



**Dos padrões internacionais de estruturação da
informação de pesquisa aos indicadores:
primeira incursão na temática**

Código do Projeto/Ação: TICS0114

Convocatória: 9ª Convocatória

Diálogo: Sociedade da Informação

Editores

Maria de Nazaré Freitas Pereira

Hélia de Sousa Chaves

Ronaldo Ferreira Araújo

Brasília, DF
2019

Instituto Brasileiro de Informação em Ciência e Tecnologia - IBICT

Diretoria

Cecília Leite Oliveira

Coordenação Geral de Pesquisa e Desenvolvimento de Novos Produtos (CGNP)

Arthur Fernando Costa

Coordenação Geral de Pesquisa e Manutenção de Produtos Consolidados (CGPC)

Bianca Amaro

Coordenação Geral de Tecnologias de Informação e Informática (CGTI)

Marcos Pereira de Novais

Coordenação de Ensino e Pesquisa, Ciência e Tecnologia Da Informação (COEPPE)

Lena Vania Ribeiro Pinheiro

Coordenação de Planejamento, Acompanhamento e Avaliação (COPAV)

José Luis dos Santos Nascimento

Coordenação de Administração (COADM)

Reginaldo de Araújo Silva

Seção de Editoração

Ramón Martins Sodoma da Fonseca

Delegação da União Europeia no Brasil - DELBRA

Embaixador

João Gomes Cravinho

Ministro Conselheiro, Chefe do Setor de Sociedade da Informação e Mercado Digital

Carlos Oliveira

Primeira Secretária - Chefe do Sector FPI - Regional Team Americas

Maria Rosa Sabbatelli

Coordenador da Iniciativa de Apoio aos Diálogos Setoriais UE-Brasil

Costanzo Fisogni

CESO Development Consultants/ WYG/Camões, I.P

Consórcio Executor





**Dos padrões internacionais de estruturação da
informação de pesquisa aos indicadores:
primeira incursão na temática**

**Código do Projeto/Ação: TICS0114
Convocatória: 9ª Convocatória
Diálogo: Sociedade da Informação**

Editores

Maria de Nazaré Freitas Pereira

Hélia de Sousa Chaves

Ronaldo Ferreira Araújo

Brasília, DF

2019





2019 Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict)
Os autores são responsáveis pela apresentação dos fatos contidos e opiniões expressas nesta obra.

Equipe técnica

Editores

Maria de Nazaré Freitas Pereira
Hélia de Sousa Chaves
Ronaldo Ferreira Araújo

Revisão Editorial

Ana Lúcia Mendes

Normalização de referências

Juliana Pinheiro Farias
Kelen Cândida Vieira Bomfim
Lilian Maria Thomé Andrade Brandão

Revisão gramatical

Margaret de Palermo Silva

Projeto Gráfico

Ramón Martins Sodoma da Fonseca/SEEDIT

Capa

Rodrigo Azevedo Moreira – Capa

Tradução

Michael Allan

D 722

Dos padrões internacionais de estruturação da informação de pesquisa aos indicadores : primeira incursão na temática / editores: Maria de Nazaré Freitas Pereira, Hélia de Sousa Chaves, Ronaldo Ferreira Araújo
[tradução de] Allan Michael. – Brasília, D.F. : Instituto Brasileiro de Informação em Ciência e Tecnologia : Diálogos Setoriais União Europeia Brasil, 2019.
220p. : il.

Inclui bibliografia.
ISBN 978-85-7013-153-9

1. Normas e padrões. 2. Representação da informação. 3. Indicadores de C&TI. 4. Sistemas de informação de pesquisas. I. Pereira, Maria de Nazaré Freitas (ed.). II. Chaves, Hélia de Sousa (ed.). III. Araújo, Ronald Ferreira (ed.). IV. Allan, Michael (trad.).

CDU: 001.891(083.7)

Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict)

Setor de Autarquias Sul (SAUS)
Quadra 05, Lote 06, Bloco H – 5º Andar
Cep: 70070-912 – Brasília, DF
Telefones: 55 (61) 3217-6360 / 55 (61) 3217-6350
www.ibict.br

Rua Lauro Muller, 455 - 4º Andar - Botafogo
Cep: 22290-160 – Rio de Janeiro, RJ
Telefones: 55 (21) 2275-0321
Fax: 55 (21) 2275-3590
<http://www.ibict.br/capacitacao-e-ensino/pos-graduacao-em-ciencia-da-informacao>
<http://www.ppgci.ufirj.br>



A Anaiza, Cecília, Arthur e Hélia e aos demais funcionários e bolsistas do IBICT e das terceirizadas que me acolheram despedaçada pela partida da minha filha Patrícia (*in memoriam*), restaurando condições afetivas e intelectuais, vitais para pensar e produzir em grupo de antepassados e contemporâneos o projeto do coração.

Maria de Nazaré (Nazinha)





Comissão Avaliadora

Eny Marcelino de Almeida Nunes

Mestre em Ciência da Informação pela Universidade de Brasília (UnB). Especialização em Administração em Sistemas de Informação pela Universidade Católica de Brasília. Tecnologista Sênior do Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT)/Ministério da Ciência, Tecnologia, Inovações e Comunicações (MCTIC), Brasília, DF. Currículo Lattes: <http://lattes.cnpq.br/5582608438159648>
E-mail: enymarcelino@gmail.com

Maria de Nazaré Freitas Pereira

Doutora em Ciências Sociais pelo Instituto Universitário de Pesquisas do Rio de Janeiro (IUPERJ), Rio de Janeiro, RJ. Consultora em Sistemas de Informação de Pesquisa e Comunicação Científica. Currículo Lattes: <http://buscatextual.cnpq.br/buscatextual/visualizacv.do?id=K4787118P6>
E-mail: marianazare@uol.com.br

Ronaldo Ferreira Araújo

Doutor em Ciência da Informação pela Universidade Federal de Minas Gerais (UFMG) – Belo Horizonte, Minas Gerais - Brasil. Professor da Universidade Federal de Alagoas (UFAL) - Maceió, AL. Currículo Lattes: <http://buscatextual.cnpq.br/buscatextual/visualizacv.do?id=K4184315Y9>
E-mail: ronaldfa@gmail.com



Sumário

Apresentação	10
Ministro Conselheiro Carlos Oliveira	

Apresentação	12
Cecília Leite	

Prefácio: Camadas de CRIS	13
Maria de Nazaré Freitas Pereira	

Parte I | Artigos

Desafios para a consolidação de indicadores em Ciência, Tecnologia e Inovação no Brasil: reflexões iniciais para a estruturação do Piloto BR-CRIS IBICT-FAPEAL	39
Victor Lemos Tenorio	
Francisca Rosaline Leite Mota	
Hans Ponfick Aragão	
Ronaldo Ferreira Araújo	

Classificação e gestão do conhecimento	60
Hagar Espanha Gomes	

Os estudos cibernéticos da informação: indicadores de estruturas web e de recursos da web social	93
Ronaldo Ferreira de Araújo	

Parte II | Sumários Executivos

Sobre a “Classification de la Recherche Universitaire” (Tabela CASRAI)	115
Hagar Espanha Gomes	

Sobre o estudo de comparação entre a Tabela do CNPq e a Tabela CASRAI	123
Hagar Espanha Gomes	



Thoughts on the Classification of University Research (CASRAI Table)	133
Hagar Espanha Gomes Tradução: Michael Allan	
Thoughts on the comparative study of the CNPq and CASRAI tables	141
Hagar Espanha Gomes Tradução: Michael Allan	
Parte III Instrumentos de Pesquisa	
Bibliografia anotada: identificadores persistentes (identificadores ORCID e RINGGOLD)	153
Juliana Pinheiro Farias Kelen Cândida Vieira Bomfim	
Parte IV Relatórios	
Relatório de viagem	177
Paulo Henrique de Assis Santana	
Piloto BR CRIS IBICT-FAPEAL: uma avaliação exploratória do DSpace-CRIS para a construção de um sistema de informação de pesquisa local	190
Kleber de Barros Alcanfôr Hélia de Sousa Chaves Lilian Maria Thomé Andrade Brandão Rogério Mendes Castilho	
Sobre os autores	216



Apresentação

As atividades de pesquisa e inovação estão no centro das prioridades da União Europeia. O conhecimento gerado por essas atividades é fulcral para a criação de novos produtos e serviços, promover o investimento, melhorar a competitividade da economia e criar empregos, contribuindo de forma decisiva para o bem-estar dos cidadãos.

A orientação estratégica da política de pesquisa e inovação da União Europeia foi definida em 2015 pelo comissário Carlos Moedas, com o lema “Inovação Aberta, Ciência Aberta, Aberta para o Mundo” (no original em inglês “Open Innovation, Open Science, Open to the World”). Essas palavras traduzem uma mudança importante no paradigma tradicional de produção e disseminação do conhecimento científico.

A União Europeia agrega ainda hoje o grupo de países que, de acordo com os indicadores bibliométricos tradicionais, lideram a produção de ciência no mundo. Todavia, para que esse investimento na produção de conhecimento seja plenamente frutuoso, é necessário ter presente que, resultado da disseminação das tecnologias digitais, a inovação se desenvolve através de um processo dinâmico de interação entre vários atores, agregados em redes de colaboração e ecossistemas de inovação, que está muito além da simplicidade linear da transferência de informação dos centros de pesquisa para as empresas – esta é a essência da abordagem Inovação Aberta (Open Innovation).

Desenvolvimentos como o Big Data, os computadores que permitem analisar grandes volumes de dados, bem como as possibilidades de colaboração oferecidas pelas tecnologias digitais, estão a mudar os paradigmas de produção de conhecimento científico. Trata-se de facto de uma nova forma de explorar o potencial conhecimento científico resultante da mobilização de um conjunto mais vasto de atores e da forma como o conhecimento é partilhado entre esses atores, envolvendo igualmente uma visão mais pluridisciplinar sobre a pesquisa científica. Entre os conceitos associados ao paradigma da Ciência Aberta (Open Science) estão o Acesso Aberto (Open Access) e o Código Aberto (Open Source), que substanciam precisamente essa dimensão colaborativa da criação de conhecimento.

A terceira dimensão da estratégia da União Europeia diz respeito precisamente a uma ciência Aberta ao Mundo (Open to the World). É em grande medida o corolário natural das grandes transformações que são ilustradas pelo facto de vivermos hoje num mundo cada vez mais multipolar, em que empresas e instituições que operam em escala global têm papel determinante na criação de conhecimento e de riqueza. É por isso essencial facultar os instrumentos e o enquadramento que facilitem e dinamizem essa colaboração.

A dimensão internacional do programa de pesquisa e inovação da União Europeia “Horizonte 2020”, que está aberto à participação de pesquisadores de todo o mundo, constitui um exemplo paradigmático do potencial colaborativo da ciência. O Brasil é um dos países com os quais essa cooperação é mais intensa e profícua.

A cooperação com o Brasil, ilustrada no presente projecto de estruturação da informação sobre pesquisa científica liderado pelo IBICT, vem dar um pequeno, mas significativo passo nesse processo de integração da ciência brasileira às redes globais de criação de conhecimento científico.

A União Europeia tem em curso um ambicioso programa de criação de um Espaço Europeu de Investigação (European Research Area), que visa melhorar a eficiência e a eficácia dos sistemas nacionais de pesquisa, facilitando a troca de ideias, a mobilidade dos pesquisadores e a circulação de ideias, num desígnio que tem como propósito último o desenvolvimento econômico e social.

Gostaríamos de ver o Brasil liderar um processo semelhante em escala regional, tirando partido do seu papel especial enquanto economia mais avançada e desenvolvida da América Latina. Estariam assim criadas condições para imaginar um grande espaço aberto de pesquisa unindo União Europeia, Brasil e outros países da América Latina.

Ministro Conselheiro Carlos Oliveira

Chefe do Setor de Sociedade da Informação e Mídia
Delegação da União Europeia no Brasil

Apresentação

O Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT), unidade de pesquisa do Ministério da Ciência, Tecnologia, Inovações e Comunicações (MCTIC) tem atuado no desenvolvimento de pesquisas, serviços e produtos de informação, com expertise na criação de sistemas de informação que incorporam padrões internacionais e promovem a interoperabilidade com outros sistemas de instituições parceiras. Com base nessa experiência e à luz de sua missão de “promover a competência, o desenvolvimento de recursos e a infraestrutura de informação em ciência e tecnologia para a produção, socialização e integração do conhecimento científico e tecnológico”, o IBICT identificou a importância da integração de sistemas sobre a atividade de pesquisa em ciência e tecnologia no País e de sua posição como líder dessa ação.

Assim, o IBICT vem investindo, desde 2014, na realização de estudos, na articulação de parcerias nacionais e internacionais e na promoção de eventos de capacitação e de tomada de decisão, no sentido de reunir esforços para a construção de um Sistema de Sistemas (SoS) da Pesquisa Brasileira, nos moldes de sistemas do tipo CRIS (do inglês Current Research Information Systems). Uma dessas ações foi a submissão de um projeto, intitulado “Dos padrões internacionais de estruturação da informação de pesquisa aos indicadores”, à Iniciativa Diálogos União Europeia Brasil (Diálogos Setoriais), no âmbito da 9ª Convocatória de 2017, no tema Sociedade da Informação.

Esta coletânea, que ora apresento, é uma das contrapartidas do IBICT ao Projeto mencionado acima, o qual deu origem a ações que estreitaram a cooperação entre entidades brasileiras e europeias, por meio da troca de experiências entre especialistas de ambas as partes, notadamente nas missões dos brasileiros a Portugal e dos portugueses ao Brasil.

O conteúdo da coletânea e a vinculação institucional dos autores atestam os avanços alcançados pelo Projeto, por meio de estudos realizados por peritos oriundos de diversas especialidades do campo da Ciência da Informação e da Tecnologia da Informação e Comunicação. A riqueza deste trabalho está espelhada nos estudos sobre tabelas nacionais e internacionais de classificação – CNPq, Capes, CASRAI, OCDE –, gestão do conhecimento, indicadores de CT&I, cibermetria e ferramentas e padrões para a construção do piloto BRCRIS IBICT-FAPEAL.

Com os agradecimentos deste Instituto a todos que se debruçaram na construção desta Coletânea, desejo que este trabalho colabore para a geração de novos conhecimentos e seja uma boa leitura aos interessados na temática.

Cecilia Leite

Diretora do Instituto Brasileiro de Informação em
Ciência e Tecnologia

CAMADAS DE CRIS

*Estranha é a nossa situação aqui na Terra.
Cada um de nós chega para uma breve visita,
sem saber por quê,
parece que às vezes por um propósito divino.
Do ponto de vista da vida cotidiana, porém,
existe uma coisa que sabemos de fato:
que estamos aqui pelo bem dos outros.*

Albert Einstein

Dos padrões internacionais de estruturação da informação de pesquisa aos indicadores é título de projeto submetido à chamada Diálogos Setoriais, da 9ª Convocatória no tema Sociedade da Informação, em julho de 2017. Nesta coletânea, contrapartida do IBICT ao projeto, encontra-se acrescido de um esclarecimento: primeira incursão na temática.

O foco é nos sistemas CRIS, sigla em inglês para Current Research Information System. A definição da euroCRIS considera CRIS “(...) qualquer ferramenta informativa dedicada a fornecer acesso e divulgar informações de pesquisa.”¹

Data de 2014 as primeiras incursões do IBICT na temática CRIS, inicialmente comissionando estudos e pesquisas sobre o tema. Os trabalhos foram posteriormente apresentados e debatidos em reuniões internas e workshops com parceiros nacionais e estrangeiros, de forma ampla, agregando-se cada vez mais conteúdos e orientações pontuais no campo político e operacional.

Em um desses eventos, o euroCRIS Days/Jornada euroCRIS, realizado no IBICT, Brasília, nos dias 16 a 18 de novembro de 2017, a Fundação de Amparo à Pesquisa do Estado de Alagoas (FAPEAL) colocou publicamente sua candidatura a projeto piloto, condição aprovada pelos presentes.

Com essa motivação em mente, apresentou-se projeto aos Diálogos Setoriais, cujos resultados no campo operacional capitalizam a partir da expertise da Fundação para a Ciência e a Tecnologia e da Universidade do Minho/Portugal com projetos de CRIS.

A execução de um projeto de pesquisa caminha da desordem para a ordem. Na comunicação científica, os primeiros passos são dados com anotações em cadernetas de laboratório, breves comunicações em periódicos científicos (servem para marcar a presença do pesquisador no campo! e, portanto, a precedência), extensos relatórios aos financiadores, dissertações, teses..., até que chega a hora de vir a público na forma de artigo, patente... Essas atividades científicas agenciam diferentes atores e demandam deles ações de informação compreendidas nos processos e fluxos da construção de fatos, artefatos e resultados de pesquisa. Quando organizados, eles se constituem em sistemas CRIS.

¹ euroCRIS. Disponível em: <<https://www.eurocris.org/>>. Acesso em: 20 de agosto de 2018.

A incursão na temática aqui proposta se apresenta na forma de estudos: artigos, sumários executivos, ferramenta de pesquisa e relatórios; independentemente da forma, estão nos limites do tema desta coletânea.

Como todo prefácio, este apresenta a obra, dialoga com o autor, esperando estimular o gosto do leitor. E sempre que há deixas, complementa o autor.

Os artigos, em número de três, discorrem sobre aspectos que estão no coração da estruturação da informação de pesquisa.

Dois autores entram em campo com os indicadores, requisitos para serem estabelecidos, ferramentas que permitem extraí-los e sua abrangência em tempos de redes sociais. E o terceiro foca nas classificações, ferramenta de trabalho sem a qual não é possível pensar em indicadores, desde os mais simples, que se originam dos metadados e seus atributos, aos mais complexos, que dependem de classificações de campos disciplinares, assuntos de pesquisa, áreas de aplicação e tantas outras dimensões.

Victor Lemos Tenório, Francisca Rosaline Leite Mota, Hans Ponfick Aragão e Ronaldo Ferreira Araújo, autores do primeiro artigo, têm vínculo com a Fundação de Amparo à Pesquisa do Estado de Alagoas (FAPEAL), parceira do IBICT para desenvolvimento de um piloto de CRIS.

O artigo, Desafios para a consolidação de indicadores em ciência, tecnologia e inovação no Brasil: reflexões iniciais para a estruturação do projeto piloto BRICRIS IBICT-FAPEAL, tem como objetivo estudar as fontes de gestão da informação de pesquisa.

Mas, antes disso, os autores se cercam de um conjunto de referências para falar de indicadores de desempenho científico e de esforço tecnológico.

Uma das citações mostra que o Brasil comparado com países em desenvolvimento é

[...] significativamente diferenciado em dimensões como escala de sua economia, inserção externa, tecnologia, qualificação da força de trabalho, investimento em pesquisa e desenvolvimento (P&D), cooperação técnica e produção científica.²

Esse diagnóstico leva Victor e colaboradores a argumentarem que um sistema de informação de pesquisa, estruturado em torno da pesquisa aqui financiada, tem papel fundamental a cumprir.

² ARBIX, G.; NEGRI, J. A.; VERMULM, R. apud TENORIO, V. L.; MOTA, F. R. L.; ARAGÃO, H.P.; ARAÚJO, R. F. Desafios para a consolidação de indicadores em Ciência, Tecnologia e Inovação no Brasil: reflexões iniciais para a estruturação do Piloto BRICRIS IBICT-FAPEAL. In: PEREIRA, M. de N. F.; CHAVES, H. de S.; ARAÚJO, R. F. de (Eds.). **Dos padrões internacionais de estruturação da informação de pesquisa aos indicadores: primeira incursão na temática.** Brasília: IBICT, DELBRA, 2019. p. 44.

Com efeito, os autores por eles citados, todos ligados à FINEP, afirmam que a inovação aqui se dá por meio da “[...] geração de conhecimento que é suportado por centros de excelência de universidades que conseguem acompanhar, quando não se conectar, com a pesquisa avançada em nível mundial.”³

Sistemas CRIS têm papel estratégico a desempenhar na geração de riqueza de um país. Para os autores das citações acima, é exatamente na expansão desse núcleo dinâmico, na sua “[...] diversificação e integração com o conjunto da economia, tecendo malhas entre grandes, médias e pequenas empresas, que está a aposta no futuro de uma economia vibrante.”⁴

Após a discussão de indicadores de desenvolvimento econômico, Victor e colaboradores relatam os resultados da modelagem de processos em notação BPMN (Business Process Modeling Notation) e o diagnóstico do fluxo informacional dos processos de submissão e acompanhamento de projetos em C&T na FAPEAL, primeira etapa do piloto de CRIS.

Dos resultados desse primeiro artigo, em seus aspectos de mapeamento das fontes, chama atenção, quando do acompanhamento da prestação de contas do projeto de pesquisa, o uso de sistema de gestão que combina registro da informação em meios eletrônico e físico (suporte em papel). Esse último está associado à entrega de resultados (publicações) decorrentes de financiamento.

³ Ibidem.

⁴ Ibidem.

Isso dificulta sua incorporação em repositórios institucionais, se considerarmos que nem todo resultado de pesquisa é dissertação ou tese, com registro obrigatório em sistema de avaliação da pós. Cabe pontuar o uso de extensos relatórios, com farta documentação sobre o projeto como um tipo muito comum de resultado, sem representação em meios eletrônicos oficiais.

No segundo artigo, Classificação e gestão do conhecimento, Hagar Espanha Gomes, consultora do IBICT para o BRICRIS, descortina a organização do conhecimento, seja ele científico ou cultural, em inúmeras camadas.

O artigo, com dezenas de páginas, enfoca o tema da classificação com finalidades que vão desde organizar bibliografias e bibliotecas até a gestão do conhecimento, ocasião em que trata dos indicadores sob o ponto de vista dos sistemas de classificação que possibilitam sua extração.

O que se depreende das densas análises teóricas e de alinhamento de fatos históricos, ao longo dos tempos, é que sem classificação não é possível haver organização do conhecimento para fins de ‘re-uso’ e acesso.

No tópico sobre Classificação e Bibliotecas, ficamos sabendo que quando do confisco das bibliotecas após a Revolução Francesa, com seu objetivo de “educação para todos”, as bibliotecas nomeadas públicas eram ainda de acesso fechado e seu acervo bastante erudito para a massa da população ainda iletrada.

Apesar disso, a classificação de Louise Brunet foi adotada como modelo para o 'sistema' das bibliotecas confiscadas.

Pode-se supor que o acesso não ocorreu de imediato. Não pela falta de catálogo, a sua época impresso, com os livros classificados por assunto, mas por ausência de pensamento de inclusão social ou de desenvolvimento sustentável e das decorrentes políticas públicas.

Em tempos de acesso aberto ao texto completo, articular uma política de acesso aberto aos conteúdos pode permitir a inclusão social de usuários outros que não os pesquisadores e os eruditos. Configurar-se-ia, assim, uma política de duplo acesso: ao conteúdo e ao texto. Outra implicação do acesso é na classificação para acervos não de bibliotecas públicas, mas de coleções que exigem especialistas intelectual e tecnologicamente mais sofisticados para organizá-las.

Não se trata do acesso, mas de sua estruturação/organização, conforme aspectos da classificação previamente adotada. Essa ferramenta, hoje, se exprime também em metadados, indexações temáticas e derivação de indicadores. Isso nos coloca no campo da semântica, do significado, do entendimento e da estatística com suas fórmulas. Definir antes e atribuir depois representam a chave da estabilidade!

Voltando ao texto de Hagar, vejamos o que se pode destacar de precioso, quando ela desdobra na Bibliografia o tema da Classificação, sob o ponto de vista das iniciativas internacionais.

Nesse tópico, cuja figura central é Paul Otlet, tudo nos interessa, pois nele podemos encontrar os fundamentos do UNISIST (1971) que deita as raízes do que viria a se chamar CRIS, conforme argumentaremos mais adiante. O mundo do conhecimento transcende as quatro paredes dos lugares que o abrigam, como as bibliotecas, os livros, as escolas, as universidades, os laboratórios de pesquisa⁵.

Se pensarmos bem, essa condição de conhecimento é a mesma gerada pela integração de diferentes sistemas CRIS.

Otlet acreditava na comunicação intelectual, na mudança social e na paz mundial pela via da informação. Consequentemente, uma de suas definições mais simbólicas, citada no texto de Hagar, afirma que

A Bibliografia é o primeiro meio de organizar as relações internacionais. Reunir as ideias dos homens, fazer uso por uns dos trabalhos dos outros. O Repertório bibliográfico universal é uma bolsa intelectual possível graças a um entreposto intelectual. A oferta e a demanda se encontram.⁶

⁵ PEREIRA, M. de N. F. Prefácio que esclarece o leitor a propósito do sonho de Otlet: aventura em tecnologia da informação e comunicação. *In*: PEREIRA, M. de N. F.; PINHEIRO, L. V. R. (orgs.).

O sonho de Otlet: aventura em tecnologia da informação e comunicação. Rio de Janeiro: IBICT/DEP, Brasília: IBICT/DDI, 2000.

⁶ OTLET apud LÉVIE, opus cit., p.63, apud GOMES, H. E, Classificação e gestão do conhecimento. *In*: PEREIRA, M. de N. F.; CHAVES, H. de S.; ARAÚJO, R. F. de (Eds.). **Dos padrões internacionais de estruturação da informação de pesquisa aos indicadores: primeira incursão na temática.** Brasília: IBICT, DELBRA, 2019. p. 73.

O caráter internacional está presente desde a ideia de uma rede de informação universal e em todas as suas realizações, como o Instituto Internacional de Bibliografia (atual Federação Internacional de Informação e Documentação - FID), Escritório Central de Associações Internacionais, Instituto Internacional de Cooperação Intelectual (precursor da UNESCO), Mundaneum, Classificação Decimal Universal, Repertório Bibliográfico Universal.

O que surpreende na obra de Otlet e de La Fontaine é a reunião de todas essas iniciativas em um único lugar, no Mundaneum, em prol de 'um modelo internacional de informação científica' para organizar o conhecimento mundial e classificá-lo de acordo com um sistema por eles desenvolvido, a Classificação Decimal Universal (CDU).

Voltando ao texto de Hagar. As principais decisões para o desenvolvimento da CDU incluíam a incorporação de outras tabelas e a mobilização de cientistas, processo interrompido com a Segunda Guerra Mundial, e depois retomado com a conferência internacional de 1948 (Royal Society, Londres) para discutir a reforma do sistema de informação científica, pela dificuldade de acompanhar seu crescimento. Um explosivo problema de quantidade!

Esses eventos formam um caldo cultural que se movimenta incorporando mais pessoas, mais realizações. Na esteira desse continuum, Hagar discorre sobre a criação do Classification Research Group (CRG), no Reino Unido, em 1952 e a proposta conjunta Unesco e ICSU (International Council of Scientific Unions), em 1971, de um sistema de informação mundial de ciências, UNISIST.

“Não se buscava implementação, mas fornecimento de diretrizes, princípios e recomendações aos países-membros e à comunidade científica.”⁷

A definição do UNISIST para informação científica tem sentido amplo.

Cientistas que são seus construtores e usuários, apenas pedem que as contribuições de uns e outros sejam verificáveis; [o artigo] é um meio através do qual os cientistas no mundo mantêm sua disciplina. É um meio de educação de futuros cientistas, e o principal reservatório de conceitos e dados onde recorrem para aplicação em programas econômicos e tecnológicos.⁸ (ênfase acrescentada)

A novidade desse programa está exatamente na informação científica, conforme relata Hagar.

Questões ligadas a políticas de ciências, financiamento, equipamentos e formação de mão de obra já eram temas familiares; mas os recursos informacionais tinham recebido até então atenção menor, e este foi o tema central do estudo: mobilização de recursos – no sentido mais amplo – de informação da ciência.⁹

Voltando ao texto original do UNISIST, para se documentar a origem da concepção de sistemas CRIS, fora do propósito do artigo em análise, mas interessante do ponto de vista do foco geral da coletânea, encontramos sintonia entre eles.

⁷ GOMES opus cit., p. 76.

⁸ UNESCO apud GOMES opus cit., p. 76.

⁹ GOMES opus cit., p.76.

A primeira delas se refere ao modelo social do UNISIST, que integra em um sistema de comunicação produtores, intermediários e usuários de conhecimento de todos os tipos de profissão. Do ponto de vista institucional, mencionam-se, a título de exemplo, institutos de pesquisa, publicadores e bibliotecas.¹⁰

A segunda concepção de sistemas CRIS tem origem no confronto entre o problema da informação científica, anteriormente quantitativo, e os novos requisitos que decorrem de novas formas de ciência e de organização.

As instituições da revista científica, dos serviços científicos de resumo e indexação, de bibliotecas científicas e centros de dados estão sob o desafio de modificar suas funções e serviços para acomodar a ciência em suas novas formas.¹¹

Entre as novas formas, o documento reconhece a pesquisa orientada para solução de problemas financiada por agências de governo. A nova forma de organizar é juntar pesquisador, agência, projeto, resultados etc., modelando-se, assim, o que viria a ser conhecido como sistema CRIS.

Há discussões ainda hoje pertinentes, e dos argumentos e problematizações é possível extrair uma série de ideias para organizá-los e definir serviços a serem prestados aos usuários.

Ainda no âmbito do UNISIST, o artigo de Hagar nos leva novamente à classificação com apresentação do Manual BSO (Broad System of Ordering), desenvolvido para uma rede mundial de informação, cobrindo todos os campos de conhecimento.

A ideia da rede, no UNISIST, requeria padrões comuns e tecnologia compatível, e é nesta perspectiva que se insere o padrão BSO, como uma linguagem de compatibilidade entre os diversos elementos da rede, podendo também funcionar como uma linguagem de comutação.¹²

Afinal, continua Hagar: “Grandes massas de dados, organizados de forma descentralizada e independente requeriam soluções para sua integração, harmonização, compatibilização ou qualquer outro meio de compartilhamento.”¹³

Os termos acima, componentes de argumentações e propostas veiculadas na época do UNISIST, são os mesmos que usamos hoje com a mesma finalidade.

A criação da euroCRIS, no começo deste século, para lidar com a informação de pesquisa, suscita a indagação sobre que elementos o estudo UNISIT/ICSU poderia oferecer. Hagar responde:

¹⁰ UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION AND THE INTERNATIONAL COUNCIL OF SCIENTIFIC UNION. **UNISIST: Synopsis of the Feasibility Study on a World Science Information System.** Paris: Unesco, 1971.

¹¹ Ibidem, p.23.

¹² GOMES opus cit., p. 77.

¹³ Ibidem, p. 76.

As questões relacionadas à divulgação e acesso às informações de pesquisa ainda são atuais, apesar dos desenvolvimentos da TI e das contribuições da Ciência da Informação, visto que o problema não é apenas tecnológico, mas envolve interesses e necessidades de vários países em seus esforços de integração e/ou harmonização de sistemas de informação.¹⁴

Interesses esses ainda não plenamente atendidos, conforme se observa na menção pela autora do diagnóstico de recente estudo sobre o desenvolvimento de um sistema de informação de pesquisa, no contexto da Ciência Aberta, realizado pela Science Europe.¹⁵

Tal documento enumera dificuldades para o desenvolvimento de um sistema de informação para pesquisa, como as diversas bases de dados de organização de pesquisa que diferem muito, em especial quanto a definições, sistemas de classificação, formatos e propósitos. E estes aspectos interferem nos sistemas/linguagens de indexação adotados como estrutura, abrangência, nível de abrangência dos termos. Interoperabilidade é uma questão recorrente.¹⁶

Ora, como sabemos, há diferentes níveis de interoperabilidade¹⁷. Para a euroCRIS, a interoperabilidade ocorre no âmbito dos metadados, com seu padrão CERIF (Common European Research Information Format).

O padrão CERIF é fundamental para a compatibilidade no nível computacional. É preciso, entretanto, abordar o problema da classificação na representação do conteúdo dos projetos e relatórios de pesquisa corrente para que ocorra adequado uso do padrão CERIF. Este tem sido um aspecto pouco abordado, fato também registrado quando da proposta do UNISIST.¹⁸

A última seção do substantivo artigo de Hagar trata de Classificação e Gestão, ocasião em que apresenta diferentes esquemas de classificação com finalidade de gestão sobre os acervos de informação, marcando antes uma importante e crucial diferença. Uma coisa são os esquemas de classificação como os de Otlet e UNISIST, orientados para resultados de pesquisa. Outra coisa são os esquemas para derivar indicadores estatísticos.

¹⁴ *Ibidem.*, p. 78.

¹⁵ SCIENCE EUROPE apud GOMES opus cit., p.78.

¹⁶ GOMES opus cit., p. 78.

¹⁷ BAKER, D.; SIMONS, E.; BROWN, J. The various aspects of Interoperability: A strategic partnership driving interoperability in research information through standards. **euroCRIS Strategic Membership Meeting Autumn 2014** (KNAW, Amsterdam, Nov 11-12, 2014). Disponível em: <http://dspacecris.eurocris.org/bitstream/11366/354/1/Introduction_to_the_theme_of_the_meetingBaker_Brown_Simons_euroCRIS_SMM2014.pptx>

¹⁸ GOMES opus cit., p. 78.

Classificações com finalidades estatísticas têm natureza agregadora, ou seja, têm nível de generalidade maior do que os esquemas bibliotecários. Em contrapartida, quando voltados para planejamento de ciência, baseiam-se igualmente em disciplinas ou grandes áreas do saber.¹⁹

Sempre que essa diferença vem à baila, penso, em termos de indexação, sobre a inesperada clonagem da ovelha Dolly – uma quebra de paradigma –, o projeto de pesquisa que a financiou e o artigo publicado na revista *Nature*, em 1997. O projeto teve financiamento do The Biotechnology and Biological Sciences Research Council (BBSRC), do Reino Unido. Em seu site não há registro de projetos, mas, se houvesse, os termos usados para descrever seu conteúdo seriam poucos, mais abrangentes. Teriam usado uma classificação ‘magra’! Aqui é fase da ciência em ação, da ciência em se fazendo!

Quando a ciência está pronta, é preciso adentrar em seus conteúdos para permitir acesso e ‘re-uso’. Uma prática comum é a da atribuição de palavras-chave pelos próprios autores. Os autores do artigo de clonagem da ovelha Dolly atribuíram cerca de 30 palavras-chave para descrever o conteúdo do artigo (quadro 1).

A indexação não para aí. Há os serviços de indexação e resumo da literatura científica, disponíveis em bases de dados, algumas delas acessíveis via Portal da CAPES. *Nature*, a revista que publicou o artigo da Dolly, é indexada em inúmeros desses serviços, cada um deles com seu próprio sistema de classificação. Aqui, o usual é distribuir os artigos em grandes áreas de conhecimento com algum nível de particularidade. As palavras-chave atribuídas pelos autores são incorporadas para ampliar a encontrabilidade.

Quadro 1 – Palavras-chave do Artigo que Relata o Experimento da Ovelha Dolly

Viable offspring derived from fetal and adult mammalian cells. I. Wilmut ; A. E. Schnieke ; J. McWhir ; A. J. Kind ; K. H. S. Campbell. *Nature*, 1997, Vol.385 (6619), p.810.
Animal Cloning ; Biotechnology ; Cells ; Embryo Transfer ; Embryonic Development ; Embryos ; Fetus ; in Vitro Culture ; Nuclei ; Oocytes ; Pregnancy Rate ; Transfer ; Adult Cells ; Cell Nuclei ; Cloning ; Dolly ; Embryo Development ; Embryo Growth ; Foetus ; Sheep ;
HYPERLINK "[http://rnp-primo.hosted.exlibrisgroup.com/primo_library/libweb/action/search.do?vl\(freeText0\)=+Ovis+&vl\(1045019533UI0\)=sub&vl\(1045019534UI1\)=all_items&fn=search&tab=default_tab&mode=Basic&vid=CAPES_V1&scps.scps=scope%3a\(%22CAPES%22\)%2cEbscoLocalCAPES%2cprimo_central_multiple_fe&ct=lateralLinking](http://rnp-primo.hosted.exlibrisgroup.com/primo_library/libweb/action/search.do?vl(freeText0)=+Ovis+&vl(1045019533UI0)=sub&vl(1045019534UI1)=all_items&fn=search&tab=default_tab&mode=Basic&vid=CAPES_V1&scps.scps=scope%3a(%22CAPES%22)%2cEbscoLocalCAPES%2cprimo_central_multiple_fe&ct=lateralLinking)"
"Encontrar todos os registros que contenham" \t "_parent" Ovis ; Bovidae ; Ruminants ;
HYPERLINK "[http://rnp-primo.hosted.exlibrisgroup.com/primo_library/libweb/action/search.do?vl\(freeText0\)=+Artiodactyla+&vl\(1045019533UI0\)=sub&vl\(1045019534UI1\)=all_items&fn=search&tab=default_tab&mode=Basic&vid=CAPES_V1&scps.scps=scope%3a\(%22CAPES%22\)%2cEbscoLocalCAPES%2cprimo_central_multiple_fe&ct=lateralLinking](http://rnp-primo.hosted.exlibrisgroup.com/primo_library/libweb/action/search.do?vl(freeText0)=+Artiodactyla+&vl(1045019533UI0)=sub&vl(1045019534UI1)=all_items&fn=search&tab=default_tab&mode=Basic&vid=CAPES_V1&scps.scps=scope%3a(%22CAPES%22)%2cEbscoLocalCAPES%2cprimo_central_multiple_fe&ct=lateralLinking)"
"Encontrar todos os registros que contenham" \t "_parent" Artiodactyla ; Mammals ; Vertebrates ; Chordata ; Animals ; Eukaryotes

Extraído de: Portal da Capes

¹⁹ Ibidem, p. 80.

Resumo, indexação, palavra-chave, bibliografias nacionais, serviços de indexação e resumo da literatura, bases de dados, temas relacionados ao parágrafo anterior, trazem à tona dois estudos brasileiros, cujos resultados atestam a falta de controle de qualidade de bases de dados nacionais para produção de indicadores de C&T, independentemente do resultado (geral ou parcial) a que cada estudo chegou^{20,21}. Essa ocorrência mostra o quanto precisamos, como país, nos debruçarmos sobre essa situação.

Com implicações no tema indicador, Hagar dá exemplos de esquemas de classificação para gestão de pesquisa, analisando generosamente os internacionais Australia-New Zealand Standard Research Classification, Fields of Science (FoS/OCDE) e International Standard Classification of Education (ISCED/UNESCO), e o brasileiro Áreas do Conhecimento, desenvolvido por CNPq e CAPES.

O Consortia Advancing Standards in Research Administration Information (CASRAI) é mencionado em linhas gerais, dado que é objeto de dois sumários executivos (português e inglês) também escritos por Hagar, publicados nesta coletânea e comentados mais adiante.

As análises das tabelas em uso e das possibilidades de desenvolvimento de um esquema de classificação para uso comum atestam o desafio a ser enfrentado pelos interessados (Science Europe, euroCRIS, CASRAI, BRICRIS...).

A adoção de um padrão – regional ou internacional – eliminaria a necessidade de estudos de compatibilização, harmonização ou outro. Mas iria requerer o desenvolvimento de instrumentos intermediários, de alto custo, acrescido da necessidade de constante atualização relacionada ao avanço do conhecimento, como também à inclusão de novo país-membro no sistema de informação.²²

O penúltimo tópico do artigo traz uma reflexão a respeito da classificação na era digital, meio que levou, segundo a autora, a sua democratização:

A Teoria da Classificação se democratizou: deixa de ser preocupação exclusiva de pesquisadores: chegou recentemente no mundo da gestão do conhecimento como taxonomia em novos produtos e serviços de informação em meio digital.²³

²⁰ PEREIRA, M. de N. F. **Por uma economia do conhecimento: avaliação de bases de dados nacionais para a produção de indicadores de C&T (Ciência e Tecnologia)**. Rio de Janeiro, 2003. Relatório Final de Bolsa de Produtividade de Pesquisa. Processo CNPq: 520416/1993-7(NV). viii, 150p., xcii.

²¹ ALVARES, L.; BERLANGA-LLAVORI, R. A pesquisa científica e tecnológica sob a ótica da biblioteca digital brasileira de teses e dissertações. *In*: ENCONTRO IBÉRICO EDICIC, 7, 2017, **Atas: Ciência Aberta o contributo da Ciência da Informação**. Coimbra: Universidade de Coimbra - Centro de Estudos Interdisciplinares do Século XX, 2017. p. 1293-1304.

²² GOMES, opus cit., p. 80.

²³ LAMBE, p. 99 apud GOMES opus cit.88, p.

A grande estrela teórica desse movimento é a classificação analítico-sintética de Ranganathan com seu Método de Faceta, o qual produz taxonomias rigorosas. O método se popularizou através dos pesquisadores do Classification Research Group do Reino Unido, em especial com a produção do Thesaurofacet.

A observação de Hagar sobre a pertinência do método para esquemas com alto grau de abrangência, como requer um esquema para produção de indicadores, destaca que:

Ao selecionar as classes gerais, os respectivos itens específicos da taxonomia de base são úteis na descrição de seu conteúdo com informação do tipo “Aqui se incluem...”. A partir do método indutivo produzem-se, desse modo, esquemas consistentes e orientação segura para os usuários.²⁴

A continuidade dos estudos de classificação no item Sumário Executivo/Executive Summary nos leva a desviar da apresentação do último artigo para contextualizar e comentar os dois sumários relacionados ao artigo de Hagar, anteriormente apresentado.

Começamos pelo contexto. A necessidade de definir ferramentas de uso comum, e no plano internacional, levou o IBICT a manifestar junto ao CASRAI sua vontade de traduzir a tabela de classificação por eles organizada. O CASRAI é outra organização internacional não governamental que atua na semântica da informação de pesquisa, com glossários, dicionários, templates diversos e um sistema de classificação.

A autorização para tradução veio acompanhada de convite para reunião de instalação da Agência do CASRAI na Europa, em abril de 2017, e a consequente participação no Comitê Gestor da citada Agência.

Nessa ocasião, a consultora do BRCRIS/IBICT Hagar Espanha Gomes, uma das representantes do Brasil, informou aos presentes sobre a autorização do CASRAI para traduzir a tabela de classificação por eles organizada. Isso suscitou interesse dos presentes, merecendo destaque o apoio da Science Europe. Em decorrência, o IBICT assumiu publicamente a tarefa de prosseguir nos estudos, o que de fato aconteceu.

Os resultados dos estudos com o sistema de classificação CASRAI, para fins de sua tradução, foram apresentados em janeiro de 2018, em reunião a distância do comitê executivo da nova agência europeia.

Agora vejamos o conteúdo. Os dois sumários executivos publicados nesta coletânea foram produzidos e traduzidos especialmente com essa finalidade: Sobre a “Classification de la Recherche Universitaire” (Tabela CASRAI)²⁵ e Sobre o estudo de comparação entre a Tabela do CNPq e a Tabela CASRAI.²⁶

²⁵ Extraído de: GOMES, H. E. **Estudo CASRAI: Compatibilidade de sistemas**. Produto 1 do projeto: Dos padrões internacionais de estruturação da informação de pesquisa aos indicadores (BRCRIS). Rio de Janeiro, 32 p. 2017. IBICT - Iniciativa de Apoio aos Diálogos Setoriais União Europeia/Brasil.

²⁶ Extraído de: GOMES, H. E. **Estudo CASRAI – Estudo com tabelas CNPq, FoS e ISCED**. Produto 2 do projeto: Dos padrões internacionais de estruturação da informação de pesquisa aos indicadores (BRCRIS). Rio de Janeiro, 32 p. 2017. IBICT - Iniciativa de Apoio aos Diálogos Setoriais União Europeia/Brasil.

²⁴ GOMES opus cit., p. 89.

O primeiro sumário trata da estrutura da Tabela CASRAI, da análise das dimensões e das propostas para sua reformulação. O segundo objetiva verificar o grau de compatibilidade entre as tabelas do CNPq e do CASRAI. A conclusão aponta para o desenvolvimento de uma metalinguagem, única maneira de tornar compatível diversos esquemas de uma rede de CRIS. Ou seja, não se trata apenas de traduzir a tabela. Há que se fazer muito mais que isso.

Voltando à seção de artigos, o terceiro e último é de autoria de Ronaldo Ferreira de Araújo, consultor da FAPEAL para o BRCCRIS: Os Estudos Cibernéticos da Informação: Indicadores de Estruturas Web e de Recursos da Web Social, o qual não faz parte do projeto financiado pelos Diálogos Setoriais.

Ronaldo é autor convidado para compor o tema dos indicadores com artigo anteriormente publicado pela Editora da Universidade Federal de Alagoas, EDUFAL, na coletânea Estudos Métricos da Informação na Web: atores, ações e dispositivos informacionais. O artigo foi cedido para publicação em nossa coletânea.

Pelo título, dá para perceber que os indicadores dos quais Ronaldo se propõe a tratar são dos tempos de Internet. Com efeito, na abertura de seu artigo, Ronaldo marca a diferença entre os indicadores tradicionais de comunicação científica e aqueles que medem a interação da ciência com a sociedade, como tweets, blogs...

Um bom exemplo para entendermos como eles são incorporados à informação científica é a altmetria (medida alternativa) do artigo da ovelha Dolly (quadro 2), anteriormente comentado.

No caso em exame, esses indicadores foram incorporados por meio de técnicas de visualização no serviço de indexação e resumo da literatura que indexa a Nature. Outros serviços operam de maneira semelhante.

Quadro 2 – Almetria do Artigo que Relata o Experimento da Ovelha Dolly

Escolhido por 38 agências de notícias
Blogado por 13
Referenciado em 4 fontes políticas
Twittado por 11
Referenciado em 146 patentes *
Mentioned by 1 peer review sites
Em 3 Facebook pages
Referenciado em 7 Wikipedia pages
Mencionado em 1 Google+ posts
Highlighted by 1 platforms
1089 Leitores em Mendeley
7 Leitores em CiteULike

Extraído de: Portal da CAPES

Qual a definição de altmetria que podemos extrair do artigo de Ronaldo?

De forma geral, a altmetria pode ser entendida como a aplicação das métricas de mídias sociais no contexto da comunicação científica. Assim, a diferença é que o investimento no marketing digital e o monitoramento da conversação e engajamento dos usuários não estão voltados para empresas ou marcas e sim para a atenção on-line que produtos científicos, como os artigos de periódicos, recebem.²⁷

²⁷ ARAÚJO, R. F. de. Os Estudos Cibernéticos da Informação: Indicadores de Estruturas Web e de Recursos da Web Social. In: PEREIRA, M. de N. F.; CHAVES, H. de S.; ARAÚJO, R. F. de (Eds.). **Dos padrões internacionais de estruturação da informação de pesquisa aos indicadores: primeira incursão na temática**. Brasília: IBICT, DELBRA, 2019. p. 101.

O exemplo do quadro 2 é apresentado tão somente para ilustrar do que se trata. Quem quer que vá analisar esses indicadores saberá que a clonagem da ovelha Dolly terá que ser contextualizada em seu tempo para se entender o porquê de tão poucos tweets, páginas de Facebook, fontes políticas, mas de muitas patentes e muitos leitores na rede social acadêmica Mendeley.

Esses indicadores não são os únicos, como mostra Ronaldo. É possível medir a popularidade de uma biblioteca, de um site, via seguidores nas redes sociais, clicks de likes e assemelhados.

Esse é um enfoque. O outro é teórico, caminho substantivo seguido pelo autor. O campo desses estudos é chamado de cibermetria (cybermetrics) “[...] ou ‘métricas web’ [...] termos genericamente utilizados para se referir à medição quantitativa da criação e utilização de conteúdos web, sejam eles de natureza social, política, científica ou empresarial.”²⁸

Seu trabalho “(...) discorre brevemente sobre o amplo campo da cibermetria, situando seus subcampos e objetos, bem como sua interface com disciplinas tradicionais da Ciência da Informação como a bibliometria, cientometria e informetria.”²⁹

A contribuição de Ronaldo se dá no plano teórico, ao esquematizar as relações entre os subcampos dos estudos métricos da informação, recorrendo a conceituações de autores consagrados na literatura desse tema.

²⁸ ARAÚJO, opus cit., p. 96.

²⁹ Ibidem, p. 93.

Sua abordagem não é apenas discursiva. Particularmente interessante é a maneira como ele se posiciona para reconfigurar as inter-relações entre os subcampos dos estudos métricos da informação, onde a sistematização do campo pelos autores ocorre na forma de diagramas.

Para seu esquema gráfico, Ronaldo argumenta que é preciso primeiro considerar o quê se pretende com o estudo a ser desenvolvido e escolher, por aproximação às métricas tradicionais, aquelas mais coerentes com o objetivo do estudo.

Seu esquema tem três quadrantes: no primeiro são situados os estudos webométricos e webmétricos; no segundo, encontram-se cyberscientometrics ou cientometria 2.0, e, no terceiro, encontram-se estudos mais próximos da cibercultura e das métricas de mídias sociais, com abordagem mais comercial.

Os indicadores altmétricos são gerados a partir de identificador persistente de documento, como DOI, sigla em inglês para Digital Object Identifier, e PubmedID. E há uma série de requisitos, um aninhado no outro, para que a ferramenta altmetric possa traduzir em indicadores o que rastreou ao longo do caminho: sites de notícias, blogs, redes sociais³⁰.

³⁰ NASCIMENTO, A. Maneiras alternativas de gerar métricas alternativas. **Bibliotecários sem fronteiras**. Disponível em: <<https://bsf.org.br/>>

O Brasil enfrenta dificuldades para gerar indicadores altmétricos. De maneira geral, as publicações brasileiras não contam com DOI, dado que requer uma assinatura paga. Segundo, quando contam, há incipiência de visualização gráfica nos repositórios institucionais.³¹

Com esse comentário, nos aproximamos da seção Ferramentas de Pesquisa, uma Bibliografia Anotada sobre Identificadores Persistentes (ORCID e RINGGOLD Identifiers), de autoria de Juliana Pinheiro Farias e Kelen Cândida Vieira Bomfim, ambas bolsistas do IBICT.

A literatura escolhida é a de indicadores persistentes com foco naqueles que são usados para pessoas (Open Researcher and Contributor ID – ORCID ID) e organizações (RINGGOLD), complementando o leque de padrões para organização de informação de pesquisa e a decorrente interoperabilidade entre sistemas afins.

Outra motivação para a escolha do foco da bibliografia foi a assinatura, pelo IBICT, de uma parceria com CAPES, CNPq, CONFAP e Scielo para formação do Consórcio Nacional ORCID, com assinatura compartilhada de serviços de consultas e atualização de dados, em dezembro de 2017. Esta ação, liderada pela CAPES, financiou a primeira anuidade das cinco instituições no ORCID.

A RNP participa também do Consórcio na condição de fornecedor de infraestrutura tecnológica e suporte técnico de alto nível.

O trabalho em tela, de autoria de Juliana e Kelen, coloca em perspectiva a ideia de identificador persistente, cuja definição pelo ORCID diz o seguinte:

Um identificador persistente é uma referência de longa duração a um objeto digital que fornece informação sobre esse objeto, independentemente do que acontece com ele. Desenvolvido para tratar de 'links quebrados', um identificador persistente pode ser definido para fornecer uma representação apropriada de um objeto, independentemente dos objetos mudarem sua localização on-line ou ficarem off-line.³²

A perspectiva adotada leva as autoras aos primórdios da estruturação de registros bibliográficos, no século XIX, onde se estabelecia “árduo trabalho manual” de controle da autoridade para os pontos de acesso em catálogos de biblioteca em livros e em fichas.

A partir desse ponto, são apresentados os principais eventos que levaram ao desenvolvimento do ORCID e do RINGGOLD no período de 2018 a 2006. São incluídos 16 consórcios ORCID, apresentados após a bibliografia.

³¹ REIS, J. E. dos; SPINOLA, A. T. P.; AMARAL, R. M. do. Incipiência da visualização de indicadores bibliométricos e altmétricos nos Repositórios Institucionais brasileiros. *Em Questão*, v. 23, 2017. Disponível em: <<https://www.redaly.c.org/html/4656/465650499012/>>

³² CASRAI. Dictionary. Disponível em: <http://dictionary.casrai.org/Persistent_identifier>. apud FARIAS, J. P.; BOMFIM, K. C. V. Bibliografia anotada: identificadores persistentes (ORCID e RINGGOLD Identifiers) In: PEREIRA, M. de N. F.; CHAVES, H. de S.; ARAÚJO, R. F. de (Eds.). **Dos padrões internacionais de estruturação da informação de pesquisa aos indicadores: primeira incursão na temática**. Brasília: IBICT, DELBRA, 2019. p. 155.

A última seção é a de relatórios. O primeiro trata da viagem a Portugal e o segundo do relato de avaliação exploratória do DSpace-CRIS.

Paulo Henrique de Assis Santana discorre sobre a missão a Portugal, na condição de consultor no âmbito do projeto financiado pelos Diálogos Setoriais. Seu minucioso e preciso relato abrange tempo e espaço dos compromissos, instituições e pessoas visitadas, agenda de trabalho, comentários sobre as visitas, explanações técnicas e uma lista com sugestões do quê fazer. Ele dá o caminho das pedras!

Como ressalta o autor, pelo adiantado estado de desenvolvimento de Portugal com o PTCRIS,

[...] seria interessante que, sob a égide do acordo de cooperação já existente entre IBICT e FCT, fosse viabilizado o desenvolvimento conjunto de ferramentas comuns ao PTCRIS e BRICRIS, principalmente no que diz respeito à inserção das informações do OASIS-BR no OpenAIRE. Alternativa ou paralelamente, poder-se-ia trabalhar para a inclusão das informações relativas ao INCTs naquele ambiente.³³

Paulo Henrique conclui seu relato listando, de maneira pontual, quase uma dezena de direções que podem acelerar, com algum aporte financeiro, a cooperação entre IBICT e FCT no desenvolvimento do BRICRIS.

Sua sugestão de incluir projetos dos INCTs (Institutos Nacionais de Ciência e Tecnologia) apoiados pelo CNPq, no desenvolvimento de um CRIS, se acolhida, representa a internacionalização dos “melhores grupos de pesquisa em áreas de fronteira da ciência e em áreas estratégicas para o desenvolvimento sustentável do país.”³⁴

Os resultados de 122 INCTs estão disponíveis na Plataforma Carlos Chagas em relatórios armazenados em seus respectivos processos. Análise de apenas dois deles mostra a criação, durante a vigência dos processos, de 879 produtos (517 artigos completos publicados em periódicos, 2 livros, 11 capítulos de livros, 168 trabalhos completos publicados em anais de congressos, 171 produtos envolvendo dissertações de mestrado e teses de doutorado, além de trabalhos aceitos para publicação em periódicos).³⁵

O próximo e último trabalho concerne ao desenvolvimento de um piloto de CRIS no outro extremo da ciência produzida pelos INCTs, com conteúdos típicos de ciência cidadã.

³³ SANTANA, P. H. de A. Relatório de viagem. In: PEREIRA, M. de N. F.; CHAVES, H. de S.; ARAÚJO, R. F. de (Eds.). **Dos padrões internacionais de estruturação da informação de pesquisa aos indicadores: primeira incursão na temática.** Brasília: IBICT, DELBRA, 2019. p. 187.

³⁴ INCT. Institutos Nacionais de Ciência e Tecnologia. Disponível em: <<http://inct.cnpq.br/>>. Acesso em: 2 de dezembro de 2018.

³⁵ SANTANA opus cit., p. 188.

A ciência cidadã refere-se ao engajamento do público em geral em atividades de pesquisa científica, quando os cidadãos contribuem ativamente para a ciência, seja com seu esforço intelectual ou o conhecimento em volta, seja com suas ferramentas e recursos.³⁶

Essa característica permeia o PPSUS³⁷,

[...] uma iniciativa de fomento à pesquisa em saúde nas Unidades Federativas (UF), que promove o desenvolvimento científico e tecnológico, visando atender as peculiaridades e especificidades de cada estado brasileiro e contribuir para a redução das desigualdades regionais.³⁷

O relatório de pesquisa Piloto BRICRIS IBICT-FAPEAL: uma avaliação exploratória do DSpace-CRIS para a construção de um sistema de informação de pesquisa local, de autoria de Kleber de Barros Alcanfôr, Hélia de Sousa Chaves, Lilian Maria Thomé Andrade Brandão e Rogério Mendes Castilho, ambienta alguns projetos do PPSUS-FAPEAL em DSpace-CRIS, uma ferramenta em código aberto e uso livre, recomendada pela euroCRIS.

Merece destaque o uso de duas ferramentas em caráter complementar: Docker e PDI (Pentaho Data Integration).

³⁶ MOTA, P. J. G. *Ciência Aberta, Ciência Cidadã. In: ConfOA. Anais...* Lisboa, 2018. Disponível em: <<https://pt.slideshare.net/ConfOA/cincia-aberta-cincia-cidad>>. Acesso em: 2 de dezembro de 2018.

³⁷ Programa Pesquisa para o SUS (PPSUS). Disponível em: <<http://portalms.saude.gov.br/acoes-e-programas/ppsus>>. Acesso em: 2 de dezembro de 2018.

A primeira possibilitou configurar o DSpace-CRIS para atendimento das particularidades do projeto, enquanto a segunda facilitou a realização de inúmeras operações de integração de dados.

O DSpace-CRIS é capaz de ambientar sistemas CRIS em sua inteireza, ou seja, não apenas o módulo tradicional dos resultados de pesquisa, geralmente informações bibliográficas, mas informações cadastrais (instituições, pesquisadores, projetos de pesquisa) e até mesmo dados brutos de pesquisa (questionários, sons de baleias, monitoramento florestal por satélite...). A ferramenta está em conformidade com os padrões internacionais relevantes (como CERIF, IIF) para facilitar a interoperabilidade e a transferência de dados.³⁸

Com vistas à elaboração do processo de coleta, tratamento e preparação dos dados para o piloto, os autores passam em revisão duas versões do formato CERIF, a de 2006 e a de 2012. O CERIF é um formato padrão internacional de metadados desenvolvido especificamente para estruturação de informação de pesquisa em suas diferentes manifestações.

Por se tratar de uma primeira aplicação prática, e tendo como base o CERIF, foram contempladas as entidades base, pesquisador, instituição e projeto, seguindo-se do módulo de resultados.

³⁸ DSpace-CRIS Home. Disponível em: <<https://wiki.duraspace.org/display/DSPACECRIS/DSpace-CRIS+Home>>.

A entidade projeto tem um campo no formulário para a classificação do CASRAI e área do conhecimento do CNPq. A título de experimentação foram classificados 32 projetos pela Profa. Hagar Espanha Gomes, ficando, assim, devidamente registrados no DSpace-CRIS.

Do ponto de vista técnico, o relatório é um exemplo de superação de dificuldades técnicas de uma ferramenta de organização de informação.

A primeira dificuldade enfrentada foi a de “uma grande complexidade nos recursos de configuração”.

Percebemos, então, que seria necessário ter a possibilidade de reconstruir completamente o ambiente a qualquer momento e garantir total liberdade para experimentar os recursos de configuração dos metadados das entidades do sistema CRIS. Assim, buscamos uma solução para gerenciar ambientes de servidores com possibilidade de ter arquivos de imagem, contendo as pré-configurações, e restaurá-los a qualquer momento.³⁹

³⁹ ALCANFÔR, K. de B.; CHAVES, H. de S.; BRANDÃO, L. M. T. de A.; CASTILHO, R. M. Piloto BR-CRIS IBICT-FAPEAL: uma avaliação exploratória do DSpace-CRIS para a construção de um sistema de informação de pesquisa local *In*: PEREIRA, M. de N. F.; CHAVES, H. de S.; ARAÚJO, R. F. de (Eds.). **Dos padrões internacionais de estruturação da informação de pesquisa aos indicadores: primeira incursão na temática**. Brasília: IBICT, DELBRA, 2019. p. 199.

Conforme relato dos autores, a versão 5.8 do DSpace-CRIS usada no projeto não dispõe de API (Application Programming Interface) para integração de dados, recorrendo-se às tradicionais rotinas de exportação e importação, em formato Excel, e arquivos individuais por entidade. Vale ressaltar que a versão DSpace 7, prevista para 2019, já incorpora integração por API em suas funcionalidades (DURASPACE, jun. 2018).

O relatório de pesquisa segue destrinchando as inúmeras particularidades técnicas requeridas para a instalação e configuração do DSpace-CRIS para finalmente

[...] alcançar a integração dos dados das entidades base. Isso porque, conforme o CERIF, fica evidente que um sistema CRIS sem que essas três entidades estejam coletadas, qualificadas e relacionadas fica sem sentido.⁴⁰

Resumindo, o estudo deve ser visto como uma experimentação da ferramenta DSpace-CRIS nos processos de configuração de metadados, carga automática e registro manual através dos formulários da interface web.

A ciência do novo inconsciente revela que escolhemos produtos pela embalagem e livros pela capa. Mas essa é apenas uma das formas de processamento mental. A outra, a do conteúdo, foi antes amplamente demonstrada. Dessa forma, estima-se que aflore sinergia entre os dois.

Pela bela capa que embala este livro, nosso reconhecimento ao Rodrigo Azevedo Moreira, designer do IBICT.

⁴⁰ Ibidem, p. 202.

A obra é de uma luminosidade ímpar. Lembra Volpi, o artista brasileiro das bandeirinhas. Lindo! Sua brasilidade é inconfundível, seja pelo colorido ou pela luz que da capa emana. Sua simbologia abarca o título da coletânea em sua plenitude: vai dos padrões aos indicadores.

Há outro logo, o do BR CRIS. Ele se alinha à brasilidade estampada na capa, com a prevalência do azul e a presença sutil do verde e amarelo para os contornos do símbolo, conferindo movimento, vida e dinamismo à pesquisa e ao sistema de informação. Sutileza e leveza são as palavras-chave que definem o logo: a distinção entre o “Br” e o “Cris” se dá com destaque para a letra “i” de informação que parece flutuar. A circulação da informação de forma integrada também é um conceito associado ao símbolo. A autora é Mariela Muruga, também designer do IBICT, a quem devo uma substantiva descrição técnica do logo, orientação para o que escrevi.

Mariela e Rodrigo são criadores de símbolos CRISMáticos! Obrigada pelo cuidado que aportam ao trabalho dos tantos Outros mobilizados para imprimir vida a esta coletânea.

Pelo conteúdo, aos autores anteriormente mencionados agradecemos pelas inúmeras idas-e-vindas de versões para avaliação das mudanças requeridas, e que não foram poucas, reconheço.

Avaliação compartilhada com Eny Marcelino de Almeida Nunes e com os editores Hélia de Sousa Chaves e Ronaldo Ferreira Araújo, que também cumpriram função de orientadores nos trabalhos em que ambos têm autoria coletiva.

A revisão opera de muitas maneiras. A revisão editorial esteve a cargo de Ana Lúcia Mendes e de Fernanda Hardman Neves, incansáveis no batimento entre versões que passaram pelo crivo meticuloso de Margaret de Palermo Silva, carro-chefe do IBICT na revisão gramatical.

Uma etapa importante em tempos de acesso aberto e de ‘re-uso’ de informação é a normalização das referências bibliográficas, que esteve a cargo de Lilian Maria Thomé Andrade Brandão, Kelen Cândida Vieira Bomfim e Juliana Pinheiro Farias.

O time de revisão do infinitamente pequeno, escondido em notas de rodapé, em fontes de gráficos e outras ilustrações, e em listas finais de referências, permite, com esse trabalho apurado, reverência aos antepassados e contemporâneos. O grupo compartilhou mudanças na ABNT, respeitando estilos de citação e possibilidades de atendimento às novas orientações, com a coletânea a pleno vapor, já em processo de editoração.

Ramón Martins Sodoma da Fonseca, chefe da Seção de editoração do IBICT, é o responsável pelo competente trabalho de montagem deste livro, ao aliar estética e usabilidade a serviço do leitor. Obrigada por isso, Ramón.

Outro agradecimento vai para o tradutor Michael Allan, escocês-canadense que adotou o Brasil e o português com tal desembaraço que vem funcionando como tradutor, contribuindo para internacionalizar aqui e ali a ciência que aqui se produz. Pela sua generosidade, ao traduzir os resumos e os sumários executivos desta coletânea, agradecemos.

Na versão piloto, Márcia Liz e Thiara Carina, cuidaram, respectivamente, da editoração do conteúdo e do tratamento de imagens, pelo que agradecemos. E essa contribuição nos remete ao financiamento da coletânea pelo PRODOC/UNESCO, 914BRZ2005, projeto que reuniu outros profissionais.

O TdR (Termo de Referência) apresentado após aprovação do projeto, do qual decorre esta coletânea, foi intensamente debatido com o ministro conselheiro Carlos Oliveira, Chefe do Setor de Sociedade da Informação e Mídia/Delegação da União Europeia no Brasil. Sua dedicação resultou em visão mais ampliada e estratégica, e devidamente articulada a nossa principal motivação, o BRCRISAs indicações recebidas do conselheiro Carlos traziam também orientações seguras do Comissário Carlos Moedas, responsável por definir, em 2015, a política de pesquisa e inovação da União Europeia, “uma mudança importante no paradigma tradicional de produção e disseminação do conhecimento científico”, como declara o ministro na apresentação de abertura desta coletânea.

Por todos os ganhos, nos sentimos em débito de gratidão para com os dois conselheiros, ambos chamados de Carlos, e também para com a expedita e competente Noelia Barriuso, orientadora técnica da equipe dos Diálogos Setoriais da União Europeia Brasil. Obrigada por todas as conquistas que nos ajudaram a alcançar.

Escrevo este Prefácio às vésperas da celebração do 65º aniversário do IBICT. Ano passado, nossa diretora Cecília Leite foi reconduzida para seu segundo mandato na direção do Instituto, contando também com excepcional apoio dos funcionários. Portanto, é hora de honrar as pessoas que tomam decisões e enfrentam desafios para que as coisas mudem de patamar.

Nossa diretora Cecília é a grande responsável por termos construído um arcabouço conceitual fundado no que existe internacionalmente de mais sólido para uma iniciativa nacional de CRIS, ao dar sinal verde, com seu jeito informal e confiante de ser, para trazer João Mendes Moreira, da Fundação para a Ciência e a Tecnologia de Portugal (FCT), líder do PTCRIS, para workshops no IBICT, ainda em 2015.

João abriu caminho para a vinda ao IBICT de Ed Simons, Presidente da euroCRIS (Amsterdam, Holanda) e membros do Board: Andrea Bollini (DSpace-CRIS), Jan Dvorak (CERIF), Michele Mennielli (relações internacionais) e Pablo de Castro (participação a distância, repositórios institucionais).

Esse movimento, dependente de expertise em cooperação internacional, encontrou em Paulo Egler, à época responsável pelo BRCRIS, a figura talhada para organizá-lo. Paulo foi secundado por um time de especialistas nessas matérias: José Luis dos Santos Nascimento, José Carlos Cordeiro, Mirian Aparecida Ribeiro e Cristiane Félix, também mobilizados em outras ocasiões. Entraram em cena igualmente Jane Gadelha, Jordana Padovani e Patrícia Osandón para organizar eventos e fazer prestações de conta. A todos, o nosso reconhecimento.

Cecília compartilhou a aliança com a euroCRIS, convidando todas as fundações de amparo à pesquisa estaduais e as federais para eventos nacionais e estrangeiros. Representantes do CNPq, FINEP, CAPES, FAPEAL, FAPESP, FAPEG, CGEE, RNP e Elsevier Brasil estiveram presentes em eventos no IBICT.

De todas as FAPs, a aliança mais próxima e que rendeu mais frutos imediatos foi com a FAPEAL. Sob sua indicação, fizemos o piloto de CRIS, aqui relatado no último trabalho; há dois artigos na coletânea com três autores ligados à instituição. Por isso, agradecimentos são devidos ao professor Fábio Guedes Gomes, seu presidente, por não ter medido esforços para garantir a presença, mais de uma vez, de seus técnicos em reuniões e workshops de trabalho no IBICT. Obrigada à professora Francisca Rosaline Leite Mota, ao professor Ronaldo Ferreira Araújo e ao mestrando em Ciência da Computação Victor Lemos Tenório, pela presteza com que sempre nos atenderam.

A motivação da FAPEAL por sua participação no piloto de CRIS levou Thiago José Tavares Ávila, Superintendente de Produção da Informação e do Conhecimento da Secretaria de Estado do Planejamento, Gestão e Patrimônio, do governo do estado de Alagoas, e um dos maiores especialistas em dados abertos de governo, a se interessar pela temática de CRIS, participando de reuniões e debates internos, pelo que também agradecemos.

Cecília ampliou alianças internacionais, ao organizar comitiva brasileira para comparecer à reunião de instalação, em Malta, da Agência do CASRAI na Europa, em 2017, atendendo ao convite de David Baker, diretor-executivo da organização.

A extensão do convite a outras organizações brasileiras resultou na presença das seguintes instituições como membros da comitiva: CAPES (Harrysson Gilgamesh de Medeiros Nóbrega), Comitê de Assessoramento do CNPq da Área de Ciência da Informação (Professora Mariângela Spotti Lopes Fujita) e Universidade Federal Fluminense (Professora Maria Luiza de Almeida Campos).

Nesse evento, o encontro da Professora Ligia Ribeiro, da Universidade do Porto, indicada para assumir a presidência do Comitê Executivo, gerou dividendos nas alianças com os portugueses, ampliando-as de tal sorte que a assistência técnica a ser prestada pelos portugueses aos brasileiros, no âmbito do projeto dos Diálogos Setoriais (DS), estava garantida.

Eloy Rodrigues, Pedro Príncipe e José Carvalho, da Universidade do Minho, passaram a integrar o projeto DS, trazendo excelência técnica aos resultados alcançados. Nossa gratidão a todos.

Os dividendos da reunião do CASRAI, em Malta, não param aí. O IBICT passou a fazer parte do Comitê Gestor do CASRAI para a Europa. O tema 'sistemas de classificação', igualmente de interesse do CASRAI, foi extensamente estudado pela Professora Hagar Espanha Gomes, membro da comitiva e perita contratada pelos Diálogos Setoriais para realizar esses estudos. Em decorrência, é bastante provável que o Brasil se veja, dentro em pouco, participante de um projeto internacional para desenvolvimento de um sistema de classificação de projetos de pesquisa que atenda a todos. Diante de toda essa grandeza, acima de tudo de espírito, só temos a usar de novo uma das primeiras palavras que aprendemos a dizer: obrigada!

É surpreendente como as dificuldades dos tempos de crise que afetaram a administração pública no nosso País, principalmente nos últimos dois anos, foram contornadas e as ações continuadas, ainda que sem o vigor necessário para significativa mudança de patamar. As pessoas no IBICT se ajudam umas às outras e isso faz delas não um "amontoado de indivíduos", mas uma equipe movida pela cooperação mútua. Tiago Braga, Coordenador de Tecnologias Aplicadas a Novos Produtos do IBICT, e Marcos Novais, Coordenador de Geral de TI, estão entre essas pessoas.

A eles agradecemos por nos permitir realizar ações que aqui se relatam, sem as quais conquistas importantes não teriam sido possíveis. Gostaríamos de um adendo a uma integrante da equipe de Tiago, Suelen Santos, a quem devemos os primeiros passos, quando da elaboração do projeto DS.

Arthur Fernando Costa, atual coordenador do BRCCRIS, é dessas pessoas com grande capacidade de mobilização de recursos para fazer acontecer aquilo que à primeira vista parece impossível. A sua precisa resposta às demandas do BRCCRIS permitiram enfrentar camadas de CRIS, título deste prefácio, que vão desde as interiores, como as de sistemas de classificação, que operam no âmago da ciência, e de sua informação, condição mandatória para que as camadas se externalizem. Ou as de organização da informação de pesquisa, como o piloto com o DSpace-CRIS, que opera na camada externa, juntando interfaces e partes internas para circulação da informação de forma integrada e particularizada ao mesmo tempo. Uma não deve existir sem a outra, sob pena de se comprometer a qualidade da informação e a conseqüente extração de indicadores para subsidiar, por exemplo, estudos de interação de redes de inovação, no estilo que se aproxima do exposto pelo Conselheiro Carlos Oliveira, da Delegação da União Europeia no Brasil, em sua apresentação de abertura desta coletânea.

Entre o coordenador Arthur e a diretora Cecília, circula Hélia, já referida anteriormente. Doutoranda em Ciência da Informação pela UnB, com pesquisa sobre sistemas CRIS, sua grande capacidade de agenciamento responde pela participação do IBICT no Consórcio Nacional ORCID. Há também o cadastro de instituições de C&T do IBICT, em fase de ajuste de modelagem e padronização, para possibilitar a incorporação de identificadores únicos internacionais. Hélia também não perde de vista a negociação junto ao CNPq para construção de um CRIS a partir dos dados de pesquisa do Programa Institutos Nacionais de Ciência e Tecnologia (INCTs). As equipes do CNPq e do IBICT já fizeram as primeiras aproximações técnicas e uma proposta de projeto conjunto está sendo elaborada.

Há mais gente para nomear e agradecer: Valéria da Silva Vieira Macêdo e Valquíria Nisia Fonseca Leite, secretárias da diretoria do Instituto, com quem contamos muitas vezes para acelerar processos de localização de pessoas, distribuição de convites, agradecimentos e recepção; Marina Pimentel Braz, a mascote da equipe do BRICRIS, cujo empenho em aprender faz dela uma trainee especial; Cláudia Regina de Araújo Mohn Souza e Daniela Abrahame Cunha, responsáveis pela cobertura dos eventos e a consequente propagação das ideias e realizações do BRICRIS; Valéria Paiva, gerente de projeto com fundação de apoio, e seu time de profissionais de primeira linha, por todas as vezes que trabalharam para nossos projetos acontecerem; Claudia Aparecida Nóbrega Franco, Ivete da Silva Vieira, Paloma Caroliny Alves Magalhães e Renata Monteiro Rodrigues.

À guisa de finalização, vamos descortinar o futuro que se aproxima. O patamar de competência em que nós do IBICT agora alcançamos, graças ao mergulho empírico que o Projeto DS nos permitiu, o que se atesta com a presente coletânea, abre as portas para uma nova aventura de conhecimento, cooperação e solidariedade.

A tragédia que se abateu sobre o Museu Nacional, em setembro de 2018, levou o IBICT a prestar solidariedade, colocando sua expertise em sistemas CRIS, preservação digital e socialização da ciência na reconfiguração do Arquivo Histórico do Museu Nacional.

O Sistema Colheita, o CRIS da SEMEAR (Seção de Memória e Arquivo do Museu Nacional) vai capitalizar na experiência da equipe responsável pela experimentação do DSpace-CRIS, no âmbito do projeto DS aqui relatado. Os trabalhos estão em andamento. A inserção de Kleber de Barros Alcanfôr, Rogério Mendes Castilho e Lilian Maria Thomé Andrade Brandão na rede de especialistas responsável pelo trabalho trará ganhos na configuração da ferramenta e no entendimento de outras facetas.

Em outro nível, os estudos para constituição de um Espaço Brasileiro de Pesquisa, a exemplo do que está em processo de criação na União Europeia, que foram momentaneamente deixados de lado, serão retomados. E com isso, enfrentado o desafio colocado pelo ministro Carlos Oliveira.

Gostaríamos de ver o Brasil liderar um processo semelhante em escala regional, tirando partido do seu papel especial enquanto economia mais avançada e desenvolvida da América Latina. Estariam assim criadas condições para imaginar um grande espaço aberto de pesquisa unindo União Europeia, Brasil e outros países da América Latina.⁴¹

Finalmente, à guisa de reverência, misturado com encantamento pela magia da vida, vamos em direção ao passado.

Sistemas CRIS não se constituem em novidade para o IBICT. Data de 1962, as primeiras incursões para o levantamento e registro das informações de pesquisas em processo no Brasil, quando da Presidência do IBBD por sua fundadora, a paraense Lydia de Queiroz Sambaquy (*in memoriam*). Dez anos depois, criou-se o CAPESQ (Cadastro de Pesquisas), integrando o Banco de Dados do IBBD (atual IBICT), as rotinas e procedimentos são automatizados e implanta-se um sistema que permitiu a inter-relação entre os dados dos diferentes subsistemas.

Sentindo o Governo Brasileiro a necessidade de criar uma infraestrutura de pesquisas, vem sendo destinado, para este fim, um montante de recursos que tem sido aplicado com maior frequência nas Universidades, propiciando a criação de Centros e Institutos de Pesquisa.

Também no setor particular, tanto as indústrias como as entidades privadas têm concentrado esforços no sentido de ampliar os seus campos de pesquisas, procurando investir mais recursos nessa área de atividade. Para acompanhar o desenvolvimento das pesquisas nessas organizações de estudos, é necessário a divulgação dos resultados e mesmo dos métodos de pesquisa, colocando, assim, à disposição de outros Centros, os conhecimentos adquiridos.⁴²

O calendário marcava 1974. Estava criado o CRIS! Na equipe, a autora do artigo acima citado, Ilse Dumpbel César, Yone Chastinet, Ida Maria Cardoso Lima (*in memoriam*) e outros profissionais. Ida Maria reunia raras habilidades, sendo capaz de corrigir no olho as referências bibliográficas perfuradas em cartões(!), empacotados para seguirem no carro do Instituto Nacional de Pesquisas Espaciais (INPE), para sua sede em São José dos Campos. Somente lá os cartões poderiam ser processados em computadores de grande porte. A presidência do IBBD estava sob a batuta de Hagar Espanha Gomes, tantas vezes nomeada neste prefácio por seus estudos de sistemas de classificação para o projeto DS.

Naqueles idos, dois acontecimentos atravessam tempo e espaço. Em São José dos Campos, um estagiário intercalava o jogo de bola com o desembarque das caixas de cartão perfurado do CAPESQ.

⁴¹ OLIVEIRA, C. Apresentação. In: PEREIRA, M. de N. F.; CHAVES, H. de S.; ARAÚJO, R. F. de (Eds.). **Dos padrões internacionais de estruturação da informação de pesquisa aos indicadores: primeira incursão na temática**. Brasília: IBICT, DELBRA, 2019. p. 11

⁴² CÉSAR, I. D. Projeto CAPESQ [comunicações]. **Ciência da Informação**, v.3, n.2, 1974. Disponível em: <<http://ridi.ibict.br/handle/123456789/358>>.

Em Belém do Pará, Abner Lelis Vicentini (*in memoriam*) e Hagar Espanha Gomes transferem tecnologia, no âmbito do Sistema Nacional de Informação Científica e Tecnológica (SNICT)⁴³ para a Rede de Bibliotecas da Amazônia (REBAM), da Superintendência de Desenvolvimento da Amazônia (SUDAM) para replicar o CAPESQ. Três jovens mães se encontram.

Vida longa para você Hagar, minha mestra e a quem devo tanto! Vida longa também para Paulo Henrique de Assis Santana, o estagiário da bola, ainda enredado no BRCCRIS, e meu dileto amigo que me faz rir e me enche de motivação. Vida longa para Lena Vania Ribeiro Pinheiro e Anaiza Caminha Gaspar, ambas companheiras de REBAM. A primeira, integrante da coordenação do projeto, responsável pelo Sistema de Informação da Amazônia (SIAMAZ), um cadastro de projetos de pesquisa a ser relacionado às publicações do Catálogo Coletivo da REBAM. A segunda, representante do Banco de Desenvolvimento do Maranhão (BDM), a quem devo minha inserção no campo operacional do IBICT e onde até hoje me encontro.

Hagar, Lena e Anaiza são companheiras dos inesquecíveis tempos da REBAM e do nosso querido IBICT, a Casa das Raízes. Ela, nos seus 65 anos, nos dá sentido de pertencimento, equilíbrio e harmonia. Ela é a prova de que fazemos as coisas nutridas por grupos de antepassados e contemporâneos: um empreendimento coletivo que atravessa a História.

Não só a racionalidade, a individualidade também é um mito. Humanos raramente pensam por si mesmos. E sim, pensamos em grupos. Assim como é preciso uma tribo para criar uma criança, é preciso uma tribo para inventar uma ferramenta, resolver um conflito ou curar uma doença. Nenhum indivíduo sabe tudo o que é preciso para construir uma catedral, uma bomba atômica ou uma aeronave. O que deu ao Homo sapiens uma vantagem em relação a todos os outros animais e nos tornou os senhores do planeta não foi nossa racionalidade individual, mas nossa incomparável capacidade de pensar juntos em grandes grupos.⁴⁴

Sistemas CRIS podem ser desenhados de maneira criativa para pensar juntos em grandes grupos. Uma revisão de literatura, por exemplo, compulsiva incontáveis referências bibliográficas de incontáveis autores. Pode ser destaque entre os serviços para os usuários, principalmente jovens cientistas. No outro extremo, redes de inovação computam incontáveis dados e esquadrinham possibilidades de colaboração entre humanos.

Todo debate que requer mais ciência e toda ciência que requer mais debate requer ser ancorado no mesmo ideal para se tornar legítimo.

Vida longa ao IBICT e a todos que fazem dele a sua missão do bem-servir. Feliz aniversário!

Boa leitura!

⁴³ GOMES, H. E.; VICENTINI, A. L. C.; POMPEU, A. L.; TEIXEIRA, C.; SAMBAQUY, L. de Q. Diretrizes básicas para a implantação do Sistema Nacional de Informação Científica e Tecnológica (SNICT). *Ciência da Informação*, v.1, n.2, 1973. Disponível em: <<http://hdl.ibict.br/handle/123456789/353>>.

⁴⁴ NOAH, Yuval Harari. **21 lições para o século 21**. Rio de Janeiro: Companhia das Letras. 2018.

REFERÊNCIAS

ALCANFÔR, Kleber de Barros; CHAVES, Hélia de Sousa; BRANDÃO, Lilian Maria Thomé Andrade; CASTILHO, Rogério Mendes. Piloto BRICRIS IBICT-FAPEAL: uma avaliação exploratória do DSpace-CRIS para a construção de um sistema de informação de pesquisa local *In*: PEREIRA, Maria de Nazaré Freitas; CHAVES, Hélia de Sousa; ARAÚJO, Ronaldo Ferreira de (Eds.). **Dos padrões internacionais de estruturação da informação de pesquisa aos indicadores: primeira incursão na temática**. Brasília: IBICT, DELBRA, 2019.

ALVARES, Lillian; BERLANGA-LLAVORI, Rafael. A pesquisa científica e tecnológica sob a ótica da biblioteca digital brasileira de teses e dissertações. *In*: ENCONTRO IBÉRICO EDICIC, 7, 2017, **Atas: Ciência Aberta o contributo da Ciência da Informação**. Coimbra: Universidade de Coimbra - Centro de Estudos Interdisciplinares do Século XX, 2017.

ARAÚJO, Ronaldo Ferreira de. Os Estudos Cibernéticos da Informação: Indicadores de Estruturas Web e de Recursos da Web Social. *In*: PEREIRA, Maria de Nazaré Freitas; CHAVES, Hélia de SOUSA; ARAÚJO, Ronaldo Ferreira de (Eds.). **Dos padrões internacionais de estruturação da informação de pesquisa aos indicadores: primeira incursão na temática**. Brasília: IBICT, DELBRA, 2019.

BAKER, David; SIMONS, Edward; BROWN, Josh. **The various aspects of Interoperability: A strategic partnership driving interoperability in research information through standards**. euroCRIS Strategic Membership Meeting Autumn 2014 (KNAW, Amsterdam, Nov 11-12, 2014). Disponível em: <http://dspacecris.eurocris.org/bitstream/11366/354/1/Introduction_to_the_theme_of_the_meetingBaker_Brown_Simons_euroCRIS_SMM2014.pptx>. Acesso em: 20 de novembro de 2018.

CÉSAR, Ilse Dümpel. Projeto CAPESQ [comunicações]. **Ciência da Informação**, v.3, n.2, 1974. Disponível em: <<http://ridi.ibict.br/handle/123456789/358>>. Acesso em: 05 de março de 2019.

FARIAS, Juliana Pinheiro; BOMFIM, Kellen Cândida Vieira. Bibliografia anotada: identificadores persistentes (ORCID e RINGGOLD Identifiers) *In*: PEREIRA, Maria de Nazaré Freitas; CHAVES, Hélia de SOUSA; ARAÚJO, Ronaldo Ferreira de (Eds.). **Dos padrões internacionais de estruturação da informação de pesquisa aos indicadores: primeira incursão na temática**. Brasília: IBICT, DELBRA, 2019.

GOMES, Hagar Espanha. **Estudo CASRAI – Estudo com tabelas CNPq, FoS e ISCED**. Produto 2 do projeto: Dos padrões internacionais de estruturação da informação de pesquisa aos indicadores (BRICRIS). Rio de Janeiro. 2017. 32p. IBICT - Iniciativa de Apoio aos Diálogos Setoriais União Europeia/Brasil.

GOMES, Hagar Espanha. **Estudo CASRAI: Compatibilidade de sistemas**. Produto 1 do projeto: Dos padrões internacionais de estruturação da informação de pesquisa aos indicadores (BRICRIS). Rio de Janeiro. 2017. 32p. IBICT - Iniciativa de Apoio aos Diálogos Setoriais União Europeia/Brasil.

GOMES, Hagar Espanha; VICENTINI, Abner L.C.; POMPEU, Angela Lerche; TEIXEIRA, Cesar; SAMBAQUY, Lydia de Queiroz. Diretrizes básicas para a implantação do Sistema Nacional de Informação Científica e Tecnológica (SNICT). **Ciência da Informação**, v.1, n.2, 1973. Disponível em: <<http://ridi.ibict.br/handle/123456789/353>>. Acesso em: 05 de março de 2019.

MOTA, Paulo Jorge Gama. Ciência Aberta, Ciência Cidadã. *In: ConfOA. Anais...* Lisboa, 2018. Disponível em: <<https://pt.slideshare.net/ConfOA/cincia-aberta-cincia-cidad>>. Acesso em: 2 de dezembro de 2018.

NASCIMENTO, Andréa. **Maneiras alternativas de gerar métricas alternativas**. Bibliotecários sem fronteiras. Disponível em: <<https://bsf.org.br/>>. Acesso em: 2 de dezembro de 2018.

NOAH, Yuval Harari. **21 lições para o século 21**. Rio de Janeiro: Companhia das Letras. 2018.

OLIVEIRA, Carlos. Apresentação. *In: PEREIRA, Maria de Nazaré Freitas; CHAVES, Hélia de Sousa; ARAÚJO, Ronaldo Ferreira de (Eds.). Dos padrões internacionais de estruturação da informação de pesquisa aos indicadores: primeira incursão na temática*. Brasília: IBICT, DELBRA, 2019.

PEREIRA, Maria de Nazaré Freitas. **Por uma economia do conhecimento: avaliação de bases de dados nacionais para a produção de indicadores de C&T (Ciência e Tecnologia)**. Rio de Janeiro, 2003. Relatório Final de Bolsa de Produtividade de Pesquisa. Processo CNPq: 520416/1993-7(NV). viii, 150p., xcii.

PEREIRA, Maria de Nazaré Freitas. Prefácio que esclarece o leitor a propósito do sonho de Otlet: aventura em tecnologia da informação e comunicação. *In: PEREIRA, Maria de Nazaré Freitas, PINHEIRO, Lena Vania Ribeiro (orgs.). O sonho de Otlet: aventura em tecnologia da informação e comunicação*. Rio de Janeiro: IBICT/DEP, Brasília: IBICT/DDI, 2000.

REIS, José Eduardo dos; SPINOLA, Adriana Tahereh Pereira; AMARAL, Roniberto Morato do. Incipiência da visualização de indicadores bibliométricos e altmétricos nos Repositórios

Institucionais brasileiros. **Em Questão**, v. 23, 2017. Disponível em: <<https://www.redalyc.org/html/4656/465650499012/>>

SANTANA, Paulo Henrique de Assis. Relatório de viagem. *In: PEREIRA, Maria de Nazaré Freitas; CHAVES, Hélia de Sousa; ARAÚJO, Ronaldo Ferreira de (Eds.). Dos padrões internacionais de estruturação da informação de pesquisa aos indicadores: primeira incursão na temática*. Brasília: IBICT, DELBRA, 2019.

UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION AND THE INTERNATIONAL COUNCIL OF SCIENTIFIC UNION. **UNISIST: Synopsis of the Feasibility Study on a World Science Information System**. Paris: Unesco, 1971.

SITES CONSULTADOS

euroCRIS. Disponível em: <<https://www.eurocris.org/>>. Acesso em: 2 de dezembro de 2018.

DSpace-CRIS Home. Disponível em: <<https://wiki.duraspace.org/display/DSPACECRIS/DSpace-CRIS+Home>>. Acesso em: 2 de dezembro de 2018.

INCT. Institutos Nacionais de Ciência e Tecnologia. Disponível em: <<http://inct.cnpq.br/>>. Acesso em: 2 de dezembro de 2018.

Programa Pesquisa para o SUS (PPSUS). Disponível em: <<http://portalms.saude.gov.br/acoes-e-programas/ppsus>>. Acesso em: 2 de dezembro de 2018. .

Parte I

ARTIGOS



Desafios para a consolidação de indicadores em Ciência, Tecnologia e Inovação no Brasil: reflexões iniciais para a estruturação do Piloto BRCRIS IBICT-FAPEAL

VICTOR LEMOS TENORIO

FRANCISCA ROSALINE LEITE MOTA

HANS PONFICK ARAGÃO

RONALDO FERREIRA ARAÚJO

RESUMO

Considerações preliminares para a estruturação do Projeto BRCRIS: Piloto BRCRIS IBICT-FAPEAL, em estágio de desenvolvimento pelas equipes do Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) e da Fundação de Amparo à Pesquisa do Estado de Alagoas (FAPEAL). Para tanto reflete sobre a importância do monitoramento dos indicadores de ciência, tecnologia e inovação no Brasil, situando o papel da FAPEAL para o estado de Alagoas. Como esforço empírico realiza, por meio de modelagem de processos em notação BPMN, o diagnóstico do fluxo informacional dos processos de submissão e acompanhamento de projetos em C&T na FAPEAL, com levantamento dos metadados necessários aos diferentes subsistemas de informação, acompanhados de definições de padrões terminológicos e da identificação dos recursos informacionais utilizados como fonte de dados para popular os subsistemas CRIS.

Palavras chave: Indicadores de CT&I. BRCRIS. Sistemas de informação de pesquisa corrente. Modelagem BPMN.

Challenges to the consolidation of indicators for Science, Technology and Innovation in Brazil: initial reflections for the structuring of the BRCRIS IBICT-FAPEAL Pilot Project

ABSTRACT

Preliminary considerations on the structuring of the BRCRIS IBICT-FAPEAL Pilot Project, currently in development by teams from the Brazilian Institute for Information in Science and Technology (IBICT) and the Foundation for Research Funding of the State of Alagoas (FAPEAL). Reflections on the importance of monitoring science, technology and innovation indicators in Brazil, pointing out the role of FAPEAL for Alagoas. As an empirical endeavor, it presents, by modelling processes in BPM notation, a diagnosis of the information flow of submission and monitoring processes with regard to S&T projects in FAPEAL, with collection of the metadata necessary for the various information subsystems, accompanied by definitions of terminological standards and identification of the informational sources used to populate CRIS Subsystems.

Keywords: S&T indicators. BRCRIS. Current research information systems. BPMN modelling.

Desafíos para la consolidación de indicadores para la Ciencia, Tecnología e Innovación en Brasil: reflexiones iniciales para la estructuración del Proyecto Piloto BRCRIS IBICT-FAPEAL

RESUMEN

Consideraciones preliminares sobre la estructuración del Proyecto Piloto BRCRIS BICT-FAPEAL, actualmente en desarrollo por equipos del Instituto Brasileño de Información en Ciencia y Tecnología (IBICT) y la Fundación para el Financiamiento de la Investigación del Estado de Alagoas (FAPEAL). Reflexiones sobre la importancia de monitorear los indicadores de ciencia, tecnología e innovación en Brasil, destacando el papel de FAPEAL para Alagoas. Como un esfuerzo empírico, presenta, mediante el modelado de procesos en notación BPM, un diagnóstico del flujo de información de los procesos de presentación y monitoreo con respecto a los proyectos de C&T en FAPEAL, con la recopilación de los metadatos necesarios para los diversos subsistemas de información, acompañados de las definiciones de estándares terminológicos e identificación de las fuentes informativas utilizadas para poblar los subsistemas CRIS.

Palabras clave: *Indicadores de C&T. BRCRIS. Investigación actual de los sistemas de información. Modelado BPMN.*

INTRODUÇÃO

O desenvolvimento científico, tecnológico e de inovação (CT&I) é condição primordial para que um país se destaque ante os demais e tenha condições de responder às problemáticas sociais, educacionais, culturais, políticas e econômicas que enfrenta. Pensar ciência e tecnologia está longe de se circunscrever meramente ao campo das “ideias”, uma vez que boa parte dos problemas reais é resolvida no emprego do conhecimento técnico-científico no desenvolvimento de pesquisas que possuem grande impacto. Quando se refere à realidade brasileira, por mais que existam projetos e pesquisas com alcance internacional, ainda há inúmeros desafios ligados à gestão de recursos públicos e privados destinados a alavancar a CT&I no país.

No Brasil, os principais entes públicos responsáveis pela gestão orçamentária dos investimentos destinados à pesquisa são, respectivamente, o Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), a Financiadora de Estudos e Projetos (FINEP) e as Fundações Estaduais de Amparo à Pesquisa (FAPs). Para todos eles, o registro e controle das informações decorrentes do financiamento de projetos de pesquisa e a gestão adequada de indicadores de C&T se tornam indispensáveis ao processo de tomada de decisão, no que se refere à formulação de políticas públicas e à definição de áreas prioritárias para fomento.

Esse cenário não é uma peculiaridade exclusiva do país, sendo uma demanda de qualquer nação preocupada com a melhor alocação de recursos e investimentos em C&T, sobretudo em contextos adversos de crise econômica e cortes orçamentários. No âmbito internacional - com vistas à confluência de questões locais e globais de atores envolvidos com a atividade de pesquisa científica, seu fomento, resultados e impactos - têm surgido organizações, como a euroCRIS, associação internacional sem fins lucrativos, que reúne especialistas em informação de pesquisa. A euroCRIS atua como um fórum para que pessoas e organizações entrem em diálogo e resolução de assuntos relacionados ao uso da tecnologia da informação no desenvolvimento dos sistemas de informação de pesquisa (Current Research Information Systems - CRIS). A euroCRIS suporta o desenvolvimento do formato padrão para intercâmbio de informações em todos os aspectos do ciclo de vida da informação de pesquisa.

A missão da euroCRIS é promover a cooperação e compartilhar conhecimentos entre as comunidades de informação de pesquisa em âmbito internacional. Com essa finalidade, promove o contínuo aperfeiçoamento do Common European Research Information Format (CERIF), formato europeu comum para intercâmbio de informações de pesquisa. As áreas de interesse também abrangem bancos de dados de pesquisa, dados relacionados ao CRIS, como conjuntos de dados científicos, repositórios institucionais (acesso aberto), bem como mecanismos de acesso e intercâmbio de dados, padrões e diretrizes e melhores práticas para CRIS¹. O principal objetivo do CERIF é promover a interoperabilidade entre diferentes sistemas de informação CRIS.

O presente trabalho apresenta considerações e entendimentos iniciais para a estruturação do Projeto Piloto BRCRIS IBICT-FAPEAL, em estágio de desenvolvimento pelas equipes do Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) e da Fundação de Amparo à Pesquisa do Estado de Alagoas (FAPEAL).

As considerações e entendimentos têm por base a análise e diagnóstico do fluxo informacional dos processos de submissão e acompanhamento de projetos em C&T na FAPEAL, com levantamento dos metadados necessários aos diferentes subsistemas de informação, acompanhados de definições de padrões terminológicos e da identificação dos recursos informacionais utilizados como fonte de dados para popular os subsistemas CRIS.



¹ EuroCris. 2017. Disponível em: <<https://www.eurocris.org/what-eurocris>>. Acesso em: 20 jul. 2018.

INDICADORES DE CT&I NO BRASIL

O acompanhamento do avanço científico e tecnológico, a geração e o monitoramento dos indicadores de CT&I são indispensáveis para o desenvolvimento socioeconômico e para o aumento de produtividade, inclusive em países em desenvolvimento e de industrialização tardia, como o Brasil.

Indicadores de desempenho científico e de esforço tecnológico mostram que, na comparação com países em desenvolvimento, o Brasil é

[...] significativamente diferenciado em dimensões como escala de sua economia, inserção externa, tecnologia, qualificação da força de trabalho, investimento em pesquisa e desenvolvimento (P&D), cooperação técnica e produção científica.²

Essa característica de heterogeneidade da economia brasileira faz com que seja desafiador refletir sobre os rumos do desenvolvimento na superação de barreiras e entraves. O Brasil possui um núcleo de empresas que inova por meio da “[...] geração de conhecimento que é suportado por centros de excelência de universidades que conseguem acompanhar, quando não se conectar, com a pesquisa avançada em nível mundial”³.

Para os autores das citações acima, é exatamente na expansão desse núcleo dinâmico, na sua “[...] diversificação e integração com o conjunto da economia, tecendo malhas entre grandes, médias e pequenas empresas, que está a aposta no futuro de uma economia vibrante”⁴.

²ARBIX, G.; NEGRI, J. A.; VERMULM, R. Novos e velhos desafios do desenvolvimento. *Revista USP*, São Paulo, n. 93, p. 6-10, mar./maio, 2012. Disponível em: <<https://www.revistas.usp.br/revusp/article/view/44997/48609>>. Acesso em: 27 ago. 2018.

³Ibidem, p.8.

⁴Ibidem.

Em um possível cenário de expansão para esses núcleos, pensar em políticas de apoio à CT&I é primordial, sobretudo para o desenvolvimento de economias promovidas por um processo de desenvolvimento associado às capacitações tecnológicas, com ganhos com inovações, aumento da participação no mercado internacional, ampliação e fortalecimento do mercado interno.⁵

Podemos ressaltar três razões específicas da necessidade de sistemas de indicadores de CT&I:

- (i) a científica, relacionada com a busca da compreensão dos fatores determinantes dos processos de produção;
- (ii) a política, associada com as necessidades e possibilidades de utilização dos indicadores de CT&I como instrumentos para a formulação, acompanhamento e avaliação de políticas públicas; e a
- (iii) pragmática, que se refere ao uso dos indicadores como ferramenta auxiliar na definição e avaliação de estratégias tecnológicas de empresas, bem como na orientação das atitudes e ações de trabalhadores, instituições e do público em temas relacionados com CT&I.⁶

Geralmente, os indicadores de CT&I são pensados a partir de dimensões básicas constitutivas de um sistema de inovação estadual, tais como: prioridade governamental no que concerne à área de ciência e tecnologia; produção científica e tecnológica; base educacional e disponibilidade de recursos humanos qualificados, e amplitude e difusão da inovação no âmbito das empresas localizadas no estado.⁷

No primeiro, a prioridade é medida levando em consideração os investimentos públicos e gastos no campo científico e tecnológico (gasto per capita governamental em ciência e tecnologia; percentual de gasto em CT&I).

⁵ OLIVEIRA, F. C. B. Indicadores de Ciência, Tecnologia e Inovação: uma análise dos membros do BRIC - Brasil, Rússia, Índia e China. In: ENCONTRO DE ECONOMIA CATARINENSE, V, Florianópolis, 2011. *Anais...* Florianópolis, p. 1-25. Disponível em: <<https://goo.gl/VqX4Ak>>. Acesso em: 27 ago. 2018.

⁶ SARTORI, R.; PACHECO, R. C. S. Indicadores de ciência, tecnologia e Inovação: a interação humana nos grupos de pesquisa. *Revista textos de la CiberSociedad*, 2008. Disponível em: <<http://www.cibersociedad.net/textos/articulo.php?art=186>>. Acesso em: 27 ago. 2018.

⁷ ROCHA, E. M. P.; FERREIRA, M. A. T. Indicadores de ciência, tecnologia e inovação: mensuração dos sistemas de CTel nos estados brasileiros. *Ci. Inf.*, Brasília, v. 33, n. 3, p. 61-68, Dec. 2004. Disponível em: <<http://www.scielo.br/pdf/ci/v33n3/a08v33n3.pdf>>. Acesso em: 27 ago. 2018.

O segundo relaciona-se à capacidade criativa de seu corpo de pesquisadores, para o que contribuem o volume, o nível e a qualidade da produção científica e tecnológica (livros, artigos, patentes).

No terceiro item importa a formação e a capacitação de profissionais para a CT&I (taxa de escolarização de jovens; número de cursos de pós-graduação e de alunos matriculados; número de mestres e doutores titulados; número de bolsas para mestrado e doutorado, dentre outros). No último item é considerado o grau de participação e envolvimento das empresas na condução e financiamento das atividades de pesquisa.^{8,9}

Para o bom acompanhamento desses e de outros indicadores de CT&I que deles derivam, e melhor compreensão do estágio de desenvolvimento científico e tecnológico do país, são necessários sistemas e fontes de informação estruturada por normas e padrões internacionais, para fins de gestão e interoperabilidade.



⁸ Ibidem.

⁹ SILVEIRA, Aline Dario et al. Análise do Sistema Nacional de Inovação no setor de energia na perspectiva das políticas públicas brasileiras. *Cad. EBAPE.BR*, Rio de Janeiro, v. 14, n. spe, p. 506-526, jul. 2016. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1679-39512016000700506&lng=pt&nrm=iso>. Acesso em: 27 ago. 2018.

A FUNDAÇÃO DE AMPARO À PESQUISA DO ESTADO DE ALAGOAS: CONSIDERAÇÕES NECESSÁRIAS

A Fundação de Amparo à Pesquisa do Estado de Alagoas (FAPEAL) surgiu com a inserção do capítulo IV da Constituição Estadual de 1988, intitulado “Da Ciência e da Tecnologia”. O artigo 216 do referido capítulo estabelece a destinação orçamentária de 2% da receita estimada para o financiamento da Ciência e Tecnologia, a criação de uma fundação de amparo ao desenvolvimento científico e tecnológico, e regras gerais de funcionamento da instituição.¹⁰

Um ano depois, em 27 de setembro de 1990, a Assembleia Legislativa promulgou a lei complementar número 5, levando a FAPEAL a sair do estado de projeto para se tornar instituição, assumindo a condição de entidade jurídica de direito privado. No processo de reforma administrativa do estado, sua estrutura jurídica foi alterada para fundação de direito público, através da lei complementar nº 20, de 04 de abril de 2002.¹¹

Em quase 27 anos de atividade, a FAPEAL destaca-se cada vez mais como instituição que promove ciência, tecnologia e inovação, fomentando e induzindo a pesquisa para o desenvolvimento sustentável do estado de Alagoas, através da concessão de bolsas de estudo, financiamento a pesquisas e apoio a programas e projetos de desenvolvimento do estado.¹²

A FAPEAL exerce atividades de fomento à pesquisa e indução tecnológica, além de ser gestora do Ponto de Presença (PoP) da Rede Nacional de Pesquisa (RNP), que ligou pioneiramente o estado de Alagoas à internet, oferecendo serviços de conectividade à comunidade científica, instituições governamentais e ONGs.

¹⁰ FUNDAÇÃO DE AMPARO À PESQUISA DO ESTADO DE ALAGOAS. Disponível em: <<http://www.fapeal.br/>>. Acesso em: 21 fev. 2018.

¹¹ Ibidem.

¹² Ibidem.

A Fundação viabiliza projetos de pesquisa induzidos que priorizam áreas estratégicas para o desenvolvimento do estado. Dentre suas realizações encontram-se o programa Pró-Ciências, que capacitou mais de 600 professores das disciplinas de matemática, física, química e biologia, e o projeto Nordeste de pesquisa e pós-graduação, que estimulou a realização de cursos de pós-graduação em Alagoas. Em 2017 foi destaque, recebendo as seguintes menções honrosas: gestão compartilhada do programa pesquisa para o SUS – PPSUS, na categoria “Produtos e Inovação em Saúde e Experiência Exitosa do Programa Pesquisa para o SUS: gestão compartilhada em saúde – PPSUS”, com a segunda menção pelo “Incentivo e Fomento à Ciência, Tecnologia e Inovação”.¹³ Acrescente-se ainda ao portfólio da Fundação as centenas de bolsas e auxílios que oferece anualmente, melhor qualificando e contribuindo para a formação de recursos humanos de alto nível para Alagoas.



¹³FUNDAÇÃO DE AMPARO À PESQUISA DO ESTADO DE ALAGOAS. Nosso PPSUS é modelo para o Brasil. *Fapeal em Revista*, v.2, n. 4, p. 14-17, 2018. Disponível em: <https://issuu.com/fapealemrevista1/docs/revista_hd_online>. Acesso em: 28 ago. 2018.

MATERIAL E MÉTODO

Trata-se de uma pesquisa qualitativa de cunho exploratório, desenvolvida por meio da técnica de estudo de caso com pesquisa de campo, para análise e diagnóstico do fluxo informacional dos processos de submissão e acompanhamento de projetos em C&T na FAPEAL. Essa abordagem metodológica foi a mais adequada, tendo em vistas as fases que envolvem o processo de estruturação do BRCRIS, com levantamento dos metadados necessários aos diferentes subsistemas de informação, acompanhados de definições de padrões terminológicos e da identificação dos recursos informacionais utilizados como fonte de dados para popular os subsistemas CRIS, como geração de metodologia para orientar estudos semelhantes em outras FAPs.

Os procedimentos metodológicos adotados seguiram os seguintes passos: utilização da ferramenta Bizagi Modeler para modelagem do processo existente e definição e especificação dos metadados a serem utilizados no desenvolvimento do projeto piloto.

O Bizagi Modeler é um software de gerenciamento de processos de negócio que permite criar representação gráfica através de um gráfico de fluxo, além de gerar a documentação e simular processos do negócio em análise. Permite também a publicação dos processos em uma ampla gama de formatos. O sistema usa a notação *Business Process Modeling Notation* (BPMN) por meio de elementos gráficos, o que facilita o desenvolvimento de diagramas, com leitura simples e intuitiva, e de fácil identificação, aumentando a expressividade dos gráficos.¹⁴

Para definição dos metadados, procedeu-se inicialmente a uma pesquisa bibliográfica, por meio de consultas a fontes, direta ou indiretamente, relacionadas à elaboração de mapeamento de fontes de informação, utilizando o padrão de notação BPMN. O padrão permite o registro da execução do processo do fluxo documental, da submissão ao acompanhamento do projeto, na forma de mapa de processo de trabalho. Foram realizadas entrevistas complementares com integrantes do quadro funcional, sempre considerando pessoas-chave nos processos internos, do ponto de vista de como funciona atualmente o seu mecanismo operacional.



¹⁴ BUSINESS PROCESS MODEL AND NOTATION – BPMN. 2018. Disponível em: <<http://www.bpmn.org/>>. Acesso em: 27 ago. 2018.

ENTIDADES SELECIONADAS

Baseado nos estudos realizados para a criação dos fluxos de dados e no conhecimento gerado através do projeto Arquivo FAPEAL, além de análise do sistema ORCID e da Plataforma Lattes, conseguimos identificar cinco entidades que se destacaram: pesquisador, projeto, produção em CT&I, instituição e edital. E após definir as entidades, selecionamos os dados que permitirão melhor integração entre as bases de dados para cada entidade (figura 1).

Figura 1 - Entidades



No contexto dessa análise, o pesquisador é a entidade proponente responsável pela submissão de projetos de pesquisa. Cada pesquisador está vinculado a uma instituição. Os projetos são documentos técnico-científicos que descrevem o teor das propostas apresentadas ao edital; por sua vez, esses dispõem de todas as diretrizes que normatizam e condicionam formas de participação no financiamento. A produção em CT&I é aquela publicação oriunda dos resultados alcançados na execução dos projetos de pesquisa financiados pela Fundação.

Os dados das entidades consideradas foram obtidos e ou validados das fontes listadas no quadro 1. Trata-se de sistemas e ou serviços de informação que apoiaram a identificação e consolidação de dados de cada entidade.

Quadro 1 - Fontes de informação das entidades selecionadas

SIGLA	Descrição
SIGFAP	Sistema de informação e gestão de projetos das fundações estaduais de amparo à pesquisa.
BASE ACCESS	Sistema de gerenciamento de banco de dados da Microsoft.
ORCID	Código alfanumérico não proprietário para identificar exclusivamente cientistas e outros autores acadêmicos e contribuidores.
LATTES	Criada e mantida pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), a plataforma integra as bases de dados de currículos, grupos de pesquisa e instituições, em um único sistema de informação das áreas de ciência e tecnologia atuando no Brasil.

O SIGFAP é adotado pela FAPEAL para cadastro de pesquisadores, sendo necessário possuir login e senha para concorrer em qualquer edital da Fundação. O sistema é fruto de cooperação técnica entre as Fundações de Amparo à Pesquisa e a Fundação de Apoio ao Desenvolvimento do Ensino, Ciência e Tecnologia do Estado do Mato Grosso do Sul (FUNDECT).

O SIGFAP tem como objetivo desenvolver soluções computacionais e tecnológicas e de inovação, tornando a gestão nas FAPs mais eficiente e eficaz, por meio de planejamento, acompanhamento e avaliação de dos programas e projetos de CT&I. O sistema proporciona transparência pública nas suas ações, a facilitação da comunicação entre comunidades científicas, democratizando as informações e indicadores de CT&I. Permite a divulgação, submissão, avaliação e acompanhamento da execução de projetos de pesquisa, de extensão, apoio a eventos e bolsas de diferentes modalidades.¹⁵

¹⁵ FUNDECT - Fundação de Apoio ao Desenvolvimento do Ensino, Ciência e Tecnologia do Estado de Mato Grosso do Sul: Sistema de Informação e Gestão de Projetos das Fundações Estaduais de Amparo à Pesquisa. 2014. Disponível em: <<https://sigfundect.ledes.net/>>. Acesso em: 14 set. 2016.

A Base Access agrega todas as informações dos pesquisadores, desde às pessoais e de vínculo institucional registradas no SIGFAP, bem como dados gerais sobre a relação do pesquisador com a Fundação, como, por exemplo, número de propostas submetidas a quantos e quais editais, resultado das submissões (aprovadas, rejeitadas), prestações de contas realizadas e pareceres emitidos. O ORCID e o Lattes são fontes auxiliares na validação de informações prestadas pelos pesquisadores em seu cadastro no SIGFAP e posterior agregação na Base Access.



FLUXOS INFORMACIONAIS E METADADOS PARA O PILOTO BRCRIS IBICT-FAPEAL

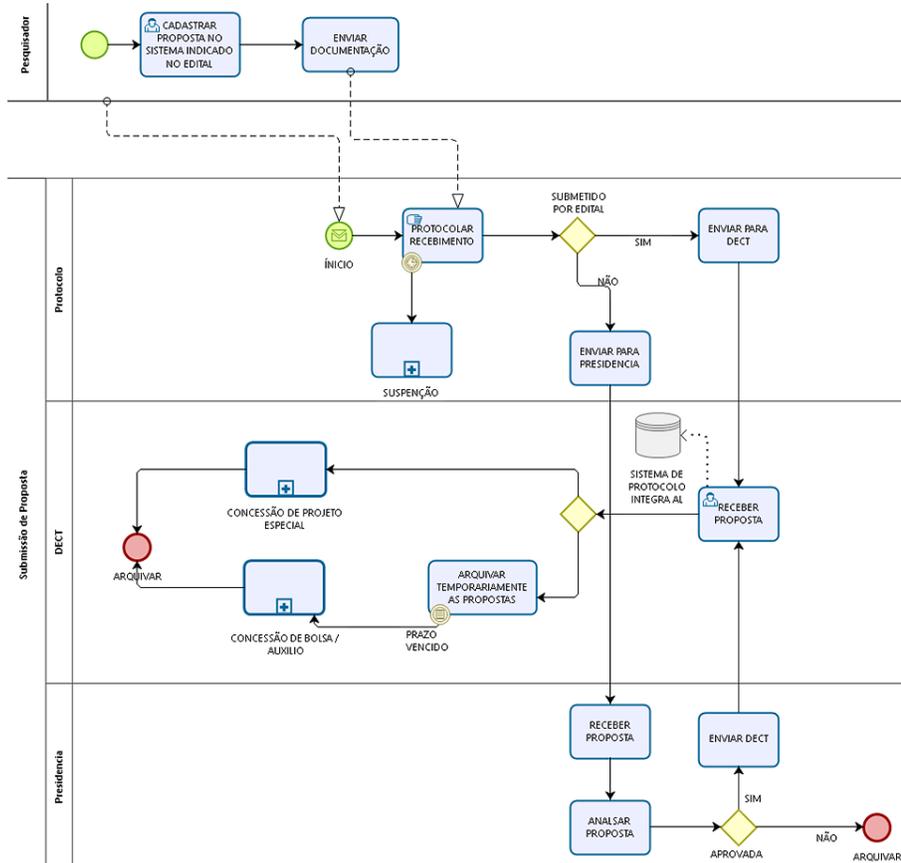
Primeiramente foi feito o levantamento de todos os processos da instituição, o que se deu mediante da análise dos documentos que continham os organogramas, normas, manuais e procedimentos. Através dessa análise foi criado um esboço do fluxo de funcionamento da instituição. Em sua maioria, os documentos datavam de antes de antes de 2012 e algumas informações se mostravam defasadas.

A partir desse ponto iniciou-se análise para obtenção de dados primários sobre o funcionamento da instituição com os profissionais responsáveis pela execução das atividades de cada processo, com origem em diversos setores da FAPEAL, o que nos permitiu expandir e aperfeiçoar o fluxo até sua versão atual. Foram necessárias várias incursões pela Fundação, várias entrevistas com seus representantes, grande número de mensagens trocadas com eles, até o fluxo se encontrar completo.

O primeiro fluxo é o de lançamento dos editais após assinatura do convênio. Em resumo, a presidência da FAPEAL solicita à Diretoria Executiva de Ciência e Tecnologia (DECT) a elaboração do edital. Após sua conclusão, ele é enviado ao setor jurídico para emitir parecer; se favorável, segue para o setor financeiro para reserva de recurso, continua para a presidência para ser homologado e, após tudo isso, é lançado pela DECT.

Tendo em vista a envergadura do conjunto das etapas de processos analisados e a dimensão das imagens geradas no mapeamento da ferramenta utilizada, optou-se neste artigo por apresentar de forma mais detalhada dois processos: o de submissão de propostas (figura 2) e o de prestação de contas (figura 3).

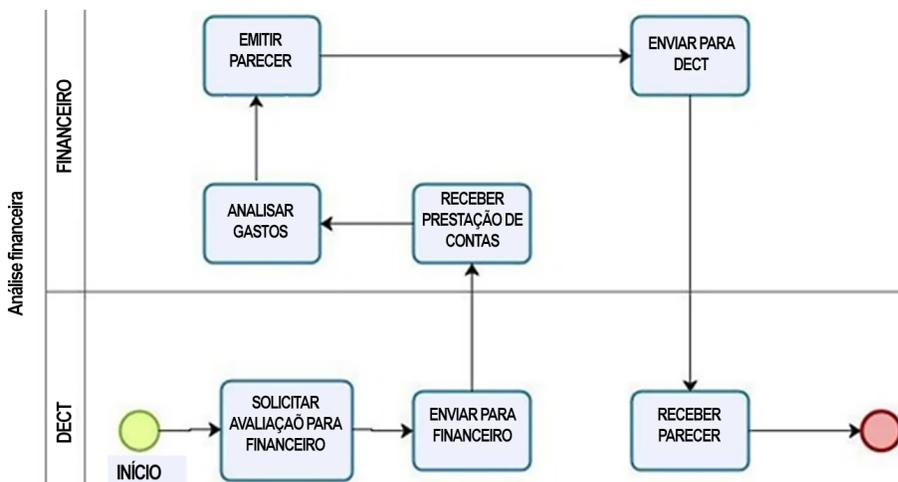
Figura 2 - Processo de submissão de propostas - padrão



A figura 2 mostra o processo de submissão de propostas utilizando a notação BPMN e a ferramenta Bizagi, tanto para concessão de auxílios como para obtenção de bolsas.

O fluxo de tarefas do processo consiste das seguintes etapas: a) **Cadastro de Proposta no Sistema Indicado pelo Edital** – Após conhecimento do edital vigente e de interesse, o pesquisador submete sua proposta ao sistema informatizado indicado pelo edital, podendo, na maioria dos casos, ser o SIGFAP b) **Envio da Documentação** – etapa na qual destinada a entrega de documentação complementar especificada de acordo com edital; c) **Protocolar Recebimento (Protocolo FAP)** – Nesse ponto é necessário o registro de recebimento em um sistema informatizado do governo do estado, o INTEGRA, para tramitação física do processo; d) **Encaminhamento para o setor competente** - Gateway de uma ou mais saídas de propostas para encaminhamento do processo físico; e) **Recebimento de Proposta** - Nesse momento, a assessoria da DECT tem a opção de arquivar temporariamente a proposta, até a data de encerramento de recebimento de propostas; f) **Subprocesso de Concessão de Auxílio ou Bolsa** - as propostas seguem caminhos de acordo com edital e sua coordenação.

Figura 3 - Etapas do processo de prestação de contas



Na figura 3, é possível mostrar o subprocesso de análise financeiro de prestação de contas de um pesquisador. Neste momento a) **Solicitação de Avaliação Financeira** – A DECT faz a solicitação ao setor responsável para que seja feito o comparativo de gastos que foram propostos no projeto de pesquisa; b) **Recebimento da Prestação**; c) **Análise de Gastos** – Neste ponto é necessário que o setor faça um comparativo com que o pesquisador propôs no seu plano de trabalho e se está de acordo com as normas de prestação de contas da instituição; d) **Emissão de parecer** – Necessário para registrar a atual situação financeira do projeto; e) **Envio para DECT** – O subprocesso termina e retorna para o processo geral de acompanhamento de auxílios e bolsas;

O mapeamento permitiu observar que, dependendo do tipo de auxílio e até mesmo da bolsa, os setores responsáveis registram informações sobre os processos em sistemas diferentes, e/ou através de planilhas eletrônicas.

Os resultados preliminares mostram que é possível desenvolver a estruturação de um componente estadual de um Sistema Nacional de Informação em Pesquisa, que contemple os padrões internacionais CERIF e ORCID, o uso de vocabulários controlados, as tabelas CASRAI e de áreas de conhecimento do CNPq.



CONSIDERAÇÕES FINAIS

O estudo teve por objetivo a realização de um mapeamento de processos típicos em fundações de fomento à pesquisa para identificar o fluxo informacional, geração de dados, seleção de sistemas de informação e seus respectivos metadados, no âmbito do projeto piloto BRCRIS IBICT-FAPEAL. Utilizou-se como objeto de estudo todo o processo de submissão e acompanhamento de propostas submetidas às diferentes modalidades de auxílios e bolsas disponibilizados pela instituição. O método aplicado foi o da modelagem de processos em notação BPMN, utilizando o sistema BizAgi® para representação gráfica. Como resultados desse estudo, foi possível analisar, através dos fluxogramas obtidos, a sequência completa de atividades da instituição, e com isso visualizar quais processos poderiam apresentar geração de dados pertinentes para popular um sistema CRIS.

Cabe ressaltar que com a tramitação do dossiê entre os setores, alguns documentos importantes que compõem dados necessários para eventuais buscas de indicadores são adicionados em meio físico. Infelizmente, não se registram, em meio eletrônico, quais documentos são adicionados; por exemplo, informações sobre publicações de artigos em revistas ficam registradas apenas em papel, dentro de um relatório científico! Reforça-se, assim, a extrema necessidade de se identificar no fluxo onde os dados são gerados e de compreender a presença desses “gargalos de produção”, os quais, provavelmente, devem ocorrer em outras instituições de fomento.

A FAPEAL representa hoje uma instituição inteiramente comprometida com o desenvolvimento e com o destino do estado de Alagoas. Sua missão é a de promover o intercâmbio e a formação de pesquisadores através da concessão de bolsas de estudo, financiamento a pesquisas e apoio a programas e projetos de desenvolvimento.

O alto índice de organização documental atualmente existente na FAPEAL, somado ao diálogo estreito com setores-chave da instituição-, um diferencial sobre outras FAPs do mesmo porte-, são condições necessárias para desenvolvimento de um piloto do Projeto BRCCRIS, com representação dos subsistemas essenciais para o ciclo de vida da pesquisa. Ao mesmo tempo, o estágio alcançado pelos estudos de classificação, no âmbito do projeto Diálogos Setoriais (DS), também origem do estudo que resultou no presente artigo, representa concretamente a inserção do piloto na geração de indicadores de C&T&I, cuja semântica está presente desde a definição dos metadados, e não apenas na indexação de diferentes aspectos dos registros que formam os subsistemas de informação de pesquisa.



REFERÊNCIAS

ARBIX, G.; NEGRI, J. A.; VERMULM, R. Novos e velhos desafios do desenvolvimento. *Revista USP*, São Paulo, n. 93, p. 6-10, mar./maio, 2012. Disponível em: <<https://www.revistas.usp.br/revusp/article/view/44997/48609>>. Acesso em: 27 ago. 2018.

BUSINESS PROCESS MODEL AND NOTATION – BPMN. 2018. Disponível em: <<http://www.bpmn.org/>>. Acesso em: 27 ago. 2018.

FUNDAÇÃO DE AMPARO À PESQUISA DO ESTADO DE ALAGOAS. Nosso PPSUS é modelo para o Brasil. *Fapeal em Revista*, v.2, n. 4, p. 14-17, 2018. Disponível em: <https://issuu.com/fapealemrevista1/docs/revista_hd_online>. Acesso em: 28 ago. 2018

FUNDECT - Fundação de Apoio ao Desenvolvimento do Ensino, Ciência e Tecnologia do Estado de Mato Grosso do Sul: Sistema de Informação e Gestão de Projetos das Fundações Estaduais de Amparo à Pesquisa. 2014. Disponível em: <<https://sigfundect.ledes.net/>>. Acesso em: 14 set. 2016.

OLIVEIRA, F. C. B. Indicadores de Ciência, Tecnologia e Inovação: uma análise dos membros do BRIC - Brasil, Rússia, Índia e China. In: ENCONTRO DE ECONOMIA CATARINENSE, V, Florianópolis, 2011. *Anais...* Florianópolis, p. 1-25. Disponível em: <<https://goo.gl/VqX4Ak>>. Acesso em: 27 ago. 2018.

ROCHA, E. M. P.; FERREIRA, M. A. T. Indicadores de ciência, tecnologia e inovação: mensuração dos sistemas de CTel nos estados brasileiros. *Ci. Inf.*, Brasília, v. 33, n. 3, p. 61-68, Dec. 2004. Disponível em: <<http://www.scielo.br/pdf/ci/v33n3/a08v33n3.pdf>>. Acesso em: 27 ago. 2018.

SARTORI, R.; PACHECO, R. C. S. Indicadores de ciência, tecnologia e Inovação: a interação humana nos grupos de pesquisa. *Revista textos de la CiberSociedad*, 2008. Disponível em: <<http://www.cibersociedad.net/textos/articulo.php?art=186>>. Acesso em: 27 ago. 2018.

SILVEIRA, Aline Dario et al. Análise do Sistema Nacional de Inovação no setor de energia na perspectiva das políticas públicas brasileiras. *Cad. EBAPE.BR*, Rio de Janeiro, v. 14, n. spe, p. 506-526, jul. 2016. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1679-39512016000700506&lng=pt&nrm=iso>. Acesso em: 27 ago. 2018.

Classificação e gestão do conhecimento

Hagar Espanha Gomes

RESUMO

A classificação sempre esteve presente na organização do conhecimento. No campo dos registros do conhecimento, catálogos, bibliografias e esquemas bibliotecários foram os meios utilizados. A introdução da computação e do Método de Faceta tornou possível a produção de esquemas dinâmicos – em oposição à escrita linear no papel – suscitando novos produtos e usos, em especial na produção de indicadores para a gestão do conhecimento. A abordagem indutiva preconizada pelo método – contrariando a tradição da abordagem dedutiva na construção de esquemas de classificação –, produz esquemas consistentes, sejam eles genéricos ou não.

Palavras-chave: Classificação. Método de Faceta. Organização do conhecimento. Gestão do conhecimento.

Knowledge classification and management

ABSTRACT

Classification has always been present in the organization of information, with catalogues, bibliographies and library schemes being the most used means of recording knowledge. The introduction of computing and the Facet Method have made possible the creation of dynamic schemes – as opposed to paper-based linear writing – which make possible new products and uses, especially in the production of indicators for knowledge management. The inductive approach recommended by the method – contradicting the traditional deductive approach in the creation of classification schemes – generates consistency in these schemes whether they are generic or not.

Keywords: *Classification. Facet Method. Organization of knowledge. Knowledge management.*

Clasificación y gestión del conocimiento

RESUMEN

La clasificación siempre ha estado presente en la organización de la información, y los catálogos, bibliografías y esquemas de bibliotecas son los medios más utilizados para registrar el conocimiento. La introducción de la computación y el Método de faceta han hecho posible la creación de esquemas dinámicos, a diferencia de la escritura lineal basada en papel, que posibilita nuevos productos y usos, especialmente en la producción de indicadores para la gestión del conocimiento. El enfoque inductivo recomendado por el método, que contradice el enfoque deductivo tradicional en la creación de esquemas de clasificación, genera coherencia en estos esquemas, ya sean genéricos o no.

Palabras clave: *Clasificación. Método de la faceta. Organización del conocimiento. Conocimiento administrativo.*

INTRODUÇÃO

As várias atividades que envolvem o profissional de informação requerem, de um modo ou outro, algum modo de organização e raciocínio lógico – catalogação, indexação, recuperação – que encontram na classificação apoio para decisão. Numa exposição *a vol d'oiseau* pode-se observar não apenas sua presença, mas as diferentes funções e formas no tempo.

A Classificação esteve sempre presente na atividade humana. Como processo, é o ato de reunir entidades de acordo com características comuns. Nesta visão atual, rompe com a dicotomia que caracterizou os primeiros movimentos de classificação na ciência. Assim, um objeto pode ser classificado de vários modos, segundo as várias características, os diferentes pontos de vista, por exemplo, filosófico, científico, histórico ou outro. Quando apenas uma arte, a classificação foi aplicada de diversos modos e formas à medida que nosso conhecimento se desenvolvia.¹

A classificação tem grande relevância nas ciências, porém os maiores desenvolvimentos no campo teórico têm ocorrido na Biblioteconomia/Ciência da Informação. Com a introdução da Tecnologia da Informação (TI) na sociedade, a classificação se mostrou mais uma vez indispensável na produção de diferentes produtos e serviços, em especial por estar intimamente ligada à Lógica, como na TI. O que a seguir se apresenta é um resumo dos diferentes usos e modos da classificação, do ponto de vista da Ciência da Informação.

A Ciência da Informação se ocupa não apenas de questões teóricas, mas lida com classificações práticas seja para organização de bibliotecas e de bibliografias, seja para a organização de outros meios de registro como discos, filmes, objetos de museus e assemelhados. Atualmente, a necessidade de gerenciar grandes massas de registros informacionais trouxe novos desafios, e a classificação está no cerne desta necessidade como solução para a organização de tais registros para os diversos usos desejados, cada vez mais presente na atividade humana, em especial em meio digital.

¹ DAHLBERG, I. Teoria da Classificação ontem e hoje. Tradução do inglês por Henry B. Cox. Palestra apresentada à Conferência Brasileira de Classificação Bibliográfica, Rio de Janeiro, 12-17 de setembro de 1972. *Anais...* Brasília, IBICT/ABDF, 1979. v. 1, p. 352-370. Disponível também em <www.conexaoario.com/bit/>.

Organização do conhecimento científico ou cultural não prescinde da classificação. Filósofos no início e profissionais de informação posteriormente têm-se debruçado no estudo dos princípios de classificação. A literatura registra essa preocupação manifestada desde a Antiguidade, atravessando a Idade Média, os tempos modernos até a introdução da tecnologia da informação. No cerne das realizações bibliográficas e bibliotecárias, cientistas e profissionais de informação têm-se ocupado da organização dos registros do conhecimento, propostas de ação visando disseminação e reuso de tais registros surgem em âmbito nacional e internacional e nos dias de hoje a classificação está presente também nas atividades de planejamento e gestão da atividade científica manifestada nos esquemas de classificação para indicadores.

A oportunidade de desenvolver estudos sobre classificações para indicadores de pesquisa científica em curso mostra novas perspectivas para os sistemas práticos de classificação². Nesta área, como em muitas outras da atividade científica, existe um continuum no desenvolvimento dos princípios de classificação, mantendo-se alguns, abandonando-se outros. Existe um fio condutor no tempo nos movimentos dos cientistas a respeito de suas necessidades de informação, que é o acesso no final do fluxo: é este que vai possibilitar a reutilização do conhecimento produzido para novos avanços. Vale lembrar a primeira Lei da Informação: A Informação é para ser usada.

Quais foram essas necessidades, o que se conseguiu avançar o que ainda falta é preciso saber para não repetir experiências.

Um breve percurso nas iniciativas de classificações nos auxilia a compreender os problemas e as soluções propostas por seus desenvolvedores. Mas para melhor visão crítica a respeito das várias propostas de classificação, é necessário conhecer princípios que norteiam a construção de sistemas de classificação.

²Sistemas práticos são aqueles desenvolvidos com alguma finalidade prática, em oposição às classificações científicas desenvolvidas por filósofos e cientistas.

De início, vale registrar aspectos divergentes entre os sistemas de classificação teóricos e os sistemas práticos. Enquanto, no primeiro caso, cientistas e filósofos concebem seu modelo de organização da ciência, nos sistemas práticos o grande complicador tem sido organizar sistemas de classificação para representar o assunto, o tema tratado, em suma, o conteúdo dos registros.³ Se, no primeiro caso, o modelo é um exercício intelectual, no segundo, a problemática de representar assunto requer ajustes e concessões na organização de esquemas que possuem finalidades práticas, uma vez que assuntos cobrem em geral múltiplos tópicos ou categorias dificultando muitas vezes ao classificador encontrar no esquema o melhor lugar. No entanto, as bases teóricas da Classificação, mesmo voltadas para a produção de esquemas práticos, não devem se fixar no assunto. Pelo contrário: Richardson⁴ registra a confusão entre aqueles que consideram que a classificação se refere a assuntos e não a objetos e ele está convicto que este é um 'profundo engano teórico e prático que leva à confusão interminável'. Neste aspecto, seu estudo é pioneiro.



³Mais recentemente, representações de objetos têm adquirido grande importância em produtos e serviços de informação, porém nesse âmbito a problemática do 'assunto', está ausente.

⁴RICHARDSON, E.C. *Classification: theoretical and practical*. New York: Scribner, 1912. p. ix.

CLASSIFICAÇÃO E BIBLIOGRAFIA

A construção de Sistemas de classificação em diversas esferas da atividade humana tem longo percurso. Pode-se afirmar que as contribuições mais antigas procedem de filósofos e cientistas. Richardson, em seu livro de 1912, registra mais de 150 (cento e cinquenta) iniciativas do que ele chama de 'sistemas teóricos': inicia com Platão e Aristóteles passando por filósofos como Francis Bacon, Locke e Spencer. Na mesma obra, registra 176 (cento e setenta e seis) 'sistemas práticos', ou seja, aqueles voltados à bibliografia e à biblioteca começando com Calímaco, no século III, até a classificação de Bliss em 1910.

Mas a história mostra que os esquemas de classificação bibliográfica se estruturaram em grandes áreas do saber. A primeira e grande iniciativa de aplicar esquema de classificação à bibliografia deve-se a Gesner (1516-1565) com sua *Bibliotheca Universalis*, publicada em 1545, na qual registra cerca de 12 mil obras classificadas alfabeticamente pelo prenome dos autores de livros publicados em latim, grego e hebraico durante o primeiro século da Imprensa, completando sua obra em 1548 com um quadro sistemático que contém a classificação dos saberes em 20 áreas do conhecimento: Grammatica, Dialectica, Rhetorica, Poetica, Arithmetica, Geometria, Música, Astronomia, Astrologia, De Divinatione et Magia, Geographia, De Historiis, De diversis artibus, De naturali Philosophia, De oeconomica Philosophia, Política, De Jus civili et pontifico, Theologia. Sua obra contém atualizações mesmo depois de sua morte, até 1730⁵.

De algum modo, alguns sistemas práticos foram influenciados, até o início do século XX, por sistemas filosóficos. Uma das grandes influências foi o modelo escolástico da universidade medieval que organizava o saber em *trivium* e *quadrivium*. Assim, em vários sistemas práticos percebe-se esta influência em sua estrutura. No entanto, ao aplicar às bibliotecas algumas concessões seriam feitas. Por exemplo, o filósofo e matemático Leibniz (1646-1716), que atuou como bibliotecário para diversos nobres, inclusive para Pedro, o Grande, da Rússia, desenvolveu vários esquemas de classificação e, embora sendo filósofo, rejeitou uma classificação para biblioteca como simplesmente rigorosa e lógica, em favor de um esquema prático baseado no ensino nas faculdades.

⁵MALCLÈS, Louise-Noëlle. *La Bibliographie*. Paris: Presses Universitaires de France, 1956. p. 26.

Nos esquemas que desenvolveu, ele esteve atento ao princípio da garantia literária⁶ e incluiu uma classe *Divinatoria* que incluía Astrologia, Fisionomia, Quiromancia e tópicos semelhantes, embora não acreditasse em tais coisas, mas porque havia uma considerável quantidade de literatura sobre aqueles assuntos, e então era necessário providenciar-lhes um lugar. E percebeu ainda a dificuldade de classificação de livros em uma biblioteca, pois nem sempre se sabe em que lugar colocar certos livros, que poderiam muito bem ficar igualmente entre dois ou três lugares na estante⁷. A linearidade da classificação levou-o a organizar um catálogo alfabético inclusive com remissivas e, por todas as suas iniciativas, Leibniz deixou sua marca na História da Classificação.

Muitas iniciativas poderiam ser incluídas aqui⁸, mas o alvo são as obras que contêm algum esquema de classificação em suas atividades sistemáticas como bibliografias correntes e bibliotecas.

No capítulo de seu opúsculo *La Bibliographie*⁹ -'L'époque Artisanale'- Malclès registra o surgimento de bibliografias especializadas correntes e, no campo da classificação, importa registrar o resumo da obra de Brunet (1780-1867) e o esquema de classificação que ele desenvolveu para o *Manuel du libraire et de l'amateur de livres*¹⁰, repertório bibliográfico publicado de 1810 a 1860. O volume VI do Manual contém a estrutura detalhada de seu plano de classificação, em geral desenvolvido com três níveis de divisão, cujos saberes estão assim distribuídos: Teologia, Jurisprudência, Ciências e Artes, Literatura, História. Brunet 'abandona as reuniões arbitrárias de diferentes ciências em uma única classe, o que seduzira os eruditos e que, segundo nós, foram a principal dificuldade que levou ao fracasso de todos os outros sistemas bibliográficos'.¹¹ A classificação de Brunet foi adotada como modelo para o 'sistema' das bibliotecas confiscadas na França após a Revolução Francesa.

⁶SCHULTE-ALBERT, Hans. Gottfried Wilhelm Leibniz and library classification. *The Journal of Library History*, v. 6, n. 2, p. 133-152, 1971. p.149.

⁷Ibidem, p. 143.

⁸Vale a leitura deliciosa do capítulo Ordenadores do Universo de Alberto Manguel em seu livro *Uma História da Leitura*, São Paulo, Companhia das Letras, 1997.

⁹MALCLÈS, opus cit.

¹⁰BRUNET, J.-C. *Manuel du libraire et de l'amateur des livres*. Tomo VI. Table méthodique. Paris: Didot, 1865.

¹¹Ibidem, Tomo vi, Introduction p. vii.

Tais bibliotecas nomeadas públicas eram ainda de acesso fechado, e seu acervo bastante erudito para a massa da população ainda iletrada. O acesso ao acervo se dava através do catálogo classificado impresso de cada biblioteca, geralmente de periodicidade anual. De fato, como acentuam Létoublon e Sgard¹², o catálogo representava a biblioteca, daí a importância da classificação.



¹² LÉTOUBLON, F.; SGARD, J. *Le catalogue comme représentation de la bibliothèque*. Disponível em: <www.vox-poetica.org/sflgc/biblio/bibliofin/letoublongard.html>. Acesso em: 20 abril 2018.

CLASSIFICAÇÃO E BIBLIOTECAS

O final do século XIX e a primeira década do século XX viram o surgimento de vários esquemas de classificação bibliotecária em vários países, em especial nos Estados Unidos e no Reino Unido, resultado do acesso do público diretamente às estantes, ainda um reflexo da Revolução Francesa em seu objetivo de educação para todos.

A mais difundida nacional e internacionalmente é a Dewey Decimal Classification, desenvolvida por Melvil Dewey (1851-1931). Sua proposta de organização da ciência foi adaptada de um esquema de William T. Harris (1835-1909) eminente estudioso e educador norte-americano e um hegeliano bastante conhecido em seu tempo¹³, reconhecido como ‘uma das maiores mentes filosóficas dos tempos modernos’, além de educador de grande influência.¹⁴ Desde jovem se interessava pela classificação de livros, tendo desenvolvido uma espécie de esquema para sua biblioteca pessoal de 367 títulos. Quando superintendente de ensino para as bibliotecas públicas de St. Louis, observou que alguma coisa tinha que ser feita para organizar a quantidade crescente de livros que chegavam à biblioteca. Àquela época, a maioria dos livros era organizada alfabética ou cronologicamente.

À mesma época, Melvil Dewey estava muito empenhado no estudo de classificação para biblioteca. Em 1873 ele escreve para Harris solicitando um catálogo da Biblioteca Pública de St. Louis e outras matérias sobre o assunto. Em 1876 ele reconhece que o plano da biblioteca de St. Louis parecia com o seu, mas que ainda não tinha sido divulgado porque nem todos os aspectos essenciais haviam sido ainda decididos. Mas, ao comparar o esquema de Harris com o Decimal de Dewey, fica aparente a semelhança, inclusive ao manter a mesma ordem invertida das disciplinas. O grande sucesso e aceitação da Classificação Decimal foi a notação decimal adotada, em contraposição a propostas anteriores e contemporâneas. Mas isto foi também uma limitação, pois restringiu a divisão de qualquer área do saber em 10 tópicos, com subdivisões decimais também.

¹³ GRAZIANO, E.E. Hegel's Philosophy as basis for the Dewey Classification schedule. *Libri*, v. 9, n. 1, p. 45-52, 1959.

¹⁴ LEIDECKER, K. F. Debt of Melvil Dewey to William Torrey Harris. *Library Quarterly*, XV, p. 139-142, April, 1945.

A estrutura inicial da Classificação Decimal de Dewey divide o saber em 10 classes: Obras Gerais, Filosofia, Religião, Sociologia, Filologia, Ciências Naturais, Ciências Aplicadas, Belas Artes, Literatura, História e Geografia. A única possibilidade de expansão do esquema era a de acrescentar códigos para representar Formas bibliográficas e Lugar.

Outros esquemas de classificação desenvolvidos no período merecem citação: a Classificação da Library of Congress iniciada em novas bases no final do século XIX após o incêndio da biblioteca. Embora com raízes da Expansive Classification de Charles Ami Cutter (1837-1903), terminou por adotar o princípio da garantia literária: a Tabela se desenvolve a partir das características e relevância dos assuntos presentes na coleção.

A Classificação Expansiva de Cutter tem como característica sua possibilidade de se expandir, conforme seja adotada por pequenas coleções ou não. A morte de Cutter antes de terminar o desenvolvimento de seu esquema pode ter determinado seu abandono. Cutter tomou por base o esquema de Brunet e o aplicou da seguinte maneira: começou com a classe A – Obras gerais, onde colocou o tudo do Universo, isto é, os assuntos gerais demais para serem determinados como assuntos específicos. Depois, tomando consciência da existência, o Homem desenvolveu a Razão, coberta pela Filosofia, dando origem à classe B. Então, perguntou-se: de onde vim? E encontrou resposta na existência de Deus, surgindo, assim, as classes BR-C – Religião. Interessou-se pela sua própria vida como indivíduo e nasceu a classe E – Biografia. Depois, considerou sua existência como passando a pertencer a uma raça e a um lugar, manteve relações com os companheiros e o meio, criando, então as classes H – K – Ciências Sociais. Sentiu as forças que governam a natureza e imaginou as classes L – Q – Ciências. Orientou-se para as artes que lhe davam o sustento, surgindo as classes R – U – Tecnologia (na época chamada de Artes Úteis). Então desenvolvendo as artes originou a classe W – Belas Artes, atingindo finalmente a maior expressão de sua vida intelectual através da Literatura – Classes X-Z.¹⁵

¹⁵ BARBOSA, A.P. Teoria e prática dos sistemas de classificação bibliográfica. Rio de Janeiro: IBICT, 1969. p.104.

James Duff Brown (1862-1914), bibliotecário inglês, tomou por base do sistema a evolução das coisas no tempo e assim estruturou Subject Classification: Matéria e Força (Ciências físicas), Vida (Ciências Biológicas, Etnologia, Medicina, Biologia), Inteligência (Filosofia e Religião; Ciência política e Social), Registro (Lingua e literatura, Formas literárias, História e Geografia, Biografia). Criou tabelas auxiliares a serem usadas como subdivisão dos assuntos. Brown criou um esquema que pudesse representar assuntos interdisciplinares, ao contrário da Classificação Decimal de Dewey, que era rígida¹⁶. Cria tabelas 'categóricas, geográficas, cronológicas para formar assuntos. A Classificação de Brown aparece como uma tentativa de possibilitar maior flexibilidade até então não apresentada pelos outros sistemas da época.¹⁷

No início do século XX, Henry E. Bliss (1870-1955) produz Bibliographic classification, precedida de estudos sobre a classificação que resultaram em importantes textos sobre a teoria da classificação das ciências.¹⁸ Estabelece 11 (onze) princípios para a organização do conhecimento nas ciências.¹⁹ Faz análises críticas e teóricas dos principais esquemas bibliotecários vigentes, e em relação aos princípios para classificação nas bibliotecas acrescenta mais 28 (vinte e oito) princípios. Bliss organiza sua tabela em grandes grupos: Filosofia, Ciência, História, Tecnologia e Arte. Classificacionistas ingleses retomaram seu esquema, com a criação em 1967 da Bliss Classification Association (BCA), formado por um grupo de pesquisadores do Reino Unido com o objetivo de atualizar o esquema, tornando-o inteiramente facetado, conhecido então como BC2.²⁰ Bliss foi um estudioso da Classificação e buscou desenvolver teoria própria.

¹⁶ Atualmente, a CDD é considerada 'facetada' por seus responsáveis, por permitir combinação de códigos para formar mar assuntos, embora não o faça de maneira sistemática.

¹⁷ BARBOSA, opus cit.

¹⁸ BLISS, H.E. *The Organization of knowledge and the system of the sciences*. New York: Holt, 1929.

¹⁹ BLISS, H.E. *The Organization of knowledge in libraries and the subject approach to books*. New York: Wilson, 1933.

²⁰ BROUGHTON, V. A faceted classification as the basis of a faceted terminology: Conversion of a classified structure to thesaurus format in the Bliss Bibliographic Classification. *Axiomathes*, v. 18, n. 2, p. 193-210, 2008.

De modo geral, os esquemas de classificação bibliotecária incluíram tabelas adicionais para melhor representação dos assuntos como forma bibliográfica, tempo, espaço, local e, como visto anteriormente, foram influenciados por sistemas filosóficos.

Com Ranganathan a Classificação alcança seu status de Teoria, desenvolvida no âmbito da Biblioteconomia, ficando o resultado de seus estudos restritos à Índia durante algum tempo, mas estudada e discutida na Escola de Biblioteconomia em Londres.

Considerado o pai da Moderna Teoria da Classificação, o professor de Matemática Ranganathan recebe, no início do século XX a incumbência de estudar biblioteconomia em Londres, para implantar um sistema de bibliotecas na Índia. Classificação é um assunto bastante desenvolvido na Escola de Biblioteconomia, ao contrário do que ocorria à época nos Estados Unidos. Ranganathan estuda os esquemas de classificação vigentes e percebe que é preciso fazer ajustes às especificidades de seu País.

Percebe a dificuldade de classificar livros, pois estes frequentemente cobrem múltiplos tópicos. Observa que os esquemas permitem a representação de assuntos existentes, mas era preciso criar um esquema que possibilitasse a construção de representação de assuntos ainda por vir, pois o universo do conhecimento está em constante expansão. Sua formação matemática influenciou a formulação de sua Teoria procurando, inclusive, dar um status de ciência à Biblioteconomia formulando cânones e princípios, não apenas no âmbito da Classificação, mas em todas as outras áreas da atividade bibliotecária, como seleção, catalogação, referência. Ranganathan estudara os esquemas de classificação bibliotecária desenvolvidos à época, mas o desenvolvimento de sua Teoria deu-se durante os dez anos que trabalho 'forjando e polindo a Colon Classification', além de oito anos de atividade no ensino comparativo da Colon e da Classificação Decimal.²¹

Sua abordagem é diferente de todas as anteriores: ele não usava uma estrutura conceitual ou de assunto previamente estabelecidas, ou seja, não havia organização por disciplinas, com divisões, subdivisões e algo do tipo. Introduziu os conceitos de *faceta* e de *categorias* fundamentais.

²¹ RANGANATHAN, S. R. *Prolegomena to library classification*. Bombay: Asia publishing house, 1967, p. 22.

Segundo o método de faceta e o princípio de categorização qualquer entidade pode ser classificada por suas várias características (facetas) e, numa abordagem indutiva construir classes de mesma natureza.

Faceta é usado para denotar qualquer componente de um assunto (substância, língua, propriedade, gênero literário, religião...) que pertenceria a uma das cinco Categorias Fundamentais: Personalidade, Matéria, Energia, Espaço e Tempo (PMEST). Devido a seu alto nível de abstração, elas são de difícil compreensão. Vickery (1918-2009)²² as desdobra em: Coisas, substâncias, entidades que ocorrem naturalmente, produtos, instrumentos, constructos mentais. Suas partes constituintes, órgãos. Sistemas de coisas. Atributos de coisas: qualidades, propriedades, incluindo estrutura, medidas, processo, comportamento. Objeto da ação (paciente). Relações entre coisas, interações, efeitos, reações. Operações sobre coisas, experimentos, ensaios, operações mentais. Propriedades de atributos, relações e operações. Lugar, condição. Tempo.

Todas as iniciativas relacionadas a esquemas de classificação bibliotecária desenvolvidos anteriormente adotaram uma abordagem dedutiva, que vai perdurar até os dias de hoje seja para bibliotecas, para diferentes esquemas de classificação voltados para planejamento, gestão e pesquisa, e demais usos. Mas a introdução da Tecnologia da Informação e o surgimento de novos tipos de serviços e produtos de informação requerem novas abordagens na classificação e organização de recursos digitais. Embora sua Teoria continuasse a ser estudada no Reino Unido, foram necessários mais de 70 anos para que ela encontrasse seu verdadeiro meio para o sucesso na solução de conteúdo digital.²³ Isso significa que sua Teoria extrapolou o âmbito da Biblioteconomia.



²²VICKERY, B.C. *Faceted classification: a guide to construction and use of special schemes*. London: Aslib, 1960.

²³LAMBE, P. *Organising Knowledge: Taxonomies, Knowledge and Organisational Effectiveness*. Oxford UK: Chandos Publishing Ltd, 2007. p. 33-34.

CLASSIFICAÇÃO E INICIATIVAS INTERNACIONAIS: AINDA A BIBLIOGRAFIA

Artigo científico é o meio por excelência para comunicar resultados de pesquisa. Nesse contexto, surgem iniciativas de classificação com objetivo inteiramente bibliográfico ao final do século XIX, com Otlet (1868-1944) e La Fontaine (1854-1943), que propõem uma Bibliografia Universal corrente, de sorte a permitir reuso e disseminação dos resultados de pesquisa. Tal iniciativa resultou em contribuições relevantes para a atividade de informação. A primeira delas foi a proposta de uma nova área de estudos, a Documentação, que incluía coleta, tratamento, organização e disseminação de registros de qualquer natureza sobre qualquer suporte além do livro, objeto privilegiado até então. Não se pode esquecer que o final do século XIX marca a segunda revolução tecnológica, e tecnologias de comunicação estão presentes, e sempre que possível Otlet as incorporou às técnicas de reprodução e disseminação à literatura científica. Mas o grande esforço se concentrou ainda na bibliografia.

A ideia de uma Rede de informação universal levou Otlet a considerar as tecnologias de registro em filme e microfilme e a organização dos registros segundo um sistema de classificação internacional como padrão internacional. Ao analisar os textos de Otlet sobre classificação, Lévie²⁴ observa que para ele o sistema de classificação era uma espécie de imenso quadro do conhecimento. E ela transcreve um texto de Otlet encontrado nos escritos sobre classificação:

A Bibliografia é o primeiro meio de organizar as relações internacionais. Reunir as ideias dos homens, fazer uso por uns dos trabalhos dos outros. O Repertório bibliográfico universal é uma bolsa intelectual possível graças a um entreposto intelectual. A oferta e a demanda se encontram²⁵.

Para desenvolver sua proposta, foi criado o Instituto Internacional de Bibliografia em 1895, depois Instituto Internacional de Documentação e finalmente Federação Internacional de Documentação.

²⁴ LÉVIE, F. *L'Homme qui voulait classer le monde*: Paul Otlet et le Mundaneum. Bruxelles: Les Impressions Nouvelles, [s.d.] p. 63.

²⁵ OTLET apud LÉVIE, opus cit., p.63.

Com base na classificação Decimal de Melvil Dewey, Otlet introduziu várias tabelas adicionais para representar os vários aspectos tratados nos documentos, transformando assim o esquema de classificação em um modelo dinâmico, ao contrário dos esquemas desenvolvidos até então – seja para bibliografias, seja para catálogos-, como meio de classificar os registros em um Repertório Bibliográfico Universal. Assim, além do corpo principal do esquema, voltado para assuntos, Otlet criou tabelas diversas como forma bibliográfica, tempo, lugar, língua; e adotou para a representação destes aspectos diferentes códigos. Com uma ordem de citação para os diversos elementos da classificação, criou uma ordenação de precedência de cada tabela complementar, ou seja, uma ‘sintaxe’, algo também original.

Para poder representar assuntos especializados publicados em artigos científicos, Otlet buscou aprofundar seu esquema. Contou com a participação de cientistas diversos, como zoólogos e fisiologistas, de sorte a apoiar as bibliografias especializadas bem como a tradução e a classificação das ciências médicas e as tabelas abreviadas de anatomia e de fisiologia que comporiam a *Bibliographia Universalis*. Inicialmente conhecida como Classificação de Bruxelas, veio a ser disseminada como Classificação Decimal Universal, padrão internacional para organização dos registros. A Segunda Guerra Mundial interrompeu o processo.

A ideia de um modelo internacional de informação científica perdurou e ressurgiu em outro momento e em outras bases. Assim, o problema continuou sendo debatido por cientistas que organizam após a Segunda Guerra Mundial a primeira Conferência Internacional sobre informação científica em Londres, 1948²⁶, The Royal Society à frente, movidos por grandes eventos científicos e políticos.

O tema central da conferência foi a necessidade de reformar o sistema de informação científica. Os organizadores da conferência afirmaram que “[...] a tarefa de acompanhar a literatura científica está se tornando

²⁶VICKERY, B.C. The Royal Society Scientific Information Conference of 1948. *Journal of Documentation*, v. 54, n.3, p. 281-283, 1998.

impossível [...], levando a alguma perda absoluta de conhecimento, causada por atrasos, à inconveniência e a custos crescentes. Alguma reforma substancial era necessária”²⁷.

Entre os temas debatidos estavam as publicações primárias, ou seja, os periódicos, que são meio por excelência para o registro e a disseminação do conhecimento científico; os serviços e resumos (abstracting services); as questões relativas à tradução, classificação e indexação; a introdução do tema indexação mecânica. Durante a conferência, uma incipiente tecnologia eletrônica para o processamento de informação – o computador UNIVAC da IBM – foi apresentada. Para facilitar o acesso aos artigos, Bernal, o conhecido Cientista Vermelho, propõe a criação de uma central de depósito dos artigos, mas o que resultou foi uma biblioteca internacional de empréstimo (The National Lending Library). Dez anos depois, realiza-se em Washington a segunda conferência²⁸, que guarda algum paralelo com a primeira, ao abordar temas como necessidades dos cientistas de literatura e referência, avaliação dos periódicos de revisão e indexação, organização da informação para armazenagem e recuperação, estrutura de sistemas de informação, os dois últimos itens fortemente computacionais. Aqui prevalecem pesquisadores e demais profissionais de informação e computação.

Um relevante acontecimento foi a criação do Classification Research Group (CRG), no Reino Unido, em 1952. Dentre seus criadores estavam alguns pesquisadores participantes da Conferência de 1948, e as atividades de seus membros deram origem ao surgimento de inúmeros teóricos da classificação, trazendo as contribuições de Ranganathan com sua classificação analítico-sintética em seus estudos e pesquisas. Vale citar a publicação do *Thesaurfacet*²⁹, que tornou evidente a base classificatória nos tesouros e, atualmente, nas taxonomias.

²⁷ Ibidem, p. 282.

²⁸ INTERNATIONAL CONFERENCE ON SCIENTIFIC INFORMATION (1958). *Proceedings*. Washington, The National Academies press, 1959. 2 v.

²⁹ ATICHISON, J. The Thesaur facet: a multipurpose retrieval language tool. *Journal of Documentation*, v. 26, n. 3, p. 187-203, 1970.

Na década de 60 do século XX, a UNESCO cria um grupo para desenvolver um programa voltado para estudar a viabilidade de um sistema mundial de informação para a ciência (UNISIST), o que resultou em um relatório e resumo amplamente divulgados.³⁰ UNESCO e ICSU (International Council of Scientific Unions) se reuniram para propor um sistema de informação mundial de ciências. Não se buscava implementação, mas fornecimento de diretrizes, princípios e recomendações aos países-membros e à comunidade científica.

Informação científica naquele estudo tem sentido amplo, incluindo ciência básica, aplicada, engenharia e tecnologia, e incorpora a herança do conhecimento científico.

Cientistas que são seus construtores e usuários, apenas pedem que as contribuições de uns e outros sejam verificáveis; [o artigo] é um meio através do qual os cientistas no mundo mantêm sua disciplina. É um meio de educação de futuros cientistas, e o principal reservatório de conceitos e dados onde recorrem para aplicação em programas econômicos e tecnológicos.³¹

Fazia-se apelo aos países para se comprometerem tanto com os recursos intelectuais como materiais, necessários à condução da ciência. Questões ligadas a políticas de ciências, financiamento, equipamentos e formação de mão de obra já eram temas familiares; mas os recursos informacionais tinham recebido até então atenção menor, e este foi o tema central do estudo: mobilização de recursos – no sentido mais amplo – de informação da ciência. Grandes massas de dados, organizados de forma descentralizada e independente requeriam soluções para sua integração, harmonização, compatibilização ou qualquer outro meio de compartilhamento.

Em 1963, estudiosos do Groupe d'Études sur l'Information Scientifique (GEIS), com sede em Marselha, desenvolvem Intermediate Lexikon,³² um sistema para os centros que lidavam com o mesmo assunto, e mais tarde desenvolvida na área de Ciência da Informação, na School of Librarianship em Londres.

³⁰ UNESCO. *Unisist Synopsis of the feasibility study on a World Science Information System*. Paris: Unesco, 1971.

³¹ Idem, p. 4.

³² COATES, E.; LLOYD, G.; ŠIMANDL, D. *The BSO manual: the development, rationale and use of the Broad System of Ordering*. Hague: Fédération Internationale de Documentation, 1979.

O esquema GEIS constituía em uma tabela de conversão, na qual o código de um dado conceito em uma linguagem de indexação era acoplado ao código do mesmo conceito noutra linguagem de indexação. Um sistema com código intermediário era introduzido para levar de uma linguagem a outra. No caso de envolver três linguagens de indexação, então uma linguagem de mediação ou comunicação era estabelecida. Esta 'linguagem de comutação' requeria dos usuários que utilizassem a linguagem de indexação do seu centro (A) para a linguagem de comutação que, através de uma tabela de equivalência com o Centro B atingiria o destino. Neste estudo a combinação seria feita de um termo a outro, via tabela de equivalência, mas uma busca hierárquica não seria possível e alguma orientação seria obrigatoriamente necessária. Um código estruturado, no entanto, possibilitaria esta comutação. A característica deste modelo é a preservação de cada centro de informação com sua linguagem/ seu sistema de classificação.

Em 1979, como reflexo dos estudos realizados na elaboração do Programa Unisist, da UNESCO, a FID publica o *Manual BSO – Broad System of Ordering*³³ (BSO). BSO foi desenvolvido para uma rede mundial de informação, cobrindo todos os campos de conhecimento. Coates, renomado classificacionista, participou de sua elaboração sem adotar princípios de sistemas de classificação bibliotecária, como necessidades de informação ou garantia literária, embora tal esquema guarde semelhança com tais classificações. A última versão identificada, de 1978, parte de Disciplinas que contemplam metodologias e pontos de vista especiais focalizando um conjunto de entidades ou fenômenos, incluindo mecanismos que permitem que entidades possam ser classificadas em qualquer aspecto estudado. A ideia da rede, no UNISIST, requeria padrões comuns e tecnologia compatível, e é nesta perspectiva que se insere o padrão BSO, como uma linguagem de compatibilidade entre os diversos elementos da rede, podendo também funcionar como uma linguagem de comutação.

Em 2002 surge euroCRIS, associação internacional sem fins lucrativos que reúne especialistas em informação de pesquisa em geral e sistemas de informação de pesquisa em particular. Nesta perspectiva, o estudo da UNESCO/ICSU poderia oferecer elementos para a euroCRIS?

³³ Ibidem.

As questões relacionadas à divulgação e acesso às informações de pesquisa ainda são atuais, apesar dos desenvolvimentos da TI e das contribuições da Ciência da Informação, visto que o problema não é apenas tecnológico, mas envolve interesses e necessidades de vários países em seus esforços de integração e/ou harmonização de sistemas de informação. Uma peculiaridade da euroCRIS é o foco nas informações sobre projetos de pesquisa em curso. Em seu âmbito de atuação, então, incluem-se organização e divulgação de projetos de pesquisa, seus produtos, pesquisadores, organizações envolvidas (acadêmicas, financiadoras).

Em novembro de 2016, Science Europe divulga sua posição a respeito de Sistemas de Informação de Pesquisas,³⁴ no contexto da Ciência Aberta, voltada para a divulgação de um conhecimento que, por sua natureza, é público.³⁵ Tal documento enumera dificuldades para o desenvolvimento de um sistema de informação para pesquisa, como as diversas bases de dados de organização de pesquisa que diferem muito, em especial quanto a definições, sistemas de classificação, formatos e propósitos. E estes aspectos interferem nos sistemas/linguagens de indexação adotados como estrutura, abrangência, nível de abrangência dos termos. Interoperabilidade é uma questão recorrente.

Vários encontros e treinamentos da euroCRIS têm enfatizado o padrão CERIF, que é dedicado a metadados, e é neste contexto que a interoperabilidade tem sido discutida. O padrão CERIF é fundamental para a compatibilidade no nível computacional. É preciso, entretanto, abordar o problema da classificação na representação do conteúdo dos projetos e relatórios de pesquisa corrente para que ocorra adequado uso do padrão CERIF. Este tem sido um aspecto pouco abordado, fato também registrado quando da proposta do UNISIST. Devido à descentralização na coleta de dados e à participação de inúmeras organizações com suas características como anteriormente indicadas, estudos de interoperabilidade entre vocabulários ou sistemas de classificação são requeridos, quando se almeja a participação de várias organizações em uma rede de informação.

³⁴ SCIENCE EUROPE. SCIENCE EUROPE WORKING GROUP ON RESEARCH POLICY AND PROGRAMME EVALUATION. Position Statement on Research Information Systems: D/2016/13.324/11. Nov. 2016. Disponível em: <https://www.scienceeurope.org/wp-content/uploads/2016/11/SE_PositionStatement_RIS_WEB.pdf>. Acesso em: 20 maio 2018.

³⁵ SOLLÀ PRICE, D. *Science Since Babylon*. New Haven: Yale University Press, 1961.

No entanto, não fica claro de que tipo de classificação se trata: para recuperação de informação ou para indicadores estatísticos. No primeiro caso, o nível de detalhamento dos conceitos é alto, no segundo, ao contrário, o que se requer é uma tabela de alto nível de agregação dos conceitos.

De qualquer modo, quando se tem um modelo de participação com entrada de dados descentralizada, em algum momento surge a necessidade de criar algum instrumento de harmonização, mediação, comutação ou o que quer que venha a ser decidido. A compatibilidade é um problema

[...] altamente complexo multi-facetado, pois abrange desde capacidades de equipamentos, linguagens utilizadas, seleção e ordenação dos elementos dos registros bibliográficos e variedade de ferramentas conceituais como tesouros ou vocabulários científicos controlados e sistemas de classificação.³⁶

No ambiente de euroCRIS a compatibilidade entre os vários sistemas de classificação adotados por seus membros torna-se uma questão crucial, qualquer que seja o tipo de sistema de classificação: tesouros, taxonomias, esquemas de classificação. O tema é abordado em dois estudos promovidos pelo IBICT^{37 38}.



³⁶ UNISIST, opus cit., p. 31.

³⁷ GOMES, H.E. *Estudo CASRAI: Compatibilidade de sistemas*. IBICT - Iniciativa de Apoio aos Diálogos Setoriais União Européia/Brasil. Brasília 2017.

³⁸ GOMES, H.E. *Estudo CASRAI: Estudo com tabelas CNPq, FoS e ISCED*. IBICT - Iniciativa de Apoio aos Diálogos Setoriais União Europeia-/Brasil. Brasília, 2017.

CLASSIFICAÇÃO E GESTÃO

Iniciativas anteriores, como a de Otlet ou da UNISIST, visavam disseminação e acesso a resultados de pesquisa, via esquema de classificação. As iniciativas estão consolidadas através de inúmeras bases de dados especializadas.

A segunda metade do século XX traz iniciativas mais complexas, ou seja, de planejamento e gestão de pesquisa, as quais requerem classificações adequadas à produção de indicadores estatísticos. Classificações com finalidades estatísticas têm natureza agregadora, ou seja, têm nível de generalidade maior do que os esquemas bibliotecários. Em contrapartida, quando voltados para planejamento de ciência, baseiam-se igualmente em disciplinas ou grandes áreas do saber.³⁹

Alguns esquemas são analisados a seguir, aí incluí dos sistemas regionais como a euroCRIS, Australia-New Zealand Standard Research Classification, internacionais, FoS e ISCED, e até mesmo de um dado país, como o Brasil, todos eles no âmbito de Pesquisa & Desenvolvimento. Diversos modelos são apresentados e sua descrição evidencia a organização e a estrutura classificatória adotadas, seus princípios lógicos ou não.

CASRAI E EUROCRIS

Um dos grandes problemas de sistemas como euroCRIS, se decidir por uma classificação padrão, é a compatibilização dos diversos esquemas de classificação adotados por seus membros para obter dados de pesquisa de todos os países-membros. A situação é complexa, conforme estudo desenvolvido em 2014⁴⁰: de modo geral, cada país adota o próprio esquema de classificação, alguns baseados em disciplinas, outros em áreas do conhecimento, outros incluindo aspectos econômicos; alguns países apresentam apenas sites de busca, e assim não ficam evidentes as ferramentas adotadas (esquemas de classificação, tesouros ou outras). A adoção de um padrão – regional ou internacional – eliminaria a necessidade de estudos de compatibilização, harmonização ou outro.

³⁹Ibidem. (estudo mais detalhado).

⁴⁰PEREIRA, M. de N.F. Documento técnico: levantamento e análises das ferramentas de indexação (produto 03). Brasília: 2015. Projeto 914BRA2015 – IBICT. Edital N° 020/2014. Entrevista com a Professora Hagar Espanha Gomes.

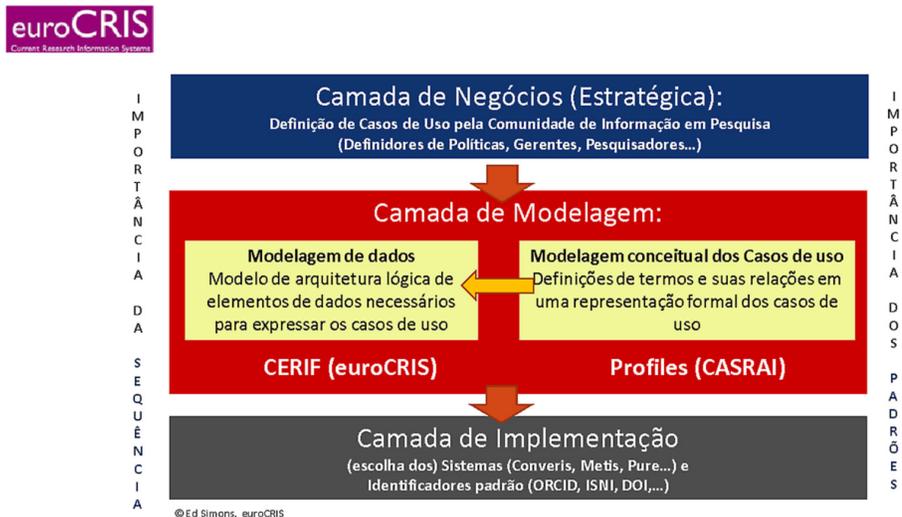
Mas iria requerer o desenvolvimento de instrumentos intermediários, de alto custo, acrescido da necessidade de constante atualização relacionada ao avanço do conhecimento, como também à inclusão de novo país-membro no sistema de informação.

Chama atenção o sistema de informação da Eslovênia, que adota vários esquemas de classificação: Fields of Research Classification (ARRS-FR-Classification.pdf), Common European Research Classification Scheme (Cerif.pdf), Fields of Research and Development (FORD.pdf)⁴¹, Statistical Classification of Socio-Economic Objectives (NABS.pdf), Nomenclature of Research and Impacts (ARRS-RR-Impacts.pdf). O último não tem ponto de contacto com os demais: trata-se de uma classificação que inclui publicação, conferências científicas, corpo editorial, chairing, prêmios. Os demais esquemas às vezes estão estruturados por áreas de conhecimento, por vezes por problemas e outros princípios. Se a complexidade está nas atividades de um único país, como harmonizar os esquemas de classificação dos demais países? Como euroCRIS pode harmonizar tais modelos?

De acordo com o modelo de implementação da euroCRIS (Figura 1) é na camada de modelagem que se encontram as atividades aqui discutidas. O padrão CERIF parece receber aceitação internacional, o mesmo não ocorrendo, ainda, no que se refere à modelagem conceitual.

⁴¹ Corresponde à Fields of Science (FoS) da OCDE (Frascati Manual).

Figura 1 – Implementação da informação de pesquisa: modelo em três camadas



Fonte: Traduzido de: SIMONS, Ed. *Research Information Implementation: 3-layer model*. Jornada euroCRIS, euroCRIS/IBICT, Brasília, 16 a 18 de novembro e 2016.

Em 2006, por iniciativa de pesquisadores, CASRAI (Consortia Advancing Standards in Research Administration Information) surge no Canadá como organização internacional sem fins lucrativos, visando adaptar princípios e práticas de padrões abertos que sirvam de ligação entre usuários de informação de pesquisa. Os pesquisadores desenvolvem uma tabela de classificação para pesquisa universitária com base no estudo e análise de 32 classificações análogas de pesquisa existentes dentro e fora do Canadá, cujo resultado foi validado e aprovado por especialistas. Seu diferencial é a possibilidade de representar pesquisas correntes por vários aspectos ou dimensões, o que a tornaria dinâmica. Três são as dimensões da Tabela: Disciplinas, Objeto de pesquisa e Áreas de aplicação. As áreas cobertas pela Tabela são: Ciências da Saúde; Ciências Naturais e Engenharias; Ciências Humanas e Sociais; Artes e Literatura.⁴² Cada área contém, em sua grande maioria, dois níveis de detalhamento.

⁴² GAUDREAU, J.; BAKER, D. *Research classification: a standard, more coherent and more complete classification of research*. (CASRAI Connecting Research). Disponível em: <<http://cssip.org/docs/>>

O primeiro e o segundo níveis são bastante genéricos. O terceiro nível é bastante detalhado, mas sem estrutura classificatória, ou seja, não agrupa os elementos deste terceiro nível segundo aspectos comuns: os elementos são ordenados alfabeticamente, o que não é recomendado para um sistema voltado para indicadores estatísticos.

A documentação da tabela CASRAI é escassa, dificultando análise profunda. Em linhas gerais, o que se observa é a ausência de uma notação que revele a estrutura dos níveis. A tabela CASRAI encontra-se em estudos de revisão.

FOS E OECD

Padrão internacional é FoS - Fields of Science -, iniciativa da OCDE que se insere nas diretrizes para coleta e divulgação de dados sobre recursos de P&D entre os membros da OCDE, conforme explicitado no *Manual Frascati*⁴³. A classificação está organizada segundo os principais campos de ciência e tecnologia recomendados no Manual, e deve ser usado para gastos de P&D do governo, do ensino superior, setor privado e não privado e, se possível, pelo setor de empresas comerciais. É nessa perspectiva que a estrutura de FoS deve ser entendida. A Tabela está organizada em dois níveis que contemplam áreas de P&D: Matemática, Engenharia e Tecnologia, Ciências médicas e da Saúde, Ciências agrárias e Humanidades. O segundo nível ainda é geral, dividido em vários grupos, no total de 36.

A última versão revista de FoS, no entanto, data de 2007 e é resultado de ampla discussão iniciada em 2002.

Devido a diferentes perspectivas dos sistemas administrativos da comunidade científica, bem como de usuários da classificação e à dinâmica própria da ciência (como a emergência de ciências interdisciplinares) não foi possível desenvolver uma classificação que satisfaça a todos os atores envolvidos.⁴⁴

CASRAI_New_classification_deck_-_pdf.pdf>. Acesso em: 20 março 2018.

⁴³ OECD. *Frascati manual 2015: guidelines for collecting and reporting data on resources and experimental developments, the measurement of scientific technological and innovation activities*. Paris: OECD Publishing, 2015. Disponível em: <DOI <http://dx.doi.org/101787/978926439012-en>>. Acesso em: 12 março 2018.

⁴⁴ OECD. Committee for Scientific and Technological Policy. Working Party of National Experts in Science and Technology Indicators: revised Field of Science and Technology (FOS) classification in the Frascati Manual. DSTI/EAS/STP/NESTI (2006 19 /FINAL) 2007.

Na revisão foram considerados entre outros “[...] o propósito da classificação, os possíveis diferentes atores usando a classificação, o relacionamento entre governo e instituições, a carga administrativa da instituição interessada”.⁴⁵

Uma observação relevante sobre FoS, a qual não se pode perder de vista, é que a classificação não pretende harmonizar dados classificados por campos de Ciência e Tecnologia para fins nacionais, mas visa um nível mínimo de comparabilidade de dados de P&D no nível internacional.⁴⁶ Isso explica seu alto grau de generalidade e aceitação.

Devido a esse nível de generalidade das subdivisões, a Tabela inclui detalhes da abrangência de conteúdo de cada uma, complementando assim a ausência de definições. Para não haver dúvida sobre o classificador, a Tabela inclui, em cada grupo, o que ali não se contempla, remetendo o usuário ao grupo adequado.

ISCED E UNESCO

Outra Classificação voltada para indicadores científicos é a classificação da UNESCO, que visa estatísticas de ensino de Ciência e Tecnologia: ISCED (International Standard Classification of Education), instrumento oficial do Instituto de Estatísticas da UNESCO (UIS)⁴⁷. O propósito de ISCED-F é o de compilar dados nacionais dos Estados membros da UNESCO e, conforme seus organizadores, a Tabela pode ser usada em contextos nacionais, ‘especialmente nos países que ainda não desenvolveram os próprios padrões nacionais de classificação para educação.

A última versão ISCED-F 2013 está organizada em 10 grandes áreas: 00 Programas e Qualificações gerais, 01 Ensino, 02 Artes e Humanidades, 03 Ciências sociais, Jornalismo e informação, 04 Negócios, Administração e Direito, 05 Ciências naturais, Matemática e Estatística, 06 Tecnologia da Informação e da Comunicação, 07 Engenharia, Manufatura e Construção,

⁴⁵ Ibidem, p. 99.

⁴⁶ Ibidem, p. 33-34.

⁴⁷ UNESCO. Institute for Statistics. ISCED Fields of Education and Training 2013 (ISCED-F 2013). Manual to accomplish the International Standard Classification of Education, 2011. Québec UNESCO-UIS, 2014. Disponível em: DOI <<http://dx.doi.org/10.15220/978-92-9189-150-4-en>>. Acesso em: 25 março 2018.

08 Agricultura, Silvicultura, Pesca e Veterinária, 09 Saúde e Bem-estar, 10 Serviços. Como se observa, foge ao padrão das áreas de conhecimento, focalizando áreas de ensino.

As áreas seguem uma abordagem temática e os níveis de classificação se baseiam na semelhança do assunto. Assim é possível classificar 'com base no conteúdo do programa e não nas características dos participantes pretendidos'.

Trata-se de uma tabela enxuta: nas 11 (onze) áreas amplas reúne 29 campos 'estritos' e cerca de 80 (oitenta) campos detalhados. Não inclui 'outros', mas fornece orientação quando o assunto não está aparente na Classificação, fornecendo, então, elementos para futuras revisões. A área com maior número de campos detalhados é Engenharia, Manufatura e Construção, que tem 12 campos detalhados distribuídos em 3 campos estreitos: Engenharia e Profissões – 6 campos detalhados; Manufatura e processos – 4 campos detalhados; Arquitetura e construção – 2 campos detalhados. Uma das áreas com menor número de detalhes é Línguas, com apenas 2 campos detalhados: Aquisição de Linguagem e Literatura e linguística.

As áreas de ensino estão organizadas em três níveis: 1º. Nível – Área ampla, 2º. Nível – Área estreita, 3º. Nível – Área detalhada. No último nível incluem-se principalmente o terceiro nível de ensino, o ensino profissional e programas de treinamento e qualificação nos níveis secundário e pós-secundários.

A estrutura classificatória contempla então dois aspectos - áreas e níveis de ensino - os quais, combinados, permitem obter resultados específicos.

AUSTRÁLIA-NOVA ZELÂNDIA

Australia-New Zealand Standard Research Classification (ANZSRC) 2008 é uma interessante iniciativa regional desenvolvida a partir dos setores de Estatística da Austrália e da Nova Zelândia, em cooperação com as principais organizações acadêmicas e de pesquisa e de usuários de informação de pesquisa daqueles países. O esquema é bem documentado, com informação sobre os princípios de sua construção, informação sobre versão anterior e metodologia empregada.

A classificação é constituída de três partes: Tipo de atividade, Campos de pesquisa e objetivo socioeconômico. A primeira é independente. Esse tipo de classificação permite que a atividade de P&D seja categorizada de acordo com o tipo de pesquisa: pesquisa básica pura, pesquisa básica estratégica, pesquisa aplicada e desenvolvimento experimental.

A classificação de campos de pesquisa “[...] considera a metodologia empregada em P&D [...]” e a classificação para objetivo socioeconômico foi desenvolvida de acordo com o “[...] propósito ou resultado de P&D conforme a percepção do fornecedor dos dados (o pesquisador)”.⁴⁸

A Seção relativa a campos de pesquisa tem três níveis hierárquicos, a saber, Divisões, Grupos e Campos. A Seção relativa a objetivo sócio econômico tem quatro níveis hierárquicos: Setor, Divisões, Grupos e Objetivos. A parte relativa a Campos de pesquisa esá distribuída em 22 (vinte e duas) divisões e a Seção de Objetivo socioeconômico está estruturada em 17 (dezessete) divisões.

Cada Grupo remete a remete ao respectivo campo fornecendo orientação para itens não incluídos, remetendo-os aos respectivos códigos. A Tabela é detalhada, mas guarda considerável nível de agregação para os objetivos dos países envolvidos.

ÁREAS DO CONHECIMENTO CNPQ/CAPES

A tabela Áreas do Conhecimento é utilizada tanto pelo CNPq como pela CAPES, de acordo com os propósitos de cada instituição. De modo geral, é utilizada pelas Fundações de Amparo à Pesquisa dos diversos estados da Federação. Não se tem conhecimento de grandes análises ou indicadores em nível nacional, mas a Tabela parece ser empregada para avaliar e conceder bolsas de pesquisa ou para aprovar projetos, conforme a natureza de cada instituição federal. Não há documentação disponível para melhor avaliação da classificação, seus propósitos, princípio adotado na estruturação, orientação, abrangência de conteúdo.

⁴⁸ Australian and New Zealand Standard Research Classification (ANZSRC). Disponível em: <<http://www.abs.gov.au/ausstats/abs@.nsf/0/4AE1B46AE2048A28CA25741800044242?opendocument>>. Acesso em: 28 dez.2017.

A tabela do CNPq cobre as seguintes Grandes Áreas do Conhecimento: A Exatas e da Terra, B Ciências Biológicas, C Engenharias, D Ciências da Saúde, E Ciências Agrárias, F Ciências Sociais Aplicadas, G Ciências Humanas. Percebe-se aí uma certa organização: a ordenação dos itens nas Grandes Áreas mantém certa tradição: primeiro as exatas e da vida, em seguida áreas de aplicação das exatas, da vida (do homem e do campo, da vida na sociedade; por último as Humanas. Nas Áreas, que são o segundo nível das Grandes Áreas, nem sempre se percebe qual o princípio de ordenação.

Não existe uma notação estruturada que refletisse agrupamentos no segundo e terceiro níveis, o que dificultaria a produção de indicadores. A Tabela apresenta classificações inadequadas, por exemplo, nomear Ciências Sociais Aplicadas como uma das grandes áreas e, assim, classificar Sociologia em Ciências Humanas; ou, classificar Design – que é arte industrial, atividade criativa – em Ciências Sociais Aplicadas.

Existem iniciativas para participação do Brasil na euroCRIS e, se isso frutificar, a Tabela deve ser reestruturada, deve incluir notação expressiva para refletir a estrutura dos níveis Áreas e Especialidades. E isso poderá ser útil, também, se a Tabela for utilizada no planejamento das atividades de pesquisa no país.



CLASSIFICAÇÃO NA ERA DIGITAL

A Teoria da Classificação se democratizou: deixa de ser preocupação exclusiva de pesquisadores: chegou recentemente no mundo da gestão do conhecimento como taxonomia em novos produtos e serviços de informação em meio digital.⁴⁹ Em seu livro *Information Architecture for the World Wide Web*⁵⁰, seus autores, bibliotecários, apresentam as bases para essa nova forma de atuação profissional e destacam o papel da classificação em ambiente digital, em especial a classificação facetada, ambiente no qual as taxonomias são as bases de tal organização.

A classificação analítico-sintética de Ranganathan com seu Método de Faceta se popularizou através dos pesquisadores do Classification Research Group do Reino Unido, em especial com a produção do *Thesaurfacet*⁵¹. Com este tesouro, produzido como índice de um esquema de classificação especializada, ficou evidente que o Método de Faceta produz taxonomias rigorosas.

Taxonomias são importantes para encontrabilidade de conteúdo: a possibilidade de inclusão de um mesmo conceito, em tantas classes de conceitos quantos sejam suas características, aumenta o potencial de encontro de uma informação a partir dos vários aspectos desejados por um usuário. Cada faceta se restringe apenas um aspecto do conteúdo; desse modo cada faceta e sua taxonomia podem ser incorporados a um registro de metadados, pois estes guardam com os dados estreita relação, ou seja, referem-se à mesma faceta de base. “Trabalhar com esquemas de classificação facetada implica simplesmente trabalhar com um grupo de taxonomia de base na qual cada documento ou peça de documento é analisado de acordo com uma ou mais taxonomias de base no mesmo tempo”⁵² Taxonomias facetadas evitam ambiguidade, uma vez que cada faceta é diferente das demais.

⁴⁹ LAMBE, opus cit. p. 99.

⁵⁰ Ibidem, p. 33-34.

⁵¹ AITCHISON, opus. cit.

⁵² LAMBE, opus cit. p. 37.

Adotar o Método de Faceta na revisão de esquemas de classificação, quaisquer que sejam os tipos ou propósitos de tais esquemas, produz resultados coerentes: agindo de modo indutivo chega-se a classes de mesma natureza (por exemplo, coisa, processo, substância, instrumento ou o que for) e não a temas, que são de conteúdo múltiplo e formam conjuntos não mutuamente exclusivos. Mesmo que outros recortes de uma área do conhecimento possam ser adotados, como recortar por função, ainda temos o método indutivo, como base para ordenar os elementos no interior de cada função.

Em síntese, a classificação segundo o método de faceta atua de modo indutivo chegando à taxonomia de base, permite que um mesmo elemento seja encontrado por diferentes características presentes nas respectivas taxonomias de base, sem ambiguidade, e cada taxonomia de base encontra seu registro no metadado adequado. O método é dinâmico, pois novas características podem ser incorporadas criando-se novas taxonomias de base de modo a atender a dinâmica do conhecimento. E está na base da organização de serviços e produtos na era digital.

O detalhamento obtido nas taxonomias de base não é empecilho para a construção de esquemas com alto grau de abrangência, como requer um esquema para produção de indicadores, pelo contrário. Ao selecionar as classes gerais, os respectivos itens específicos da taxonomia de base são úteis na descrição de seu conteúdo com informação do tipo “Aqui se incluem...”. A partir do método indutivo produzem-se, desse modo, esquemas consistentes e orientação segura para os usuários.



CLASSIFICAÇÃO SEMPRE

Organização e gestão da informação e do conhecimento são conceitos próximos. Em qualquer das opções anteriores, a classificação está no cerne, presente em manifestações tão díspares como relatadas. Uma evidência forte é a transformação do periódico científico *International Classification*, que passa a se chamar *Knowledge Organization*. A literatura comprova isto e nos tempos atuais tem consagrado o Método de Faceta.

Teoria da Classificação não é prerrogativa de profissionais da informação: outros profissionais são convidados a estudá-la.



REFERÊNCIAS

- ATICHISON, J. The Thesauro facet: a multipurpose retrieval language tool. *Journal of Documentation*, v. 26, n. 3, p. 187-203, 1970.
- Australian and New Zealand Standard Research Classification (ANZSRC). Disponível em: <<http://www.abs.gov.au/ausstats/abs@.nsf/0/4AE1B46AE2048A28CA25741800044242?opendocument>>. Acesso em: 28 dez.2017.
- BARBOSA, A.P. *Teoria e prática dos sistemas de classificação bibliográfica*. Rio de Janeiro: IBICT, 1969.
- BLISS, H.E. *The Organization of knowledge and the system of the sciences*. New York: Holt, 1929.
- BLISS, H.E. *The Organization of knowledge in libraries and the subject approach to books*. New York: Wilson, 1933.
- BROUGHTON, V. A faceted classification as the basis of a faceted terminology: Conversion of a classified structure to thesaurus format in the Bliss Bibliographic Classification. *Axiomathes*, v. 18, n. 2, p. 193-210, 2008.
- BRUNET, J.-C. *Manuel du libraire et de l'amateur des livres*. Tomo Vi. Table méthodique. Paris: Didot, 1865.
- COATES, E.; LLOYD, Ge.; ŠIMANDL, D. *The BSO manual: the development, rationale and use of the Broad System of Ordering*. Hague: Fédération Internationale de Documentation, 1979.
- DAHLBERG, I. *Teoria da Classificação ontem e hoje*. Tradução do inglês por Henry B. Cox. Palestra apresentada à Conferência Brasileira de Classificação Bibliográfica, Rio de Janeiro, 12-17 de setembro de 1972. Anais. Brasília, IBICT/ABDF, 1979. v. 1, p. 352-370. Disponível também em <www.conexaorio.com/bit/>.
- GAUDREAU, J.; BAKER, D. *Research classification: a standard, more coherent and more complete classification of research*. (CASRAI Connecting Research). Disponível em: <http://cssip.org/docs/CASRAI_New_classification_deck_-_pdf.pdf>. Acesso em: 28 dez. 2017.
- GOMES, H.E. *Estudo CASRAI – Compatibilidade de sistemas*. IBICT – Iniciativa de Apoio aos Diálogos Setoriais União Européia/Brasil. Brasília 2017.
- GOMES, H.E. *Estudo CASRAI – Estudo com tabelas CNPq, FoS e ISCED*. IBICT - Iniciativa de Apoio aos Diálogos Setoriais União Européia-/Brasil. Brasília, 2017.
- GRAZIANO, E.E. Hegel's Philosophy as basis for the Dewey Classification schedule. *Libri*, v. 9, n. 1, p. 45-52, 1959.
- PEREIRA, M. de N.F. *Documento técnico: levantamento e análises das ferramentas de indexação (produto 03)*. Brasília: 2015. Projeto 914BRA2015 – IBICT. Edital N° 020/2014. Entrevista com a Professora Hagar Espanha Gomes.
- INTERNATIONAL CONFERENCE ON SCIENTIFIC INFORMATION (1958). *Proceedings*. Washington, The National Academies press, 1959. 2 v.
- LAMBE, P. *Organising Knowledge: Taxonomies, Knowledge and Organisational Effectiveness*. Oxford UK: Chandos Publishing Ltd, 2007.

LEIDECKER, K. F. Debt of Melvil Dewey to William Torrey Harris. *Library Quarterly*, XV, p. 139-142, April, 1945.

LÉTOUBLON, F.; SGARD, J. *Le catalogue comme representation de la bibliothèque*. Disponível em: <www.vox-poetica.org/sflgc/biblio/bibliofin/letoublonsgard.html>. Acesso em: 20 abril 2018.

LÉVIE, F. *L'Homme qui voulait classer le monde: Paul Otlet et le Mundaneum*. Bruxelles: Les Impressions Nouvelles, [s.d.].

MALCLÈS, L.-N. *La Bibliographie*. Paris, Presses Universitaires de France, 1956.

OECD. *Committee for Scientific and Technological Policy. Working Party of National Experts in Science and Technology Indicators: revised Field of Science and Technology (FOS) classification in the Frascati Manual*. DSTI/EAS/STP/NESTI (2006 19 /FINAL) 2007.

OECD. *Frascati manual 2015: guidelines for collecting and reporting data on resources and experimental developments, the measurement of scientific technological and innovation activities*. Paris: OECD Publishing, 2015. Disponível em: <dx.doi.org/10.1787/978926439012-en>. Acesso em: 12 março 2018.

RANGANATHAN, S.R. *Prolegomena to library classification*. Bombay: Asia publishing house, 1967.

RICHARDSON, E.C. *Classification: theoretical and practical*. New York: Scribner, 1912.

SCHULTE-ALBERT, H. Gottfried Wilhelm Leibniz and library classification. *The Journal of Library History*, v. 6, n. 2, p. 133-152, 1971.

SCIENCE EUROPE. SCIENCE EUROPE WORKING GROUP ON RESEARCH POLICY AND PROGRAMME EVALUATION. *Position Statement on Research Information Systems: D/2016/13.324/11*. Nov. 2016. Disponível em: <https://www.scienceeurope.org/wp-content/uploads/2016/11/SE_PositionStatement_RIS_WEB.pdf>. Acesso em: 20 maio 2018.

SIMONS, E. *Research Information Implementation: 3-layer model*. Jornada euroCRIS, euroCRIS/IBICT, Brasília, 16 a 18 de novembro e 2016.

SOLLA PRICE, D. *Science Since Babylon*. New Haven: Yale University Press, 1961.

UNESCO. Institute for Statistics. ISCED Fields of Education and Training 2013 (ISCED-F 2013). *Manual to accomplish the International Standard Classification of Education*, 2011. Québec UNESCO-UIS, 2014. Disponível em: DOI <<http://dx.doi.org/10.15220/978-92-9189-150-4-en>>. Acesso em: 25 março 2018.

UNESCO. *Unisist Synopsis of the feasibility study on a World Science Information System*. Paris: Unesco, 1971.

VICKERY, B.C. *Faceted classification: a guide to construction and use of special schemes*. London: Aslib, 1960.

VICKERY, B.C. The Royal Society Scientific Information Conference of 1948. *Journal of Documentation*, v. 54, n.3, p. 281-283, 1998.

Os estudos cibernéticos da informação: indicadores de estruturas web e de recursos da web social

Ronaldo Ferreira de Araújo

RESUMO

O avanço tecnológico, o advento da internet e a evolução da websocial que conhecemos hoje mudaram drasticamente as maneiras de produção, consumo e disseminação da informação sendo necessário avançar na discussão dos tipos de estudos métricos da informação web e compreender seus contextos de aplicação. A presente reflexão discorre brevemente sobre o amplo campo da cibermetria, situando seus subcampos e objetos, bem como sua interface com disciplinas tradicionais da Ciência da Informação como a bibliometria, cientometria e informetria. Propõe-se um arranjo que dispõe quadrantes com as indicações de seus subcampos em dois eixos: dos objetos ou recursos de análise que se dedicam; da natureza da informação que tratam e os atores – público ou grupo social que cobrem em seus estudos.

Palavras-chave: Cibermetria. Webometria. Webmetria. Altméria. Métricas de mídias sociais.

Cybermetric information studies: web structures and social web resources indicators

ABSTRACT

Technological advances, the advent of the internet and the social web with which we are familiar today have drastically changed the means of production, consumption and dissemination of information, making it necessary to make progress in the discussion of the types of web information metrics studies and to understand their application contexts. The present analysis touches briefly on the wide field of cybermetrics, identifying its subfields and objects, as well as its interface with traditional Information Science disciplines such as bibliometrics, scientometrics and informetrics. A model is proposed which offers quadrants with the indications of their subfields in two axes: the objects or analytical resources they are concerned with, and the type of information they deal with, including the actors – the public or social groups which are covered in their studies.

Keywords: Cybermetrics. Webometrics. Altmetrics. Social media metrics.

Los estudios cibernéticos de la información: indicadores de estructuras web y de recursos de la web social

RESUMEN

El avance tecnológico, el advenimiento de Internet y la evolución de la websocial que conocemos hoy han cambiado drásticamente las maneras de producción, consumo y diseminación de la información siendo necesario avanzar en la discusión de los tipos de estudios métricos de la información web y comprender sus contextos de aplicación. La presente reflexión discurre brevemente sobre el amplio campo de la cibermetría, situando sus subcampos y objetos, así como su interfaz con disciplinas tradicionales de la Ciencia de la Información como la bibliometría, cientometría e informetría. Se propone un arreglo que dispone cuadrantes con las indicaciones de sus subcampos en dos ejes: de los objetos o recursos de análisis que se dedican; de la naturaleza de la información que tratan y los actores - público o grupo social que cubren en sus estudios.

Palabras clave: Cibermetría. Webometrics. Webmetria. Altmetria. Métricas de medios sociales.

INTRODUÇÃO

As áreas de Biblioteconomia e Ciência da Informação, dedicadas ao estudo da informação quanto ao seu ciclo, gestão, bem como os produtos e serviços que gera, têm estado muitas vezes na vanguarda da adoção de novas tecnologias da web, mas em geral, pouca atenção é dada para mensurar a apropriação e uso destas tecnologias.

O avanço tecnológico, o advento da internet e a evolução da web social que conhecemos hoje mudaram drasticamente as maneiras de produção, consumo e disseminação da informação, sendo necessário avançar na discussão da mediação como apropriação e atentar aos estudos que possam medir seus efeitos.

Hoje em dia já é possível saber a avaliação intra ou interinstitucional de universidades ou centros de pesquisa pela análise de seu website no conjunto de seus weblinks; ou o desempenho de produtos e serviços de informação (repositório institucional, website de periódicos) pelo número de visitas, cliques e downloads de itens.

Estudos sobre a popularidade on-line e engajamento de uma organização, uma marca ou um artigo científico são possíveis de ser realizados a partir da análise de compartilhamentos e comentários que recebem no Facebook; ou é possível ainda, verificar a visibilidade alcançada por uma unidade de informação (biblioteca, museu) pelo número de seguidores de sua conta no Twitter.

As situações apresentadas, longe de esgotarem as inúmeras possibilidades que o contexto da web oferece, podem ser analisadas no âmbito dos estudos cibernômicos da informação e carecem de um corpo teórico-metodológico que reconheça cada vez mais seu potencial bem como suas limitações.

A presente reflexão discorre brevemente sobre o amplo campo da cibermetria, situando seus subcampos e objetos, bem como sua interface com disciplinas tradicionais da Ciência da Informação como a bibliometria, cientometria e informetria.

Propõe-se um arranjo que dispõe quadrantes com as indicações de seus subcampos em dois eixos: dos objetos ou recursos de análise que se dedicam; da natureza da informação que tratam e os atores – público ou grupo social que cobrem em seus estudos.



CIBERMETRIA (CYBERMETRICS): DEFINIÇÕES E ESCOPO

Os estudos métricos da informação na web são em geral estudos de abordagens quantitativo-descritivas e realizados por meio de aplicações matemáticas e estatísticas no ciberespaço, seja na gestão, fluxos ou atividades relacionadas às informações.

A cibermetria (cybermetrics) ou ‘métricas web’ são termos genericamente utilizados para se referir à medição quantitativa da criação e utilização de conteúdos web, sejam eles de natureza social, política, científica ou empresarial.

Para Sen (2004, p.116) a cibermetria, ao combinar as partes da palavra ‘cyber’ (cibernético, ambiente eletrônico, virtual) e ‘metrics’ (cálculo, medida) pode ser compreendida como a “ciência da mensuração dos objetos cibernéticos”. E o elevado número de objetos (páginas, imagens, vídeos, *hyperlinks*, etc.) na web oferece oportunidade para uma grande variedade de análises quantitativas, campo de trabalho de cibermetria (BJÖRNEBORN; INGWERSEN, 2004; AGUILLO et al., 2006)

Esse campo é bastante amplo e abrange os estudos quantitativos de toda a Internet, incluindo seus aspectos conversacionais e de interatividade como *chats*, *mailing lists*, *new groups* e as tecnologias de estrutura da própria WWW, a partir de abordagens informétricas, bibliométricas e cientométricas (BJÖRNEBORN, 2002; VANTI, 2005).

Na cibermetria dispositivos como websites, páginas pessoais ou corporativas, desempenham o mesmo papel que os documentos nos estudos bibliométricos ou cientométricos. Mas o campo não se restringe apenas aos estudos sobre a ciência ou sobre comunicação científica, sua concepção é mais ampla e se estende a outros contextos.

As definições de Björneborn (2002) e de Sen (2004) indicam essa abrangência. Para o primeiro a cibermetria compreenderia a aplicação das tradicionais técnicas informétricas a qualquer tipo de informação disponível na web. Para o segundo autor a cibermetria no âmbito do ciberespaço é

[...] o ramo do conhecimento que utiliza técnicas e conceitos matemáticos e estatísticos; mede o crescimento, estabilidade, propagação e utilização; examina a autenticidade do conteúdo; estabelece leis que regem esses fatores; estuda a eficiência dos sistemas, serviços e produtos cibernéticos de informação; e avalia o impacto da era cibernética na sociedade (SEN, 2004, p.117, tradução livre).

Com tais considerações o escopo da cibermetria não poderia ser mais amplo. Abrange estudos métricos que vão da compreensão das estruturas e tecnologias web ao conteúdo e recursos da web social, aplicado a qualquer tipo de informação (científica ou social), voltado a qualquer grupo de instituições e indivíduos (não apenas dos acadêmicos) e ainda preocupado com a avaliação do seu impacto na sociedade. Os estudos cibernétricos podem ser usados para compreender os aspectos info-comunicacionais nos ambientes digitais, seja das estruturas (hiperlinks, domínios, weblinks, websites), bem como

[...] dos usuários produtores/receptores (pessoal ou institucional, gênero, faixa etária, região geográfica, comportamento); das mensagens – o conteúdo das mensagens (área, assunto, temática); dos canais ou dispositivos (blogs, microblogs, sites de redes sociais) com fins de mensuração da comunicação e interação online (ARAÚJO; TEIXEIRA, 2013, p. 956)

A cibermetria é considerada o grande campo de estudos métricos da informação na web, seus subcampos podem ser identificados de acordo com a interface que possuem com os estudos métricos tradicionais (bibliometria, cientometria e informetria), bem como com os objetos ou recursos de análise a que se dedicam; a natureza da informação que tratam, bem como o público ou grupo social a que se destina.



OS ESTUDOS SOBRE AS ESTRUTURAS DA WEB: WEBOMETRIA E WEBMETRIA

A webometria consiste no “estudo dos aspectos quantitativos da construção e uso dos recursos da informação, estruturas e tecnologias na Web” (BJÖRNEBORN; INGWERSEN, 2004), ou seja, na aplicação de métodos informétricos à World Wide Web, que serve para medir, dentre outros aspectos, a frequência da distribuição das páginas no ciberespaço.

A “medição aponta para o estudo ou análise comparativa da presença dos diversos países na rede, das proporções de páginas pessoais, comerciais e institucionais” (VANTI, 2002, p. 155). As páginas podem ser analisadas por tipo (setor público, privado), classificações (pessoais, institucionais), categorias (páginas-documento, páginas-índice, páginas-recurso) e por medição temporal para fins de comparação de crescimento e evolução da rede sobre determinado assunto ou matéria.

Para autores como Thelwall (2003, p.3), a webometria teria uma relação bem próxima da bibliometria, uma vez que “os documentos web – sejam eles textos ou multimídia – não deixam de ser informação registrada e armazenada, mesmo que em servidores web”, bem como da cientometria, tendo em vista que muitas atividades acadêmicas e científicas são cada vez mais apoiadas por recursos da web (VANTI, 2011).

Os instrumentos fundamentais para a realização de estudos webométricos têm sido os motores de busca, que permitem trabalhar com grandes volumes de informação, que facilitam as tarefas de quantificação e avaliação dos fluxos de intercâmbio de dados e informação na Web. Para Vanti (2002, p. 157), tais buscadores permitem contabilizar o número “total de páginas em um espaço web e os links a tais espaços, entendendo o termo espaço web no sentido de domínio (seja um domínio de país ou um domínio institucional)”.

Pode ser calculado nestas análises, por exemplo, “o tamanho médio de uma página expressado em bytes, o número médio de links por página e a densidade média por link. Para tal fim, há certos elementos que devem ser considerados, como o Uniform Resource Locators - URL, o título, as palavras-chave, o tipo de home page, o domínio, o tamanho e o número de links” (VANTI, 2002, p.157).

E similar ao fator de impacto para periódicos científicos considera-se como Fator de Impacto Web (FIW) a soma do número de links contidos em páginas web externas e internas de determinado país ou site dividido pelo número de páginas encontradas nesse país ou site da web em um dado momento. Este “indicador serve para medir e comparar a atratividade e influência que possam alcançar distintos espaços na Web” (VANTI, 2002, p.158). Para Gouveia (2013), com a constante ampliação das métricas de acesso, uso e inter-relação disponíveis na Internet, outros indicadores poderão ser incluídos.

Um subconjunto da webometria, que tem sido considerado segmento particularmente valioso e de pesquisa comercialmente relevante (CALERO; RUIZ; PIATTINI, 2005) é a webmetria (‘webmetrics’, ‘web metrics’ ou ‘web analytics’). Trata-se de um campo voltado para análises métricas de tráfego de vistas em websites (acesso e uso de informação na web: logs, page tagging) que auxilia no controle de qualidade de processos e recursos da web.

De acordo com Khoo et al. (2008), essa métrica pode ser utilizada para estimar se os objetivos dos usuários estão sendo alcançados ou não, além de apoiar estudos de usabilidade e web design, fornecendo feedbacks para desenvolvedores, gerentes e outros interessados. Podendo ser empregada em qualquer tipo de website, comercial ou acadêmico, a webmetria pode auxiliar no projeto e otimização de desempenho de websites e atrair bons resultados para a arquitetura da informação (WEISCHEDEL; HUIZINGH, 2006).

Para Gouveia e Lang (2013, p. 174), uma analogia que pode ser feita para melhor compreensão é a de que os estudos webmétricos seriam como “o levantamento do acesso a uma biblioteca, desde sua visita até o número de consultas feitas ao acervo, bem como o retorno ou não à mesma biblioteca”. Conhecer o caminho percorrido pelo usuário, o que o trouxe, quanto tempo permaneceu e com que grau de satisfação saiu não é uma preocupação apenas de uma biblioteca.

Tais informações são essenciais para a gestão de um website. No ambiente digital e com os recursos da web 2.0 os usuários deixam rastros dessas ações, sendo possível monitorá-las, inclusive o aumento de visitas a determinados websites a partir de links dispostos em blogs e sites de redes sociais fez surgir uma crescente e comercial área, ligada atualmente ao marketing digital, conhecida como social media metrics.

ESTUDOS DOS RECURSOS DA WEB SOCIAL: MÉTRICAS DE MÍDIAS SOCIAIS, ALTMETRIA E CIENTOMETRIA 2.0

As primeiras discussões sobre '*social media métrics*' ou no português, métricas de mídias sociais (OWYANG; TOLL, 2007; LEVENSHUS, 2007; POP, 2008) apresentam uma relação próxima com a abordagem webométrica. Nessa aproximação estudos métricos de mídias sociais podem ser definidos como o uso de dados webométricos para mensurar a influência digital e o engajamento.

Métricas de mídias sociais podem fornecer dados da mediação da informação sobre acesso e uso de recursos em ambientes digitais, como participação em grupos on-line, aumento (do tempo) de engajamento, visualizações de página, downloads de podcasts (LEVENSHUS, 2007), e o monitoramento dessas ações serve para avaliar como elas impactam ou são impactadas por campanhas de marketing (ou análises de crise) de organizações, marcas, produtos ou serviços (OWYANG; TOLL, 2007).

Dos primeiros trabalhos sobre tema Pop (2008) é quem mais evidencia a relação entre a webmetria e as métricas de mídias sociais. Para o autor influência digital é o ponto central da mensuração de mídias sociais. Vistas como atributos ou fatores importantes para influência digital tais métricas podem determinar resultados ou fazer melhorias. Para Pop (2008) a influência digital pode ser medida por métricas básicas (relacionadas às visitas recebidas e à relação entre o tráfego no website e o conteúdo) ou avançadas (audiência, engajamento: atenção e interação, fidelização, influência, dentre outras).

Ambientes privilegiados para obtenção de dados sobre a influência digital, os blogs e as mídias sociais estão cada vez mais presentes em todos os níveis e segmentos da sociedade, e o mesmo serve para a ciência.

É exatamente na sua aplicação para análise da informação científica que pesquisadores da Ciência da Informação têm considerado um novo subcampo da cibermetria. Trata-se da altmetria ou '*altmetrics*', que mantém uma afinidade direta com os estudos cientométricos e bibliométricos, pode-se valer de dados webométricos e webométricos, mas cujo interesse de análise é o da comunicação científica (comunidades ou assuntos) no contexto da web social e dos recursos da web 2.0 (GOUVEIA, 2013; ARAÚJO, 2015). Para Souza (2014, p. 105), a altmetria se apresenta

[...] não apenas como uma evolução técnica dos estudos métricos da informação (bibliometria, informetria, webometria, cientometria etc), agora aplicados à web social; mas também como uma reação política à crise do sistema de comunicação científica.

Estas novas métricas poderiam auxiliar estudiosos e acadêmicos a encontrar artigos importantes e talvez também avaliar o impacto de seus próprios artigos (THELWALL, 2014).

De forma geral a altmetria pode ser entendida como a aplicação das métricas de mídias sociais no contexto da comunicação científica. Assim, a diferença é que o investimento no marketing digital e o monitoramento da conversação e engajamento dos usuários não estão voltados para empresas ou marcas e sim para a atenção on-line que produtos científicos, como os artigos de periódicos, recebem.

Para Araújo (2014, p.5), “saber se os artigos passam a circular na web social e compreender as métricas que sustentam seu fluxo na rede contribui para desenhar seu quadro altmétrico”. E com esse quadro é “possivelmente verificar, a partir da atenção que recebem se resulta em futuras citações” (ARAÚJO, 2015, p.4).

A altmetria e a bibliometria, em especial, a análise de citações mensuram impactos distintos, mas que se relacionam, e que isoladas, nenhuma delas é capaz de descrever o quadro completo da comunicação científica (PRIEM; PIWOWAR; HEMMINGER, 2012). Assim, a altmetria indicaria “indícios anteriores de impactos que o artigo gera antes de obter citações podendo predizê-las ou não” (ARAÚJO, 2015, p.4).

Em alguns casos os limites entre as métricas não são tão claros, em outros surgem novas denominações. A própria altmetria, por exemplo, ao ser considerada métrica alternativa para comunicação científica lança o olhar para questões sobre a circulação de uma produção científica (onde um artigo está sendo lido, compartilhado e discutido), o que amplia o olhar para a visibilidade e o alcance dos resultados de investigação.

Uma vez preservado o contexto web, mas ampliado o escopo de análise de ‘um artigo’ para ‘uma ciência’, como na abordagem cientométrica, por exemplo, que consiste na aplicação de métodos quantitativos ao estudo da história da ciência e do progresso científico e tecnológico na medição dos incrementos de produção e produtividade de uma disciplina, ou de um grupo de pesquisadores de uma área (TAGUE-SUTCKIFFE, 1992; VANTI, 2002; ARAÚJO; ALVARENGA, 2011) já seria outro subcampo em questão.

Trata-se da ‘Cyberscientometrics’ (LEYDESDORFF & WOUTERS, 1999; AGUILLO, 2002a, 2002b) ou ainda ‘cientometria 2.0’ (PRIEM; HEMMINGER, 2010; ARAÚJO, 2014). Para Araújo (2014, p.34), a cientometria 2.0 é uma concepção coerente aos moldes de uma cultura digital quando o método cibernético “é empregado no contexto de análise da ciência e da comunicação científica, repensando ações, fluxos e canais de informação”.

Este subcampo também tem sido pensado tendo como foco análises da presença on-line de Instituições de Pesquisa & Desenvolvimento, bem como estudos do processo formal (revistas eletrônicas) e informal de comunicação científica na Internet e suas interfaces.

Acrescentam-se ainda a esse escopo avaliações quanto à presença on-line de pesquisadores e a visibilidade e o impacto dos usos que eles fazem da web para atividades científicas (ARAÚJO, 2014).

Assim, percebe-se que na Ciência da Informação encontramos várias definições e interfaces entre os subcampos da cibermetria e com outras métricas tradicionais da informação. Para contribuir com a discussão, torna-se necessário refletir cada vez mais sobre cada um deles tendo em vista a aplicabilidade que apresentam.



ESTUDOS CIBERMÉTRICOS E A CIÊNCIA DA INFORMAÇÃO

Os vários autores da Ciência da Informação lançaram seu olhar para as inter-relações entre os subcampos dos estudos métricos da informação (THELWALL; VAUGHAN; BJORNEBORN, 2003; BJÖRNEBORN; INGWERSEN, 2004; BJÖRNEBORN, 2002; VANTI, 2002; 2005; 2011; GOUVEIA, 2013) e procuraram sistematizar por meio de diagramas suas visões quanto à abrangência e limites que cada subcampo apresentava.

Por não haver consenso entre os autores, percebemos algumas divergências, principalmente quanto a assertivas de que determinados subcampos estejam completamente absorvidos por outros.

Consideramos que não é necessariamente a delimitação clara dessas inter-relações que vai reduzir as incertezas sobre esse campo, e sim (i) a identificação do propósito do estudo que se pretende desenvolver com (ii) a aproximação e emprego de métricas tradicionais (bibliometria, cientometria e infometria) coerentes ao propósito.

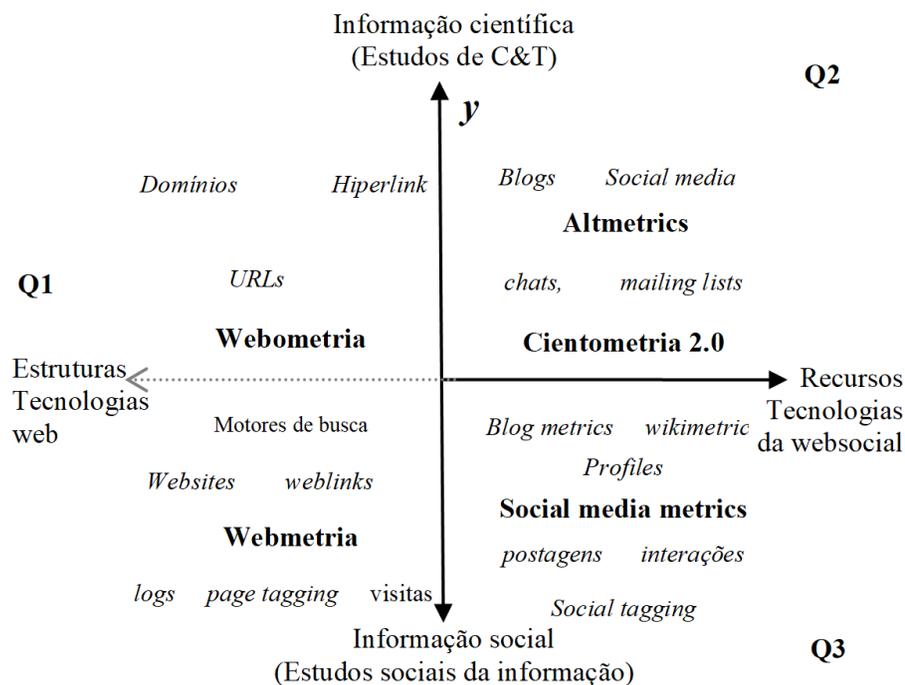
Assim, propõe-se que o arranjo entre os subcampos não seja representado por sobreposições entre eles e sim compreendido entre dois eixos. O primeiro deles (x) relaciona-se aos objetos e (ou) recursos de análise que consideram, indo das tecnologias e estruturas da web (webometria; webmetria) e seus elementos mais estáticos até os conteúdos e recursos dinâmicos da web social (métricas de mídias sociais).

No segundo eixo (y) está em consideração a natureza da informação que tratam, bem como o público ou grupo social que investigam. Dessa forma ele pode ser, em uma das pontas, voltada para estudos tradicionais da informação científica e destinadas à análise da Ciência e Tecnologia, e em outra, ligada à informação social (não científica), direcionado a outros setores e grupos sociais que não os acadêmicos.

Para facilitar a visualização do arranjo proposto, na Figura 1 temos a $f(x,y)$ que representa a amplitude estudos cibernômicos com três quadrantes (Q1, Q2 e Q3), cada um direcionado a uma combinação dos eixos.

No primeiro quadrante (Q1), temos os estudos webométricos e webmétricos. Embora instituições de ensino superior e centros de pesquisa (públicas e privadas) sejam protagonistas recorrentes nos estudos webométricos, contando com inúmeros trabalhos que retratam o desempenho que alcançam em diversos continentes, países e regiões – conforme revisão ampla descrita por Thelwall, Vaughan e Björneborn (2003) e a publicação desde 2004 de um ranking que cobre mais de 17 mil instituições de ensino superior por todo o mundo, intitulado *Webometrics Ranking of World Universities* (VANTI, 2010) – a definição e aplicação do método webométrico não o restringe à informação científica ou a um grupo social único ou setor privilegiado.

Figura 1 – Quadrantes dos estudos métricos da informação na web



No entanto, o número de estudos com abordagens webométricas que se dedicam a outros contextos que não os acadêmicos é bem reduzido, como o trabalho de Thelwall e Vaughan (2004) que considerou a análise de websites comerciais para comparar a presença on-line e desempenho de países (China, Singapura, Tailândia e estados Unidos) ou o estudo de Vaughan e You (2010), que analisa a co-ocorrência de weblinks e termos de páginas web de setores da indústria de telecomunicações ao redor do globo.

Os estudos webmétricos, por sua vez, fornecem indicadores essenciais para avaliação de desempenho de websites refletindo como se dá sua utilização pelos usuários que o acessam (WEISCHEDEL; HUIZINGH, 2006).

Assim, em Q1, ambas as abordagens (webometria e webmetria) representam análises de 'estruturas da web' do eixo x , mas sem divisão quanto ao eixo y , por terem emprego tanto comercial quanto acadêmico. A representação da falta da divisão está expressa na seta pontilhada.

Quando analisamos o eixo x no extremo de análises das 'tecnologias e recursos da web social', lado de ampla cultura participativa de comunicação mais intensiva, colaborativa e dinâmica, percebemos uma divisão clara pelos quadrantes Q2 e Q3. No segundo quadrante temos o estudo cibernômico empregado para análise cientométrica de áreas de conhecimento e comunidades científicas, bem como aos atores (pesquisadores e cientistas) sendo chamados de Cyberscientometrics ou cientometria 2.0.

Quando se trata de produtos da ciência (artigos, resultados de pesquisa – mais próxima dos estudos bibliométricos) e sua circulação em blogs e sites de redes sociais é tratada como altmetria.

O terceiro quadrante (Q3) é pouco explorado pela CI, tendo mais repercussão na Comunicação social, sendo geralmente pautado nos estudos de cibercultura e nas métricas de mídias sociais tendo um apelo mais comercial, encabeçado originalmente por profissionais e pesquisadores do marketing digital. Ele nasce com interfaces com a informação social e de aproximação com os estudos informétricos e webmétricos.



CONSIDERAÇÕES FINAIS

Os estudos cibernétricos estão em expansão e tendem a acompanhar a própria evolução da web, seja em sua estrutura e discussões que envolvam, por exemplo, sua web semântica como nas implicações do dinamismo e empoderamento dos recursos da web social.

A Ciência da Informação possui uma responsabilidade social – da democratização, ao acesso, bem como do tratamento dos recursos informacionais – que a coloca, senão no centro dessa discussão, em posição bem próxima a ele e pode contribuir com os aspectos teóricos e metodológicos desse amplo campo dos estudos métricos da informação na web.

Como dito, em alguns casos as fronteiras entre as métricas não são tão claras, em outros surgem novas denominações. Tentativas de determinação de onde começa um subcampo e termina outro pode mais confundir do que contribuir para entendimento de cada um deles. Um novo campo, que precisa de maior atenção, com denominação *Article-Level Metrics* ou ‘Métricas em nível de artigo’ (NEYLON; WU, 2009), por exemplo, combinaria tanto dados webmétricos (de acesso, cliques e *downloads*), alométricos (de atenção online de mídias sociais), como de estudos tradicionais bibliométricos (de citação) para traçar o quadro completo do impacto social de um artigo científico.

Assim o arranjo proposto não pretende determinar a delimitação dos subcampos, mas apenas reconhecê-los quanto ao propósito do estudo que se pretende desenvolver com o emprego de métricas (tradicionais ou não) coerentes ao propósito.

Por fim, vale ressaltar que estudos métricos da informação, sejam os clássicos ou os emergentes que surgem com a evolução da web, recebem constantemente críticas sobre seu caráter altamente quantitativo. A CI pode contribuir na superação ou mesmo reflexão dessas críticas explorando a complementariedade que abordagens qualitativas oferecem tanto aos estudos da informação científica quanto ao seu compromisso com a informação social.

Em estudos das estruturas e tecnologias web, análises de razões de ‘situação’, similares aos realizados nos estudos tradicionais de citação e abordagens netnográficas, podem se constituir caminhos promissores para qualificar e contextualizar socialmente os dados quantitativos levantados pelas aplicações métricas.

Os estudos de mediação e apropriação da informação, bem como de regimes de informação, por exemplo, podem contribuir e muito para as análises da ciência pautada em abordagens alométricas e ou os impactos sociais de seu uso em outros contextos socioculturais, possibilitando a identificação de atores sociais, ações de informação, dispositivos e artefatos de informação.



REFERÊNCIAS

AGUILLO et al. Scientific research activity and communication measured with cybermetrics indicators. *Journal of the American Society for Information Science and Technology*, v.57, n.10, p. 1296-1302, Ago. 2006. Disponível em: <<https://onlinelibrary-iley.ez106.periodicos.capes.gov.br/doi/full/10.1002/asi.20433>>. Acesso em: 28 ago. 2018. <https://doi.org/10.1002/asi.20433>

AGUILLO, I.F. *Cybermetrics, definitions and methods for an emerging discipline*. In: SÉMINAIRES DE L'ADEST, Paris, 14 fev., 2002. *Annales... l'ADEST*, Paris, 14 fev. 2002a.

AGUILLO, I.F. *Measuring informal scientific publication in the Web*. Conference, Responsibility under Uncertainty, York (UK), EASST 2002. (CCHS-IEDCYT) *Proceedings...* York (UK), EASST, 2002b. Disponível em: <<http://digital.csic.es/bitstream/10261/4163/1/R-8.pdf>>. Acesso em: 28 ago. 2018.

ARAÚJO, R. F. Da altmetria à análise de citações: uma análise da revista datagramazero. *DataGramaZero*, v. 16, n. 1, p. 1-20, 2015. Disponível em: <<http://www.brapci.inf.br/v/a/18494>>. Acesso em: 28 ago. 2018.

_____. Cientometria 2.0, visibilidade e citação: uma incursão altmétrica em artigos de periódicos da ciência da informação. In: *Encontro Brasileiro de Bibliometria e Cientometria*, v.4, p. 1-7, 2014, Recife. *Anais...* Recife, 2014. Disponível em: <<http://www.brapci.inf.br/index.php/article/download/27221>>. Acesso em: 28 ago. 2018.

_____; TEIXEIRA, J. C. S. Biblioteconomia conectada: uma análise da biblioblogosfera brasileira. *Revista ACB*, Florianópolis, v. 18, p. 949-978, jul./dez 2013. Disponível em: <<https://revista.acbsc.org.br/racb/article/view/924/pdf>>. Acesso em: 28 ago. 2018.

_____; ALVARENGA, L. A bibliometria na pesquisa científica da pós-graduação brasileira de 1987 a 2007. *Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação*, v. 16, n. 31, p.51-70, 2011. Disponível em: <<http://www.brapci.inf.br/v/a/10158>>. Acesso em: 28 ago. 2018. DOI:10.5007/1518-2924.2011v16n31p51.

BJÖRNEBORN, Lennart; INGWERSEN; Peter. Towards a basic framework of webometrics. *Journal of American Society for Information Science and Technology*, v. 55, n. 14, p. 1216-1227, 2004. Disponível em: <<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.94.1691&rep=rep1&type=pdf>>. Acesso em: 28 ago. 2018

_____. Small-world link structures on the web. Copenhagen, DK: *School of Library and Information Science*, 2002. Disponível em: <www.db.dk/lb/2002smallworld.pps>. Acesso em: 28 ago. 2018.

CALERO, C.; RUIZ, J.; PIATTINI, M. Classifying web metrics using the web quality model. *Online Information Review*, v. 29, n.3, p-227-248, 2005. Disponível em: <<https://www.emeraldinsight-com.ez106.periodicos.capes.gov.br/doi/full/10.1108/14684520510607560>>. Acesso em: 28 ago. 2018

GOUVEIA, F. C. Altmetria: métricas de produção científica para além das citações. *Liinc em Revista*, Rio de Janeiro, v. 9, n. 1, p. 214-227, maio 2013. Disponível em: <<http://revista.ibict.br/liinc/article/view/3434/3004>>. Acesso em: 29 ago. 2018.

GOUVEIA, F. C.; LANG, P. Da webometria à altmetria: uma jornada por uma ciência emergente. In: ALBAGLI, S. (Org.) *Fronteiras da Ciência da Informação*. Brasília: IBICT, 2013. Disponível em: <https://edisciplinas.usp.br/pluginfile.php/263580/mod_resource/content/1/Livro_Fronteiras%20da%20Ci%C3%AAncia%20da%20Informa%C3%A7%C3%A3o.pdf>. Acesso em: 29 ago. 2018

KHOO, M. et al.. Using Web Metrics to Analyze Digital Libraries. 8th ACM/IEEE-CS Joint Conference on Digital Libraries, ACM: New York, 2008. *Proceedings...* JCDL, n.8, p. 375-384, 2008. Disponível em: <<https://dl.acm.org/citation.cfm?id=1378956>>. Acesso em: 29 ago. 2018

LEVENSHUS, A. B. *Public radio's social media experiments: Risk, opportunity, Challenger*. Public Communication Division, School of Communication, American University. Disponível em: <<http://hdl.handle.net/1961/4608>>. Acesso em: 29 ago. 2018

LEYDESDORFF L.; WOUTERS P. Between texts and contexts: advances in theories of citation? *Scientometrics*, v. 44, n. 2, p.173-92, 1999. Disponível em: <<https://link-springer-com.ez106.periodicos.capes.gov.br/content/pdf/10.1007%2FBF02457378.pdf>>. Acesso em: 29 ago. 2018.

NEYLON, C.; WU, S. Article-Level Metrics and the Evolution of Scientific Impact. *PLoS Biol*, v.7, n. 11, 2009. Disponível em: <https://www.researchgate.net/publication/38094377_Article-Level_Metrics_and_the_Evolution_of_Scientific_Impact>. Acesso em: 29 ago. 2018. Doi:10.1371/journal.pbio.1000242.

OWYANG, J.; TOLL, M. Tracking the influence of conversations. A Roundtable Discussion on Social Media Metrics and Measurement. *Proceedings...* Dow Jones & Company, 2007. Disponível em: <<http://www.web-strategist.com/blog/wp-content/uploads/2007/08/trackingtheinfluence.pdf>>. Acesso em: 29 ago. 2018.

POP, I. 8th WSEAS international conference on Distance learning and web engineering. *Proceeding...* Cantabria, España, p. 169-173, set. 2008. Disponível em: <<https://dl.acm.org/citation.cfm?id=1519423>>. Acesso em: 29 ago. 2018.

PRIEM, Jason; PIWOWAR, Heather A.; HEMMINGER, B. M. Altmetrics in the wild: using social media to explore scholarly impact. *CoRR*, p. 1-17, mar. 2012. Disponível em: <<https://arxiv.org/abs/1203.4745>>. Acesso em: 29 ago. 2018.

PRIEM, J.; HEMMINGER, B.M. Scientometrics 2.0: Toward New Metrics of Scholarly Impact on the Social Web. *First Monday*, v.15, n. 7, jul. 2010. Disponível em: <<https://uncch.pure.elsevier.com/en/publications/scientometrics-20-toward-new-metrics-of-scholarly-impact-on-the-s>>. Acesso em: 29 ago. 2018.

SEN, B. K. Cybermetrics: meaning, definition, scope and constituents. *Annals of Library and Information Studies*, v. 51, n.3, p.116-120, set. 2004. Disponível em: <<http://nopr.niscair.res.in/bitstream/123456789/4064/1/ALIS%2051%283%29%20116-120.pdf>>. Acesso em: 29 ago. 2018.

TAGUE-SUTCKIFFE, J. An introduction to informetrics. *Information Processing & Management*, v. 28, n. 1, p. 1-3, 1992. Disponível em: <<https://www.sciencedirect.com/science/article/pii/030645739290087G>>. Acesso em: 29 ago. 2018.

THELWALL, M.; VAUGHAN, L. A fair history of the Web? Examining country balance in the Internet Archive. *Library & Information Science Research*, v.26, n.2, p.162-176, 2004. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0740818804000246>>. Acesso em: 29 ago. 2018.

_____. What is this link doing here?: beginning a fine-grained process of identifying reasons for academic hyperlink creation. *Information Research*, v.8, n.3, abr. 2003. Disponível em: <<http://www.informationr.net/ir/8-3/paper151.html>>. Acesso em: 29 ago. 2018.

_____. VAUGHAN, L.; BJORNEBORN, L. Webometrics. *Annual Review of Information Science And Technoy*, v.39, n.1, 2006. Disponível em: <<https://onlinelibrary-wiley.ez106.periodicos.capes.gov.br/doi/full/10.1002/aris.1440390110>>. Acesso em: 29 ago. 2018.

VANTI, Nadia. *Links hipertextuais na comunicação científica: uma análise webométrica*. Natal: EDUFRN, v. 1, p. 188, 2011.

_____. Mapeamento das Instituições Federais de Ensino Superior da Região Nordeste do Brasil na Web. *Informação & informação* (UEL. Online), v. 15, n. 1, p. 55-67, 2010. Disponível em: <<https://doaj.org/article/ea9fb7b9e8b94414baf917a1d9f922e8>>. Acesso em: 29 ago. 2018.

_____. Análise Webométrica dos Sítios Acadêmicos Latino-Americanos em Ciências Sociais. In: *VIII Encontro Nacional de Pesquisa em Ciência da Informação*, Salvador, out., 2007. *Anais...* Salvador, 2007. Disponível em: <<http://www.enancib.ppgci.ufba.br/artigos/GT7--081.pdf>>. Acesso em: 29 ago. 2018.

_____. Os links e os estudos webométricos. *Ci. Inf.*, Brasília, v. 34, n. 1, p.78-88, jan./abr. 2005. Disponível em: <<https://lume.ufrgs.br/handle/10183/94417>>. Acesso em: 29 ago. 2018.

_____. Da bibliometria à webometria: uma exploração conceitual dos mecanismos utilizados para medir o registro da informação e a difusão do conhecimento. *Ci. Inf.*, Brasília, v. 31, n. 2, p. 152-162, maio/ago. 2002. Disponível em: <<http://www.scielo.br/pdf/ci/v31n2/12918.pdf>>. Acesso em: 29 ago. 2018.

VAUGHAN, L.; YOU, J. Word co-occurrences on Webpages as a measure of the relatedness of organizations: A new Webometrics concept. *Journal of Informetrics*, v.4, n.4, p.483-491, out. 2010. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1751157710000386?via%3Dihub>>. Acesso em: 29 ago. 2018.

WEISCHDEL, B.; HUIZINGH, E. K. R. E. Website optimization with web metrics: a case study. 8th international conference on Electronic commerce. Proceeding. Canada, p. 463-470, 2006. Disponível em: <<https://www.bibsonomy.org/bibtex/2a157e7c2d98ced0aa076566e62b8ed48/ewomant>>. Acesso em: 29 ago. 2018.



RECONHECIMENTO

Artigo originalmente publicado na coletânea Estudos Métricos da Informação na Web: atores, ações e dispositivos informacionais, organizada por *Ronaldo Ferreira de Araújo* em 2015. Agradecimentos à Lidia Ramires e à *Fernanda Lins*, da Editora da Universidade Federal de Alagoas, EDUFAL, pela presteza na autorização para republicar o artigo na presente coletânea.





Parte II

SUMÁRIOS EXECUTIVOS





Sobre a “Classification de la Recherche Universitaire” (Tabela CASRAI) Hagar Espanha Gomes

Por ocasião da reunião de instalação do Comitê Executivo do CASRAI para a Europa, realizada na República de Malta em maio de 2017, foi sugerido que houvesse estudo do uso da tabela CASRAI nos serviços de informação de pesquisa no Brasil, tendo em vista a iniciativa do País de participar do euroCRIS. Os resultados desta análise devem ser considerados nos estudos de compatibilização ou harmonização que venham a ser feitos com vistas ao alcance daquele objetivo.

Os propósitos de uma tabela de classificação explicam, de algum modo, decisões adotadas em sua elaboração. Quando tais classificações visam à produção de indicadores, o primeiro exercício é saber que tipos de indicadores se pretende obter.

Os comentários a seguir contemplam a estrutura da Tabela, análise das dimensões e propostas para sua reformulação.¹

Em geral, tais indicadores resultam da combinação de mais de um aspecto para permitir várias análises, e isto se reflete na estrutura das classificações. Desse modo, consegue-se especificidade, mesmo partindo de classes gerais. Algum nível de generalidade deve caracterizar tais sistemas, entre outros motivos, porque neste nível consenso e aceitação geral são mais fáceis de alcançar.

A análise da Classification de la Recherche Universitaire – daqui em diante, tabela CASRAI – leva em conta o nível de generalização e a possibilidade de arranjo e combinação das diferentes dimensões, a característica que a torna dinâmica. Os comentários a seguir bem como as propostas visam a contornar algumas deficiências ou impropriedades da Tabela como um todo, e não sua substituição.



¹ Extraído de: GOMES, Hagar Espanha. *Estudo CASRAI - Compatibilidade de sistemas*. Produto 1 do projeto: Dos padrões internacionais de estruturação da informação de pesquisa aos indicadores (BRICRIS). Rio de Janeiro, 32 p. 2017. IBICT - Iniciativa de Apoio aos Diálogos Setoriais União Europeia/Brasil.

ESTRUTURA DA TABELA

A tabela CASRAI contempla três dimensões: disciplinas, objetos de pesquisa e campos de aplicação.

É importante considerar, logo de início, que tabelas de classificação para produzir indicadores guardam alto nível de generalidade, como ocorre na tabela CASRAI. O que se vai analisar são os princípios adotados na organização dos elementos, tanto no primeiro nível (Secteur) como no segundo (Disciplinary field).

ANÁLISE DA PRIMEIRA DIMENSÃO: DISCIPLINAS

A dimensão Disciplina está organizada por setor/disciplinas:

Natural Sciences and Engineering

- Engineering
- Applied Sciences
- Natural Sciences
- Mathematical Sciences

Health Sciences

- Medical Sciences
- Nursing
- Physical Education and Rehabilitation

Social Sciences and Humanities

- Education
- Management
- Humanities
- Social Sciences

Arts and Literature

- Writing and Fine Arts
- Arts and Literature Studies

O primeiro nível desta dimensão - Setor - não é tão diferente de outras classificações voltadas para C&T, as quais em geral adotam disciplinas como primeiro princípio de divisão do conhecimento.

Duas ‘disciplinas’ chamam atenção: Applied Sciences e Nursing. No primeiro caso, não se percebe o princípio de divisão adotado e, no segundo, apenas um item é contemplado, o que não justificaria tal subdivisão, principalmente quando ‘Nursing’ poderia integrar a disciplina Medical Sciences, uma classificação canônica.

No que se refere à análise do conteúdo de cada disciplina, não fica claro o princípio de divisão adotado: por vezes, disciplinas, por vezes, aspectos. Uma ordenação no seu interior, feita segundo alguns aspectos/facetas, evidenciaria alguma estruturação do conjunto; isto contribuiria para a criação de uma notação; neste caso, qualquer que fosse a língua adotada, a ordenação seria mantida e isso permitiria agregação de dados quando da produção de indicadores.

Ao analisar disciplinary field, observa-se que também ali não existe uma ordenação sistemática, ou seja, uma reunião dos elementos segundo aspectos comuns. Por exemplo, medical sciences arrola 40 elementos em ordem alfabética. Uma tentativa de reuni-los segundo aspectos comuns mostraria a justeza ou não da inclusão deste ou daquele item e produziria conjuntos consistentes. Se fosse o caso de atribuir uma notação, esta possibilitaria agregação maior ou menor no momento da obtenção de indicadores.

Não se pode perder de vista a possibilidade de integração de dados ou compartilhamento com outras organizações. Em relação à compatibilização da tabela CASRAI com a tabela FoS (Fields of Science and Technology, da OCDE), verifica-se que ela não oferece problema, à exceção da ‘disciplina’ Applied Sciences da Casrai, que reúne elementos de natureza muito diversa. De fato, este exercício sugere que esta ‘disciplina’ mereceria análise detalhada e possível redistribuição dos itens para outras classes.

ANÁLISE DAS DEMAIS DIMENSÕES

Em relação às demais dimensões – Objeto e Campo de Aplicação –, existem sugestões no Concept Paper de Patrick Lambe, que poderiam melhorar a tabela CASRAI. Naquele texto, algumas considerações merecem destaque, **no que concerne às dimensões** que venham a ser adotadas, ou seja, quais aspectos ou facetas seriam contemplados (as).

Generally, to avoid placing an undue cognitive burden on users, it is advisable to select only three facets [...] The selection and design of the facets in a classification system should be done with particular care, to ensure that the facets are both relevant to organizational needs and a useful representation of the activities.²

Esta observação significa rever aquelas dimensões. A partir da decisão sobre os aspectos selecionados, pode-se analisar a tabela CASRAI e verificar o que convém manter e o que convém descartar naquelas dimensões, se seriam mantidas ou substituídas por facetas selecionadas após discussão do problema. No preâmbulo, o conteúdo da dimensão objetos seriam fenômenos, mas uma análise dos itens mostra inconsistências. A proposta de facetação sugerida anteriormente significa que a Tabela continuaria a manter sua característica dinâmica, porém com seus elementos agrupados segundo algum princípio ou aspecto.

Em relação a Campo de Aplicação, talvez a classificação adotada atenda ao setor econômico dos integrantes da UE. Também neste caso os princípios de classificação para o conteúdo de cada campo deveriam ser os mesmos empregados, a saber, a organização por facetas/aspectos.

Uma observação geral: tendo em vista que as diversas dimensões da Tabela contêm alto nível de generalidade, é imprescindível que cada classe seja acompanhada de um conjunto de informação orientando sobre o conteúdo específico de cada item, indicando a abrangência de cada uma. Na sua ausência, isso significa uma carga de trabalho para o usuário. Como exemplo, cita-se a tabela FoS, que inclui informação básica sobre o que contempla cada classe; isto pouparia o esforço do pesquisador no momento de indexar seu documento na base de dados.



² ALEXANDER, J. ; LAMBE, P. *Concept Paper. Developing organizational and statistical taxonomies for the Classification of Research and Development Activities: a best practice study and guide.* Arlington, V.A, SRI International, June 15, 2015. p. 6.

PROPOSTAS DE REFORMULAÇÃO DA CASRAI: PARA DISCUSSÃO

Para iniciar, alguns princípios de classificação que poderiam ser adotados na reformulação da Tabela.

Facetas ou aspectos guardam nível de generalidade: coisas, instrumentos, técnicas, processos, e assim por diante; são constituídas de elementos de natureza específica que compõem as taxonomias. Além de contribuir para a construção das classes e subclasses, facilitaria a integração ou compatibilização com outras bases de dados.

Many R&D activities combine or merge aspects of multiple existing fields and disciplines. Thus, interdisciplinary R&D is especially challenging, since many taxonomies are designed to apply one and only one label to an activity. Despite these difficulties, developing consistent and interpretable R&D taxonomies will become even more important in the future.³

Como harmonizar taxonomias e classes? Taxonomias não fazem parte, como tal (por exemplo, árvores hierárquicas) de uma tabela do tipo CASRAI. Taxonomias são úteis para a indexação de projetos, relatórios e assemelhados, que requerem representação específica do conteúdo e poderiam ser oferecidas como padrão para a indexação daqueles tipos de documento. A elaboração de uma taxonomia é resultado do método indutivo. Cada elemento é definido segundo sua essência – processo, propriedade, objeto... -, e seguindo cada passo na divisão, chega-se ao termo mais geral, ou seja, a faceta.

O recorte das disciplinas tem objetivos práticos e é resultado de um processo dedutivo; talvez por isso mesmo seja mais difícil chegar a um consenso quando se busca maior detalhamento. O maior nível de generalidade talvez facilite um acordo entre os participantes de uma rede. A reformulação da Tabela requer maior discussão no (s) nível (eis) de detalhamento que venham a ser adotados. Em outras palavras: quais critérios adotar na organização dos elementos de cada um destes níveis?

³Ibidem, p. 6.

No processo de estruturação das classes e subclasses, as taxonomias oferecem orientação segura para a decisão dos limites de abrangência de cada uma, facilitando sua consulta e uso, a exemplo da tabela FoS. Elas podem ter, assim, duplo papel: contribuem para a organização interna das classes e subclasses e podem atuar – como instrumento à parte – como vocabulário controlado para uso na indexação.

A maior qualidade da Tabela pode estar em sua estrutura multidimensional, que não deve ser abandonada: até onde foi possível verificar, esta característica é ‘única nos países europeus’ e bastante apreciada pelos classificacionistas, em especial aqueles seguidores do método de faceta.



ATIVIDADES PROPOSTAS DE REVISÃO DA TABELA CASRAI

- a) analisar o primeiro nível de divisão - Disciplinas - adotado em geral como primeira dimensão;
- b) levantar os tipos de indicadores esperados para, em um processo indutivo, identificar as facetas a serem adotadas;
- c) desenvolver o conteúdo das facetas.

Observação final: não é boa estratégia partir do zero. Assim, a análise de outras tabelas pode fornecer elementos que, à luz dos critérios que venham a ser adotados, sejam aproveitados, devidamente compatibilizados.



REFERÊNCIAS

ALEXANDER, J. ; LAMBE, P. *Concept Paper. Developing organizational and statistical taxonomies for the Classification of Research and Development Activities: a best practice study and guide*. Arlington, V.A, SRI International, June 15, 2015.

GOMES, H.E. *Estudo CASRAI – Compatibilidade de sistemas*. Produto 1 do projeto: Dos padrões internacionais de estruturação da informação de pesquisa aos indicadores (BRCRIS). Rio de Janeiro, 32 p. 2017. IBICT - Iniciativa de Apoio aos Diálogos Setoriais União Europeia/Brasil.

Sobre o estudo de comparação entre a Tabela do CNPq e a Tabela CASRAI

Hagar Espanha Gomes

O estudo teve por objetivo verificar o grau de compatibilidade entre as tabelas do CNPq e da CASRAI.¹

A análise levou em consideração aspectos comuns não apenas no que se refere à estrutura, de modo geral, mas também a consistência na organização dos itens no interior das classes.

Tendo em vista que a tabela do CNPq está organizada em apenas uma dimensão – áreas do conhecimento – que a tabela CASRAI tem três dimensões: disciplinas, objetos de pesquisa e campos de aplicação, apenas a primeira delas foi considerada, por reunir tópicos de natureza semelhante à tabela do CNPq.



¹Extraído de: GOMES, Hagar Espanha. *Estudo CASRAI – Estudo com tabelas CNPq, FoS e ISCED*. Produto 2 do projeto: Dos padrões internacionais de estruturação da informação de pesquisa aos indicadores (BRCRIS). Rio de Janeiro, 32 p. 2017. IBICT - Iniciativa de Apoio aos Diálogos Setoriais União Europeia/Brasil.

TABELA DE ÁREAS DO CONHECIMENTO (CNPq)

Não foi possível obter documentação sobre o planejamento e a realização da Tabela. Dado o nível de detalhamento dos níveis inferiores, ela não se caracteriza como um esquema para classificações estatísticas, que requerem maior grau de agregação dos tópicos. De fato, tem sido usada pelas instituições de planejamento e pesquisa para orientar e avaliar grupos de pesquisa e não para planejamento nacional.

A tabela do CNPq está organizada em quatro níveis – grande área, área, subárea e especialidade. Em relação à última não existe clareza sobre o que seria ‘especialidade’. No primeiro nível, o conhecimento está dividido em 6 grandes áreas: a) Exatas e da Terra; b) Ciências Biológicas; c) Engenharias; d) Ciências da Saúde; e) Ciências Agrárias; f) Ciências Sociais Aplicadas; Ciências Humanas.

No segundo nível os tópicos se aproximam de disciplinas tradicionais com alto nível de desagregação, como se pode observar na quantidade de áreas presentes em cada grande área: Exatas e da Terra- 8 áreas; Ciências Biológicas – 11 áreas; Engenharia – 14 áreas; Ciências da Saúde – 8 áreas; Ciências Agrárias – 7 áreas; Ciências Sociais Aplicadas – 13 áreas; Linguística, Letras e Artes – 3 áreas.

Os demais níveis apresentam um tratamento desigual: nem todos os tópicos são desenvolvidos nos níveis inferiores.

A tabela não tem notação: os itens são distribuídos sem uma ordem lógica.

Dois pontos devem ser ressaltados em relação a essa análise: a ausência de princípios para a estruturação das áreas de conhecimento e a definição ou mesmo breve descrição dos termos. Princípios lógicos produzem taxonomias nas quais os elementos são mutuamente exclusivos, contribuindo para a formação de grupos consistentes no interior daqueles níveis. Fenômenos ou categorias podem ser uma base de classificação conforme a natureza do conhecimento. Isto daria consistência à Tabela bem como condições de atualização permanente, exigindo, provavelmente, alteração mínima na estrutura/notação.

Tendo em vista a dificuldade de acessar estudos realizados para a atualização da Tabela, não existem dados que justifiquem os critérios adotados na elaboração das áreas, subáreas e especialidade. A organização lógica no interior das áreas e, conseqüentemente, nas subáreas e especialidades (hierarquização dos conjuntos e subconjuntos) em geral não é observado se os princípios adotados deveriam estar presentes, pois facilitariam iniciativas de harmonização, compatibilização ou outro tipo de procedimento, e também algum processamento automático entre tabelas de outras organizações.

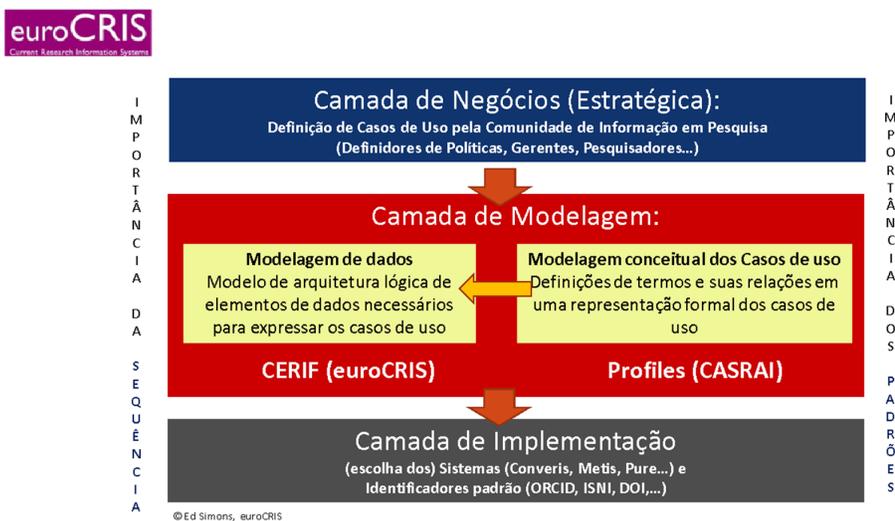
Como já referido, a análise aqui realizada a respeito da Tabela objetivou a derivação de indicadores. A partir desta perspectiva, verifica-se certo desequilíbrio na estruturação da Tabela. Ela registra 853 (oitocentos e cinquenta e três) especialidades em 93 (noventa e três) subáreas. Dentre estas, várias não registram especialidade, ressaltando as Ciências Biológicas e Linguística e Letras. Existe desequilíbrio na organização dos níveis.



TABELA DE CLASSIFICATION DE LA RECHERCHE (TABELA CASRAI)

Esta tabela se insere na modelagem conceitual do projeto euroCRIS proposta por Ed Simons, presidente da euroCRIS (Figura 1), no que se refere a vocabulários, tesouros e demais instrumentos semânticos para tratamento e recuperação de informação, para uso na gestão de pesquisa. A estreita relação com CERIF guarda paralelo com metadados e dados, ou seja, enquanto CERIF cuida da semântica de metadados, CASRAI cuida das ferramentas semânticas para representar conteúdos de documentos (projetos, relatórios e assemelhados) em seus respectivos metadados.

Figura 1 – Implementação da informação de pesquisa: modelo em três camadas



Fonte: SIMONS, E. *Research Information Implementation: 3-layer model*. Jornada EuroCRIS, EuroCRIS/IBICT, Brasília, 16 a 18 de novembro e 2016.

CASRAI (Consortia for Advancing Standards in Research) é uma organização sem fins lucrativos, e atua como padrão para registrar os assuntos dos documentos.

As áreas cobertas pela Tabela são: Ciências Naturais e Engenharias; Ciências da Saúde; Ciências Humanas e Sociais; Artes e Literatura. O conteúdo do segundo nível é bastante agregado, como convém a uma classificação para estatísticas, concentrado em 13 campos disciplinares.

O terceiro nível – disciplinas, é bastante detalhado e requer alguma ordenação para torná-lo sistemático. Por exemplo, ao analisar a lista das disciplinas do campo disciplinar Ciências Médicas, observa-se que ela não apresenta uma ordenação sistemática, ou seja, não fica claro o princípio de divisão adotado: por vezes, disciplinas, por vezes, aspectos. A lista contém 40 disciplinas médicas em ordem alfabética. Uma tentativa de reuni-las segundo aspectos comuns mostraria a justeza ou não da inclusão deste ou daquele item e produziria conjuntos consistentes. Se fosse o caso de atribuir uma notação, ela possibilitaria agregação no momento da obtenção de indicadores e, além disso, sua ordem não se alteraria nas traduções em outras línguas. E facilitaria estudo para conversão de tabelas.

Os comentários detalhados até aqui podem ser generalizados, em seus aspectos teóricos, para os demais setores, a saber: ausência de princípio de divisão, ausência de princípio de ordenação. Ausência de notação expressiva.



COMPARAÇÃO DAS TABELAS

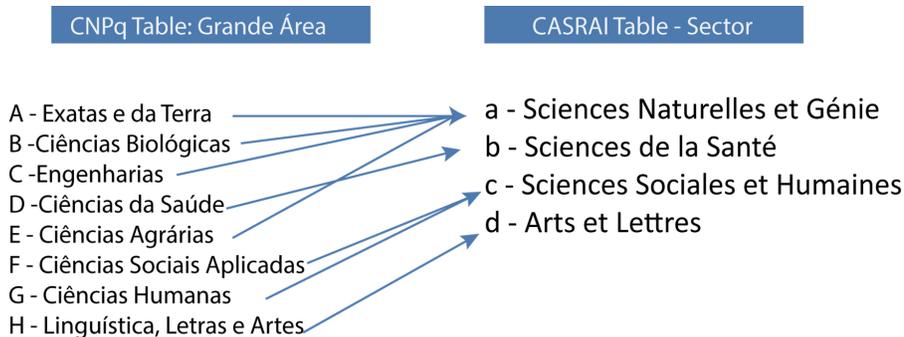
Ambas as classificações estão estruturadas nas disciplinas tradicionais, o que facilita a comparação. O estudo de comparação se dará em três níveis, a saber:

- comparação entre Grande Área do CNPq e Secteur da CASRAI;
- comparação Área do CNPq e Champ disciplinaire da CASRAI;
- comparação Área (CNPq) – Discipline (CASRAI).

Os demais níveis apresentam grande dispersão e, sem definição ou explicação, sobre a abrangência de conteúdo de cada item não há segurança sobre o resultado da análise. As demais dimensões da tabela CASRAI não têm correspondência com a tabela CNPq. Esta se limita a áreas científicas enquanto as demais dimensões da CASRAI tratam de assuntos e setores de aplicação. Este fato não prejudica a análise, porque não existem elementos comparáveis. O estudo de comparação se dará em três níveis, a saber:

- comparação entre Grande Área do CNPq e Secteur da CASRAI;
- comparação Área do CNPq e Champ disciplinaire da CASRAI;
- comparação Área (CNPq) – Discipline (CASRAI).

Figura 2 – COMPARAÇÃO GRANDE ÁREA (CNPq) – SECTEUR (CASRAI)



Neste nível fica evidente o nível de agregação entre as duas classificações.

COMPARAÇÃO ÁREA (CNPQ) - CHAMP DISCIPLINAIRE (CASRAI)

A diferença entre o segundo nível das duas tabelas é bastante grande. Na tabela do CNPq o detalhamento é alto, enquanto na tabela CASRAI ainda é bastante agregado. Assim, em vários momentos, foi necessário buscar a comparação no terceiro nível - disciplinas, da tabela CASRAI. Deste modo o grau de compatibilidade foi grande (cerca de 80%). Mas a estrutura dessas duas tabelas é bastante diferente, e isso poderia prejudicar estatísticas futuras ao reunir dados de ambas.



COMENTÁRIO GERAL

Apesar da diferença na organização das tabelas, quando se procedeu à comparação dos conceitos, independentemente dos níveis em que se encontrassem nas tabelas, foi alto o grau de compatibilidade. Mas para classificações estatísticas isto não é suficiente: as tabelas deveriam ter conteúdo mais agregado, se destinadas a produzir indicadores.

A tabela do CNPq, neste nível, apresenta discrepâncias em sua organização lógica: tópicos de Farmácia estão presentes tanto em Ciências Biológicas como em Ciências da Saúde. A ausência de definição poderia mostrar a justeza ou não desta classificação. Os conteúdos de Ciências Sociais Aplicadas e Ciências Humanas apresentam sério problema de classificação, a começar pelo nome 'Ciências Sociais Aplicadas', que exclui Sociologia, classificando-a em Ciências Humanas; e classificando Design em Ciências Sociais Aplicadas.

Em relação à CASRAI, algum ordenamento lógico seria necessário, em especial nas classes com dezenas de itens.

Ajustes nas classificações e no desenvolvimento de notação expressiva dos agrupamentos, a exemplo da tabela Fields of Science (FoS) da OCDE, tornariam ambas as tabelas consistentes, porque iriam exigir organização lógica nos agrupamentos, em vez de ordenação alfabética.

A definição é um requisito essencial para a ordenação lógica dos elementos. A definição pode estar ausente em uma tabela, mas se foi considerada quando de sua elaboração e/ou atualização, ela estaria minimamente presente nas taxonomias resultantes; isto não fica evidente na estrutura de ambas as tabelas, pois nem sempre se percebe uma relação hierárquica. Esta 'definição mínima' – do tipo é-um – poderia fornecer, por exemplo – na tabela CNPq – argumentos para a presença ou não de Farmacologia e Farmácia em diferentes áreas. Ou de Sociologia na área de humanidades. A definição é responsável não apenas na ordenação dos elementos na estrutura das classes, mas fornece elementos para uma notação expressiva que revelaria seu conteúdo e nível de agregação; facilitaria ainda a produção de indicadores.

Na ausência de definições, a clareza do termo na comunicação² não deve deixar dúvida sobre o conteúdo que se espera dele.

De fato, qualquer decisão sobre a necessidade de adotar também a tabela CASRAI no sistema de informação CNPq/CAPES irá demandar a construção de um instrumento de compatibilização, de uma metalinguagem, como plataforma de compatibilização dos diversos esquemas de uma rede como euroCRIS. Esta metalinguagem levaria em consideração outros esquemas de classificação para estatística e teria como vantagem garantir que cada centro de informação continuasse a manter os próprios esquemas.



²GLÄNZEL, Wolfgang et al. Data collection and use in research funding and performing organisations. General outlines and first results of a project launched by Science Europe. *Scientometrics*, v. 106, n. 2, p. 825-835, 2016.

REFERÊNCIAS

GLÄNZEL, W. et al. Data collection and use in research funding and performing organisations. General outlines and first results of a project launched by Science Europe. *Scientometrics*, v. 106, n. 2, p. 825-835, 2016.

GOMES, H.E. *Estudo CASRAI – Estudo com tabelas CNPq, FoS e ISCED*. Produto 2 do projeto: Dos padrões internacionais de estruturação da informação de pesquisa aos indicadores (BRCRIS). Rio de Janeiro, 32 p. 2017. IBICT - Iniciativa de Apoio aos Diálogos Setoriais União Europeia/Brasil.

SIMONS, E. *Research Information Implementation: 3-layer model*. Jornada EuroCRIS, EuroCRIS/IBICT, Brasília, 16 a 18 de novembro de 2016.

Thoughts on the Classification of University Research (CASRAI Table)

Hagar Espanha Gomes
Tradução: Michael Allan

On the occasion of the appointment of the CASRAI Executive Committee for Europe, which took place in Malta in 2017, it was suggested that a study should be undertaken of the use of the CASRAI table in research information services in Brazil, given the country's initiative to participate in euroCRIS. The results of such an analysis should be considered in the studies of compatibility and harmonization which may be done with a view to achieving that objective. The following comments follow from the meeting held recently in Malta.

The purposes of a classification table explain, to a certain extent, the decisions taken during its elaboration. When these classifications are aimed at producing indicators, the first step is to find out what types of indicators one hopes to obtain.

The comments below address the structure of the table, the analysis of its dimensions and proposals for its redrafting¹.

In general, such indicators are the result of the combination of more than one aspect in order to allow for various analyses, and this is reflected in the classification structure. In this way, specificity is achieved, even when starting off from general classes. A certain level of generality should be a characteristic of this kind of system because, amongst other reasons, it is easier at this level to arrive at consensus and general acceptance.

Analysis of the Classification of University Research – henceforth the CASRAI table – takes into account the level of generalization and the possibilities for arranging and combining its various dimensions, which render it more dynamic. The comments and suggestions offered below are aimed at finding ways to work around some deficiencies and inadequacies of the Table, and not at replacing it.



¹ Extracted from: GOMES, H.E. *Estudo CASRAI – Compatibilidade de sistemas*. Produto 1 do projeto: Dos padrões internacionais de estruturação da informação de pesquisa aos indicadores (BRCRIS). Rio de Janeiro, 32 p. 2017. IBICT - Iniciativa de Apoio aos Diálogos Setoriais União Europeia/Brasil.

STRUCTURE OF THE TABLE

The CASRAI Table presents three dimensions: disciplines, research objects and application fields.

It is important to keep in mind from the start that classification tables designed to produce indicators contain a high level of generalization. This is the case with the CASRAI Table. To be analyzed are the principles adopted in the organization of its elements, at both the first (sector) and second (disciplinary field) levels.

The CASRAI table presents three dimensions: disciplines, research objects and application fields.

It is important to keep in mind from the start that classification tables designed to produce indicators contain a high level of generalization. This is the case with the CASRAI table. To be analyzed are the principles adopted in the organization of its elements, at both the first (sector) and second (disciplinary field) levels.

ANALYSIS OF THE FIRST DIMENSION: DISCIPLINES

The Discipline dimension is organized by sector/disciplines:

Natural Sciences and Engineering

- Engineering
- Applied Sciences
- Natural Sciences
- Mathematical Sciences

Health Sciences

- Medical Sciences
- Nursing
- Physical Education and Rehabilitation

Social Sciences and Humanities

- Education
- Management
- Humanities
- Social Sciences

Arts and Literature

- Writing and Fine Arts
- Arts and Literature Studies

The first level of this dimension – sector- differs little from other classifications oriented towards Science and Technology, which generally adopt Disciplines as the primary principle of the division of knowledge.

Two “disciplines” deserve attention: Applied Sciences and Nursing. In the first of these, the adopted principle of division is not clear, and in the second, only one item is presented, which does not justify this subdivision, especially when ‘Nursing’ could be subsumed into the Medical Sciences discipline, a canonical classification.

With regard to analysis of the content of each discipline, it is also not clear as to the principle of division adopted: at times disciplines, and at others, aspects. An internal ordering undertaken according to some aspects or facets would provide evidence of some structure for the whole. This would contribute to the creation of a notation; in this case, no matter what language adopted, the ordering would be maintained and this would allow for the aggregation of data when generating indicators.

When analyzing Disciplinary Field, it can be observed that here as well there is no systematic ordering, that is, a connecting of elements based on shared aspects. Medical Sciences, for example, lists 40 elements in alphabetical order. An attempt to group these based on common aspects would demonstrate the rightness or not of including this or that item and would produce consistent groupings. If it were possible to assign a notation, this would facilitate a greater or smaller aggregation at the moment of generating indicators.

One should not lose sight of the possibility of the integration of data or of sharing it with other organizations. In terms of the compatibility of the CASRAI table with the OECD’s FoS (Fields of Science and Technology) table, it can be observed that there is no problem, with the exception of the Applied Sciences ‘discipline’, which combines very diverse elements. In fact, it is suggested that this ‘discipline’ merits a detailed analysis and a possible redistribution of its items to other classes.

ANALYSIS OF THE REMAINING DIMENSIONS

In relation to the remaining dimensions – Object and Field of Application – there exist suggestions in Patrick Lambe’s Concept Paper which could improve the CASRAI table. Within this paper, some considerations deserve to be underlined in regard to the dimensions which may be adopted, that is, which aspects or facets may be contemplated.

Generally, to avoid placing an undue cognitive burden on users, it is advisable to select only three facets [...] The selection and design of the facets in a classification system should be done with particular care, to ensure that the facets are both relevant to organizational needs and a useful representation of the activities.²

This observation implies a review of these dimensions. Starting from the decision on the selected aspects, the CASRAI table can be analyzed, and decisions made in regard to which dimensions should be maintained or discarded, and if they should be kept or replaced by facets chosen after discussion of the problem. In the Preamble, the contents of the Objects are meant to be phenomena, but an analysis of the items reveals inconsistencies. The faceting proposal made above would mean that the Table would keep its dynamic character, with its elements, however, grouped according to a certain principle or aspect.

With regard to the Field of Application, perhaps the classification adopted serves the economic sector of the EU members. In this case, also, the classification principles used for the content of each Field should be the same, namely, organized by facets/aspects.

A general observation: given that the various dimensions of the table contain a high level of generality, it is indispensable that each class be accompanied by information covering the specific content of each item, and indicating the scope of each. A lack of this information makes more work for the user. As a positive example, the FoS table may be cited, which includes basic information about what each class addresses; this saves effort on the part of the researcher at the moment of indexing his or her document in the database.



²ALEXANDER, J. ; LAMBE, P. Concept Paper. *Developing organizational and statistical taxonomies for the Classification of Research and Development Activities: a best practice study and guide*. Arlington, V.A, SRI International, June 15, 2015. p. 6.

PROPOSALS FOR REFORMULATING CASRAI: FOR DISCUSSION

To begin with, some thoughts on several classification principles which could be adopted in the reformulation of the table.

Facets or aspects contain a level of generality: things, instruments, techniques, processes and so on; they are constituted of elements of a specific nature which comprise taxonomies. As well as contributing to the construction of Classes and subclasses, they facilitate integration and compatibility with other databases.

Many R&D activities combine or merge aspects of multiple existing fields and disciplines. Thus, interdisciplinary R&D is especially challenging, since many taxonomies are designed to apply one and only one label to an activity. Despite these difficulties, developing consistent and interpretable R&D taxonomies will become even more important in the future.³

How can taxonomies and Classes be harmonized? Taxonomies do not belong, as such (with, for example, their hierarchical trees) in a table of the CASRAI type. Taxonomies are useful for indexing projects, reports and the like, which require specific representation of content and may be offered as a standard for those types of document. The elaboration of a taxonomy is the result of the inductive method. Each element is defined according to its essence – process, properties, object – and, following each step of the division, the most general term, the facet, is arrived at.

Cutting through disciplines has practical objectives and is the result of a deductive process; perhaps for this very reason it may be more difficult to arrive at a consensus when aiming for greater detail. The greater level of generality might contribute to agreement between members of a network. The reformulation of the Table requires further discussion about the degrees of detail which come to be adopted. In other words, which criteria to adopt in organizing the elements in each of these levels?

³ Ibidem, p. 3.

In the process of structuring classes and subclasses, taxonomies offer firm orientation in defining the limits of the scope of each, facilitating its consultation and use, as exemplified in the FoS table. Thus, they can serve a dual purpose: they can contribute to the internal organization of Classes and subclasses and can also function – as a separate instrument – in providing a controlled vocabulary to be used in indexing.

The greatest quality of the table may reside in its multi-dimensional structure, which should not be abandoned: as far as it was possible to check, this characteristic appears to be unique in Europe and is much appreciated by the classificationists, especially those proponents of the facet method.



SUGGESTED REVISION ACTIVITIES FOR THE CASRAI TABLE

- a) Analyze the first division level – Disciplines – generally adopted as the first dimension.
- b) Survey the types of indicators expected, in order to, by an inductive process, identify the facets to be adopted.
- c) Develop the content of the facets.

Final note: It is not a good strategy to start from zero. Analysis of other tables may supply elements which, in light of the criteria which come to be adopted, may be of use, once fittingly made compatible.



REFERENCES

ALEXANDER, J. ; LAMBE, P. Concept Paper. *Developing organizational and statistical taxonomies for the Classification of Research and Development Activities: a best practice study and guide*. Arlington, V.A, SRI International, June 15, 2015.

GOMES, H.E. *Estudo CASRAI – Compatibilidade de sistemas*. Produto 1 do projeto: Dos padrões internacionais de estruturação da informação de pesquisa aos indicadores (BRCRIS). Rio de Janeiro, 32 p. 2017. IBICT - Iniciativa de Apoio aos Diálogos Setoriais União Europeia/Brasil.

Thoughts on the comparative study of the CNPq and CASRAI tables

Hagar Espanha Gomes

Tradução: Michael Allan

The objective of the study was to assess the level of compatibility between the CNPq and CASRAI tables.¹

This analysis took into account shared features, not only with respect to structure in general, but also consistency in the organization of the items within each class.

As the CNPq table is organized in only one dimension – areas of knowledge – while the CASRAI table has three dimensions: Disciplines, Research Objects and Application Fields, only the first of these was considered, because it contains topics similar to those in the CNPq table.



¹ Extracted from: GOMES, H.E. *Estudo CASRAI – Estudo com tabelas CNPq, FoS e ISCED*. Produto 2 do projeto: Dos padrões internacionais de estruturação da informação de pesquisa aos indicadores (BRCRIS). Rio de Janeiro, 32 p. 2017. IBICT - Iniciativa de Apoio aos Diálogos Setoriais União Europeia/Brasil.

TABLE OF AREAS OF KNOWLEDGE (CNPq)

It was not possible to find documentation of the planning and creation of the table. Given the level of detail in its lower levels, this table does not present itself as a tool for statistical classification, which requires a higher level of topic aggregation. It has been used, in fact, by planning and research institutions to orientate and evaluate research groups and not for planning at the national level.

The CNPq table is organized in four levels – Wide Area, Area, Sub-Area and Specialty. At the last of these, there is little clarification of what “specialty” might entail. At the first level, knowledge is divided into six large areas: A. Exact and Earth Sciences; B. Biological Sciences; C. Engineering; D. Health Sciences; E. Agricultural Sciences; F. Applied Social Sciences; Human Sciences.

At the second Level, the topics come close to traditional disciplines, with a high degree of disaggregation, as can be seen in the number of areas present in each Wide Area: Exact and Earth Sciences – 8 areas; Biological Sciences – 11 areas; Engineering – 14 areas; Health Sciences – 8 areas; Agricultural Sciences – 7 areas; Applied Social Sciences – 13 areas; Human Sciences – 3 areas.

The remaining levels present an unequal treatment of topics, in that not all are developed in the lower levels.

The table lacks notation: items are not distributed in logical order.

Two points should be emphasized about this analysis: the absence of principles for structuring the knowledge areas, and the definition, or at least a short description of terms. Logical principles produce taxonomies within which the elements are mutually exclusive, thus contributing to the formation of consistent groups inside those levels. Phenomena or categories can form the basis for classification according to the type of knowledge. This would endow the table with consistency, as well with the conditions for ongoing updates requiring, probably, minimal changes in structure and notation.

Given the difficulty in accessing studies undertaken to update the table, there is no data which justifies the criteria adopted in the elaboration of the Areas, Subareas and Specialties. Logical organization is lacking within the Areas and, consequently, in the Subareas and Specialties (ranking of sets and subsets). In general, the principles adopted, which should be present, as they facilitate processes such as harmonization, compatibility, as well as automatic processing between other organizations' tables, are not followed.

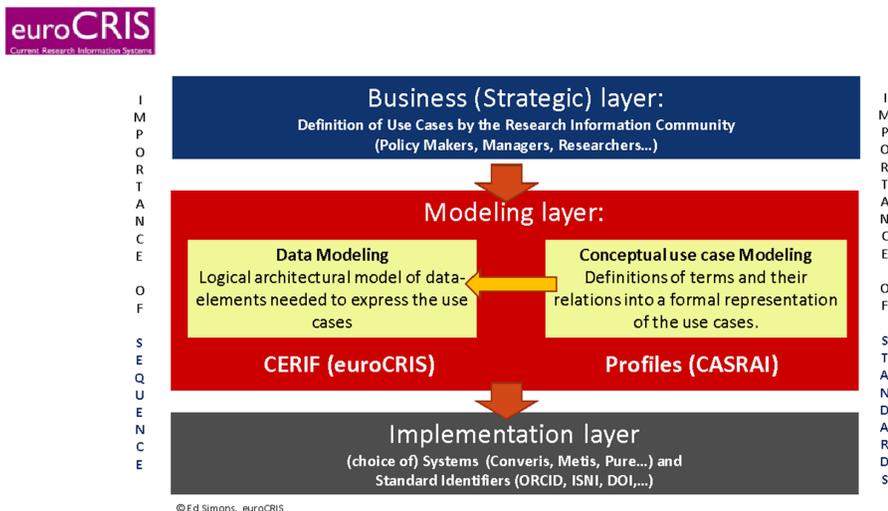
As already mentioned, this analysis of the CNPq table is focused on the derivation of indicators. From this perspective, a certain imbalance in the structure of the table is noted. It lists 853 Specialties in 93 Subareas. Amongst these, various do not list Specialties, particularly Biological Sciences and Linguistics and Literature. There is an imbalance in the organization of the levels.



THE CASRAI RESEARCH CLASSIFICATION TABLE

This table falls within the scope of the conceptual modeling proposed by Ed Simons, euroCRIS President (Figure 1), in which reference is made to vocabularies, thesauruses and other semantic instruments used for information processing and recovery in research management. The relationship with CERIF takes place in the semantics of metadata while CASRAI deals with semantic tools to represent the content of documents (projects, reports and the like) in their respective metadata.

Figure 1 – Research information implementation: 3-layer model



Source: SIMONS, E. *Research Information Implementation: 3-layer model*. Jornada EuroCRIS, EuroCRIS/IBICT, Brasília, 16 a 18 de novembro e 2016.

CASRAI (Consortia for Advancing Standards in Research) is a non-profit organization which sets the standard for recording document topics.

The areas covered by the table are: Natural Sciences and Engineering; Health Sciences; Human and Social Sciences; Art and Literature. The content of the second level is quite aggregated as befits a classification for statistics, concentrated in 13 disciplinary fields.

The third level – Disciplines – is quite detailed and needs some ordering to make it systematic. As an example, analysis of the list of disciplines contained in its Medical Sciences field produces the observation that the table does not offer systematic ordering. It is not clear which principle of division has been adopted: sometimes disciplines, sometimes aspects. The list contains 40 medical disciplines in alphabetical order. An attempt to connect them according to shared aspects would show the rightness or not of including this or that item and would produce consistent groupings. If it were the case to assign a notation, this would make possible aggregation at the moment of obtaining indicators and, moreover, the order would not change in foreign translations. It would also facilitate studies aimed at conversion of tables.

The comments presented to this point (lack of principles of division and of ordering ; absence of expressive notation) may be generalized, in their theoretical aspects, to the other sectors.



COMPARISON OF THE TABLES

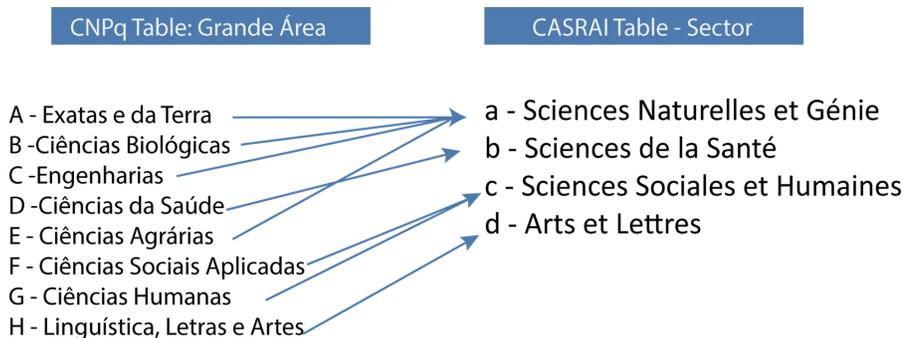
Both classifications are structured within the traditional disciplines, which facilitates comparison.

The comparison comprises three levels:

- Comparison between the CNPq Grande Area and CASRAI sector.
- Comparison between CNPq Area and CASRAI disciplinary field
- Comparison of Area (CNPq) and Discipline (CASRAI)

The remaining levels are spread wide, without definition or explanation, and, in regard to the scope of each item, there can be no confidence in the results of the analysis. The other dimensions of the CASRAI table do not correspond to the CNPq table, which limits itself to scientific areas while the other dimensions of the CASRAI table deal with application topics and sectors. This fact does not affect the analysis, because there are no comparable elements.

Figure 2 – Comparison grande área (cnpq) – sector (casrai)



At this level the degree of aggregation between the two classifications is evident.

COMPARISON BETWEEN CNPQ AREA AND CASRAI DISCIPLINARY FIELD

The differences between the second level of both tables are quite substantial. The CNPq table shows a high degree of detail, while the CASRAI table is already quite aggregated. For this reason, it was sometimes necessary to consult the third level – Disciplines – of the CASRAI table in order to make the comparison. This approach revealed a high degree of compatibility (around 80%). The structures of both tables, however, are considerably different and in the future this could hamper the generation of statistics when uniting data from both tables.



GENERAL COMMENTS

In spite of their differences in organization, the comparison of the tables' concepts, independently of the level in which they are found, revealed high degree of compatibility. But for the purposes of statistical classification this is not adequate: the tables must have more aggregated content if they are intended to produce indicators.

The CNPq table, at this level, presents discrepancies in its logical organization: Pharmaceutical topics are present in Biological Sciences as well as in Health Sciences. The absence of definition might reveal the rightness or not of this classification. The contents of Applied Social Sciences and Human Sciences present serious classification problems, starting with the name "Applied Social Sciences", which excludes Sociology, classifying it in Human Sciences, and classifying Design in Applied Social Sciences.

The CASRAI table requires logical ordering, especially in the classes with dozens of items.

Adjustments to the classifications and in the development of expressive notation of groupings, as exemplified in the Fields of Science (FoS) table of the OECD, would make both tables consistent, as they would require a logical organization of the groupings, in place of alphabetical order.

Definition is an essential prerequisite for the logical ordering of elements. The definition may appear absent from a table, but if it was taken into consideration during the drafting and updating of the table, it will be minimally present in the resulting taxonomies; this is not evident in in the structure of either table, as any kind of hierarchical relationship is not always perceptible. A minimal definition of the "is_a" type might provide, in the case of the CNPq table for example, arguments for the placement, or not, of Pharmacology and Pharmacy in different areas, or of Sociology in the Humanities area. Definition is responsible, not only in ordering elements in the structuring of classes, but it also provides for an expressive notation which would reveal its content and level of aggregation. It would also assist

in the production of indicators. Without definitions, the clarity of the term when communicating² should not leave any doubt about the content which is expected of it.

Any decision made about the necessity of adopting the CASRAI table within the CNPq/CAPES information system will require the creation of a compatibilization instrument, of a metalanguage, as a platform to render compatible the varied systems of a network such as euroCRIS. Such a metalanguage would take into account other classification schemes for statistics, and would offer the advantage that each information centre would continue to maintain its own scheme.



² GLÄNZEL, W. et al. Data collection and use in research funding and performing organisations. General outlines and first results of a project launched by Science Europe. *Scientometrics*, v. 106, n. 2, p. 825-835, 2016.

REFERENCES

GLÄNZEL, W. et al. Data collection and use in research funding and performing organisations. General outlines and first results of a project launched by Science Europe. *Scientometrics*, v. 106, n. 2, p. 825-835, 2016.

GOMES, H.E. *Estudo CASRAI – Estudo com tabelas CNPq, FoS e ISCED*. Produto 2 do projeto: Dos padrões internacionais de estruturação da informação de pesquisa aos indicadores (BRCRIS). Rio de Janeiro, 32 p. 2017. IBICT - Iniciativa de Apoio aos Diálogos Setoriais União Europeia/Brasil.

Parte III

INSTRUMENTOS DE PESQUISA





Bibliografia anotada: identificadores persistentes (Identificadores ORCID e RINGGOLD)

Juliana Pinheiro Farias
Kelen Cândida Vieira Bomfim

RESUMO

Identificadores persistentes são importantes para a vinculação dos pesquisadores e instituições às suas respectivas produções científicas. O identificador persistente de pessoas Open Researcher and Contributor ID (ORCID) e o identificador persistente de instituições (RINGGOLD) possibilitam a eliminação de ambiguidades de informações em registros, facilitando a recuperação e acesso aos dados de pesquisadores e instituições. O Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) tem desenvolvido estudos a respeito desses identificadores no âmbito do projeto BRCRIS, como complemento ao uso de padrões. A bibliografia anotada é antecedida de revisão de literatura.

Palavras-chave: Bibliografia anotada. Identificadores persistentes. Consórcio ORCID.

Annotated Bibliography: Persistent Identifiers (ORCID and RINGGOLD Identifiers)

ABSTRACT

Persistent identifiers are important for the linking of researchers and institutions with their respective scientific products. Open Researcher and Contributor ID (ORCID), the persistent identifier for individuals, and RINGGOLD, the persistent identifier for institutions, make it possible to eliminate information ambiguities in records, easing recovery and access to data from researchers and institutions. The Brazilian Institute for Information in Science and Technology (IBICT) has developed studies of these identifiers within the context of the BRCRIS project, as a compliment to the use of standards. The annotated bibliography is preceded by a literature review.

Keywords: *Annotated bibliography. Persistent identifiers. ORCID Consortium.*

Bibliografía anotada: identificadores persistentes (Identificadores ORCID y RINGGOLD)

RESUMEN

Identificadores persistentes son importantes para la vinculación de los investigadores e instituciones a sus respectivas producciones científicas. El identificador persistente de personas Open Researcher y el ID de identificación (ORCID) y el identificador persistente de instituciones (RINGGOLD) posibilitan la eliminación de ambigüedades de información en registros, facilitando la recuperación y acceso a los datos de investigadores e instituciones. El Instituto Brasileño de Información en Ciencia y Tecnología (IBICT) ha desarrollado estudios acerca de estos identificadores en el marco del proyecto BRCRIS, como complemento al uso de estándares. La bibliografía anotada es precedida de revisión de literatura.

Palabras clave: *Bibliografía anotada. Identificadores persistentes. Consorcio ORCID.*

INTRODUÇÃO

O Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) deu início, em 2014, ao desenvolvimento de estudos e pesquisas para a implantação, no país, de um sistema de informação de pesquisa, conhecido pela sigla CRIS (Current Research Information System). Na sequência, na segunda metade de 2017, o IBICT teve aprovado um projeto submetido aos Diálogos Setoriais para desenvolvimento de consultorias para aprofundamento de aspectos relacionados aos padrões internacionais para estruturação de informação de pesquisa em suas várias vertentes. Os identificadores ORCID (Open Researcher and Contributor ID) e RINGGOLD são complementares aos padrões estudados. Eles integram a categoria de identificadores persistentes.

Um identificador persistente é uma referência de longa duração a um objeto digital que fornece informação sobre esse objeto, independentemente do que acontece com ele. Desenvolvido para tratar de 'links quebrados', um identificador persistente pode ser definido para fornecer uma representação apropriada de um objeto, independentemente dos objetos mudarem sua localização on-line ou ficarem off-line.¹

Na história da estruturação de registros bibliográficos, a experiência dos identificadores persistentes é precedida por árduo trabalho manual, complementar às atividades de catalogação do acervo de uma biblioteca, conhecido como controle de autoridade para os pontos de acesso, seja dos antigos catálogos em livros e em fichas ² (datam do século XIX), seja, mais recentemente, dos modernos catálogos digitais³.

¹ CASRAI. *Dictionary*. Disponível em: <http://dictionary.casrai.org/Persistent_identifier>. Acesso em: 08 maio de 2018.

² ASSUMPÇÃO, F.S.; et. al. O controle de autoridade no domínio bibliográfico: os catálogos em livros e em fichas. *Biblios*, n. 67, p. 84-98, 2017. Disponível em: <http://www.scielo.org.pe/scielo.php?pid=S1562-47302017000200007&script=sci_abstract>. Acesso em: 08 maio 2018.

³ SILVA, A.F.; SANTOS, P.L.V.A.C. A importância do controle de autoridade: uma abordagem baseada nos objetivos e nas funções dos catálogos. In: *I Encontro Nacional de Catalogadores, III Encontro de Estudos e Pesquisas em Catalogação*, 2012, Rio de Janeiro. Anais... Rio de Janeiro: Biblioteca Nacional, 2012. Disponível em: <<http://eprints.rclis.org/18843/1/assumpcao-santos-a-importancia-do-controle-de-autoridade.pdf>>. Acesso em: 01 jul. 2013

[...] o controle de autoridade possibilita a existência de uma ferramenta para acesso à informação (o catálogo) em que o usuário consiga navegar e encontrar um único recurso informacional ou todos os recursos associados a uma determinada entidade **sem a necessidade de ele conhecer todos os nomes pelos quais tal entidade é conhecida** (Ênfase acrescentada)⁴.

O conceito de controle de autoridade deriva da tradição da Biblioteconomia, em sua vertente de catalogação. E permanece atual, em sintonia com as mudanças tecnológicas requeridas pelas recomendações do W3C (World Wide Web Consortium), resultando "(...) na criação de uma infraestrutura para o intercâmbio, o compartilhamento e a vinculação de dados de autoridade na Web de dados, além da construção de vocabulários de valor de alto nível"⁵.

Essa infraestrutura decorre da capitalização do trabalho de controle de autoridade empreendido pelas grandes bibliotecas, como a British Library, a Library of Congress, o Consórcio OCLC, cujo resultado é a criação de um novo consórcio envolvendo bibliotecas em diversos lugares do mundo para integrar seus catálogos de autoridades, agregando-se vocabulários de valor em arquivos de autoridade, devidamente publicados em Linked Open Data, um requisito da Web Semântica. Esse consórcio é o VIAF (Virtual International Authority File) "(...) uma iniciativa democrática de cooperação em nível internacional e que pode ser utilizado tanto como uma fonte confiável de arquivos de autoridade para instituições bibliotecárias, quanto para as comunidades da Web de dados"⁶.

O desenvolvimento do VIAF permitiu o escalonamento da norma ISNI (International Standard Name Identifier- ISO 27729), a qual permite identificação dos milhões de contribuidores do trabalho criativo.

⁴Ibid., p.12.

⁵ROMANETTO, M.L. et al. O Virtual International Authority File-VIAF e a agregação de valores por metadados de autoridade. *RDBCi: Revista Digital de Biblioteconomia e Ciência da Informação*, v. 15, n. 3, p. 571-590. Disponível em: <<https://periodicos.sbu.unicamp.br/ojs/index.php/rdbci/article/view/8647488/pdf>>. Acesso em: 08 maio 2018.

⁶Ibid., p.572.

O objetivo ambicioso desse novo padrão está sendo construído diretamente no legado funcional dos arquivos de autoridade de biblioteca, conforme representado nos bancos de dados contribuintes que compõem o Virtual International Authority File hospedado pela OCLC.” [...] “Embora o ISNI comece com fortes raízes nos arquivos de autoridade da biblioteca, seu escopo e seu contexto para desenvolvimento são verdadeiramente globais e multi-domínio. O ISNI visa abordar tantas comunidades quanto possível, na medida em que os objetos de identificação são as identidades públicas atrás do identificador (MACEWAN; ANGJELI; GATENBY, 2013).⁷

O ISNI faz parte de uma família de identificadores de padrões internacionais que inclui identificadores de obras, gravações, produtos e detentores de direitos em todos os repertórios, por ex. DOI (Digital Object Identifier), ISAN (International Standard Audiovisual Number), ISBN (International Standard Book Number), ISRC (International Code System Record), ISSN (International Standard Serial Number), ISTC (Incunabula Short Title Catalogue) e ISWC (International Standard Musical Work Code)⁸.

Em 2014, a Agência Internacional ISNI, encarregada de administrar, governar e desenvolver continuamente o banco de dados e serviços ISNI, assinou um memorando de entendimento com o ORCID, uma organização internacional não governamental que atua na atribuição e gerenciamento de serviços de identificadores persistentes de pessoas do ambiente de pesquisa, em estreita sintonia com o ISNI. Interoperabilidade técnica, bem como coordenação das comunicações externas é um dos pontos do memorando de entendimento.

⁷ MACEWAN, A.; ANGJELI, A.; GATENBY, J. The International Standard Name Identifier (ISNI): The evolving future of name authority control. *Cataloging & Classification Quarterly*, v. 51, n. 1-3, p. 55-71, 2013. Disponível em: Disponível em: <<https://www.tandfonline.com/doi/full/10.1080/01639374.2012.730601?scroll=top&needAccess=true>>. Acesso em: 08 maio 2018 [Google Scholar].

⁸ INTERNATIONAL STANDARD NAME IDENTIFIER. Isni: (ISO 27729). Disponível em: <<http://www.isni.org/>>. Acesso em: 08 maio.

A data de 17 de novembro de 2017 marca a criação de um consórcio com cinco membros brasileiros para filiação ao ORCID a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), o Conselho de Desenvolvimento Científico e Tecnológico (CNPq), o Conselho Nacional das Fundações Estaduais de Amparo à Pesquisa (CONFAP), o Instituto Brasileiro de Informação em Ciência e Tecnologia – IBICT, a Scientific Electronic Library OnLine (SciELO) e a Rede Nacional de Ensino e Pesquisa. Esse consórcio viabilizará a inserção dos dados nacionais nas iniciativas internacionais de cooperação⁹.

Outro identificador de interesse dessa bibliografia comentada é o RINGGOLD, uma Agência do ISNI para organizações, funcionando desde maio de 2012¹⁰.

Os identificadores persistentes de nomes de pessoas e de organizações mostram-se importantes na construção da infraestrutura do CRIS brasileiro, seguindo as tendências internacionais. ORCID e RINGGOLD permitem a vinculação inequívoca de trabalhos e funções a pessoas e à organizações que atuam no mundo da pesquisa. Do ponto de vista da constituição dos sistemas de informação de pesquisa com seus subsistemas de pesquisador, projetos, resultados (repositórios institucionais), organizações, seu futuro...

[...] depende fortemente da consistência e estabilidade dos links entre recursos digitais que eles disponibilizam online, via rede de computadores. A identificação persistente é uma característica chave para os repositórios digitais - sejam eles bibliotecas ou arquivos - que precisam funcionar segundo os paradigmas de interoperabilidade. Não se pode mais prescindir dessa característica quando se pensa na integridade dos objetos digitais e nas várias faces do acesso contínuo a eles.¹¹



⁹SANTANA, P.H. O papel do IBICT no desenvolvimento de uma estrutura do tipo CRIS: a proposta do BR-CRIS (Current Research Information System). Brasília: IBICT, 2017. *Ringgold Identifier*. Disponível em: <<https://www.ringgold.com/ringgold-identifier/>>. Acesso em: 08 maio 2018.

¹⁰RINGGOLD IDENTIFIER. Disponível em: <<https://www.ringgold.com/ringgold-identifier/>>. Acesso em: 08 maio 2018.

¹¹SAYÃO, L.F. Interoperabilidade das bibliotecas digitais: o papel dos sistemas de identificadores persistentes-URN, PURL, DOI, Handle System, CrossRef e OpenURL. *Transinformação*, v. 19, n. 1, 2007. p. 79. Disponível em: <<http://www.redalyc.org/html/3843/384334745006/>>. Acesso em: 08 maio 2018.

MATERIAL E MÉTODO

As fontes usadas para levantamento das referências bibliográficas foram o Google acadêmico e o portal da CAPES, usando os seguintes termos de busca: persistent identifiers (identificadores persistentes), institutional identifier (identificador institucional), authority control (controle de autoridade), ORCID, ORCID Consortium (Consórcio ORCID), RINGGOLD.

Na fase de levantamento dos consórcios ORCID, o procedimento de busca ocorreu em duas etapas: primeiro no Google e depois no site do ORCID. Atualmente são 17 consórcios, formados por cinco ou mais instituições sem fins lucrativos ou governamentais. Desses, apenas quatro disponibilizam seus documentos de constituição na internet. O restante apenas disponibiliza informações gerais nos seus respectivos sites. Encontrou-se a documentação do consórcio da Dinamarca, porém não há registro dele na lista do ORCID. Para o refinamento da pesquisa realizou-se busca booleana no Google, com os mesmos nomes registrados na citada lista. O resultado do levantamento para a bibliografia anotada é de 19 artigos. A data de publicação da maioria deles é dos últimos cinco anos (2018-2014, quadro 1).

Quadro 1 – Resultado do levantamento para a bibliografia anotada: por data de publicação

ANO	Número de Referências
2018	1
2017	2
2016	3
2015	2
2014	3
2013	1
2011	4
2008	2
2006	1

Foram localizados 16 consórcios ORCID que seguem apresentados após a bibliografia anotada. O método utilizado é o da bibliografia comentada que permite construir um instrumento de pesquisa com pequenas anotações/comentários em lista de referências bibliográficas do tema identificador persistente de pessoas (ORCID) e de instituições (RINGGOLD), de interesse do projeto BR CRIS do IBICT.



BIBLIOGRAFIA ANOTADA

2018

RINGGOLD IDENTIFIER. In: WIKIPÉDIA: a enciclopédia livre. Wikipédia, 2018. Disponível em: <https://en.wikipedia.org/wiki/Ringgold_identifier>. Acesso em: 14 mar. 2018.

Apresenta características e conceitos gerais do identificador institucional RINGGOLD.

2017

DAPPERT, A. et al. Connecting the persistent identifier ecosystem: building the technical and human infrastructure for Open Research. *Data Science Journal*, n.16, p. 28, 2017. Disponível em: <<https://datascience.codata.org/articles/10.5334/dsj-2017-028/>>. Acesso em: 14 mar. 2018. [Google Scholar]

O artigo trata do ecossistema de identificadores persistentes. A relação dos identificadores ao ciclo de pesquisa é estabelecida. Em seguida, os autores destacam a maneira de construí-los.

SERRA, L. G.; SILVA, L. C.; SANTAREM SEGUNDO, J. E. O Orcid como aplicação de Linked Data no catálogo de bibliotecas. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 17, 2017, Marília, SP. Disponível em: <<http://enancib.marilia.unesp.br/index.php/xviiienancib/ENANCIB/paper/viewFile/52/1265>>. Acesso em: 30 maio 2018.

Após revisão da literatura de temas implicados no objetivo do estudo, há o relato de experiência bem sucedida de uso da tecnologia Linked Data, parte da web semântica, para relacionar o ORCID aos dados de autoridade dos catálogos das bibliotecas.

2016

BILDER, G.; BROWN, J.; DEMERANVILLE, T. *Organisation identifiers: current provider survey*. 2016. Disponível em: <<https://orcid.org/sites/default/files/ckfinder/userfiles/files/20161031%20OrgIDProviderSurvey.pdf>>. Acesso em: 14 mar. 2018. [Google Scholar]

Ainda sobre identificadores organizacionais. O trabalho é fruto de levantamento, durante o segundo trimestre de 2016, a respeito dos identificadores que estão disponíveis no mercado. Seus autores são Geoffrey Bilder (Crossref), Josh Brown e Tom Demeranville (ORCID). O documento é um complemento aos requisitos que um identificador organizacional deve ter, conforme estudo realizado pelo Crossref, Datacite e ORCID, em março de 2016.

AKERS, K. et al. ORCID Author Identifiers: A primer for librarians. *Medical Reference Services Quarterly*, v. 35, n. 2, p. 135-144, 2016. Disponível em: <<https://www.tandfonline.com/doi/full/10.1080/02763869.2016.1152139?scroll=top&needAccess=true>>. Acesso em: 15 mar. 2018.

Descrição das vantagens para as instituições e pesquisadores que adotaram o ORCID. Lista as ferramentas de promoção institucional do ORCID e que podem ser úteis para os bibliotecários.

AMANTE, M. J.; DUARTE, J.; GODINHO, J. et al. PTCRIS_OrgID: Portuguese Organisation Identifiers Authoritative System. In: *3th International Conference on Current Research Information Systems, CRIS2016, 9-11 June 2016, Scotland, UK. Conference EuroCRIS. Scotland, UK: Procedia Computer Science, 2017. P.138-145*. Disponível em: <https://dSPACECRIS.eurocris.org/bitstream/11366/518/1/CRIS2016_paper_15_Amante.pdf>. Acesso em: 18 maio 2018.

Um dos principais objetivos do PTCRIS é a criação da primeira versão de um banco de dados de organizações nacionais autoritativo (com altos níveis de controle das entradas), para desenvolver um serviço de reconciliação de registros de organizações, com seus diferentes nomes, hierarquias/subordinações. Duas coisas são mandatórias: a criação de regras, princípios e processos de registro de organizações em Portugal e uso de identificadores exclusivos que eliminem a ambiguidade dos registros institucionais. O resultado é a melhor qualidade da informação com dados normalizados, atualizados e disponíveis para acesso público. O trabalho é ilustrado com um grupo de registro de organizações e com um banco de dados de organizações nacionais que sincroniza sistemas como Ringgold e ISNI.

2015

FERGUSON, N.; MOORE, R.; SCHMOLLER, S. *Review of selected organisational IDs and development of use cases for the Jisc CASRAI-UK Organisational Identifiers Working Group*. 2015. Disponível em: <<http://repository.jisc.ac.uk/5853/1/Review-of-orgIDs-usecases-clax.pdf>>. Acesso em: 14 mar. de 2018.

Estudo realizado pelo Jisc, uma organização que provê soluções digitais para o ensino e a pesquisa no Reino Unido. Os autores relatam pesquisa a respeito de soluções de registro e de provimento de serviços de valor agregado para identificador organizacional no Reino Unido. O Ringgold é apresentado como o candidato mais forte.

MEADOWS, A. ORCID: What, Why, How? Harvard-Smithsonian Center for Astrophysics, 2015. Slide presentation. Disponível em: <<http://library.cfa.harvard.edu/files/wolbach/files/20151023harvardcentforastrophysicsorcid.pdf?m=1445779276>>. Acesso em 15 mar. 2018.

Apresentação em slide das principais definições a respeito do ORCID e suas possibilidades de uso e aplicações.

2014

ASSOCIATION OF RESEARCH MANAGERS AND ADMINISTRATORS - JISC-ARMA. *Institutional ORCID implementation and cost benefit analysis report*. 2014. Disponível em: <https://repository.jisc.ac.uk/6025/2/Jisc-ARMA-ORCID_final_report.pdf>. Acesso em: 14 mar. 2018.

Mais um estudo realizado pelo Jisc, uma organização que provê soluções digitais para o ensino e a pesquisa no Reino Unido. É uma pesquisa sobre como as instituições de ensino superior (IES) devem organizar a implementação do ORCID e explorar a incorporação de IDs do ORCID em sistemas e fluxos de trabalho institucionais. Avaliar custos, benefícios e riscos da implementação do ORCID é outro objetivo do trabalho.

FENNER, M. et al. ODIN: the ORCID and DataCite interoperability network. *International Journal of Knowledge and Learning* v. 9, n. 4, p. 305-325, 2014. Disponível em: <<https://pdfs.semanticscholar.org/6898/332ff203d8231ff81580137837e78cc33fa2.pdf>>. Acesso em: 14 maio 2018.

A interoperabilidade para identificação de conjuntos de dados e seus autores representa um grande obstáculo. O trabalho propõe modelo conceitual para resolução da problemática da interoperabilidade entre diferentes identificadores de dados e pessoas.

HAAK, L. ORCID: *Persistent identifiers for researchers and contributors*. SlideShare 2014. Disponível em: <http://www.resourcenter.net/images/CSE/Files/2014/AnnMtg/Handouts/03_Haak.pdf>. Acesso em: 15 mar. 2018.

Apresentação em slides sobre o ORCID e outros identificadores padrões e seus principais benefícios.

2013

HAAK, L. *ORCID plans to launch affiliation module using ISNI and Ringgold Organization Identifiers*. 2013. Disponível em: <<https://orcid.org/blog/2013/06/27/orcid-plans-launch-affiliation-module-using-isni-and-ringgold-organization>>. Acesso em: 15 mar. 2018.

Discorre sobre o planejamento do ORCID para trabalhar com o Ringgold com o objetivo de obter identificadores organizacionais para uso público de registro. A proposta inclui o Identificador institucional do Ringgold, o número ISNI, o nome da instituição, os dados de localização, a URL e o nome alternativo.

2011

DERIDDER, J. *I² and ISNI: Improving the Information Supply Chain with Standard Institutional Identifiers*. *Information Standards Quarterly*, Summer v. 23, n. 3, p. 26-29, 2011. Disponível em: <<https://groups.niso.org/publications/isq/2011/v23no3/deridder>>. Acesso em: 15 mar. 2018. [Google Scholar]

Discorre sobre a importância do grupo NISO Institutional Identifier, criado para gerar um padrão de identificador de fluxos de trabalho mais eficientes para a cadeia de fornecimento de informações.

GATENBY, J.; MACEWAN, A. ISNI: A new system for name identification. *Information Standards Quarterly*, v. 23 n. 3, p. 4-9, 2011. Disponível em: <<https://groups.niso.org/publications/isq/2011/v23no3/gatenby>>. Acesso em: 15 mar. 2018. [Google Scholar]

O artigo discute o desenvolvimento do padrão ISNI, a arquitetura do identificador ISNI e como a ISNI International Agency está criando seu banco de dados inicial.

PRESENTERS, T.F.; PRESENTERS H. H; RECORDER, D.E. *One Identifier: Find Your Oasis with NISO's I2 (Institutional Identifiers) Standard*. 2011. Disponível em: <<https://www.tandfonline.com/doi/full/10.1080/0361526X.2011.556037>>. Acesso em: 14 mar. 2018. [Google Scholar]

Outro artigo a tratar dos identificadores institucionais discorre sobre a implementação do padrão NISO (National Information Standards Organization), para melhoria dos processos de gerenciamento de recursos eletrônicos. Sua implantação permitirá a criação de um identificador único padrão para cada instituição, podendo ser usado por editores, agentes e hosts on-line.

SPECIAL topic: organization and people identifiers. *Information Standards Quarterly Summer*, v. 23, n.3, 2011. Disponível em: <https://groups.niso.org/apps/group_public/download.php/7250/isq_v23no3.pdf>. Acesso em: 15 mar. de 2018. [Google Scholar]

Visão geral dos principais padrões e iniciativas sobre o tema dos identificadores de pessoas e organizações.

2008

CARPENTER, T. Standards column - the problems of institutional identification: toward a universal institutional ID. *Against the Grain*, v. 20, n. 1, 2008. DOI: 10.7771/2380-176X.2722. Disponível em: <<https://docs.lib.purdue.edu/cgi/viewcontent.cgi?article=2722&context=atg>>. Acesso em: 14 mar. 2018. [Google Scholar]

O identificador universal de instituições é objeto desse artigo, focalizando a sua importância e os problemas causados pela sua ausência.-

HENDERSON, H. Why do You Lose Access to Your Electronic Resources? How an Institutional Identifier can Help. *Serials*, v. 21, n. 2, p.134-139, 2008. Doi: 10.1629/21134. Disponível em: <<https://serials.uksg.org/articles/abstract/10.1629/21134/>>. Acesso em: 14 mar. 2018. [Google Scholar]

Mais um artigo a tratar de identificador institucional único, destacando perdas de acesso aos recursos eletrônicos da instituição, na ausência desse identificador.

2006

DAVIDSON, J. *Persistent Identifiers*. DCC Briefing Papers: introduction to curation Edinburgh: Digital Curation Centre. 2006. Disponível em: <<http://www.dcc.ac.uk/sites/default/files/documents/resource/briefing-papers/persistent-identifiers.pdf>>. Acesso em: 14 mar. 2018. [Google Scholar]

Artigo de fundamental importância sobre os diferentes aspectos que devem ser considerados para que a identificação persistente de recursos digitais seja bem-sucedida. A decisão de usar a identificação digital é apenas uma parte de uma estratégia geral de preservação digital. Sem um compromisso institucional adequado e papéis e responsabilidades claramente definidos, os identificadores não podem oferecer garantias de persistência, de capacidade do recurso continuar sendo localizado (locatability) ou capacidade de estar pronto para ser usado (actionability) em longo ou curto prazo.



INFORMAÇÕES DE CONSÓRCIOS

AUSTRÁLIA. *Australian Access Federation (AAF)*. Disponível em: <<https://aaf.edu.au/>, http://www.and.s.org.au/___data/assets/pdf_file/0019/417340/consortium-model.pdf>. Acesso em: 11 maio 2018.

Como todos os consórcios, o australiano destina-se a maximizar o benefício do ORCID para as instituições participantes e contribuir para a flexibilidade dos serviços de identificação. Foi desenvolvido por um grupo de trabalho composto por representantes de instituições de ensino e pesquisa: Australian Universities (AU), Australian University Librarian Council (CAUL), Australasian Research Management Society (ARMS), Australian National Data Service (ANDS) Australian Federation of Access (AAF) and Australian University Information Technology Board (CAUDIT). Dois conselhos de pesquisa, o The Australian Research Council e o The National Health and Medical Research Council (NHMRC) desempenharam função importante durante o desenvolvimento do Modelo de Consórcio. O termo inicial da afiliação é estabelecido em três anos (36 meses).

BELGIUM. *Elektron VZW*. Disponível em: <<https://elektronvzw.be/>>. Acesso em: 11 maio 2018.

O consórcio das instituições de ensino e pesquisa da Bélgica ocorre no âmbito de uma associação que promove a cooperação de fontes de informação eletrônicas para educação, pesquisa, serviços e políticas, com o objetivo de prover livre acesso às publicações científicas de seus membros. Fazem parte do consórcio as seguintes instituições: Fonds Wetenschappelijk Onderzoek - Vlaanderen (FWO), Ghent University, KU Leuven, University of Antwerp, University of Hasselt e Vrije Universiteit Brussel. É composto por uma assembleia geral, onde cada membro efetivo tem um representante com capacidade de voto; membros associados têm o direito de participar da assembleia geral sem direito a voto.

BRASIL. RNP - *Rede Nacional de Ensino e Pesquisa*. Disponível em: <<https://orcid.org/members/0010f00002DzH58AAF-rede-nacional-de-ensino-e-pesquisa-rnp>>. Acesso em: 11 maio 2018.

O primeiro consórcio ORCID no Brasil data de novembro de 2017, sendo formado por cinco instituições: CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior), CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico), CONFAP (Conselho Nacional das Fundações Estaduais de Amparo à Pesquisa), IBICT (Instituto Brasileiro de Informação em Ciência e Tecnologia) e SciELO (Scientific Electronic Library Online). A coordenação da implementação está a cargo da RNP (Rede Nacional de Ensino e Pesquisa), que opera com tecnologia avançada de infraestrutura de rede para facilitar a pesquisa colaborativa em diversas áreas do conhecimento.

Obs: O lançamento oficial do consórcio ocorreu no dia 22 de maio de 2018, na CAPES com a presença de todas as instituições participantes. “Entre uma vez, use sempre”, conforme Geraldo Nunes Sobrinho, diretor de Programas e Bolsas no País da CAPES, explica a funcionalidade do ORCID, “um identificador digital persistente que permitirá a conexão de pesquisadores às suas afiliações e atividades por meio da integração com editoras, agências de financiamento e bases de dados.” NOVAIS, Gisele. Lançado consórcio para assinatura de identificador digital de pesquisadores. CCS/CAPES, 2018. Disponível em: <<http://www.capes.gov.br/sala-de-imprensa/noticias/8882-lancado-consorcio-para-assinatura-de-identificador-digital-de-pesquisadores>>. Acesso em: 24 maio 2018.

FINLAND. CSC IT Center for Science Ltd. Disponível em: <<https://www.csc.fi/>>. Acesso em: 14 maio 2018.

O consórcio da Finlândia também funciona como instituição sem fins lucrativos, especializada em TICs com serviços de qualidade para pesquisa, educação, cultura, administração pública e empresas. O consórcio realizou estudos preliminares para uma possível implantação do ORCID nacional.

GERMANY. TIB - *Technische Informationsbibliothek*. Disponível em: <<https://orcid.org/members/001G000001EAnAvIAL-technische-informationsbibliothek-tib>>. Acesso em: 14 maio 2018.

Na Alemanha, o consórcio ORCID é financiado pela Fundação Alemã de Pesquisa (Deutsche Forschungsgemeinschaft, DFG). Atualmente participam 39 instituições; seus milhões de autores têm suas publicações em todo o mundo, e dados de pesquisa decorrentes relacionados a seus identificadores únicos. A gestão administrativa do consórcio é realizada pela Biblioteca Nacional Alemã de Ciência e Tecnologia (Technische Informationsbibliothek - TIB), situada em Hannover. A TIB também é responsável pela filiação ao ORCID das instituições de pesquisa alemãs. Sua política é a do acesso aberto com acesso irrestrito e gratuito às publicações científicas.

ITALY. CINECA. Disponível em: <<https://www.cineca.it/it>>. Acesso em: 11 maio 2018.

O consórcio ORCID italiano é administrado no âmbito de um antigo consórcio, sem fins lucrativos, o CINECA, fundado em 1969. Ao todo são 74 instituições, 70 universidades e quatro instituições de pesquisa, além do Ministério Italiano de Educação. O CINECA tem o maior centro de computação da Itália, desenvolvendo serviços para a indústria e administração pública, funcionando também como ponto de referência para o sistema acadêmico nacional.

NETHERLANDS. SURFMARKET. Research Institute. Disponível em: <<https://orcid.org/members/001G000001rkowJIAQ-surfmarket>>. Acesso em: 14 maio 2018.

Outra rede de ensino e pesquisa, a exemplo da brasileira RNP, lidera projeto piloto para filiação ao ORCID nos Países Baixos. A implementação do consórcio ocorre juntamente com o projeto de inovação de acesso aberto. São oito as instituições participantes.

NEW ZEALAND. *Royal Society Te Apārangi*. Disponível em: <<https://orcid.org/members/001G000001xTSfuIAG-royal-society-te-aparangi>>. Acesso em: 11 de maio 2018.

A Royal Society Te Apārangi administra o consórcio ORCID na Nova Zelândia com 48 instituições de pesquisa. Essa sociedade científica gerencia a tomada de decisão em torno de diversos fundos de pesquisa, sendo também responsável pela edição de quase uma dezena de revistas científicas.

NORWAY. CERES - *National Center for systems and services for Research and Studies*. Disponível em: <<http://www.ceres.no/english/>>. Acesso em: 11 maio 2018.

O consórcio da Noruega conta com 90 membros. Sua coordenação está a cargo do National Center for Research and Studies Systems and Services (CERES) que gerencia, desenvolve, mantém e opera sistemas e serviços de administração e pesquisa de estudantes, em nome do setor de ensino superior norueguês e das comunidades de pesquisa norueguesas.

PORTUGAL. *University of Aveiro*. Disponível em: <<https://orcid.org/members/0010f00002DmXadAAF-university-of-aveiro>>. Acesso em: 14 maio 2018.

Em Portugal o consórcio é liderado pela Universidade de Aveiro (UA), reunindo cinco instituições. A UA é frequentada por cerca de 15 mil alunos em programas de graduação e pós-graduação.

SOUTH AFRICA. TENET - *Tertiary Education and Research Network of South Africa*. Disponível em: <<https://orcid.org/members/001G000001zVNxWIAW-tenet>>. Acesso em: 14 maio 2018.

A TENET (Tertiary Education and Research Network of South Africa), operadora da rede nacional de pesquisa e educação, SANReN, lidera o consórcio ORCID na África do Sul, com 10 instituições-membro. Podem se filiar ao consórcio universidades sul-africanas, conselhos de pesquisa e outras organizações sem fins lucrativos que trabalham em apoio à pesquisa e educação, bastando retornar um pedido de adesão. Há informação sobre preço na página do consórcio. Como operador de consórcio, a TENET fornece suporte de primeira linha para membros que fazem uso de integrações ORCID.

SWEDEN. SUNET - *Data communication and infrastructure for research and education Swedish Federation of Identity*. Disponível em: <<https://orcid.org/members/001G000001qTsKuIAK-sunet>>. Acesso em: 14 maio 2018.

Mais uma rede de ensino e pesquisa que abriu caminho para a internet lidera um consórcio ORCID. A SUNET é a rede de computadores da Swedish University, unidade do Swedish Research Council. Atualmente são 11 instituições que integram o consórcio.

TAIWAN. *National Chiao Tung University*. Disponível em: <<https://orcid.org/members/001G000001d4StWIAU-national-chiao-tung-university>>. Acesso em: 11 maio 2018.

As instituições que fazem parte do consórcio ORCID na National Chiao Tung University, em Taiwan, são: Chang Gung University, National Chengchi University, National Chiao Tung University, National Taiwan Normal University e National Taiwan University.

UNITED KINGDOM. JISC. Disponível em: <<http://ukorcidsupport.jisc.ac.uk/>>; <<https://www.research-consulting.com/wp-content/uploads/2016/11/Institutional-ORCID-Implementation-and-Cost-Benefit-Analysis-Report.pdf>>. Acesso em: 14 maio 2018.

No Reino Unido, o processo de adoção de identificadores persistentes para maior integração aos sistemas CRIS inicia em maio de 2014. São oito instituições piloto que participaram do Projeto ORCID do Jisc-ARMA: Aston University; Imperial College of London; Northumbria University; University of Southampton; University of Swansea; University of Kent; University of Oxford; e York University. As instituições desenvolveram projetos pilotos para integração dos identificadores persistentes e o sistema CRIS. O custo de associação para o primeiro ano, atualmente US\$ 4.000, provavelmente cairá no futuro de um acordo de consórcio nacional.

UNITED STATES. *Health Research Alliance*. Disponível em: <<https://www.healthra.org/>>. Acesso em: 14 maio 2018.

O consórcio americano data de 2017 e reúne instituições sem fins lucrativos que são financiadas em projetos na área de saúde e pesquisa biomédica. Os membros do consórcio da Health Research Alliance (HRA) pretendem maior integração com o ORCID no processo de gerenciamento de subsídios, facilitando a submissão e o relato do progresso para seus beneficiários e para a obtenção dos resultados das pesquisas gerenciadas pelo programa. O fluxo de informação aprimorado ajudará os financiadores a entender melhor o impacto de seus programas de financiamento. As instituições participantes do consórcio são: Alzheimers' Association, American Cancer Society, Autism Speaks, Burroughs Wellcome Fund, Conquer Cancer Foundation, Damon Runyon Cancer Research Foundation, Fondation Leducq, Howard Hughes Medical Institute (HHMI), Kenneth Rainin Foundation, Simons Foundation e St. Baldrick's Foundation.

UNITED STATES. ORCID *us community - Big ten academic alliance, gwla, lyrasis, nerl*. Disponível em: <<https://orcid.org/members/001G000001wND7XIAW-lyrasis>>. Acesso em: 11 maio 2018.

A comunidade ORCID dos EUA forma um dos maiores consórcios no mundo, com 95 instituições-membro. A Big Ten Academic Alliance juntamente com a LYRISIS são parceiros para compor um consórcio nacional ORCID para instituições de pesquisa no país.



AGRADECIMENTOS

Trabalho elaborado sob orientação da professora Maria de Nazaré Freitas Pereira, a quem agradecemos pelos ensinamentos. Agradecemos também à Eny Marcelino de Almeida Nunes, tecnologista sênior do IBITC, pela primeira revisão deste artigo.



REFERÊNCIAS

ASSUMPÇÃO, F. S.; et al. O controle de autoridade no domínio bibliográfico: os catálogos em livros e em fichas. *Biblios*, n. 67, p. 84-98, 2017. Disponível em: <http://www.scielo.org.pe/scielo.php?pid=S1562-47302017000200007&script=sci_abstract>. Acesso em: 08 maio 2018.

MACEWAN, Andrew; ANGJELI, Anila; GATENBY, Janifer. The International Standard Name Identifier (ISNI): The evolving future of name authority control. *Cataloging & Classification Quarterly*, v. 51, n. 1-3, p. 55-71, 2013. Disponível: Google Scholar (rascunho final em word do artigo de acesso pago). International Standard Name Identifier (ISO 27729). Disponível em: <<http://www.isni.org/>>. Acesso em: 08 maio 2018.

ROMANETTO, M. L. et al. O Virtual International Authority File–VIAF e a agregação de valores por metadados de autoridade. *RDBC: Revista Digital de Biblioteconomia e Ciência da Informação*, v. 15, n. 3, p. 571-590. Disponível em: <<https://periodicos.sbu.unicamp.br/ojs/index.php/rdbci/article/view/8647488/pdf>>. Acesso em: 08 maio 2018.

SANTANA, Paulo Henrique. *O papel do IBICT no desenvolvimento de uma estrutura do tipo CRIS: a proposta do BR-CRIS (Current Research Information System)*. Brasília: IBICT, 2017. RINGGOLD IDENTIFIER. Disponível em: <<https://www.ringgold.com/ringgold-identifier/>>. Acesso em: 08 maio 2018.

SAYÃO, Luís Fernando. *Interoperabilidade das bibliotecas digitais: o papel dos sistemas de identificadores persistentes-URN, PURL, DOI, Handle System, CrossRef e OpenURL*. Transinformação, v. 19, n. 1, 2007. Disponível em: <<http://www.redalyc.org/html/3843/384334745006/>>. Acesso em: 08 maio 2018.

SILVA, A. F. et al. O controle de autoridade no domínio bibliográfico: os catálogos em livros e em fichas. *Biblios*, n. 67, p. 84-98, 2017. Disponível em: <http://www.scielo.org.pe/scielo.php?pid=S1562-47302017000200007&script=sci_arttext&lng=pt>. Acesso em: 08 maio 2018.

SILVA, A. F.; SANTOS, P. L. V. A. C. A importância do controle de autoridade: uma abordagem baseada nos objetivos e nas funções dos catálogos. In: *I Encontro Nacional de Catalogadores, III Encontro de Estudos e Pesquisas em Catalogação*, 2012, Rio de Janeiro. Anais... Rio de Janeiro: Biblioteca Nacional, 2012. Disponível em: <<http://eprints.rclis.org/18843/1/assumpcao-santos-a-importancia-do-controle-de-autoridade.pdf>>. Acesso em: 01 jul. 2013.

SITES CONSULTADOS

CASRAI. Dictionary. Disponível em: <http://dictionary.casrai.org/Persistent_identifier>. Acesso em: 08 maio 2018.

INTERNATIONAL STANDARD NAME IDENTIFIER. Disponível em: <<http://www.isni.org/>>. Acesso em: 08 maio 2018.

RINGGOLD IDENTIFIER. Disponível em: <<https://www.ringgold.com/ringgold-identifier/>>. Acesso em: 08 maio 2018.

Parte IV

RELATÓRIOS





Relatório de viagem
Paulo Henrique de Assis Santana
Brasília
Dezembro 2017

INTRODUÇÃO

Este documento tem por objetivo o relato das visitas realizadas em Portugal e as possibilidades de aplicação, no desenvolvimento do BRCRIS, dos conceitos e ferramentas discutidas nas referidas visitas, bem como indicar datas e roteiros da viagem. Embora tenha sido escrito em dezembro de 2017, alguns dados mencionados foram atualizados em 24 de setembro de 2018.

PERÍODO E ROTEIRO DA VIAGEM

Início: 11/11/2017

Final: 18/11/2017

- ✓ Saída de Brasília às 18:55 horas (hora local) do dia 11/11/2017, sábado, no voo TP 58, com chegada a Lisboa às 6:00 horas (hora local) do dia 12/11/2017.
- ✓ Visita à FCT/FCCN (Fundação para Ciência e Tecnologia/Fundação para a Computação Científica Nacional) no dia 13/11/2017.
- ✓ Saída de Lisboa para Braga no comboio 137 às 19:09 horas do dia 13/11/2017, segunda-feira, com chegada a Braga por volta de 23:40 horas.
- ✓ Saída de Braga para Porto em comboio às 07:50 do dia 14/11/2017, terça-feira, com chegada às 08:50 horas.
- ✓ Visita à Faculdade de Engenharia da Universidade do Porto no dia 14/11/2017.
- ✓ Saída de Porto para Braga em comboio às 22:30 horas do dia 14/11/2017, terça-feira, com chegada às 23:40 horas.

- ✓ Visita ao serviço de documentação da Universidade do Minho (câmpus de Braga), dias 15, 16 e 17/11/2017
- ✓ Saída de Braga para Lisboa no comboio 134 às 18:07 horas do dia 17/11/2017, sexta-feira, com chegada às 22:37 horas.
- ✓ Saída de Lisboa às 09:50 horas (hora local) do dia 18/11/2017, sábado, no voo TP 59, com chegada a Brasília às 17:25 horas (hora local) do mesmo dia.
- ✓ Visitas
 - Dia 13/11/2017, visita à FCT/FCCN, onde foram discutidos o conceito do PTCRIS, seus componentes, estratégias de implementação e possibilidades de cooperação.
 - Dia 14/11/2017, visita à Universidade do Porto para conhecer seu sistema de gestão acadêmica (SIGARRA – Sistema de Informação para Gestão Agregada dos Recursos e Registos Acadêmicos) e sua estratégia de integração com o PTCRIS a partir das informações registradas neste sistema interno.
 - Dias 15, 16 e 17/11/2017, visita ao Serviço de Documentação da Universidade do Minho para conhecer a estratégia de desenvolvimento dos componentes do PTCRIS e as ferramentas desenvolvidas para a operação do RCAAP (Repositório Científico de Acesso Aberto de Portugal), sua integração com o OpenAIRE (Open Access Infrastructure for Research in Europe) e com a rede La Referencia e repositório OASIS-BR (Portal Brasileiro de Acesso Aberto à Informação Científica).



DISCUSSÕES

FCT/FCCN

Na FCCN fomos apresentados às estratégias de desenvolvimento do PTCRIS, projeto que visa promover a integração de vários sistemas de informação de suporte à atividade científica e assim criar um ecossistema aberto do qual todos os agentes podem facilmente beneficiar-se, com vantagens para todos. Sua concepção, estrutura e padrões parecem-nos excelente exemplo para seguir na tarefa de estruturação do BRCRIS.

O PTCRIS não é um sistema convencional, mas sim uma plataforma, um sistema de sistemas (SoS – System of Systems). Por meio dele pretende-se facilitar os processos de gestão, produção e acesso à informação fidedigna, completa e atualizada, sobre a atividade científica nacional.

Sua concepção está baseada em quatro dimensões basilares: (a) um instrumento de registro de currículos de pesquisadores, (b) um repositório nacional de acesso aberto para produções científicas e tecnológicas, (c) um cadastro de instituições e (d) uma base nacional de financiamento à pesquisa. Da integração destas dimensões resulta o PTCRIS.

INSTRUMENTOS BASILARES DO PTCRIS

✓ Currículo – Ciência Vitae

O currículo hoje vigente em Portugal denomina-se Plataforma DeGóis e deriva diretamente da Plataforma Lattes, oferecida à FCT em 2001 pelo CNPq, no escopo da construção de uma rede ibero-americana denominada Rede Scienti. Entretanto, no escopo do PTCRIS, tal plataforma cederá lugar a novo instrumento denominado “Ciência Vitae”, alinhado às normativas CASRAI (Consortia Advancing Standards in Research Administration Information).

Uma das razões que levaram ao desenvolvimento do Ciência Vitae em substituição ao DeGóis é o baixo índice SUS (System Usability Scale) deste último. De acordo com o relato da FCCN, o SUS associado ao DeGóis é de apenas 27, quando a média deste índice é 68. Tal relato deve servir-nos de alerta em relação à usabilidade do Lattes, pois entende-se que SUS abaixo de 50 é sinal de que investimentos em design e usabilidade devem ser priorizados dentro do plano de negócios.

Além disto, a filosofia de operação do “Ciência Vitae” é a de atualização automática do currículo tendo por base conteúdos do ORCID e do RCAAP, conteúdos estes oferecidos automaticamente ao pesquisador para que este autorize (ou não) sua inclusão em seu currículo. No contexto da integração, destaca-se o serviço PTCRISync, que usa o ORCID como hub para autenticação e intercâmbio de informação entre os diversos sistemas do PTCRIS, permitindo que seus usuários insiram ou atualizem um item de informação em um dos seus sistemas constituintes, propagando-o para os outros sistemas, contribuindo assim para a manutenção da coerência de informações no PTCRIS.

✓ **Repositório RCAAP - Repositório Científico de Acesso Aberto de Portugal.**

Este repositório utiliza como ferramenta o software livre DSpace versão 5 e congrega as publicações científicas e tecnológicas de Portugal em acesso aberto. Faz coleta, com uso do protocolo OAI-PMH (Open Archives Initiative Protocol for Metadata Harvesting), dos registros dos repositórios portugueses e também do OASIS-BR repositório brasileiro.

É interessante observar o conceito de acesso aberto utilizado em Portugal: todas as publicações derivadas de financiamento público devem obrigatoriamente ser depositadas no RCAAP. Se tiverem restrições de acesso, por serem publicações em revistas de acesso restrito ou estarem associadas a temas sensíveis, são depositadas com acesso aos seus metadados e a informação relativa ao tempo em que o acesso a seu texto completo é proibido. Com esta medida, o RCAAP passa a ser a fonte de toda a produção científica nacional, com acesso aberto a todos os seus metadados e aos textos não restritos.

Segue as diretrizes OpenAIRE 4.0 e expõe seus dados na infraestrutura OpenAIRE (The OpenAIRE Project - Open Access Infrastructure for Research in Europe). Tal infraestrutura está se abrindo para exibição dos dados da América Latina, o que sugere a possibilidade de exposição dos dados do OASISBR e a construção de estrutura similar para os dados dos Institutos Nacionais de Ciência e Tecnologia (INCT) brasileiros.

✓ **Cadastro de Instituições**

Tendo por base o cadastro de instituições do Ringgold, está sendo construído um cadastro nacional de instituições, com a atribuição de identificadores às instituições nacionais e internacionais ligadas à pesquisa.

Esta base é fundamental para identificação única das instituições que participam de pesquisas em Portugal ou que se relacionam com pesquisas portuguesas.

Interessante observar que, na criação da plataforma Lattes, pretendia-se estabelecer um cadastro de instituições a partir das informações curriculares registradas pelos pesquisadores. Infelizmente, tal ideia não foi implementada, tendo em vista o esforço de análise envolvido. Entretanto, uma cooperação com a FCT poderia reativar esta dimensão da plataforma Lattes.

É interessante ressaltar que o CNPq dispõe hoje de um Diretório de Instituições de Pesquisa integrado a uma estrutura de coleta de informações sobre as infraestruturas de pesquisa existentes. Tal sistema poderia ser um dos pilares para a construção do cadastro de instituições do BRCRIS, com aplicação imediata ao projeto piloto com a Fapeal.

Outra possibilidade que parece ser interessante é a utilização deste diretório para mapear as instituições e suas uingraestruturas de pesquisa no âmbito dos INCTs.

✓ **Base Nacional de Projetos**

As informações sobre projetos de pesquisa e seus resultados está sendo construída a partir dos sistemas de informação existentes. Nesta base, os projetos também terão identificadores únicos, de forma a permitir cruzamentos seguros.

No que diz respeito a tal base, um piloto interessante seria a organização dos INCTs como projetos de pesquisa nesta base, relacionando-os com seus pesquisadores, instituições sede, laboratórios associados, pesquisadores.

Considerando que a cardinalidade do conjunto de institutos é da ordem de centena, o universo a ser trabalhado é limitado em número, embora os projetos sejam de alta relevância, bem como as instituições, os pesquisadores e os produtos derivados.

POSSIBILIDADES DE COOPERAÇÃO

Na discussão, foram identificados pontos onde parece ser possível e desejável a cooperação entre os atores brasileiros do BRCRIS e portugueses do PTCRIS.

São eles:

- ✓ Base de dados de financiamento
 - Intercâmbio de experiência e boas práticas
 - Intercâmbio de informação (OpenAIRE, BD Financiamento)
- ✓ Sistema de gestão de curricular
 - Intercâmbio de informações curriculares (possibilidade de carregar um currículo Ciência Vitae no Currículo Lattes e vice-versa).

Identificou-se, também, a possibilidade de cooperação no contexto do projeto ORBIT (ORCID Reducing Burden an Improving Transparency Project).

O ORBIT tem como público alvo a comunidade de financiadores e tem dois componentes: uma rede de financiadores e projeto demonstrativos. É uma iniciativa do US National Institutes of Health (NIH), que deseja popular sua plataforma ScienCV com dados vindos dos registros ORCID. O Swiss National Science Fund aderiu à ideia e o NIH estabeleceu um pequeno financiamento para ajudar a desenvolver esta colaboração. Há também discussões em andamento com a European Research Council, Research Councils UK e Welcome Trust.

O consórcio ORCID, que está sendo formado com participação do CNPq, CAPES, IBICT, CONFAP e SciELO, poderia agregar-se a este projeto juntamente com a FCT de Portugal.

É interessante observar que já existe Acordo de Cooperação Técnica Nº 1/2015, entre IBICT e FCT, cujo objeto é a “realização de projetos conjuntos na área de informação em Ciência e Tecnologia”, com vigência de 5 anos a partir de fevereiro de 2015, e que possibilita o desenvolvimento de ações conjuntas, o que poderia acelerar o estabelecimento de cooperação mais efetiva entre as instituições no desenvolvimento de suas plataformas PTCRIS e BRKRIS, envolvendo também Projetos P-o-P (CERIF-XML) e participação na infraestrutura OpenAIRE 4.0 (CERIF-XML e DC).

FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO

Nesta visita, o Sistema de Informação para Gestão Agregada dos Recursos e Registos Acadêmicos – SIGARRA nos foi apresentado em detalhe, com discussão da estratégia de sua integração com o PTCRIS.

A discussão nos pareceu fundamental, pois é instância exemplar dos problemas a resolver no âmbito do BRKRIS relativamente à integração de grandes universidades cujos sistemas de informação estejam em avançado estágio de desenvolvimento, não podendo se desviar dos caminhos traçados em seu planejamento.

Foram-nos descritos os problemas e as interfaces em desenvolvimento para atender às exigências da plataforma nacional PTCRIS.

SERVIÇO DE INFORMAÇÃO DA UNIVERSIDADE DO MINHO EM BRAGA

Foram dispendidos três dias de discussões extremamente proveitosas com equipe técnica da Universidade do Minho, quando foram apresentados os desenvolvimentos relativos às plataformas de acesso aberto das quais Portugal participa.

Duas linhas principais foram apresentadas:

- ✓ Desenvolvimentos relativos ao RCAAP e sua integração com o OASIS-BR e a rede La Referencia.
- ✓ Desenvolvimentos para integração à infraestrutura OpenAIRE.

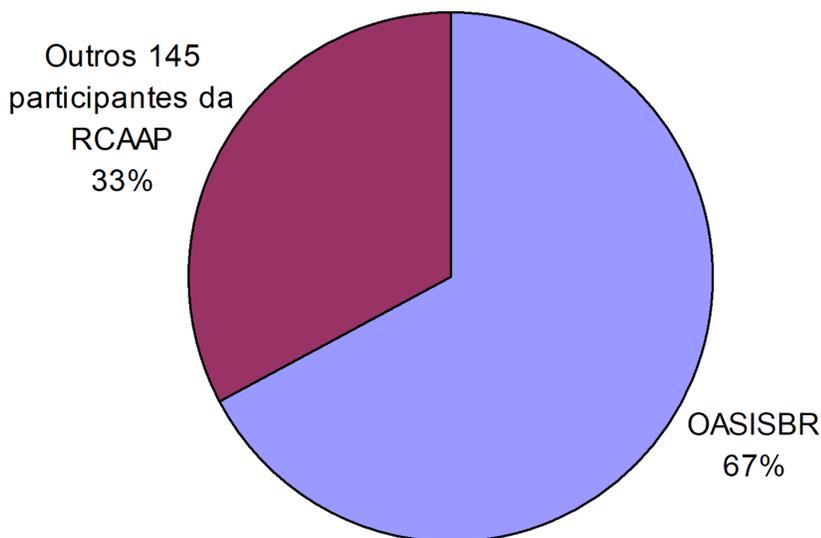
RCAAP

Fazendo uso do software livre DSpace versão 5, a Universidade do Minho desenvolveu e aprimora o repositório RCAAP, integrando as informações científicas de Portugal e Brasil (via OASIS-BR) em acesso aberto.

Está em curso trabalho de integração com a rede latino-americana em acesso aberto, La Referencia, e com a infraestrutura Open AIRE.

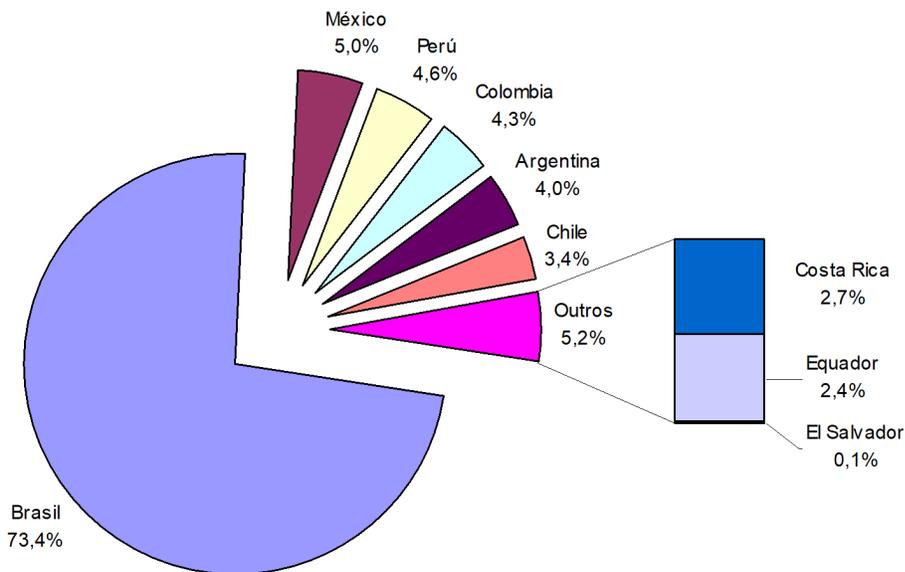
É interessante observar que, com relação à rede La Referencia, as informações brasileiras em acesso aberto representam quase 75% dos registros da rede, percentual que se excede os 67% de sua participação na rede RCAAP, conforme pode ser visto nas figuras 1a e 1b.

Figura 1a - Percentual de registros do OASISBR na rede RCAAP



Extraído de: Repositório de Acesso Aberto em Portugal. Disponível em:<<https://www.rcaap.pt/directory.jsp>>.

Figura 1b – Percentual de registros dos países participantes da rede La Referencia



Extraído de: La Referencia: página de busca. Disponível em: <<http://www.lareferencia.info/vufind/Search/Results?lookfor=&type=Title&limit=20&sort=relevance>>.

A política estabelecida em Portugal, de obrigatoriedade do depósito na rede de acesso aberto RCAAP, para as publicações derivadas de projetos financiados com dinheiro público, mesmo que o texto completo da publicação não esteja disponível ainda para acesso aberto (publicações em revistas que não autorizam o acesso aberto durante certo período, publicações que envolvem informações sensíveis, etc.), dão enorme força e significado ao RCAAP, política que poderia ser reproduzida no Brasil no contexto do OASIS-BR.

Por conseguinte, por sua importância e porte, a estrutura OASIS-BR poderia ser um dos pilares da plataforma BRCRIS, sendo que sua relevância seria ainda maior se fosse estabelecida, pelas agências de fomento brasileiras, a política de depósito obrigatório das informações oriundas de projetos financiados com dinheiro público.

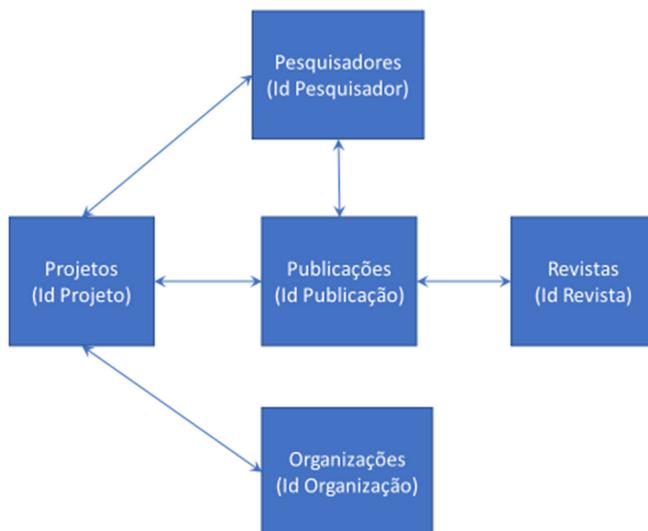
ESTRUTURAÇÃO DAS BASES PARA INTEGRAÇÃO DAS INFORMAÇÕES

O esforço para criação de bases de organizações, projetos, pesquisadores, publicações e revistas implica a adoção de identificadores únicos e persistentes para cada uma dessas entidades.

No âmbito do PTCRIS, a Universidade do Minho trabalha com DOI para publicações, Ciência ID (identificador nacional de pesquisador) e ORCID para pesquisadores, Identificadores baseados no ISNI (International Standard Name Identifier) compatíveis com o Ringgold para Organizações, ISSN e ISBN para revistas e livros.

A atribuição de identificadores persistentes viabiliza a integração das informações, conforme mostrado na figura 2, viabilizando cruzamentos que são base para a gestão adequada dos resultados de projetos de C&T, objetivo de um CRIS.

Figura 2 – Integração de bases por meio de atribuição de identificadores únicos e persistentes a seus componentes



Não vimos, durante as discussões, avaliação definitiva sobre o uso ou não da ferramenta DSpaceCRIS, parecendo-nos que a tendência é continuar o uso do DSpace.

INTEGRAÇÃO COM A INFRAESTRUTURA DO OPENAIRE

Trabalhos de integração do RCAAP ao OpenAIRE estão sendo desenvolvidos, inclusive com uma versão beta que testa a integração das informações da rede La Referencia, o piloto sendo desenvolvido com informações do CONICYT, do Chile, selecionado por La Referencia para este experimento. Experiência semelhante poderia ser desenvolvida no contexto do BR CRIS, envolvendo, por exemplo, os dados oriundos dos INCTs.

POSSIBILIDADES DE COOPERAÇÃO

Entende-se que a Universidade do Minho está fortemente envolvida na construção de ambientes que viabilizarão a plataforma PTCRIS.

Assm, seria interessante que, sob a égide do acordo de cooperação já existente entre IBICT e FCT, fosse viabilizado o desenvolvimento conjunto de ferramentas comuns ao PTCRIS e BR CRIS, principalmente no que diz respeito à inserção das informações do OASIS-BR no OpenAIRE. Alternativa ou paralelamente, poder-se-ia trabalhar para a inclusão das informações relativas ao INCTs naquele ambiente.



CONCLUSÃO

Dos contatos realizados, parece-nos indubitável que deveríamos estreitar laços de cooperação com a FCT envolvendo desenvolvimento conjunto das plataformas PTCRIS e BRKRIS.

Pelo adiantado estado de desenvolvimento daquela primeira, acreditamos que a velocidade de desenvolvimento da segunda seria fortemente acelerada se tal cooperação fosse consagrada de maneira formal e com algum aporte financeiro nas seguintes direções:

- ✓ Desenvolvimento conjunto de uma base de instituições
- ✓ Alinhamento das políticas de acesso aberto do RCAAP e do OASIS-BR
- ✓ Integração dos dados brasileiros no OpenAIRE
- ✓ Integração da plataforma Lattes e do Ciência Vitae (uso do XML CERIF)
- ✓ Participação na iniciativa ORBIT
- ✓ Intercâmbio de boas práticas
- ✓ Compartilhamento de ferramentas com a Universidade do Minho e a rede La Referencia
- ✓ Desenvolvimento de uma base de projetos
- ✓ Esquema comum de identificação de revistas e livros

Além do projeto piloto que o IBICT está desenvolvendo com a FAPEAL (Fundação de Amparo à Pesquisa do Estado de Alagoas), uma iniciativa interessante, que poderia ser feita com a participação do CNPq, seria, como já foi dito, trabalhar com os resultados dos Institutos Nacionais de Ciência e Tecnologia (INCT), projetos estes apoiados pelo CNPq e que representam volume significativo de pesquisas e resultados em áreas de impacto no Brasil.

Os resultados de 122 INCTs estão disponíveis na Plataforma Carlos Chagas em relatórios armazenados em seus respectivos processos. Análise de apenas dois deles, mostra a criação, durante a vigência dos processos, de 879

produtos (517 artigos completos publicados em periódicos, 2 livros, 11 capítulos de livros, 168 trabalhos completos publicados em anais de congressos, 171 produtos envolvendo dissertações de mestrado e teses de doutorado, além de trabalhos aceitos para publicação em periódicos).

A compilação estruturada dessas produções poderia ensejar o desenvolvimento piloto das bases necessárias à estruturação do BR CRIS: pesquisadores (Lattes com ORCID, cpf e IdLattes); publicações (DOI); organizações (aposição de identificadores nas organizações participantes dos INCTs); projetos (identificação dos projetos representados pelos INCTs).

Para esse trabalho, poderiam ser convidados os próprios INCTs, para os quais a identificação de seus trabalhos, pesquisadores, revistas utilizadas e organizações participantes seria trivial (cada um deles teria que trabalhar apenas com relativamente pequeno número de entidades, em vez de um grupo apenas tentar organizar os dados de todos os 122 INCTs).

O resultado desse trabalho, a nosso ver, seria de grande utilidade para criação de ambiente dinâmico de acompanhamento e avaliação dos INCTs pelo CNPq (de forma que nos parece inédita), além de viabilizar a exposição internacional de seus resultados, dando visibilidade e significância aos trabalhos de desenvolvimento do BR CRIS.

Todo o material utilizado nas discussões na FCT/FCCN e Universidade do Minho está disponível na rede do CNPq.

Piloto BRCRIS IBICT-FAPEAL: uma avaliação exploratória do DSpace-CRIS para a construção de um sistema de informação de pesquisa local

**Kleber de Barros Alcanfôr
Hélia de Sousa Chaves
Lilian Maria Thomé Andrade Brandão
Rogério Mendes Castilho**

INTRODUÇÃO

São chamados de CRIS (Current Research Information Systems)¹ os sistemas de informação criados para estruturar e permitir a gestão de informações sobre todo o ciclo da pesquisa científica, abrangendo desde o projeto de pesquisa submetido pelo pesquisador até os resultados finais, frutos da pesquisa financiada – como publicações, patentes e outros produtos –, passando pelas instituições e pessoas envolvidas na pesquisa, pela infraestrutura – laboratórios e equipamentos – utilizada para sua realização, pelos dados gerados pela pesquisa, além de editais e eventos de divulgação da pesquisa.

Um sistema tipo CRIS tem por principal conceito ser um sistema de sistemas, que operam cada um em sua finalidade específica, com a integração necessária para a informação sobre a pesquisa. Essa integração se dá por meio de padrões internacionais utilizados, por exemplo, para identificação de pessoas, instituições, projetos e documentos, para compatibilização da terminologia utilizada na gestão da pesquisa e para definição de metadados. Esses padrões garantem a visibilidade internacional da produção científica de um país e, por consequência, de todos os envolvidos em sua realização. Sistemas CRIS podem ser utilizados para a organização de sistemas de interesse nacional, local, institucional ou de uma área específica, os chamados CRIS temáticos.

¹ Traduzido para o português como “Sistema de Informação de Pesquisa Corrente”. Nesse contexto, o termo “corrente” define pesquisa de interesse atual, estejam elas em andamento ou já tenham sido concluídas.

São diversos os públicos beneficiários desse tipo de sistema, a começar pelos próprios pesquisadores, gestores de ciência e tecnologia (incluindo agências de fomento), empresários, organizações de transferência de tecnologia, bibliotecas, mídia e o público em geral. Isso porque, por meio de sistemas CRIS é possível:

- ✓ ter acesso a pesquisas financiadas com recursos públicos ou privados;
- ✓ identificar novas oportunidades de financiamento de pesquisas;
- ✓ identificar potenciais parcerias para o desenvolvimento de pesquisas;
- ✓ fazer o melhor acompanhamento das pesquisas financiadas, evitando-se duplicações;
- ✓ otimizar a distribuição de verbas para pesquisa;
- ✓ avaliar tendências de pesquisa em âmbitos local, nacional e internacional;
- ✓ fazer gestão dos dados de pesquisa oriundos dessas pesquisas e, conseqüentemente, promover o reuso desses dados e a gestão de resultados de forma eficiente;
- ✓ criar redes de pesquisa e compartilhamento;
- ✓ identificar potencial de inovação das pesquisas desenvolvidas; e
- ✓ preservar a memória da pesquisa do País.

O principal fundamento dos sistemas CRIS é a padronização do formato de estruturação da informação, por meio do CERIF (Common European Research Information Format), padrão europeu desenvolvido, nos seus primórdios, pela Comissão Europeia para definição de metadados de informação de pesquisa. O CERIF foi transferido à euroCRIS, uma associação internacional sem fins lucrativos, fundada em 2002, que tem por missão “[...] promover a colaboração no âmbito da comunidade que lida com informação de pesquisa e promover a interoperabilidade avançada através do CERIF.”² O uso do CERIF é incentivado pela União Europeia para que seus países membros alcancem competitividade máxima em todas as frentes da atividade de pesquisa.

² euroCRIS. What is euroCRIS? Disponível em: <<https://www.eurocris.org/what-eurocris>>.

O Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT), unidade de pesquisa do Ministério da Ciência, Tecnologia, Inovações e Comunicações (MCTIC), é o órgão nacional de informação dedicado a oferecer informação de qualidade na *web* com a missão de “[...] promover a competência, o desenvolvimento de recursos e a infraestrutura de informação em ciência e tecnologia para a produção, socialização e integração do conhecimento científico e tecnológico”³. O IBICT tem atuado no desenvolvimento de pesquisas, serviços e produtos de informação, com expertise na criação de sistemas de informação que incorporam padrões internacionais e promovem a interoperabilidade com outros sistemas de instituições parceiras. Com base nessa experiência e à luz de sua missão, o IBICT identificou a importância dos sistemas CRIS para organização e acesso a informações sobre a atividade de pesquisa em ciência e tecnologia no País e sua capacidade de contribuir com essa ação.

Nesse sentido, entre 2014 e 2017, o Instituto investiu na realização de estudos, levantamentos e análises sobre os sistemas CRIS, abrangendo, entre outros assuntos, o modelo de dados, o conjunto de ferramentas utilizadas para administrar sistemas CRIS, iniciativas estrangeiras em sistemas CRIS, ferramentas de indexação, quadro referencial teórico e da metodologia a ser adotada para integrar sistemas de informação⁴. Promoveu a realização de *workshops* e capacitações, em Brasília. Formalizou parcerias com instituições brasileiras de fomento à pesquisa e com organizações europeias líderes no desenvolvimento de sistemas CRIS: Fundação de Ciência e Tecnologia (FCT) de Portugal, (2015) e a euroCRIS (2016).

Em maio de 2017, o IBICT submeteu proposta de projeto à 9ª Convocatória da Iniciativa de Apoio aos Diálogos Setoriais União Europeia Brasil, no âmbito do Diálogo Sociedade da Informação, para o desenvolvimento de um projeto piloto BRICRIS com dados sobre pesquisas financiadas pela Fundação de Amparo à Pesquisa do Estado de Alagoas (FAPEAL).

³ INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA – IBICT. Missão. Brasília, DF: 2018.

⁴ Foram sete os documentos técnicos sobre os temas relacionados a sistemas CRIS elaborados pela consultora do IBICT Maria de Nazaré Freitas Pereira. Ver lista de referências no final deste relatório.

Em dezembro de 2017, IBICT, Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Conselho Nacional das Fundações Estaduais de Amparo à Pesquisa (CONFAP), Scientific Electronic Library Online (SciELO) e Rede Nacional de Ensino e Pesquisa (RNP) se uniram para formação de um consórcio nacional com a finalidade de propiciar a inserção do código identificador ORCID ID nas bases de dados dessas instituições, de forma a conectar os dados relacionados aos pesquisadores e suas produções científicas a outros sistemas internacionais. Essa iniciativa representou mais um passo em direção à concretização do projeto BR CRIS.

O projeto TICS0114, financiado pela Iniciativa de Apoio aos Diálogos Setoriais União Europeia-Brasil, proporcionou o aprofundamento dos estudos conceituais até então realizados pelo IBICT sobre o padrão CERIF, o principal fundamento para a padronização da informação a ser compartilhada e integrada em um sistema CRIS. Trata-se de um assunto complexo, que apresenta vasta documentação, com abordagem conceitual sobre as entidades que integram o sistema, suas relações, características de categorização e classificação, além da especificação de interoperabilidade.

Neste trabalho abordamos a criação de um espaço de experimentação de um conjunto de informações sobre pesquisas financiadas com recursos públicos, que pudessem abranger a maior quantidade de características típicas de um sistema do tipo CRIS.

O universo de pesquisas definido para este experimento é no âmbito do fomento federal. Na primeira exploração dos dados, foi observado que seria mais consistente escolher um programa que contivesse as características desejadas em termos de conteúdo. Nesse sentido, a FAPEAL, que já vinha participando das iniciativas CRIS conduzidas pelo IBICT, se candidatou à condição de piloto, por ocasião da Jornada euroCRIS em novembro de 2016. Desta forma, a FAPEAL sugeriu como fonte de dados o Programa Pesquisa para o SUS (PPSUS), uma iniciativa de fomento do Governo Federal, operado em parceria direta com as agências de fomento estaduais, cofinanciado por agências como CNPq, CAPES e FINEP.

Esse recorte contemplou as principais características desejadas para o piloto e abrange as pesquisas do estado de Alagoas operadas pela FAPEAL no âmbito do Programa PPSUS⁵.

O resultado esperado nesse caso foi construir um sistema CRIS local e ao mesmo tempo especializado em saúde, na modalidade projetos de extensão. O foco foi em aplicação prática dos conceitos estudados em uma avaliação exploratória dos aspectos e características particulares de uma agência de fomento estadual e suas relações com o sistema nacional de fomento.



⁵ Os dados foram extraídos do portal PPSUS. (BRASIL. Ministério da Saúde).

MATERIAL E MÉTODO

O material para popular o DSpace-CRIS, ferramenta escolhida para o piloto, consistiu das fontes de dados descritas a seguir.

Inicialmente, as fontes de dados eleitas para exame foram a base de dados do SIGFAPEAL⁶ e uma base de dados de registros internos da FAPEAL, em Access, contendo informações não disponíveis no SIGFAP. Após minuciosa análise desse conteúdo, constatamos que esses dados continham informações basicamente voltadas à gestão dos projetos de pesquisa, sob o ponto de vista de prestação de contas dos recursos financiados, sem grandes aprofundamentos com relação à pesquisa propriamente dita. Adotar essa fonte limitaria bastante o processo de experimentação, não sendo possível alcançar nem mesmo as informações das entidades base previstas no modelo CERIF. Assim, as principais fontes de informação utilizadas na coleta de dados foram o arquivo, em formato Access, contendo a base de dados fornecida pela FAPEAL, o sistema Pesquisa Saúde do Ministério da Saúde no âmbito do PPSUS no estado de Alagoas, e arquivos em formato Excel fornecidos pela CAPES. Para complementação e enriquecimento dos dados, foram feitas consultas na Internet (Plataforma Lattes, Google Scholar e ISNI).

O método adotado para este estudo considera como base os recursos oferecidos pela ferramenta de sistematização CRIS, DSpace-CRIS, para estabelecer a forma de organização dos dados, tratamento e registro. Nesse caso, o registro pode ocorrer de duas formas: por meio de processo automatizado ou entrada de dados através dos formulários eletrônicos. Outro aspecto que influenciou bastante a metodologia foi a especificação do formato CERIF, não somente quanto a seus conceitos, mas também devido às mudanças que ocorreram na evolução de suas versões.

⁶Sistema de gestão de projetos de pesquisa da FAPEAL, parte da plataforma computacional intitulada Sigfap (Sistema de Informação e Gestão de Projetos da FAP), criada em 2002 para permitir a gestão dos projetos de pesquisa e inovação, e para facilitar o trâmite desde a submissão da proposta até a fase de prestação de contas (parcial e final). Cf. TURINE, Marcelo A.S.; CARROMEU, Camilo; SILVA, Márcio A.I. da; CAGNIN, Maria Istela. Gestão pública flexível e ágil por meio do Sigfap. *ComCiência*, Revista Eletrônica de Jornalismo Científico. 10 jun. 2011. Disponível em: <<http://www.comciencia.br/comciencia/handler.php?section=8&edicao=67&id=851>>.

FORMATO CERIF

Compreender a evolução dos conceitos das entidades e suas relações no decorrer dos anos, expostas nas seis últimas versões do CERIF, foi a grande fonte de inspiração para elaborarmos o processo de coleta, tratamento e preparação dos dados para este piloto. Para melhor compreensão da evolução desse padrão, apresentamos a seguir duas imagens representativas da estrutura do CERIF, em versões em cujos modelos de dados foram feitas mudanças expressivas.

Na versão CERIF2006, publicada em outubro de 2007, as principais entidades foram estabelecidas como Pessoas, Unidades Organizacionais, Projetos e Publicações; os autorrelacionamentos já eram reconhecidos e o modelo enfatizava a produção científica como foco de interesse nos resultados da pesquisa (figura 1).

Figura 1 - CERIF2006-FDM, Out/2007

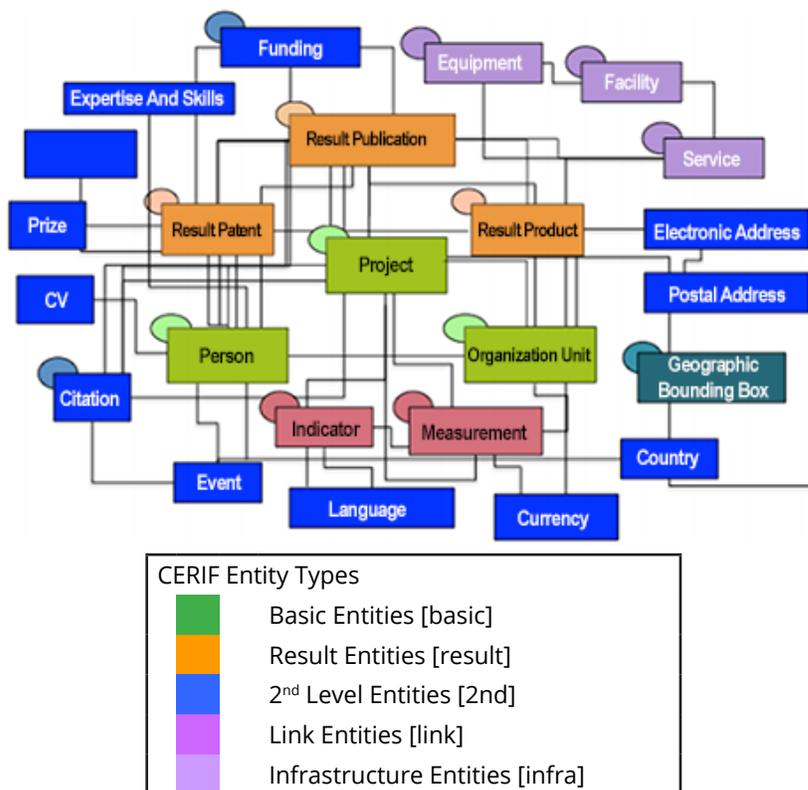


Extraído de: JÖRG, Brigitte *et al.* **CERIF2006-1.1 Full Data Model (FDM): Model Introduction and Specification.** euroCRIS, 17 out. 2007. Disponível em: <https://www.eurocris.org/Uploads/Web%20pages/CERIF2006/CERIF2006_FDM_1.1.pdf>.

Dos padrões internacionais de estruturação da informação de pesquisa aos indicadores: primeira incursão na temática Parte IV Relatórios| Piloto BRCRIS IBICT-FAPEAL: uma avaliação exploratória do DSpace-CRIS para a construção de um sistema de informação de pesquisa local

Já na versão CERIF1.3 foram incorporadas grandes mudanças em sua estrutura conceitual, onde **Pessoas**, **Unidades Organizacionais** e **Projetos** passam a ser entidades base, e foi criada uma camada de entidades de resultado, com **Publicações**, **Patentes** e **Produções Tecnológicas**. Várias outras novas camadas foram estabelecidas, conformando o modelo atual do CERIF. No entanto, tais camadas não serão objeto deste estudo. Por se tratar de uma primeira aplicação prática, foram contempladas as entidades base e de resultados relacionados às pesquisas operadas pela FAPEAL no âmbito do PPSUS. Esse novo modelo está representado na figura 2.

Figura 2 - CERIF1.3-FDM, Jan/2012



Extraído de: JÖRG, Brigitte et al. CERIF 1.3 full data model (FDM): introduction and specification. euroCRIS, 28 Jan. 2012. 52 p. Disponível em: <https://www.eurocris.org/Uploads/Web%20pages/CERIF-1.3/Specifications/CERIF1.3_FDM.pdf>.

METADADOS

A definição dos metadados para composição da estrutura de informação do piloto foi feita, inicialmente, com base nos dados disponíveis no PPSUS e, posteriormente, complementadas com aqueles fornecidos pela FAPEAL. Nesse contexto, as principais informações foram organizadas a partir dos conceitos do CERIF, iniciando pelas entidades de base: pesquisador (pessoa), instituição (unidade organizacional) e projeto; seguidas pelas secundárias, estas referentes à produção científica, conforme detalhamento a seguir.

Pesquisador: nome completo, língua natal, nome para citação, variações do nome, instituição de vínculo, grupo de pesquisa, site pessoal, e-mail, outros e-mails, biografia e formação acadêmica.

Instituição: nome, sigla, nível organizacional, nome do diretor/presidente, data de criação, conselho científico, ISNI⁷, endereço, bairro, cidade, estado e país.

Projeto de Pesquisa: título, código, coordenador, instituição, data de início, data término, área de conhecimento CNPq (área, subárea, especialidade), CASRAI (Disciplina [Setor, Campo, Nome], Assunto da pesquisa [Campo, Nome], Áreas de aplicação [Campo, Subcampo]), Descrição do projeto (Palavras-chave, Abstract, Resumo para o público, Aplicabilidade ao SUS, Objetivo da pesquisa, Metodologia, Resultados encontrados e Recomendações para o SUS). Vale observar que esses dois últimos campos decorrem da natureza dos projetos de extensão do PPSUS.

⁷ O ISNI é o número padrão internacional certificado pela "ISO 27729: International Standard Name Identifier", para identificar milhões de colaboradores de trabalhos criativos, assim como aqueles que atuam em sua distribuição, como pesquisadores, inventores, escritores, artistas, ilustradores, produtores, editores, ilustradores, entre outros. O código ISNI faz parte de uma família de identificadores padrão internacionais, como DOI, ISAN, ISBN, ISRC, ISSN, ISTC, e ISWC. Em agosto de 2017, a Agência Internacional do ISNI (ISNI-IA) anunciou mudanças em sua infraestrutura para abranger o fornecimento de identificadores para organizações atuantes no campo da comunicação acadêmica. Cf. INTERNATIONAL STANDARD NAME IDENTIFIER (ISO 27729) – ISNI. Site. Disponível em: <<http://www.isni.org/>>.

Produção científica: autores, título, subtítulo, data da publicação, editora, citação, número de série, identificadores (DOI, ISBN etc.), tipo de publicação, idioma, palavras-chave, resumo, parcerias, projeto vinculado, periódico, conferência e descrição.

FERRAMENTAS E INFRAESTRUTURA

A principal ferramenta utilizada foi DSpace-CRIS versão 5.8, Docker⁸ versão 18.06.1- e Data Integration da Suite Pentaho versão 7.1, uma solução recomendada pela euroCRIS, em código aberto e uso livre, que possui uma abordagem flexível e configurável, permitindo, assim, construir um espaço para a experimentação do conteúdo de informação de pesquisa em sistemas CRIS.

Para a instalação e configuração inicial, foi utilizado o Guia de Instalação do DSpace-CRIS publicado pelo IBICT em 2016. Após a conclusão da instalação inicial, foram feitas várias tentativas de configuração de ambiente em uma máquina local e ficou evidente que a ferramenta apresentava uma grande complexidade nos recursos de configuração. Percebemos, então, que seria necessário ter a possibilidade de reconstruir completamente o ambiente a qualquer momento e garantir total liberdade para experimentar os recursos de configuração dos metadados das entidades do sistema CRIS. Assim, buscamos uma solução para gerenciar ambientes de servidores com possibilidade de ter arquivos de imagem, contendo as pré-configurações, e restaurá-los a qualquer momento.

O próximo passo foi automatizar o processo de instalação, de forma que fosse possível criar versões de configuração e carga de dados. Para isso, utilizamos a ferramenta Docker e conseguimos construir um processo automático que possibilitou que uma instância dessa instalação, chamada “imagem”, pudesse ser criada e apagada inúmeras vezes. Fato esse que foi muito importante para que pudéssemos chegar a uma configuração que atendesse ao projeto e às suas particularidades. Além disso, essa ferramenta cria *containers* e viabiliza a instalação do *software*, de forma

⁸ Docker é uma tecnologia de software que fornece containers. Fornece uma camada adicional de abstração e automação de virtualização de nível de sistema operacional no Windows e no Linux.

que seu ambiente seja seguro, confiável, e possua um processo simples de recuperação e distribuição do projeto para ser instalado em qualquer outro ambiente para compartilhamento deste experimento ou mesmo para aproveitar sua estrutura na criação de um novo.

Na sequência, o IBICT disponibilizou um ambiente de *hardware* com uma configuração simples de uma máquina virtual dedicada (CPU 16 Core, 2.793MHz, 16Gb RAM e 250GB de HD) disponível para a *web*. Através do Docker, foram criadas várias instâncias⁹ de uma instalação padrão do DSpace-CRIS já com acesso compartilhado na *web*.

Outra ferramenta que utilizamos foi o PDI (Pentaho Data Integration), parte da Suíte Pentaho, que é um conjunto de componentes de *software* desenvolvidos originalmente para construção de soluções de *Business Intelligence* (BI). O PDI é utilizado para criar processos de extração, transformação e carga (do inglês *Extraction, Transformation and Loading* (ETL), que alimentam o banco de dados. Com essa ferramenta, é possível fazer inúmeras operações de Integração de dados, como por exemplo: migração de dados; movimentação de grandes volumes de dados; transformação de dados; limpeza de dados; conformidade de dados.

CAMPOS PERSONALIZADOS

Outro passo importante foi a criação de campos personalizados para complementação das informações constantes da instalação padrão do DSpace-CRIS, conforme o elenco de dados definido. Essa é uma das grandes características dessa ferramenta e neste estudo conseguimos realizar todas as personalizações necessárias.

Para cada entidade existe uma configuração de interface (caixas de dados e campos que podem ser criados e agrupados em diversas configurações), que não será abordada em detalhes neste documento, mas não poderia deixar de ser citada. A seguir, apresentamos a figura 3, que consiste na montagem de partes de telas extraídas do ambiente de administração do DSpace-CRIS para retratar esses recursos.

⁹ Nesse caso, instância é um processo utilizado para criar uma cópia de uma certa configuração completa.

Figura 3 - Imagem capturada da interface do DSpace-CRIS

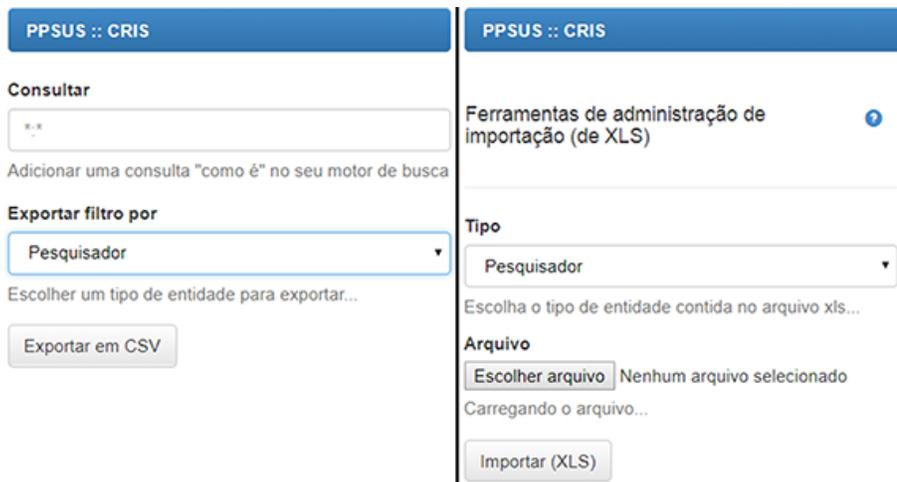


Extraído de: DURASPACE. **DSpace Release 7.0 Status:** version 7.0. Created by Ivan Masár, last modified by Tim Donohue, June 2018. Disponível em: <<https://wiki.duraspace.org/display/DSPACE/DSpace+Release+7.0+Status#DSpaceRelease7.0Status-WhatwillbeincludedinDSpace7?>>.

Outro fator determinante foi a disponibilidade de recursos pelo DSpace-CRIS. Contudo, na versão do projeto ainda não possui API (DURASPACE, abr. 2018) para integração de dados. A implementação é feita a partir de rotinas de exportação e importação (DURASPACE, 2017), em formato Excel, e arquivos individuais por entidade. Vale ressaltar que a versão DSpace 7, prevista para 2019, já incorpora integração por API em suas funcionalidades (DURASPACE, jun. 2018).

Essa limitação foi decisiva para estabelecermos o processo de carga dos dados coletados. Esse recurso está disponível na opção “Módulos CRIS” no menu principal do ambiente Administrador. Ao selecionar “exportação” ou “importação”, aparecerá a opção para escolher a entidade. É importante observar que o processo de carga ou extração ocorre individualmente, por entidade, conforme exemplificado na figura 4.

Figura 4 - Imagens capturadas da interface do DSpace-CRIS



Extraído de: DURASPACE. **DSpace Release 7.0 Status:** version 7.0. Created by Ivan Masár, last modified by Tim Donohue, June 2018. Disponível em: <<https://wiki.duraspace.org/display/DSPACE/DSpace+Release+7.0+Status#DSpaceRelease7.0Status-WhatwillbeincludedinDSpace7?>>>.

Nesse caso, o primeiro passo foi criar um registro de exemplo para cada entidade e gerar uma exportação para analisar o formato do arquivo. A importação funciona da mesma forma acrescentando-se duas colunas, uma no início, com o título ACTION, cujo preenchimento ocorre conforme a ação desejada: INSERT, DELETE, UPDATE, SHOW ou HIDE; e outra coluna com título NONE com conteúdo “#” para representar o final da linha para importação. Assim, através da inclusão de conteúdo pela interface do DSpace-CRIS, observação das mudanças no arquivo de exportação, inclusão de dados no arquivo Excel e importação no sistema, foi possível estabelecer o modelo de integração dos dados coletados.

A partir dessa experiência, o desafio foi o de alcançar a integração dos dados das entidades base. Isso porque, conforme o CERIF, fica evidente que um sistema CRIS sem que essas três entidades estejam coletadas, qualificadas e relacionadas fica sem sentido.

TRATAMENTO, REGISTRO E ENRIQUECIMENTO DOS DADOS

Neste estudo tivemos a oportunidade de experimentar a ferramenta DSpace-CRIS nos processos de configuração de metadados, carga automática e registro manual, através dos formulários da interface web.

A forma de tratamento de dados é individualizada para cada entidade de informação (Pesquisador, Instituição, Projeto e Produção Científica), e a sua ordem deve respeitar a necessidade de construção dos relacionamentos. O processo de integração de informações sobre pesquisa em um sistema CRIS inclui identificação, tratamento, pesquisa, qualificação e enriquecimento dos dados. A seguir, apresentamos como aconteceu nosso experimento.

A primeira experiência consistiu na exploração do processo de carga automática, onde foram registrados 4670 pesquisadores e 170 instituições coletadas dos arquivos da FAPEAL. Esse processo foi bastante interessante pelo aprendizado. Quanto à coleta automática e ao tratamento, utilizamos a ferramenta PDI para realizar as transformações, conforme exemplificado na figura 5, e o recurso de importação do DSpace-CRIS em formato Excel para carregar os dados. A figura 6 mostra a tela inicial desse experimento.

Figura 5 - Processo de tratamento de dados sobre pesquisadores

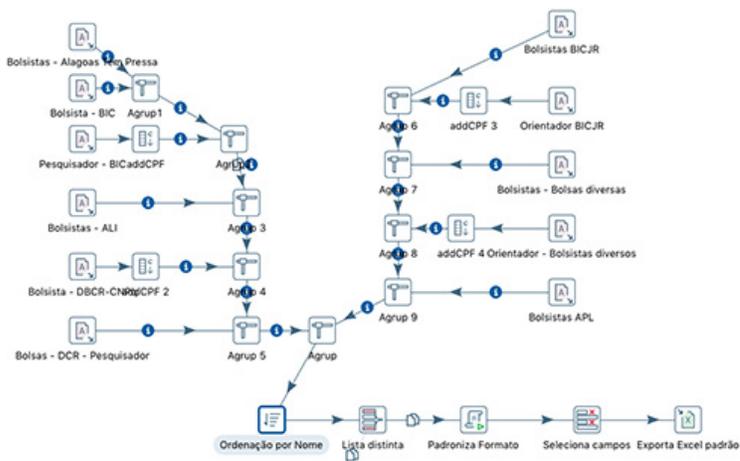
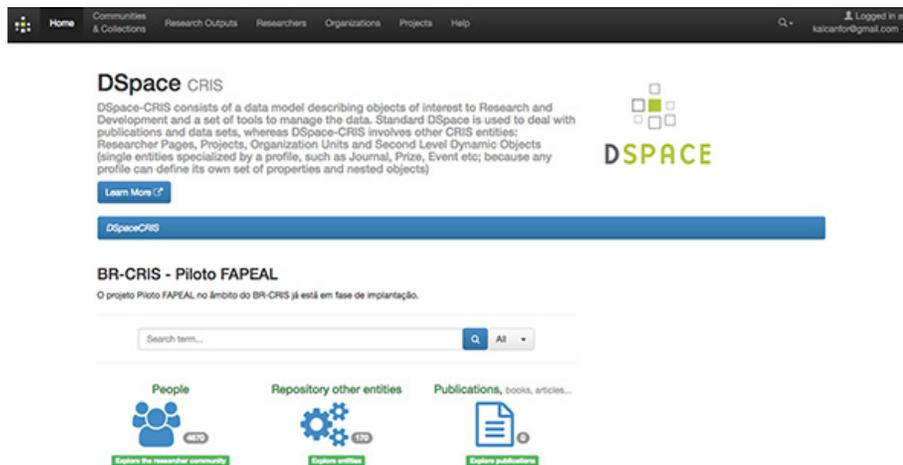


Figura 6 – Tela inicial DSpace-CRIS (primeira experiência)



Conforme citado anteriormente, durante o experimento, o foco deste estudo foi redirecionado para os projetos financiados no contexto do programa PPSUS no estado de Alagoas. Dessa forma, os dados da primeira carga não foram aproveitados.

Nesse novo contexto, todo o processo foi reiniciado a partir de uma nova instalação padrão do DSpace-CRIS. Vale observar que a forma de identificar o escopo do estudo é determinada pela entidade **Projetos**, pois a intenção é registrar instituições, equipe e os resultados decorrentes deles. O critério de seleção dos projetos foi comparar a base entregue pela Fapeal e os dados do PPSUS/Pesquisa Saúde.¹⁰

Após a seleção dos projetos, o foco de tratamento passa a ser as outras entidades base: Pesquisador e Instituição. O processo de tratamento inicialmente teve como referência o sistema Pesquisa Saúde e os arquivos da FAPEAL e foram registrados manualmente no DSpace-CRIS, através dos formulários eletrônicos; em um segundo momento, houve um processamento automatizado para tratar os dados fornecidos pela CAPES.

¹⁰ <http://pesquisasaude.saude.gov.br/pesquisas.xhtml>

A seguir apresentamos os tratamentos realizados no primeiro momento:

Pesquisador: para tratar essa entidade, o primeiro objetivo a ser alcançado foi separar uma lista distinta de todas as pessoas envolvidas de alguma forma nos projetos selecionados. No entanto, vale observar que, como o PPSUS registra apenas o coordenador do projeto de pesquisa e não contempla a equipe técnica, foram cadastrados os coordenadores de projetos, tal como registrados. Um esforço adicional diz respeito ao registro das variações dos nomes dos autores e, para isso, fizemos uso do recurso de desambiguação de nomes do DSpace-CRIS. O sistema permite o registro de até três variações de nomes. Todos os dados dos pesquisadores cadastrados no sistema foram investigados, para identificação daqueles que já possuíam ORCID ou qualquer outro identificador persistente. Além do ORCID, foram localizados Scopus, Author ID, Researcher ID e Lattes ID, e todos esses identificadores foram devidamente registrados nos respectivos perfis dos pesquisadores. Outro esforço realizado foi a complementação de e-mail e nome de citação, quando necessário.

Instituição: foram consideradas as instituições onde o projeto está sendo desenvolvido. O código ISNI teve como fonte principal seu site oficial de busca.

Projeto: os dados de projetos de pesquisa foram extraídos do sistema Pesquisa Saúde que disponibiliza muitas informações (resumo, palavras-chave, texto, resumo para o público, aplicabilidade para o SUS entre outros), e complementados com os arquivos da FAPEAL. Uma amostra com 32 projetos foi selecionada para experimentar a classificação CASRAI e área de conhecimento do CNPq. Os projetos foram classificados pela Profa. Hagar Espanha Gomes e devidamente registrados no DSpace-CRIS.

Produção Científica: a produção científica dos pesquisadores foi extraída dos dados do Sistema Pesquisa Saúde. Tais dados foram complementados com buscas na Internet. Essas informações foram as mais trabalhosas, vez que as referências das respectivas produções científicas descritas pelos pesquisadores estavam incompletas e com dados inconsistentes. Isso aconteceu porque o campo a ser preenchido era aberto, sem padronização, não exigindo nenhum critério de inclusão. Alguns pesquisadores preencheram com informações completas, outros apenas com informações quantitativas de sua produção, sem especificar nenhuma. Isso custou à equipe uma extensa pesquisa para garantir a qualidade dos dados. Além disso, para registro de uma produção científica no DSpace-CRIS, existe a obrigatoriedade de anexar o documento, tornando-se necessário localizar cada documento citado pelo autor.

Em um segundo momento, após alcançar os primeiros resultados, a CAPES decidiu cooperar cedendo dois arquivos, em formato Excel, com extração dos projetos do estado de Alagoas: um arquivo contendo projetos referentes ao período de 2008 a 2012, extraído do Coleta CAPES, e outro contendo projetos referentes ao período de 2013 a 2016, extraído da Plataforma Sucupira. Esses arquivos continham basicamente informações sobre os **projetos, equipe e financiadores**, o que proporcionou nova oportunidade para enriquecimento dos dados através de processamento automático. Nesse contexto, apresentamos a seguir os esquemas dos processos desenvolvidos com a ferramenta PDI para extração e transformação dos dados desses arquivos.

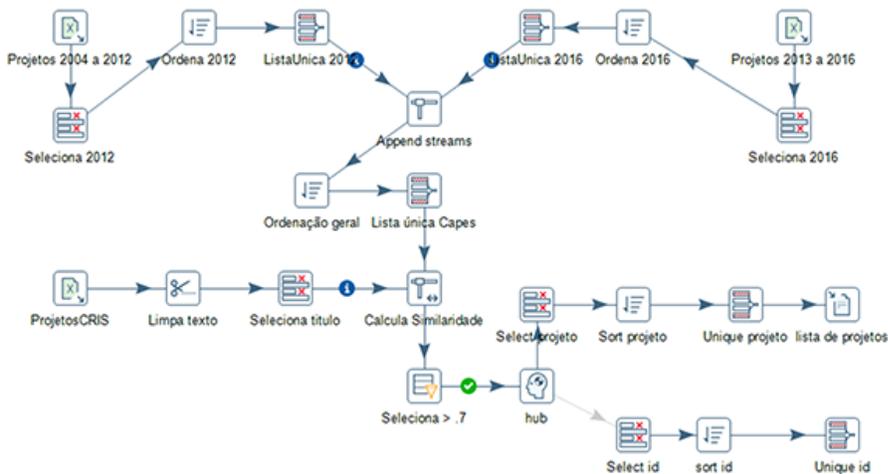
A figura 7 apresenta o processo macro, chamada de *Job* na PDI, que organiza a sequência das transformações e apresenta uma visão geral. Nesse processo foram identificados 57 projetos, 275 membros de equipe, três instituições parceiras e diversas outras instituições financiadoras. Destacamos que os dados não foram incorporadas no DSpace-CRIS por falta de validação pelos outros parceiros deste estudo.

Figura 7 - Processo Geral



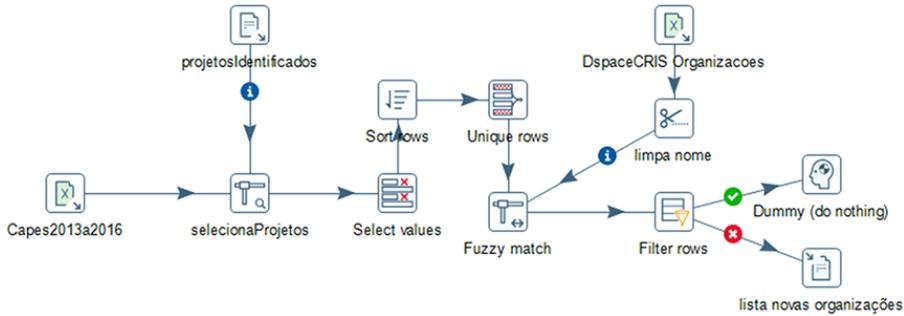
O primeiro passo para o processamento automático foi identificar quais os projetos da base CAPES que correspondiam aos projetos selecionados para este estudo. Na figura 8 podemos destacar que foi necessário tratamento para compatibilizar algumas diferenças entre os dois arquivos de dados dos projetos que retratam a migração para a Plataforma Sucupira. A forma de identificação foi realizada pelo cálculo de similaridade no campo do título comparado com os dados já registrados no DSpace-CRIS. Foram identificados 57 projetos como de interesse.

Figura 8 - Identificação dos Projetos



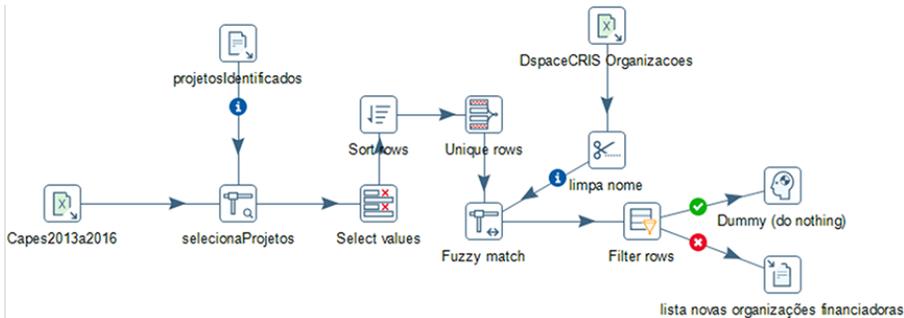
O passo seguinte consistiu em identificar quais outras instituições tiveram algum tipo de envolvimento com os projetos selecionados na base CAPES, exceto no papel de financiadores (figura 9). Apenas três instituições foram identificadas.

Figura 9 - Identificação de novas instituições



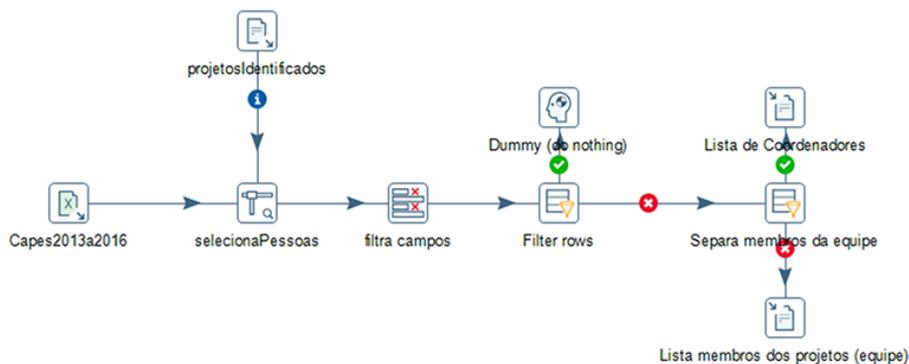
A terceira transformação, apresentada na figura 10, teve por objetivo identificar especialmente outras instituições financiadoras desses projetos. Nesse caso, foram identificadas várias instituições.

Figura 10 - Identificação de novas instituições financiadoras



Por último, o objetivo foi identificar membros de equipes dos projetos, nesse caso com grande expectativa para enriquecimentos dos dados coletados no primeiro momento, pois a única informação disponível até então era sobre o coordenador do projeto. Essa transformação, apresentada na figura 11, identificou 253 membros e destacou os coordenadores dos respectivos projetos, fato que representa expressiva base de verificação desses dados.

Figura 11 - Identificação de membros de equipe



Ao final desse processamento automatizado, tudo indica que é grande a possibilidade de aproveitamento da base CAPES para processo de enriquecimento dos dados neste estudo. Contudo, para ter real precisão sobre os quantitativos e incorporação efetiva, teremos que aguardar a validação pelos nossos parceiros.



CONSIDERAÇÕES FINAIS

Em primeiro lugar, entendemos que este estudo alcançou o objetivo desejado, qual seja o de construir um espaço de experimentação de um sistema CRIS local e temático, quando se propôs a contemplar os projetos de pesquisa financiados pelo PPSUS no âmbito da FAPEAL.

O exercício de planejar, mesmo que a título de experimento, um sistema de informação sobre pesquisa, com foco em aplicação prática do formato CERIF, constrói referência conceitual a ser alcançada. Passamos por diversas dificuldades, superamos muitas delas e não alcançamos várias expectativas que tínhamos no início. Esse é o pano de fundo de nossas considerações.

O aprendizado da ferramenta DSpace-CRIS não é trivial, quanto à infraestrutura e configuração inicial, mas especialmente quanto ao mapeamento dos metadados das entidades e o comportamento nos formulários e suas integrações. Ao mesmo tempo nos sentimos recompensados pelo aprendizado em cada esforço dedicado.

O primeiro fundamento é construir as entidades base: pesquisador, instituição e projeto. É importante definir seus próprios identificadores persistentes e registrar outros identificadores externos, para estabelecer possibilidades de integração e interoperabilidade, como por exemplo, ORCID ID, ISNI, RINGOLD, SCOPUS, ID LATTES, entre outros. Nossa percepção nos diz que quem comanda o escopo principal de um CRIS é a entidade projeto; nela estão aplicados os esforços de pessoas e instituições, constitui objeto de financiamento e a quem são atribuídos os resultados.

Os resultados quanto à interoperabilidade no DSpace-CRIS foram bastante superficiais, limitando-se à integração com o identificador único ORCID ID, a qual se resumiu exclusivamente na autenticação de usuário. A utilização do protocolo OAI-PMH e a exposição de dados por *webservices* deverão ser experimentadas em outro contexto. Outro resultado que deverá ser mais bem desenvolvido é a personalização das interfaces de busca e apresentação de dados e de painéis quantitativos. Estamos depositando muitas expectativas em sua nova versão, o DSpace 7, que prevê a implantação de API REST e interface em Angular - Javascript, além de outras grandes novidades que devem estar disponíveis em 2019.

Por fim, a experiência de estudar os modelos de dados e tentar construir as entidades e suas relações, a partir de dados coletados em bases criadas em contextos diversos e com objetivos individuais, nos faz perceber a coerência e a importância dos conceitos propostos pelo padrão CERIF. É como se fosse o reconhecimento da falta de uma cultura, algo que depois de compreendido se tornasse óbvio no mesmo instante.



REFERÊNCIAS BIBLIOGRÁFICAS

BRASIL. Ministério da Saúde. Departamento de Ciência, Tecnologia e Insumos Estratégicos. **Sistema Pesquisa Saúde**. Disponível em: <<http://www.saude.gov.br/pesquisasaude>>.

DURASPACE. **DSpace Release 7.0 Status**: version 7.0. Created by Ivan Masár, last modified by Tim Donohue, June 2018. Disponível em: <<https://wiki.duraspace.org/display/DSPACE/DSpace+Release+7.0+Status#DSpaceRelease7.0Status-WhatwillbeincludedinDSpace7?>>.

DURASPACE. **DSpace-CRIS Home**. In: Wiki Duraspace. Disponível em: <<https://wiki.duraspace.org/display/DSPACECRIS/DSpace-CRIS+Home>>.

DURASPACE. **Interoperability with external systems and databases**. Created by Michele Mennielli, last modified by Claudio Cortese (4Science), abr. 2018. Disponível em: <<https://wiki.duraspace.org/display/DSPACECRIS/Interoperability+with+external+systems+and+databases>>.

DURASPACE. **XLS Import and export tool**. Created by Andrea Bollini (4Science), last modified by Luigi Andrea Pascarelli (4Science), nov. 2017. Disponível em: <<https://wiki.duraspace.org/display/DSPACECRIS/XLS+Import+and+Export+too>>.

EUROCRIS. **What is euroCRIS?** Disponível em: <<https://www.eurocris.org/what-eurocris>>.

GOMES, Hagar Espanha. **Estudo CASRAI – Compatibilidade de sistemas**. Produto 1 do projeto: Dos padrões internacionais de estruturação da informação de pesquisa aos indicadores (BRCRIS). Rio de Janeiro, 2017. 32 p. IBICT - Iniciativa de Apoio aos Diálogos Setoriais União Europeia/Brasil.

GOMES, Hagar Espanha. **Estudo CASRAI – Estudo com tabelas CNPq, FoS e ISCED**. Produto 2 do projeto: Dos padrões internacionais de estruturação da informação de pesquisa aos indicadores (BRCRIS). Rio de Janeiro, 2017. 32 p. IBICT - Iniciativa de Apoio aos Diálogos Setoriais União Europeia/Brasil.

HITACHI VANTARA. **Data Integration – Kettle**. Disponível em: <<https://community.hitachivantara.com/docs/DOC-1009855-data-integration-kettle>>.

HITACHI VANTARA. **Pentaho**. Disponível em: <<https://sourceforge.net/projects/pentaho/>>.

INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA – IBICT. **Missão**. Brasília, DF:2018. Disponível em: <<http://www.ibict.br/sobre-o-ibict-1>>.

INTERNATIONAL STANDARD NAME IDENTIFIER (ISO 27729) – ISNI. **Site**. Disponível em: <<http://www.isni.org/>>.

JÖRG, Brigitte *et al.* **CERIF 1.3 full data model (FDM)**: introduction and specification. euroCRIS, 28 jan. 2012. 52 p. Disponível em: <https://www.eurocris.org/Uploads/Web%20pages/CERIF-1.3/Specifications/CERIF1.3_FDM.pdf>.

Dos padrões internacionais de estruturação da informação de pesquisa aos indicadores: primeira incursão na temática Parte IV Relatórios| Piloto BRCRIS IBICT-FAPEAL: uma avaliação exploratória do DSpace-CRIS para a construção de um sistema de informação de pesquisa local

JÖRG, Brigitte *et al.* **CERIF2006-1.1 Full Data Model (FDM): Model Introduction and Specification.** euroCRIS, 17 out. 2007. Disponível em: <https://www.eurocris.org/Uploads/Web%20pages/CERIF2006/CERIF2006_FDM_1.1.pdf>.

JÖRG, Brigitte. CERIF 2008 – 1.2: The common European research information format. *In:* euroCRIS Membership Meeting, Prague, Czech Republic, nov. 2010. **Presentation...** Prague, Czech Republic, 2010. 76 slides. Disponível em: <<https://slideplayer.com/slide/7566334/>>.

PEREIRA, Maria de Nazaré Freitas. **Descrição de modelo de dados, conjunto de ferramentas para administrar sistemas CRIS e melhores práticas.** Projeto 914BRA2015 – IBICT. Edital 020/2014. Produto 2. Brasília: IBICT, 2015. 136 p. (Documento Técnico).

PEREIRA, Maria de Nazaré Freitas. **Documento técnico contendo orientações de programas de trabalho para desenvolvimento do BR-CRIS de forma modular e interativa.** Projeto 914BRA2015 – IBICT. Edital 004/201. Produto 3. Brasília: IBICT, 2017. 48 p. (Documento Técnico).

PEREIRA, Maria de Nazaré Freitas. **Levantamento e análise de produtos de informação que permitam delinear o cenário atual brasileiro de componentes de sistemas de informação de pesquisa: projetos, organizações de pesquisa, currículos e resultados, com vistas a sua integração no BR-CRIS.** Projeto 914BRA2015 – IBICT. Edital 004/201. Produto 2. Brasília: IBICT, 2016. 83 p. (Documento Técnico).

PEREIRA, Maria de Nazaré Freitas. **Levantamento e análises das ferramentas de indexação.** Projeto 914BRA2015 – IBICT. Edital 020/2014. Produto 3. Brasília: IBICT, 2015. 79 p. (Documento Técnico). Entrevista com a Professora Hagar Espanha Gomes.

PEREIRA, Maria de Nazaré Freitas. **Projeto Sistema de Informação de Pesquisa – BRCRIS.** Brasília: IBICT, 2017. 8 p. (Nota Técnica).

PEREIRA, Maria de Nazaré Freitas. **Resultado do quadro referencial teórico e da metodologia a ser adotada para integrar sistemas de informação.** Projeto 914BRA2015 – IBICT. Edital 004/201. Produto 1. Brasília: IBICT, 2016. 87 p. (Documento Técnico).

PEREIRA, Maria de Nazaré Freitas. **Sistemas de informação do tipo CRIS desenvolvidos por instituições estrangeiras.** Projeto 914BRA2015 – IBICT. Edital 020/2014. Produto 1. Brasília: IBICT, 2014. 146 p. (Documento Técnico).

PEREIRA, Maria de Nazaré Freitas. **Sugestões para o mapa da competência.** Projeto 914BRA2015 – IBICT. Edital 020/2014. Produto 4. Brasília: IBICT, 2016. 83 p. (Documento Técnico).

SILVEIRA, L. A.; SHINTAKU, M.; BOLINNI, A. **Guia de instalação do DSpace-CRIS.** Brasília: IBICT, 2016. Disponível em: <<http://labcoat.ibict.br/portal/wp-content/uploads/2016/12/Guia-de-Instala%C3%A7%C3%A3o-do-DSpcaeCRIS.pdf>>.

TENÓRIO, Victor Lemos. **Especificação e definição dos metadados.** Metodologia utilizada nos próximos estudos; sugestões. Produto 2 do projeto: Dos padrões internacionais de estruturação da informação de pesquisa aos indicadores (BRCRIS). Maceió: FAPEAL, 2017. 7p. IBICT - Iniciativa de Apoio aos Diálogos Setoriais União Europeia/Brasil.

TENÓRIO, Victor Lemos. **Metodologia do estudo. Fontes de informações mapeadas conforme bizagi modeler.** Produto 1 do projeto: Dos padrões internacionais de estruturação da informação de pesquisa aos indicadores (BRCRIS). Maceió: FAPEAL, 2017. 71p. IBICT - Iniciativa de Apoio aos Diálogos Setoriais União Europeia/Brasil.

TURINE, Marcelo A.S.; CARROMEU, Camilo; SILVA, Márcio A.I. da; CAGNIN, Maria Istela. Gestão pública flexível e ágil por meio do Sigfap. **ComCiência, Revista Eletrônica de Jornalismo Científico.** 10 jun. 2011. Disponível em: <<http://www.comciencia.br/comciencia/handler.php?section=8&edicao=67&id=851>>.

Dos padrões internacionais de estruturação da informação de pesquisa aos indicadores: primeira incursão na temática
Parte IV Relatórios | Piloto BRCRIS IBICT-FAPEAL: uma avaliação exploratória do DSpace-CRIS para a construção de um sistema de informação de pesquisa local

Francisca Rosaline Leite Mota

Doutora e Mestre em Ciência da Informação pelo PPGCI/UFMG. Atua diretamente com projetos de organização documental, Gestão Eletrônica de Documentos (GED) e Tecnologia da Informação. É pesquisadora nas áreas de GED, Informação e Saúde, Sistemas de Informação em Saúde, Direito e Saúde, Telemedicina, Telessaúde e Teleassistência, Acessibilidade em Bibliotecas.

E-mail: rosemota@yahoo.com.br

Hagar Espanha Gomes

Possui graduação em Curso de Biblioteconomia pela Fundação Biblioteca Nacional (1955) e mestrado em Mestrado Em Ciência da Informação pelo Instituto Brasileiro de Bibliografia e Documentação (1972). É Livre Docente em Bibliotrafia pela UFF (1976). Tem experiência na área de Ciência da Informação, com ênfase em Organização do Conhecimento, atuando principalmente nos seguintes temas: indexação, linguagem documentária, tesouro, terminologia e taxonomia, áreas em que atua como Consultora.

E-mail: haguespanhagomes@gmail.com

Hans Ponfick Aragão

Graduando em Análise e Desenvolvimento de Sistemas na Universidade Estadual de Ciências da Saúde de Alagoas, UNCISAL, Brasil. Bolsista pela FAPEAL no Projeto Perspectivas e desafios no processo de gestão de recursos informacionais em instituições de fomento à pesquisa no estado de Alagoas.

E-mail: lisboapablo@hotmail.com

Hélia de Sousa Chaves

Doutoranda pela Universidade de Brasília, a partir do 1º semestre de 2017. Mestre em Ciência da Informação pela Universidade de Brasília (2008), graduada em Letras - Inglês pelo Centro Universitário de Brasília (1984), Bacharel em Tradução pela Universidade de Brasília (1988), especialista em Inteligência Competitiva pela Universidade Federal do Rio de Janeiro (1999) e D.E.A. pela Universidade de Marselha, França. Profissional de informação do Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT). Possui experiência no desenvolvimento de sistemas de informação de pesquisa e no estudo de conteúdos de bases de dados utilizando técnicas de tratamento automático de informação (mineração de dados e mineração de textos). Membro do Projeto BRCRIS IBICT.

E-mail: helia@ibict.br

Juliana Pinheiro Farias

Bacharel em Arquivologia pela Universidade de Brasília - UnB no ano de 2014. Bolsista do CNPq pelo Programa de Capacitação Institucional (PCI-D), no Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT). Tem experiência na elaboração e desenvolvimento de pesquisa científica e acadêmica, assim como, em estudos e projetos da área informacional. Atualmente colabora como auxiliar de pesquisas em alguns projetos do instituto, como o BRCRIS. Vigência em: 01/10/2015 à 31/12/2018.

E-mail: julianafarias@ibict.br

Kelen Cândida Vieira Bomfim

Bacharel em Biblioteconomia pela Universidade de Brasília (2015). Bolsista do Programa de Capacitação Institucional do CNPq, atuando no Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) com análise de informação no projeto “Desenvolvimento, modernização e criação de novas estratégias para organização, preservação e acesso à informação científica, tecnológica e social no Brasil” com objetivo no Projeto Institucional (Sistema Mercado de Trabalho para Docência, Pesquisa e Gestão em C&T). Atualmente colabora como auxiliar de pesquisas no projeto BRCRIS. Vigência em: 01/10/2015 à 31/12/2018.

E-mail: kelenbomfim@ibict.br

Kleber de Barros Alcanfôr

Graduado em Administração pela Universidade Católica de Brasília (1992), Especialização em Análise de Sistemas (1994) e MBA Engenharia de Software (2011). Há 18 anos atua na área de CT&I. Atualmente é assessor técnico do Centro de Gestão e Estudos Estratégicos, responsável pela unidade de desenvolvimento de sistemas e aplicações, atuando desde 2003 com proposição e desenvolvimento de ferramentas de apoio aos estudos prospectivos, avaliação e planejamento estratégico, especialmente para dinâmicas de conhecimento, consulta de percepção de especialistas e mapeamento de competências e pesquisas em ambiente web. Membro do Projeto BRCRIS IBICT.

E-mail: kalcanfor@gmail.com

Lilian Maria Thomé Andrade Brandão

Graduação em Biblioteconomia pela Universidade de Brasília (1978). Especialização em Sistema de Informação na Oxford University (1988). Trabalhou no Sistema Nacional de Informação Documental Agrícola na Biblioteca Nacional de Agricultura - BINAGRI. Responsável pelos Processos Técnicos do BRACARIS - Current Agricultural Information System no Brasil da FAO (1987). Trabalhou no CNPq/SELAP - Cadastro de Instituições de C&T e Formulário do Banco de Currículos (1990). Atuou como chefe da Biblioteca do CENAGRI - Centro Nacional de Informação Agrícola (1997).

Trabalhou no PROSSIGA - Programa de Informação para Gestão de Ciência, Tecnologia e Inovação, do CNPq atuando nos Sistemas de Fomento à Pesquisa em C&T no Brasil, Instituições de C&T, Eventos de C&T (2004). Trabalhou como analista de informação do Centro de Gestão e Estudos Estratégicos - CGEE (2016). Desde maio de 2017 atua como bolsista no Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) - Membro do Projeto piloto BRCRIS IBICT.

E-mail: lilianbrandao@ibict.br

Maria de Nazaré Freitas Pereira

Possui graduação em Biblioteconomia pela Universidade Federal do Pará (1967), mestrado em Ciência da Informação pela Universidade Federal do Rio de Janeiro (1981) e doutorado em Ciências Humanas: Sociologia pelo Instituto Universitário de Pesquisas do Rio de Janeiro (1997). Tem cursos de especialização em Transferência de Tecnologia e Desenvolvimento, Science Policy Research Unit (SPRU) - Sussex University, Brighton - Inglaterra (1989 /1990) e em Sistemas e Redes de Informação, Universidade de São Paulo/ Escola de Comunicação e Artes (1973). Atuou na implantação do Programa Sociedade da Informação no Brasil, na condição de Coordenadora de Inclusão Digital (2001/2003). Realiza estudos de consultoria para o IBICT sobre sistemas CRIS (Current Research Information System) desde 2014.

E-mail: marianazare@uol.com.br

Paulo Henrique de Assis Santana

Graduado em Engenharia Mecânica pela Escola de Engenharia de São Carlos(1969), especialização em Consultoria Organizacional pela Universidade de Brasília(1979) e mestrado em Ciência Espacial pelo Instituto Nacional de Pesquisas Espaciais(1974). Atualmente é Coordenador Geral de Tecnologia da Informação do Ministério do Meio Ambiente. Atuando principalmente nos seguintes temas: equação de Schrodinger, faixas de energia, massa efetiva, modelo Kronig-Penney, potencial periódico unidimensional. Membro do Projeto piloto BRCRIS IBICT.

E-mail: Paulo.assis@cnpq.br

Rogério Mendes Castilho

Graduação em Sistemas de Informação pela Universidade Paulista (2006) com 10 anos de experiência de trabalho no ambiente do SNCTI - Sistema Nacional de Ciência, Tecnologia e Inovação. Experiência em gestão de fontes de informação, com interoperabilidade em Sistemas, bem como transformação de dados para geração de painéis de indicadores. Desde 2017 atua como bolsista no Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) - Membro do Projeto BRCRIS.

E-mail: rogeriocastilho@ibict.br

Victor Lemos Tenório

Mestrando em Modelagem Computacional pela Universidade Federal de Alagoas, Tecnólogo em Análise e Desenvolvimento de Sistemas pela Universidade Estadual de Ciências da Saúde de Alagoas (2014) e Graduado em Gestão em Sistemas de Informação - FAA - Faculdade Alagoana de Administração (2009). Bolsista de Desenvolvimento Tecnológico no Projeto Perspectivas e Desafios no Processo de Gestão de Recursos Informacionais em Instituições de Fomento à Pesquisa no Estado de Alagoas. Membro do Projeto piloto BRCRIS IBICT/FAPEAL.

E-mail: victor.ltenorio@hotmail.com



Diálogos
UNIÃO EUROPEIA - BRASIL

BrCris

ibict 65

MINISTÉRIO DA
CIÊNCIA, TECNOLOGIA,
INOVAÇÕES E COMUNICAÇÕES

MINISTÉRIO DA
ECONOMIA

MINISTÉRIO DAS
RELAÇÕES EXTERIORES

**PÁTRIA AMADA
BRASIL**
GOVERNO FEDERAL