-la. A interação direta entre o sistema e o usuário cria um processo/conceito de “folksonomia assistida”.

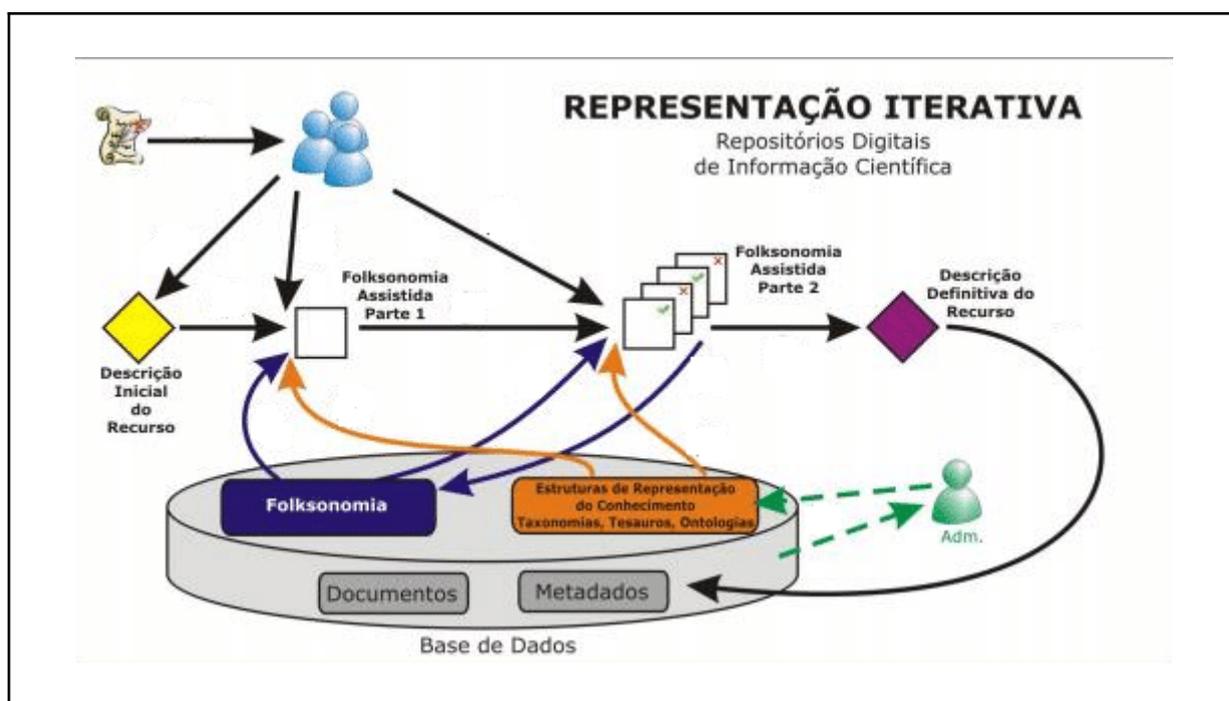
Santarém Segundo (2010, p. 209) afirma que a folksonomia assistida, empregada na representação iterativa, “vai ao encontro da necessidade de auxiliar o usuário na descrição do assunto do recurso a ser depositado, em relação ao domínio do conhecimento do qual o repositório faz parte”.

Dessa maneira, o modelo colaborativo de representação iterativa transforma um repositório digital em uma ferramenta apta a descrever, armazenar e recuperar informação considerando a análise semântica e a

construção coletiva destas estruturas relacionais de informações (SANTARÉM SEGUNDO, 2010).

O modelo proposto se consolida mediante uma interface que dispõe a apresentação de sugestões de *tags* para serem empregadas no momento em que o usuário informa as palavras-chave pretendidas. Novas palavras-chave são adicionadas pelos usuários em um processo contínuo (iteração), podendo ser reutilizadas como descritores por outros usuários. Nesse sentido, o processo de iteratividade se firma com a retroalimentação da informação, que permite o enriquecimento e a consistência do conjunto de descritores, mantendo a reconstrução da estrutura de representação do conhecimento do repositório digital, de forma contínua.

Figura 3 – Representação iterativa – dinâmica de funcionamento



Fonte: Santarém Segundo (2010, p.188).

Um estudo complementar de Santarém Segundo e Vidotti (2011), mais direcionado à seguinte questão, defende que o funcionamento da folksonomia assistida depende de uma interface de sistema distinta que deverá ser implementada ao repositório digital, visto que este recurso não

está disponível nas ferramentas atualmente utilizadas. Algumas vezes é preciso alterar a estrutura de banco de dados que as ferramentas utilizam, criando um conjunto de condições que estabelecem o armazenamento e o relacionamento dessas informações, semanticamente. Dessa forma, o estudo de processos e metodologias para melhorar o desenvolvimento de *software* é constante dentro da área de Ciência da Computação no intuito de adaptação para a construção desse modelo.

Santarém Segundo e Vidotti (2011) complementam que esse modelo tem como diferencial a alteração do paradigma do autoarquivamento de recursos digitais, criando relacionamentos entre os termos que ligam o pensamento do usuário ao material depositado no ambiente digital. Esse pode ser um quesito importante em um sistema que também envolve vocabulário “livre”²⁵. No entanto, no que concerne aos repositórios digitais, grande parte dos *softwares* ainda não oferece relacionamento de termos e busca semântica em seu conteúdo.

Com o emprego do modelo proposto, conforme termos são inseridos no sistema, com a utilização de uma estrutura controlada de representação do conhecimento (ontologia, tesouro, taxonomia), o sistema busca possíveis relações semânticas, dispostas hierarquicamente, de forma horizontal e/ou vertical, se reestruturando. Após esta busca, o sistema apresentará ao usuário um conjunto de termos que poderão ser aceitos de forma total ou parcial ou ainda descartados. Em todo esse processo, cabe ao usuário eleger as *tags* que melhor representem seu recurso digital dentro do domínio do repositório digital em que está sendo realizado o depósito (SANTARÉM SEGUNDO; VIDOTTI, 2011). O processo de gravação das informações é efetivado quando o usuário elege definitivamente os termos que gostaria de usar como *tags*.

Quanto ao registro dos termos, os seguintes níveis de interação são possíveis para o usuário dentro da representação iterativa: [1] nível

²⁵ Forma de organização e representação das informações em que não se adota regras ou políticas de indexação para o controle do vocabulário empregado na descrição dos recursos, podendo utilizar como base o senso comum e a linguagem natural.

“pré-aprovado” – o usuário realiza uma atividade que pode ser revertida ou editada por outros usuários; [2] nível “por permissão” – apenas os usuários que possuem determinado perfil podem executar a ação e [3] nível “por aprovação” – em que é necessário um fluxo de trabalho antes da publicação o produto da ação realizada. Os modelos de aprovação podem ser divididos em hierárquicos ou ainda por sistemas de votação.

No entanto, orienta-se que deve ficar a cargo de um usuário administrador (preferencialmente um profissional da informação) a avaliação e a retroalimentação periódica da infraestrutura de representação da informação e da folksonomia no sistema de informação. Lerman (2007, p. 50) alerta que “a maioria dos métodos iterativos recomenda que a duração de uma iteração seja entre duas e seis semanas”.

Desta maneira, a folksonomia assistida adota um conjunto de medidas que aumentam a probabilidade de os documentos recuperados serem compatíveis com aquilo que os usuários estão buscando em um sistema de informação, considerando o controle no nível de interação do usuário ao atribuir as *tags* ao recurso informacional. Ainda, permite ao usuário uma descrição livre para os recursos digitais, contudo, esta descrição é orientada pelo próprio conhecimento disponibilizado no sistema de informação, através de um processo de iteração. A estrutura funcional deste modelo desemboca do princípio da agregação de valores ao repositório através da inteligência coletiva.

3.2 Modelo colaborativo para catálogo *web* facetado

Alguns pesquisadores internacionais, como Maly e Zubair (2009, 2010), abordam um modelo de colaboração classificado como *web* facetado. De forma prática, alguns protótipos foram idealizados pioneiramente, oriundos dos estudos de Devadason et al. (2002), que desenvolveu o primeiro protótipo desse tipo; de Lay (2009), com um

protótipo aplicado a um *site* de compartilhamento de *bookmarks* (*links* de internet favoritos), similar ao "delicious.com", que é um ícone da web 2.0; de Resmini e Rosati (2007) e de Sigurnjornsson (2010), com uma proposta aplicada a um *site* de compartilhamento de imagens denominado "Flickr". A denominação dos protótipos estão expostos no Quadro 2, a seguir.

Quadro 2 - Protótipos que propõem uma etiquetagem facetada.

Criadores	Denominação do protótipo	Ano de criação
(DEVADASON et al., 2002)	" <i>Faceted Index Internet Gateways</i> "	2002
(LAI, 2009)	" <i>Faccete</i> "	2009
(RESMINI e ROSATI, 2007)	" <i>FaceTag</i> "	2007
(SIGURBJORNSSON, 2010)	" <i>TagExplorer</i> "	2010

Fonte: Elaborado pelo autor.

De forma geral, os protótipos listados concretizam como as etiquetas geradas pelos usuários podem efetivamente ser combinadas com um consistente esquema de classificação facetada para melhorar a arquitetura de informação de um sistema colaborativo de indexação, permitindo a busca por palavras-chave e a navegação pelas facetas.

No contexto nacional, Silva (2013), em sua tese de doutorado intitulada "Proposta de modelo de colaboração para catálogo *web* facetado", aponta um modelo de colaboração para indexação e busca de registros em um catálogo *web facetado* colaborativo, denominado "Facetlog". Seu estudo teve como objetivo "facilitar ou melhorar a qualidade da indexação, oferecendo uma estratégia complementar à etiquetagem, através da seleção de termos de uma taxonomia facetada" (SILVA, 2013, p. 9). Como ponderado pelo autor, também visou o "aperfeiçoamento na recuperação da informação em catálogos *web*, através de uma interface de busca que combina a navegação facetada com a busca palavra-chave" (SILVA, 2013, p. 9.). A pesquisa desenvolvida

teve início no ano de 2002 e foi finalizada ao decorrer de quase uma década.

O estudo se originou das dificuldades encontradas na atuação profissional do pesquisador e ilustram fortemente as motivações para a subsistência de uma taxonomia facetada. Enquanto trabalhava no setor de Pesquisa e Desenvolvimento de uma empresa de desenvolvimento de equipamentos eletrônicos, o mesmo se deparou com um cenário em que uma grande quantidade de documentos deveria ser organizada, entretanto, foi observável a complexidade desse processo, baseado na estrutura de pastas. Essas pastas, segundo orientação da gerência, deveriam ser organizadas pela versão do produto. Contudo, este método muitas vezes se mostrava ineficiente.

Como relatado pelo referido pesquisador, o trabalho era frequentemente dificultado pela excedente quantidade de requisitos (posteriormente, "facetas") a serem sistematicamente adicionados para cada versão do produto. Sucessivos testes ocasionavam a perda de controle na administração da etiquetagem dos requisitos implementados. Com efeito, era observada a ausência de mecanismos que melhor mensurassem, quando necessário, a evidência e a aglutinação dos requisitos aprovados, dos reprovados, dos requisitos vigentes e dos não vigentes (em cada versão). Eram recorrentes as dúvidas e o retrabalho que dificultavam a tomada de decisão corporativa, como nos casos de descarte, de adiamento de versão e priorização de requisitos. A carga de trabalho empenhado na elucidação de todos os requisitos possíveis na etiquetagem e na recuperação documental era intensa e desfavorável para a economia de tempo.

O modelo de colaboração para catálogo *web* facetado foi pensado como forma alternativa às estruturas consideradas pouco flexíveis, baseadas na hierarquização de termos. O estudo de Silva (2013) parte da constatação de que os sistemas enumerativos encontrados na Biblioteconomia, como a Classificação Universal Dewey (CDD) e a Classificação Decimal Universal (CDU), que comportam estrutura

hierárquica de terminologia, podem não ser considerados flexíveis o suficiente para utilização em catálogos informatizados, ao passo que podem apresentar restrições na acomodação de novos assuntos ou para a submissão de assuntos compostos (BATISTA; CARVALHO, 2003; BROUGHTON; SLAVIC, 2007).

Conforme ressaltado por Lima (2004 p. 78), “[...] em uma área do conhecimento, o número de assuntos específicos que podem ser selecionados em uma classificação é infinito”. O sistema de classificação facetada, por sua vez, pode ser considerado mais escalável e sustentável, disponibilizando ainda mecanismos mais simples para a adição e remoção de termos nas facetas. Isso significa que ao apagar ou renomear um termo somente o respectivo nó da faceta tem impacto, enquanto na taxonomia tradicional seria necessário apagar ou alterar vários termos na estrutura da árvore (TZITZIKAS *et al.*, 2002).

Neste sentido, pode-se dizer que o instrumento resultante da classificação facetada é um vocabulário controlado, que pode ser denominado “esquema de classificação facetada” (*faceted classification scheme*), “sistema de classificação facetada” (*faceted classification system*) ou “taxonomia facetada” (*faceted taxonomy*) (SILVA, 2013). Conforme Tzitzikas *et al.* (2002, p. 3),

[...] rigorosamente, uma taxonomia facetada consiste em um conjunto finito de facetas. Cada faceta consiste de uma terminologia, que possui um conjunto finito de palavras ou termos, estruturados por meio de uma relação hierárquica. Cada faceta é desenhada separadamente e modela um aspecto distinto do domínio.

O princípio da análise facetada foi idealizado por Shiyali Ramamrita Ranganathan (1892-1972) e teve início com a publicação de dois trabalhos de grande relevância para área da Ciência da Informação: o livro “Colon Classification”²⁶ (Classificação de Dois Pontos), em 1933 e

²⁶ Primeiro sistema de classificação bibliográfica com bases no princípio analítico-sintético ou análise por facetas, priorizando as categorias - *Personality, Matter, Energy, Space e Time* (PMEST). O número de classificação é formado por símbolos unidos pelo sinal dos dois pontos, relacionando as diferentes perspectivas com que um assunto de uma obra pode ser representado.

posteriormente do livro "Prolegomena", em 1937. Esse matemático e também bibliotecário, originário da Índia, propôs que os termos fossem divididos em facetas ao invés de ter uma única grande hierarquia que acomodaria todos os termos. Uma faceta viria a ser uma lista de termos com características semelhantes, porém com relação distinta. Buscando maior fluidez em comparação com os sistemas hierárquicos, os assuntos compostos poderiam ser gerados no ato da indexação ou na recuperação do documento.

Figura 4 – Esboço de categorização facetada – do tipo *menu dropdown*



Fonte: Nextews.com (2021)

A tese de doutorado de Silva (2013) apresenta uma consistente fundamentação teórica para a criação de um protótipo que incorpora a etiquetagem facetada, incluindo estudos conduzidos por Denton (2003), Fargan (2010) e Spiteri (2011). Esses autores, assim como Silva (2013), defendem a taxonomia facetada como eficiente, vantajosa e importante, na medida em que facilita a recuperação da informação.

Dessa maneira, o modelo prima pelo esforço coletivo e pela reorganização automática, na medida em que leva em conta os interesses

dos usuários e ajuda-os a navegar pela coleção de forma rápida e intuitiva. (MALY; WU; ZUBAIR, 2010). Segundo Oren, Delbru e Decker (2006), a navegação facetada pode ser considerada um paradigma visual de consultas e está apoiada na teoria de conjuntos, em que a partir do mecanismo integrado de navegação facetada, o sistema pode retornar documentos relacionados com uma ou várias classes selecionadas simultaneamente, atuando como filtros dos resultados.

Silva (2013) menciona outros protótipos posteriores que agregam melhorias às interfaces para a navegação facetada, dispostos no Quadro 3, a seguir.

Quadro 3 - Protótipos com aperfeiçoamento da interface.

Criadores	Denominação do protótipo	Ano de criação
Universidade de Berkeley, EUA	Flamenco ²⁷	2002
(OREN, DELBRU e DECKER, 2006)	BrowseRDF	2006
(TVAROZEK e BIELIKOVÁ, 2007)	FARTIC	2007
Universidade de Helsinki, Finlândia	<i>HealthFinland</i>	2007

Fonte: Elaborado pelo autor.

No geral, esses protótipos apresentaram avanços e soluções quanto ao aperfeiçoamento da interface, enfatizando que neste tipo de navegação é desejável uma interface de fácil utilização, que opere da melhor maneira possível, em atenção às questões de usabilidade. Nesse sentido, Nielsen (2003, p. 1) pondera que a usabilidade é definida como “um atributo de qualidade que avalia o quão uma interface é fácil para o usuário utilizar. A palavra usabilidade também se refere a métodos para melhorar a facilidade de uso durante o processo de design”.

Dessa forma, o estudo proposto por Silva (2013) se direciona ao desenvolvimento de um catálogo computadorizado (ou catálogo automatizado) que quando acessível por um navegador de *internet* pode

²⁷ Disponível em: <http://flamenco.berkeley.edu>

passar a ser definido por “catálogo baseado na internet”, “catálogo web” ou ainda, “diretório web”.

Em síntese, [para o estudo] define-se um catálogo *web* como um conjunto de registros que descrevem itens, que podem ser entidades do mundo real ou abstratas. Os registros são acessíveis pela Internet e organizados sob algum critério, como ordenação alfabética e classificação por autor, ou um vocabulário controlado, como uma taxonomia (SILVA, 13, p. 71).

Considera-se que a navegação facetada pode ser utilizada em um catálogo computadorizado com o objetivo de refinamento da busca e a navegação pelo catálogo. Assim, este tipo de navegação pode se consolidar de maneiras distintas, como em uma navegação “um nível por tela”, em “árvore expansível”, em “*menu dropdown*” (Figura 4) e em “caixas de seleção encadeadas” (Silva, 2013, p. 53-55). A contribuição dos usuários também é adicionável na navegação facetada, conforme estudos de Yang e Hofmann (2013). A indexação (etiquetagem) realizada pelo usuário pode ser trazida como uma sugestão que é aprovada por um membro administrador do sistema, onde, após a aprovação a etiqueta aparece vinculada ao registro para os demais usuários do sistema.

Por fim, cabe ressaltar que Silva (2013) desenvolveu o respectivo protótipo com base do processo de desenvolvimento de *software* RUP²⁸, e fez a modelagem do sistema através de diagramas UML²⁹. Em seu estudo, realizou testes de usabilidade do sistema proposto, mediante a utilização em um catálogo *web* de empresas locais (na fase de experimentação) e, por fim, demonstrou que é possível inserir a indexação através da taxonomia facetada em um catálogo tradicional que apresentava apenas a

²⁸ *Rational Unified Process* - “Processo Unificado Rational”. *Software* que objetiva ênfase na comunicação com o cliente, sugerindo um processo iterativo e incremental e utilizando amplamente a linguagem *Unified Modeling Language* (UML) durante o processo de desenvolvimento.

²⁹ *Unified Modeling Language* – “Linguagem de Modelagem Unificada”. Linguagem não proprietária, utilizada para análise, desenho e implementação de sistemas de computador, modelagem de processos de negócios ou processos em geral. A UML é uma notação que especifica diagramas cujo objetivo é representar elementos definidos na modelagem orientada a objetos.

etiquetagem social. Silva (2013) orientou que trabalhos futuros viessem a melhorar, estender ou utilizar os elementos do Facetlog, primando pela melhoria da interface de busca, das configurações do modelo de colaboração e da qualidade *web* semântica empregada.

3.3 Modelo colaborativo baseado em *tags* categorizadas

O modelo colaborativo baseado em *tags* categorizadas foi idealizado por Yoo et al. (2013) e corresponde a um sistema a ser utilizado por funcionários de empresas (usuários ativos do sistema), auxiliando a classificação e o compartilhamento de registros diversificados (manuais e relatórios, por exemplo). Entretanto, os autores ressaltam que ele pode ser aplicado em outros tipos de ambientes digitais colaborativos.

Nesse modelo, os descritores – denominados “*Categorized Tags*” – no momento da indexação, devem ser incluídos em múltiplas categorias que fazem parte. Dessa forma, as *tags* atribuídas são automaticamente integradas no sistema de informação, constituindo uma folksonomia estruturada de forma colaborativa (*Collaboratively built Structured Folksonomy* - CSF, em inglês) e permitindo também uma organização e busca de forma semântica.

Yoo et al. (2013) classificam e conceituam dois tipos de folksonomia a serem consideradas no modelo desenvolvido: [1] a “folksonomia plana”, composta por termos atribuídos livremente pelos usuários e que não possuem hierarquia entre si - correspondendo a uma representação de informações em que não se adota regras ou políticas de indexação e; [2] a “folksonomia estruturada”, em que os termos empregados pelos usuários integram uma estrutura hierárquica mediante seleção de categorias, correspondendo ao diferencial deste modelo.

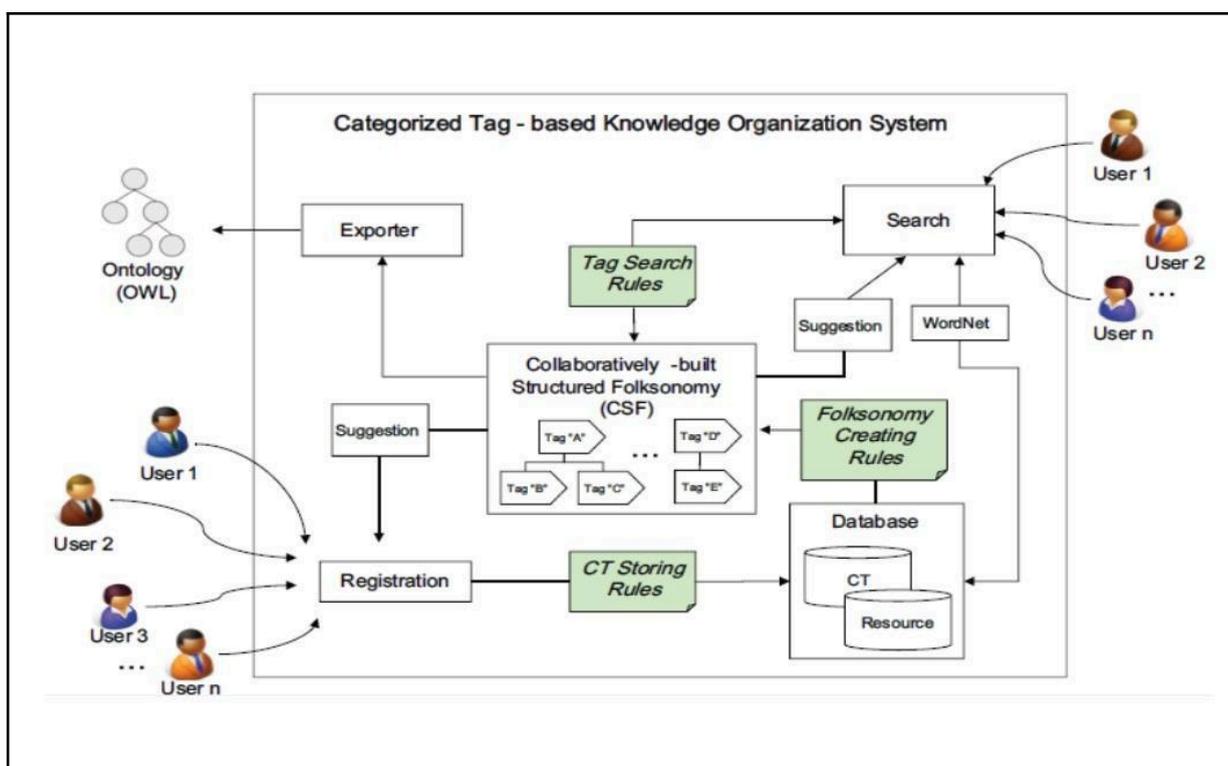
Os referidos autores concordam que o controle do nível de liberdade do usuário ao atribuir *tags* é a garantia de que o processo de iteratividade possa resultar na reconstrução do conhecimento de forma moderada e

coletiva, permitindo uma estrutura de organização e representação de informação consistente e garantindo maior relação descrição-recurso.

Assim, no modelo definido como “pré-aprovado”, qualquer usuário cadastrado pode incluir novos registros no sistema, bem como contestar e corrigir a indexação de recursos; no modelo definido como “por aprovação”, se uma *tag* for registrada de forma incompleta, ou ainda, sem escolha de categoria, o sistema exibe uma mensagem para que o usuário efetue a verificação e complete.

Esse modelo envolve a dinâmica de algumas estruturas e funções essenciais, que o conferem singularidade, conforme descrito na Figura 5, a saber:

Figura 5 – Modelo colaborativo baseada em *tags* categorizadas - dinâmica de estruturas e funcionamento



Fonte: Yoo et al. (2013, p. 597)

- a) “*Registration*” (registro): corresponde ao processo de registro e submissão de um documento. O usuário insere metadados básicos sobre o recurso informacional, como título, autor, tipo de

documento, consolidando a representação descritiva do recurso. Potenciais *tags* categorizadas, disponíveis no sistema, podem ser sugeridas neste momento, sendo ou não aceitas. O processo é composto por duas partes: a definição da *tag* e em seguida, da sua categoria, a fim de especificar a descrição do recurso, delineando ainda mais o contexto da etiqueta inserida;

- b) "*CT Storing Rules*" (Regras de armazenamento de *tags* categorizadas): duas *tags* categorizadas podem ser registradas simultaneamente ou seguidamente. O sistema verifica se existe alguma entrada para a *tag* categorizada em sua tabela de etiquetas. Se houver, ele aumenta o campo de contagem da entrada da *tag* em 1 (um), caso contrário, uma nova *tag* é adicionada na respectiva tabela;
- c) "*Database*" (Banco de dados): corresponde à ambiência onde os recursos são depositados e sistematizados em tabela, a fim de armazenar *tags* atribuídas e/ou selecionadas pelos usuários, levando em conta regras específicas;
- d) "*Folksonomy creating rules*" (Regras que criam a folksonomia): o sistema que suporta o modelo possui um mecanismo que permite a especificação de *tags* que fazem parte de categorias distintas. Equívocos de registro podem ser corrigidos mediante consulta de outros usuários que podem substituir etiquetas. Dessa maneira, admite-se que a avaliação das *tags* categorizadas fique a cargo dos próprios usuários do sistema, colaborativamente;
- e) "*CSF – Collaboratively built Structured Folksonomy*" (Folksonomia estruturada de forma colaborativa): *tags* registradas são integradas de forma automática à medida que os usuários contribuem para a indexação dos recursos informacionais, consolidando uma folksonomia estruturada e colaborativa;
- f) "*Suggestion CT*" (Sugestão de *tags* categorizadas): ao digitar uma palavra-chave essa função encontra "nós" da CSF que representam

- o termo digitado, recuperando etiquetas similares de cada "nó" encontrado, ao passo que propõe *tags* similares;
- g) "*Search*" (Pesquisa): corresponde a uma função de busca de recursos. A pesquisa por meio de palavras-chave apóia a correspondência de termos enquanto que a pesquisa baseada em *tags* categorizadas permite a busca semântica, oferecendo distintos pontos de acesso para o documento;
 - h) "*Tags rearch rules*" (Regras de pesquisas com *tags*): na pesquisa de *tags* categorizadas, cada descritor em uma etiqueta categorizada fornece um potencial ponto de entrada para a busca de documentos disponíveis no sistema. Ao selecionar uma das *tags* existentes, os diferentes significados da etiqueta selecionada são extraídos do *WordNet* (um banco de dados lexical) e listados na tela, permitindo ao usuário a seleção do significado que melhor corresponder ao conteúdo. Em seguida, todos os documentos relacionados a essa *tag* são mostrados para o usuário; e
 - i) "*Exporter*" (Exportador): função que permite exportar uma *tag* categorizada do CSF em formato estruturado (uma ontologia, por exemplo) ou em formato de arquivo padrão. É possível que se compartilhe e que se reutilize as *tags* categorizadas em outros sistemas baseados em e *Extensible Markup Language* (XML) ou aplicativos provenientes das tecnologias semânticas, tendo em vista que elas representam o consenso dos usuários no que concerne à relação entre a *tag* e a sua categoria.

A partir do exposto, parte-se para uma explanação sobre os principais avanços e perspectivas futuras das pesquisas dedicadas aos modelos colaborativos de indexação social para repositórios digitais identificados e apresentados nesta seção.

4 AVANÇOS DA TEMÁTICA NA CIÊNCIA DA INFORMAÇÃO BRASILEIRA

No estudo de descrição e análise conduzido por Santos (2016) em sua dissertação de mestrado intitulada “Modelos colaborativos de indexação social e a sua aplicabilidade na base de dados referencial de artigos de periódicos em ciência da informação (BRAPCI)”, a autora sistematizou as características análogas compreendidas pelos três modelos colaborativos de indexação social, igualmente identificados em sua pesquisa (Quadro 2).

Quadro 4 – Elementos identificados nos modelos colaborativos de indexação social analisados.

Elementos identificados e analisados	Modelo colaborativo de representação iterativa	Modelo colaborativo de indexação social Facetlog	Modelo colaborativo baseado em tags categorizadas
Domínio de aplicação recomendado	Repositório digital de publicações científicas	Catálogo <i>Web</i> facetado de empresas de um município	Sistema de organização dos documentos de uma empresa
Inclusão de recursos	Autoarquivamento	Autoarquivamento	Autoarquivamento
Recursos autoarquivados e representados no sistema	Produções científicas dos pesquisadores de determinada instituição	Informações relativas as empresas de determinada cidade	Documentos de determinada empresa (manuais e relatórios, por exemplo)
Perfis de usuários existentes no modelo	Usuário administrador e usuário autor do trabalho a ser representado e depositado no sistema	Super administrador, administrador, usuário membro e usuário visitante	Usuário funcionário da empresa
Tipos de representações a serem realizadas pelos usuários para descrever os recursos	Representação descritiva e temática	Representação descritiva e temática	Representação descritiva e temática

Características das etiquetas para a indexação dos recursos informacionais	Etiquetas provenientes da linguagem natural e/ou de termos que fazem parte de um instrumento de controle terminológico associado ao sistema	Etiquetas provenientes da linguagem natural e/ou de termos que fazem parte de um instrumento de controle terminológico associado ao sistema	Apenas etiquetas categorizadas (<i>tags</i> que denotem as suas respectivas categorias em que fazem parte)
Designação da linguagem de indexação disponibilizada pelo sistema e a seleção de suas respectivas categorias	Deve ocorrer em níveis previamente estabelecidos pelo administrador do repositório levando em consideração o seu domínio de aplicação.	Deve ocorrer em níveis previamente estabelecidos pelo administrador do catálogo, para que ele determine as facetas e as suas categorias levando em consideração o domínio de aplicação do catálogo	As próprias <i>tags</i> categorizadas atribuídas pelos usuários constituem a linguagem de indexação disponibilizada pelo sistema, uma vez que elas são automaticamente integradas em uma lista de potenciais descritores e se configurarão como sugestões de <i>tags</i>
Linguagem de indexação sugerida como auxílio para a representação temática dos documentos	Tesouro, ontologia e taxonomia	Taxonomia facetada	Lista de potenciais <i>tags</i> categorizadas decorrente de representações realizadas anteriormente pelos usuários do sistema
Sugestão de termos no processo de indexação	Intervenção do sistema (mediante a função de autocompletar) no momento em que o usuário estiver digitando os termos de indexação - com a apresentação de um conjunto de termos inseridos previamente pelo administrador do sistema compostos tanto de descritores existentes no vocabulário controlado associado ao repositório como de etiquetas já atribuídas anteriormente pelos usuários	Intervenção do sistema (mediante a função de autocompletar) no momento em que o usuário estiver digitando os termos de indexação com a apresentação de termos já registrados e de descritores provenientes da taxonomia facetada como recomendação automática de termos. Além disso, a própria taxonomia facetada é disponibilizada na interface de indexação.	Intervenção do sistema (mediante a função de autocompletar) no momento em que o usuário estiver digitando os termos categorizados de indexação - apresentando uma lista de potenciais <i>tags</i> categorizadas como recomendação automática de descritores.

Dessa forma, como ponto importante, confirma-se que o modelo colaborativo de representação iterativa foi idealizado para o autoarquivamento de produções científicas de determinada instituição. No modelo colaborativo de indexação social Facetlog os recursos a serem autoarquivados e representados no catálogo *web* facetado, originalmente, seriam informações relativas às empresas de determinada cidade. Em nítido contraste, no modelo colaborativo baseado em *tags* categorizadas, os recursos digitais que objetivaram a criação deste modelo foram os documentos de determinada empresa (como manuais e relatórios). Entretanto, de acordo com Santos (2016), de forma global, apesar dos diferentes modelos colaborativos de indexação social apresentarem domínios de aplicação originalmente distintos, ocorre a possibilidade da implantação das suas regras em diversos contextos alternativos.

Nas pesquisas direcionadas à aplicabilidade dos modelos em ambientes digitais específicos, destaca-se a supracitada, que objetivou a análise e a discussão sobre os modelos colaborativos de indexação social encontrados na literatura brasileira, expondo ainda vantagens e desvantagens quanto à possível implantação na BRAPCI. Neste caso específico de base de dados, a política de indexação impõe que o indexador faça uso dos mesmos termos informados pelos autores nas palavras-chave dos seus artigos para a descrição do conteúdo da produção intelectual a ser depositada neste sistema, a denominada “indexação automática”.

Partindo da perspectiva de que a agilidade atribuída à forma de indexação supracitada não garante necessariamente qualidade na representação efetuada, Santos (2016) discorre sobre adequações necessárias para uma possível implantação, considerando questões relacionadas às próprias políticas de indexação³⁰ da base de dados estudada. Reitera que a possível aplicação dos modelos, de forma

³⁰ O emprego de políticas de indexação é considerado uma decisão administrativa que dá a um sistema de informação condições para atividades de representação temática de recursos informacionais, de maneira normativa, formal, racional, consistente e padronizada.

adaptada, colabora para minimizar as implicações negativas da política de indexação, tais como:

[...] a ocorrência de erros ortográficos em descritores; a presença de descritores duplicados; a representação da informação sem a interferência do indexador; a participação restrita do autor na indexação do trabalho de sua autoria; a ausência de revisão dos termos existentes no metadado assunto da base de dados; e a participação do indexador apenas na atividade de categorização dos termos e definição de remissivas entre termos (SANTOS, 2016, p. 9).

Silva (2019), em sua tese de dissertação intitulada "A folksonomia assistida aplicada na construção de Corpus de Referência para avaliação de sistemas de indexação automática", visou correlacionar os conceitos de "folksonomia assistida" e "indexação automática", propondo meios para que a indexação assistida possa ser aplicada na avaliação de sistemas de indexação automática a partir da compilação de Corpus de Referência, por meio da plataforma de gerenciamento de coleção Tainacan.

Na dissertação de mestrado de Viana (2020), intitulada "Representação colaborativa de dados científicos: estudo na Rede de Repositórios de Dados Científicos do Estado de São Paulo", a autora elucida possíveis formas de operação em repositórios de dados científicos que atuam como sistemas informacionais colaborativos, mediante a participação ativa dos usuários ou grupos de pessoas na organização e ao tratamento dos registros científicos provenientes de atividades de investigação. Porém, a mesma não faz possíveis relações com os modelos descritos neste trabalho.

De forma exploratória, esses estudos contribuem para a Ciência da Informação na medida em que investigam os modelos colaborativos de indexação social anteriormente emergidos, sob os espectros de análise, correlação, aplicação e avaliação de desempenho dos modelos em determinados ambientes diversificados.

Em suma, observou-se que a literatura científica das áreas aqui investigadas não avançou na proposição de modelos colaborativos de indexação social para repositórios digitais para além dos identificados

neste estudo, sendo constatado nesta pesquisa o mesmo cenário discutido por Santos (2016), cenário que demonstra a necessidade da temática figurar na agenda de discussão da área.

5 CONSIDERAÇÕES FINAIS

Diferentes estudos defendem a hibridização/coexistência dos vocabulários controlados e da folksonomia para a representação da informação em ambientes digitais. Considerando o contexto de Ciência Aberta e o movimento de acesso aberto em que os repositórios digitais figuram como importantes sistemas de informação, esta pesquisa discorreu sobre as estratégias de inovação e tecnologia para a indexação social em sistemas híbridos de informação, a partir da investigação dos modelos colaborativos de indexação social para repositórios digitais, de forma interdisciplinar, ressaltando seus fundamentos teóricos e características.

Mediante estudo exploratório e descritivo conduzido a partir de pesquisa bibliográfica, foram identificadas na literatura científica de Ciência da Informação e Ciência da Computação três propostas de modelos colaborativos de indexação social - modelo colaborativo de representação iterativa, modelo colaborativo para catálogo *web* facetado e modelo colaborativo baseado em *tags* categorizadas. Segundo os autores que abordam a temática, esses modelos se adequam aos repositórios digitais, a depender de ajustes operacionais nos sistemas de informação ponderados e às políticas de indexação a serem direcionados aos mesmos. No contexto brasileiro, estudos sobre recuperação semântica e aplicabilidade desses modelos foram identificados como tendências.

No geral, infere-se que modelos colaborativos de indexação social agregam aos sistemas de informação vantagens como a possibilidade da participação colaborativa de diversificados atores de determinada área do conhecimento e relativa e imediata viabilidade de contestação, correção e alteração da indexação realizada por outro usuário, em um processo de iteração mediada pela inteligência coletiva.

Entretanto, para o profissional da informação, a adoção desses modelos não significa a plena terceirização de tarefas. No processo colaborativo de indexação social, considera-se importante a permanente

intervenção de um usuário administrador, preferencialmente um profissional da informação capacitado, delegando a ele as tarefas de análise holística do sistema de informação e de correção de eventuais inconsistências operacionais, concedendo melhor qualidade no processamento e aumentando a eficiência da recuperação da informação. As atividades desse ator deverão incluir apoio ao desenvolvimento, reformulação e cumprimento de estratégias operacionais a serem documentadas em uma consistente política de indexação.

Dessa maneira, prima-se pelo aperfeiçoamento das metodologias e ferramentas que considerem a integração da participação dos usuários no processo de indexação da informação, reiterando o auxílio de um instrumento de controle terminológico, a construção e atualização desses instrumentos de representação da informação, para mediar o processo e potencializar a adequada recuperação da informação em ambientes digitais.

REFERÊNCIAS

ALBAGLI, S. Ciência aberta em questão. In: ALBAGLI, Sarita; MACIEL, Maria Lúcia; ABDO, Alexandre Hannud (Ed.). **Ciência aberta, questões abertas**. Brasília: IBICT; Rio de Janeiro: UNIRIO, 2015. p. 9-26.

ALBAGLI, S.; CLINIO, A.; RAYCHTOCK, S. Ciência aberta: correntes interpretativas e tipos de ação. **Liinc**, Rio de Janeiro, RJ, v.10, n.2, p. 434-450, nov. 2014. Disponível em: <https://periodicos.ufpb.br/index.php/pbcib/article/view/23180>. Acesso em: 27 mar. 2021.

ALBAGLI, S. Ciência Aberta em questão. *In*: Seminário Internacional Ciência Aberta, Questões Abertas, 1. Rio de Janeiro, 2014. **Anais [...]**. Rio de Janeiro: Liinc; IBICT; OKF; Unirio, 2014. Disponível em: https://www.cienciaaberta.net/wpcontent/uploads/2014/05/20140820_Albagli_Ciencia_Aberta_em_questao.pdf. Acesso em: 26 mar. 2021.

AMSTEL, V. F. **Folksonomia**: vocabulário descontrolado na arquitetura da informação ou samba do crioulo doido. São Paulo, p. 15, 2007. Disponível em: http://www.guilhermo.com/ai_biblioteca/referencia.asp?referencia=317. Acesso em: 20 maio 2021.

AQUINO, M. C. Hipertexto 2.0, folksonomia e memória coletiva: um estudo das tags na organização da Web. **Revista da Associação Nacional dos Programas de Pós-Graduação em Comunicação**, v. 18, n.2, ago. 2007. Disponível em: <http://www.compos.org.br/seer/index.php/e-compos/article/view/165/166>>. Acesso em: 24 maio 2021.

BATISTA, E. A. **Uma Taxonomia Facetada para Técnicas de Elicitação de Requisitos**. 2003. 150f. Dissertação (Mestrado em Ciência da Informação) - Instituto de Comunicação, Unicamp, Campinas, 2003.

BORGES, G. S. B.; MACULAN, B. C. M. S.; LIMA, G. Â. B. O. Indexação automática e semântica: estudo da análise do conteúdo de teses e dissertações. **Inf. Soc.: Est.**, João Pessoa, v. 18, n. 2, p. 181-193, maio/ago. 2008.

BRANDT, M.; MEDEIROS, M. B. B. Folksonomia: esquema de representação do conhecimento?. **Transinformação**, Campinas, v. 22, n. 2, maio/ago. 2010.

BROOKS, T. A. Where is meaning when from is gone? Knowledge representation on the web. **Information Research**. v. 6, n. 2, jan. 2001.

BROUGHTON, V.; SLAVIC, A. Building a faceted classification for the humanities: principles and procedures. **Journal of Documentation**, London, v. 63, n. 5, p. 727-754, 2007.

BROWN, P. et al. **The democratic indexing of images. New Review of Hypermedia and Multimedia**, 2, 1996, p. 107-120.

CARLAN, E. **Sistemas de organização do conhecimento**: uma reflexão no contexto da ciência da informação. 2010. 195 f., il. Dissertação (Mestrado em Ciência da Informação), Universidade de Brasília, Brasília, 2010.

CATARINO, M. E. **Integração das folksonomias nos metadados**: identificação de novos elementos como contributo para a descrição de recursos em repositórios. 233 f. 2009. Tese (Doutorado em Tecnologias e Sistemas de Informação) – Escola de Engenharia, Universidade do Minho, Guimarães, 2009.

CATARINO, M. E.; BAPTISTA, A. A. Folksonomias: características da etiquetas na descrição de recursos da Web. **Informação & Informação**, Londrina, v. 14, n. esp., 2009.

CAVALCANTI, C. R. Indexação e tesouro: metodologia e técnicas. Brasília, ABDF. 1978.

CINTRA, A. M. et al. **Para entender as linguagens documentárias**. 2. ed. São Paulo: Polis, 2002.

COLABORATIVO. In: DICIO, Dicionário Online de Português. Porto: 7Graus, 2021. Disponível em: <https://www.dicio.com.br/colaborativo/>. Acesso em: 17/11/2021.

CREASER, C.. Scholarly communication and access to research output. *In*: EVANS, WENDY, BAKER, DAVIS. **Libraries and Society**: role, responsibility and future in an age of change. EBSCO Publishing. 2011. p. 53-66.

DAMES, K. M. Social software in the library. **LLRX**, jul. 2004. Disponível em: <http://www.llrx.com/features/socialsoftware.htm>. Acesso em: 26 maio 2021.

DENTON, W. **How to make a faceted classification and put it on the web**. Miskatonic University Press, nov. 2003. Disponível em: <http://www.miskatonic.org/library/facet-web-howto.html>. Acesso em: 26 maio 2021.

DEVADASON, F. J. et al. Faceted Indexing Based System for Organizing and Accessing Internet. *Knowledge Organization*, v. 29, p. 65-77, 2002. Disponível em: <http://www.reocities.com/devadason.geo/FactDSIS.pdf>. Acesso em: 17 nov. 2021.

DEVADASON, F. J. et al. Faceted Indexing Based System for Organizing and Accessing Internet. **Knowledge Organization**, v. 29, p. 65-77, 2002. Disponível em: <http://www.reocities.com/devadason.geo/FactDSIS.pdf>. Acesso em: 20 mar. 2021.

EDOLS, L. **Taxonomies are what?** Free Pint, Reino Unido, n. 97, out. 2001. Disponível em: <http://www.freepint.com/issues/041001.htm#eature>. Acesso em nov. 2021.

FAGAN, J. C. Usability studies of faceted browsing: A literature review. *Information Technology and Libraries*, v. 29, p. 58-66, 2010.

FERRARI, R. D. D.; PIRES, G. L. **Auto-arquivamento e acesso aberto**: deveres e direitos digitais na sociedade em rede. *RDBCI*, Campinas, SP, v. 12, n. 1, p. 22-38, fev. 2014. Disponível em: <http://www.sbu.unicamp.br/seer/ojs/index.php/rbci/article/view/3874>. Acesso em: 26 out. 2021.

FUJITA, M. S. L. **Leitura em análise documentária**. 1999. 178f. Monografia (Relatório Parcial de Pesquisa) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista; CNPq. Marília.

GALVINO, C. C. T.; ROSA, M. N. B.; OLIVEIRA, B. M. J. F.; OLIVEIRA, B. M. J. F. O movimento de acesso aberto e a ciência aberta: uma proposta de repositório de dados e memória na universidade federal de alagoas. **Ciência da Informação em Revista**, v. 7, n. 1, p. 34-45, 2020. DOI: 10.28998/cirev.2020v7n1c. Acesso em: 27 mar. 2021.

GOUVÊA, C.; LOH, S. Folksonomias: identificações de padrões na seleção de tags para descrever conteúdos. **Revista Eletrônica de Sistemas de Informação**, vol. 11, n.2, 2007. Disponível em: <http://www.periodicosibepes.org.br/index.php/reinfo/article/view/214>. Acesso em: 20 maio 2021.

GRACIOSO, L. S. Parâmetros teóricos para elaboração de instrumentos pragmáticos de representação e organização da informação na web: considerações preliminares sobre uma possível proposta metodológica. *In CID: rev. Ci. Inf. e Doc.*, Ribeirão Preto, v.1, n.1, p.138-158, 2010.

GUEDES, R. M.; DIAS, E. J. W. Indexação Social: abordagem conceitual. *In: Revista ACB: Biblioteconomia em Santa Catarina, Florianópolis*, v.15, n.1, p. 39-53, jan./jun. 2010. Disponível em: http://revista.acbsc.org.br/index.php/racb/article/view/686/pdf_17. Acesso em: 30 maio 2021.

HEY, T. **The big idea: the next scientific revolution**. **Harvard Business Review**, Brighton, Nov. 2010. Disponível em: <https://hbr.org/2010/11/the-bigidea-the-next-scientific-revolution>. Acesso em: 22 maio 2021.

HEY, T.; TANSLEY, S.; TOLLE, K. Jim Gray on e-Science: a transformed scientific method. In: [...] **The fourth paradigm: data intensive scientific discovery**. Redmond: Microsoft Research, 2009.

INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA. **Repositórios digitais**. [S. l.], 2012. Disponível em: <http://sitehistorico.ibict.br/informacao-para-ciencia-tecnologia-e-inovacao%20/repositorios-digitais/repositorios-brasileiros>. Acesso em: 29 abr. 2021.

JACOB, E. K. **Ontologies and the semantic web**. Bulletin for the American Society for Information Science and Technology, v. 29, n. 4, p. 19-22, Apr./Mayo 2003.

LAI, P. **Facette**: using facets to improve tag-based bookmarking. 2009. 52f. Dissertação (Master of Engineering in Electrical Engineering and Computer Science) - Massachusetts Institute of Tecnology, 2009.

LANCASTER, F. W. Indexação e resumos: teoria e prática. Brasília: Briquet de Lemos, 2004. 452 p. **Revista ACB**: Biblioteconomia em Santa Catarina, Florianópolis, v.15, n.1, p. 39-53 jan./jun., 2010.

LANGRIDGE, D. **Classificação: abordagem para estudantes de biblioteconomia**. Rio de Janeiro: Interciência, 2006. 126 p.

LERMAN, K. Dynamics of Collaborative Document Rating Systems. In: 9th WebKDD and 1st SNA-KDD, 2007. Workshop on Web mining and social network analysis. **Proceedings...** New York: ACM, 2007. p. 46-55.

LÉVY, P. **A inteligência coletiva: por uma antropologia do ciberespaço**. 4. ed. São Paulo: Loyola, 2003

LIMA, G. A. B. O. **Mapa hipertextual (MHTX)**: um modelo para organização hipertextual de documentos. 2004.199f. Tese (Doutorado em Ciência da Informação) - Escola de Ciência de Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2004.

MALY, K.; WU, H.; ZUBAIR, M. A collaborative faceted categorization system user interactions. In: 14th International Conference on Electronic , 2010, Helsinki, Finland **Proceedings...** Transforming the Nature of Communication: Turid Hedlund and Yasar Tonta, 2010. Disponível em: http://elpub.scix.net/data/works/att/106_elpub2010.content.pdf. Acesso em: 26 maio. 2021.

MALY, K; WU, H; ZUBAIR, M. Automated Support for a Collaborative System to Organize a Collection using Facets. *In*: Rethinking Electronic Publishing: Innovation in Communication Paradigms and Technologies - **Proceedings** of the 13th International Conference on Electronic Publishing, 187-203.ELPUB. Milano, Italy, 2009.

NATIONAL INFORMATION STANDARDS ORGANIZATION. **NISO TR02**: guidelines for indexes and related information retrieval devices. Estados Unidos: NISO, 1997.

NATIONAL INFORMATION STANDARDS ORGANIZATION. **Understanding Metadata**. Bethesda, MD: NISO Press, 2004. Disponível em: <http://www.niso.org/standards/resources/UnderstandingMetadata.pdf>. Acesso em: 26 out. 2021.

NEXTEWS. **Jak utworzyć rozwijane menu CSS**, 2021. Disponível em: <http://nextews.com/0a090e19/>. Acesso em: 21 out. 2021.

NIELSEN, J. **Usability 101**: Introduction to Usability. Useit.com. 2003. Disponível em: <http://www.useit.com/alertbox/20030825.html>. Acesso em: 19 out. 2021.

NORUZI, A. Folksonomies: Why do we need controlled vocabulary? **Webology**, v.4, n.2, 2007. Disponível em: <http://www.webology.org/2007/v4n2/editorial12.html>. Acesso em: 20 maio 2021.

OREN, E.; DELBRU, R.; DECKER, S. Extending faceted navigation for RDF data. *In*: 5th International Semantic Web Conference, 2006. **Proceedings...** ISWC, 2006. Disponível em: <http://iswc2006.semanticweb.org/items/Oren2006nx.pdf>. Acesso em: 26 maio 2021.

QUINTARELLI, E.; RESMINI, A.; ROSATI, L. FaceTag: Integrating Bottom-up and Top-down Classification in a Social Tagging System. **Bulletin of the American Society for Information Science and Technology**, p. 10-15, jun./jul. 2007.

REIS, G. **Arquitetura da Informação de websites**. 2007. Disponível em: <http://www.guilhermo.com/apresentacoes/arquiteturainformacao-fesa-guilhermoreis.pdf>. Acesso em: 20 maio 2021.

ROBREDO, J; CUNHA, M. B.. **Documentação de hoje e de amanhã**: uma abordagem informatizada da biblioteconomia e dos sistemas de informação. São Paulo: Global, 1994. 400 p.

SANTARÉM SEGUNDO, J. E. **Representação Iterativa**: um modelo para repositórios digitais. Marília, SP, 2010. 224f. Tese (Doutorado) – Curso de Ciência da Informação, Universidade Estadual Paulista Júlio de Mesquita Filho, Marília, SP, 2010. Disponível em: <https://repositorio.unesp.br/handle/11449/103346>. Acesso em: 20 maio 2021.

SANTAREM SEGUNDO, J. E.; SIQUEIRA, C. S. Aplicación teórica-conceptual de folksonomías asistidas para La recuperación de información. **Scire**, Zaragoza, v. 19, n. 2, p.77-82, 2013.

SANTARÉM SEGUNDO, J. E.; VIDOTTI, S. A. B. G. Representação Iterativa e folksonomia assistida para repositórios digitais. **Liinc em Revista**, Rio de Janeiro, v.7, n.1, mar. 2011, p. 283-300. Disponível em: <http://revista.ibict.br/liinc/article/view/3300>. Acesso em: 20 maio 2021.

SANTAREM SEGUNDO, J. E.; VIDOTTI, S. A. B. G. Rede de tags para recuperação da informação no contexto da Representação Iterativa. **InCID: Revista de Ciência da Informação e Documentação**, [S. l.], v. 2, n. 1, p. 86-109, 2011. DOI: 10.11606/issn.2178-2075.v2i1p86-109. Disponível em: <https://www.revistas.usp.br/incid/article/view/42336>. Acesso em: 15 abr. 2021.

SANTOS, R. F.; CORRÊA, R. F. Modelos Colaborativos de Indexação Social e sua aplicabilidade em Bibliotecas Digitais. **Liinc em Revista**, v.11, p.273-286, 2015.

SANTOS, R. F.; CORRÊA, R. F. A folksonomia e a representação colaborativa da informação em ambientes digitais. **Tendências da Pesquisa Brasileira em Ciência da Informação**, v.8, n.1, jan./jun. 2015. Acesso em: 20 maio 2021.

SANTOS, R. F. dos. **Modelos colaborativos de indexação social e sua aplicabilidade na Base de Dados Referencial de Artigos e Periódicos em Ciência da Informação (BRAPCI)**. Recife, 2016. Dissertação (Mestrado) – Universidade Federal de Pernambuco, Centro de Artes e Comunicação, Ciência da Informação. 2016.

SANTOS, R. F. D.; CORRÊA, R. F. Modelos colaborativos de indexação: aspectos conceituais e elementos constituintes. **Encontro Nacional de Pesquisa em Ciência da Informação**, n. XX ENANCIB, 2019. Disponível em: <<http://hdl.handle.net/20.500.11959/brapci/122633>>. Acesso em: 15 abr. 2021.

SANTOS, R. F. D.; CORRÊA, R. F. **Organização da informação em repositórios digitais**: uma abordagem sobre a política de indexação da base de dados referencial de artigos de periódicos em ciência da informação (brapci). 2017, p. 249-261, . Disponível em: <http://hdl.handle.net/20.500.11959/brapci/122077>. Acesso em: 15 abr. 2021.

SAYÃO, L. F. Padrões para bibliotecas digitais abertas e interoperáveis. 2007. **Encontros Bibli: Revista eletrônica de Biblioteconomia e Ciência da informação**, 12(1), 18-47. Disponível em: <https://periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2007v12n1esp1p18>. Acesso em: 15 abr. 2021.

SAYÃO, L. F. Modelos teóricos em ciência da informação; abstração e método científico. **Ciência da Informação**, Brasília, v. 30, n. 1, p. 82-91, jan./abr. 2001.

SIGURBJÖRNSSON, B., ZWOL, R. **Tagexplorer**: Faceted browsing of flickr photos. Yahoo! Research, 2010.

SILVA, M. F. **Proposta de modelo de colaboração para catálogo web facetado**. 269f. Tese (Doutorado em Ciência da Informação) – Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2013.

SILVA, B. F. M.; CORRÊA, R. F. Aplicação da folksonomia assistida na construção de corpus de referência em ciência da informação. **Em Questão**, v. 26, n. 2, p. 413-436, 2020. Acesso em: 26 maio 2021.

SILVA, B. F. M. **A folksonomia assistida aplicada na construção de Corpus de Referência para avaliação de sistemas de indexação automática**. 189f. Dissertação (Mestrado em Ciência da Informação). Universidade Federal de Pernambuco, 2019.

SILVA, F. C. C.; SILVEIRA, L. O ecossistema da Ciência Aberta. **Transinformação**, Campinas, v. 31, e190001, 2019. Disponível em: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0103-37862019000100302&lng=en&nrm=iso. Acesso em: 22 maio 2021.

SPITERI, L. **Faceted navigation of social tagging applications**. Dalhousie Repository. Halifax: Dalhousie University, 2011. Disponível em: <http://dalspace.library.dal.ca/handle/10222/14297>. Acesso em: 30 maio 2021.

TZITZIKAS, Y. *et al.* Extended Faceted Taxonomies for Web Catalogs. *In: Web Information Systems Engineering - WISE, Third International Conference, 2002. Proceedings...* IEEE, 2002. p. 192-201.

VEIGA, V.; MACENA, L. G. O autoarquivamento nos repositórios institucionais brasileiros: um estudo exploratório. **PontodeAcesso**, Salvador, v. 9, n. 3, p. 35-47, dez. 2015. Disponível em: <https://portalseer.ufba.br/index.php/revistaici/article/view/15107>. Acesso em: 26 mar. 2021.

VIANA, J. M. A.. **Representação colaborativa de dados científicos: estudo na rede de repositórios de dados científicos do estado de São Paulo**. 2020. Dissertação (Mestrado em Ciência da Informação) – Universidade Federal de São Carlos, São Carlos, 2020. Disponível em: <https://repositorio.ufscar.br/handle/ufscar/13027>. Acesso em: 24 maio 2021.

WAL, T. V. **Folksonomy definition and wikipedia**. 2005. Disponível em: <http://www.vanderwal.net/random/entrysel.php?blog=1750>. Acesso em: 18 jun. 2021.

WEITZEL, S. R. O papel dos repositórios institucionais e temáticos na estrutura da produção científica. **Em questão**, v. 12, n. 1, p. 51-71, 2006. Disponível em: <https://seer.ufrgs.br/EmQuestao/article/view/19/7>. Acesso em: 26 mar. 2021.

WEITZEL, S. R. MESQUYUSA, M. A. A. Preservação digital em repositórios institucionais: práticas na região Sudeste do Brasil. **Liinc em revista**, Rio de Janeiro, v. 11, n. 1, p. 181-196, maio 2015. Disponível em:

<https://revista.ibict.br/liinc/article/view/3618/3087>. Acesso em: 27 mar. 2021.

XAVIER R. F. Evolução das plataformas de acesso aberto brasileiras: propriedades e perspectivas. **Ci. Inf.**, Brasília, DF, v.48 n.3 (Supl.), p.497-500, set./dez. 2019. Disponível em: <http://revista.ibict.br/ciinf/article/view/4949/4469>. Acesso em: 27 mar. 2021.

YANG, S. Q.; HOFMANN, M. A. The next generation library catalog: A comparative study of the OPACs of Koha, Evergreen, and Voyager. **Information Technology and Libraries**, v. 29, n. 3, p. 141-150, 2013.

YOO, D. et al. Building and evaluating a collaboratively built structured folksonomy. **Journal of Information Science**, [S.l.], v.39, n.5, p.593-607, 2013.

APÊNDICE A

LISTA DE REFERÊNCIAS DOS ESTUDOS IDENTIFICADOS

DEVADASON, F. J. et al. Faceted Indexing Based System for Organizing and Accessing Internet. *Knowledge Organization*, v. 29, p. 65-77, 2002. Disponível em: <<http://www.reocities.com/devadason.geo/FactDSIS.pdf>>. Acesso em: 17 nov. 2021.

DEVADASON, F. J. et al. Faceted Indexing Based System for Organizing and Accessing Internet. **Knowledge Organization**, v. 29, p. 65-77, 2002. Disponível em: <<http://www.reocities.com/devadason.geo/FactDSIS.pdf>>. Acesso em: 20 mar. 2021.

LAI, P. **Facette**: using facets to improve tag-based bookmarking. 2009. 52f. Dissertação (Master of Engineering in Electrical Engineering and Computer Science) - Massachusetts Institute of Technology, 2009.

MALY, K.; WU, H.; ZUBAIR, M. A collaborative faceted categorization system user interactions. In: 14th International Conference on Electronic , 2010, Helsinki, Finland **Proceedings...** Transforming the Nature of Communication: Turid Hedlund and Yasar Tonta, 2010. Disponível em: http://elpub.scix.net/data/works/att/106_elpub2010.content.pdf. Acesso em: 26 maio. 2021.

MALY, K; WU, H; ZUBAIR, M. Automated Support for a Collaborative System to Organize a Collection using Facets. *In: Rethinking Electronic Publishing: Innovation in Communication Paradigms and Technologies - Proceedings of the 13th International Conference on Electronic Publishing*, 187-203. ELPUB. Milano, Italy, 2009.

QUINTARELLI, E.; RESMINI, A.; ROSATI, L. FaceTag: Integrating Bottom-up and Top-down Classification in a Social Tagging System. **Bulletin of the American Society for Information Science and Technology**, p. 10-15, jun./jul. 2007.

SANTARÉM SEGUNDO, J. E. **Representação Iterativa**: um modelo para repositórios digitais. Marília, SP, 2010. 224f. Tese (Doutorado) – Curso de Ciência da Informação, Universidade Estadual Paulista Júlio de Mesquita Filho, Marília, SP, 2010. Disponível em:

<https://repositorio.unesp.br/handle/11449/103346>. Acesso em: 20 maio 2021.

SANTAREM SEGUNDO, J. E.; VIDOTTI, S. A. B. G. Rede de tags para recuperação da informação no contexto da Representação Iterativa. **InCID: Revista de Ciência da Informação e Documentação**, [S. l.], v. 2, n. 1, p. 86-109, 2011. p. 86-109. Disponível em: <https://www.revistas.usp.br/incid/article/view/42336>. Acesso em: 15 abr. 2021.

SANTARÉM SEGUNDO, J. E.; VIDOTTI, S. A. B. G. Representação Iterativa e folksonomia assistida para repositórios digitais. **Liinc em Revista**, Rio de Janeiro, v.7, n.1, mar. 2011, p. 283-300. Disponível em: <http://revista.ibict.br/liinc/article/view/3300>. Acesso em: 20 maio 2021.

SANTOS, R. F.; CORRÊA, R. F. Modelos Colaborativos de Indexação Social e sua aplicabilidade em Bibliotecas Digitais. **Liinc em Revista**, v.11, p.273-286, 2015.

SANTOS, R. F.; CORRÊA, R. F. A folksonomia e a representação colaborativa da informação em ambientes digitais. **Tendências da Pesquisa Brasileira em Ciência da Informação**, v.8, n.1, jan./jun. 2015. Acesso em: 20 maio 2021.

SANTOS, R. F. dos. **Modelos colaborativos de indexação social e sua aplicabilidade na Base de Dados Referencial de Artigos e Periódicos em Ciência da Informação (BRAPCI)**. Recife, 2016. Dissertação (Mestrado) – Universidade Federal de Pernambuco, Centro de Artes e Comunicação, Ciência da Informação. 2016.

SANTOS, R. F. D.; CORRÊA, R. F. Modelos colaborativos de indexação: aspectos conceituais e elementos constituintes. **Encontro Nacional de Pesquisa em Ciência da Informação**, n. XX ENANCIB, 2019. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/122633>. Acesso em: 15 abr. 2021.

SIGURBJÖRNSSON, B., ZWOL, R. **Tagexplorer**: Faceted browsing of flickr photos. Yahoo! Research, 2010.

SILVA, M. F. Proposta **de modelo de colaboração para catálogo web facetado**. 269f. Tese (Doutorado em Ciência da Informação) – Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2013.

SILVA, B. F. M.; CORRÊA, R. F. Aplicação da folksonomia assistida na construção de corpus de referência em ciência da informação. **Em Questão**, v. 26, n. 2, p. 413-436, 2020. Acesso em: 26 maio 2021.

SILVA, B. F. M. **A folksonomia assistida aplicada na construção de Corpus de Referência para avaliação de sistemas de indexação automática**. 189f. Dissertação (Mestrado em Ciência da Informação). Universidade Federal de Pernambuco, 2019.

W3C CONSORTIUM. Web Semântica. 2011. Disponível em: <https://www.w3c.br/Padroes/WebSemantica>. Acesso em: 17 nov. 2021.

YOO, D. et al. Building and evaluating a collaboratively built structured folksonomy. **Journal of Information Science**, [S.l.], v.39, n.5, p.593-607, 2013.