

GESTÃO DE DADOS CIENTÍFICOS NO CONTEXTO DOS PARADIGMAS DE CAPURRO: ALGUMAS REFLEXÕES

Jane Lecardelli

Mestranda em Gestão de Unidades de Informação na Universidade do Estado de Santa Catarina (UDESC).

E-mail: jane.biblio@gmail.com

Gabriela Monteiro

Mestre em Gestão de Unidades de Informação pela Universidade do Estado de Santa Catarina (UDESC).

E-mail: gm.monteiro@hotmail.com

Elaine Rosângela de Oliveira

Lucas

Professora do Programa de Pós-Graduação em Gestão da Informação da Universidade do Estado de Santa Catarina (UDESC). Doutora em Ciência da Informação pela Universidade de São Paulo (USP).

E-mail: lanilucas@gmail.com

Recebido em: 30/05/2019

Aceito em: 19/08/2019

RESUMO

O presente trabalho tem o objetivo de apresentar reflexões sobre as relações entre o ciclo de gestão de dados científicos e os paradigmas do campo da Ciência da Informação (CI) propostos por Rafael Capurro, com base no ciclo de gestão de dados adotado pelo Repositório Científico de Acesso Aberto de Portugal (RCAAP). Caracteriza-se como uma pesquisa de revisão bibliográfica para contextualização do campo da CI, seus paradigmas, a gestão de dados científicos, o ciclo de gestão de dados, para então, no contexto dos paradigmas físico, cognitivo e social de Capurro, analisar a convergência do processo de gestão de dados, especificamente por meio do modelo do ciclo de gestão de dados adotado pelo RCAAP, por ser este um modelo em uso e já consolidado no programa da União Europeia *Horizon 2020*. As reflexões apontam que a gestão de dados tem convergência com o paradigma social, pois busca atender novas demandas da sociedade para o reuso de dados, onde os indivíduos influenciam e são influenciados, em um contínuo processo de construção coletiva de saberes.

Palavras-chave: Gestão de Dados Científicos. Paradigmas da Ciência da Informação. Reuso de Dados.

SCIENTIFIC DATA MANAGEMENT IN THE CONTEXT OF THE PARADIGMS OF CAPURRO: A FEW REFLECTIONS

ABSTRACT

The present work has the objective of presenting reflections on the relations between the cycle of scientific data management and the paradigms of the field of Information Science (IC) proposed by Rafael Capurro, based on the data management cycle adopted by the Portuguese Open Access Scientific Repository [Repositório Científico de Acesso Aberto de Portugal, RCAAP]. It is characterized as a bibliographic review research for the contextualization of the field of IC, its paradigms, the management of scientific data, the cycle of data management, in order to subsequently analyze the convergence of the data management process in the context of the physical, cognitive and social paradigms of

Capurro, specifically by means of the data management cycle model adopted by the RCAAP, as this model is in use and already consolidated in the *Horizon 2020* program of the European Union. The reflections point out that data management converges with the social paradigm, since it seeks to meet new demands of society for the reuse of data, where individuals influence and are being influenced, in a continuous process of a collective construction of knowledge.

Keywords: Scientific data management. Paradigms of information Science. Reuse of data.

1 INTRODUÇÃO

No campo das pesquisas científicas, as questões relacionadas à gestão de dados científicos têm recebido atenção especial devido às agências de fomento iniciarem um movimento no sentido de solicitar os dados relativos às pesquisas, como forma de validação de estudos para maior visibilidade aos pesquisadores, possibilidade de reuso dos dados em novos estudos e contextos, além da economia de esforços e custos com a coleta de dados.

Essas iniciativas vão influenciar as discussões sobre o direito do cidadão, seja ele pesquisador ou não, a ter acesso ao conhecimento produzido na sociedade em que vive, principalmente quando esse conhecimento é gerado a partir de dinheiro oriundo dos cofres públicos. A partir destas iniciativas, fica visível que as discussões sobre o acesso aberto à produção científica começam a atingir um número cada vez maior de instituições, pesquisadores, governos e agências de fomento (CAVALCANTI; SALES, 2017, p.79).

Esse movimento tem várias iniciativas na Europa e EUA e, apesar de não ter sensibilizado completamente todos os envolvidos com pesquisas, já registra publicações, eventos, seminários, além de estabelecer a criação de repositórios de dados, no intuito “de mostrar a importância acadêmica, científica, econômica e as diversas possibilidades de desenvolvimento nas mais diferentes áreas a partir do reconhecimento da necessidade de se adotar o paradigma de dados abertos”. (CAVALCANTI; SALES, 2017, p.81).

No Brasil, foi publicado em setembro de 2018 o Relatório “Acesso Aberto a Dados de Pesquisa no Brasil – 2018”, documento este:

que faz parte da primeira etapa do projeto Rede de Dados de Pesquisa Brasileira (RDP Brasil), o qual objetivou tanto identificar repositórios brasileiros que disponibilizam dados de pesquisa, quanto levantar características relacionadas à abrangência, à temática, aos padrões de metadados e aos softwares utilizados. (PAVÃO, et al., 2018, p.03).

Esse relatório cita alguns aspectos quanto à gestão de dados no país e demonstra que o tema ainda carece de estudos e análises mais detalhadas.

O levantamento revela a existência de poucas iniciativas nacionais relacionadas ao compartilhamento dos dados de pesquisa brasileiros. As universidades – responsáveis pela maior parte da pesquisa desenvolvida no Brasil – não têm protagonismo no que diz respeito aos repositórios de dados de pesquisa: apenas UFPR e PUC Rio foram identificadas no levantamento. Outro detalhe interessante é a existência de alguns repositórios identificados nos websites institucionais, mas não cadastrados no Re3data, o que revela a incipiência das iniciativas. (PAVÃO, et al., 2018, p.19).

Esse artigo de revisão bibliográfica, objetiva contextualizar o ciclo da gestão de dados científicos, conforme a tríade de paradigmas da Ciência da Informação (CI), que se constitui do paradigma físico, cognitivo e o social, definidos por Rafael Capurro (2003) em seu artigo “Epistemologia e ciência da informação”. Esses paradigmas foram apresentados no cenário brasileiro, no 5º Encontro Nacional de Pesquisa em Ciência da Informação, realizado na cidade de Belo Horizonte, em 10 de novembro de 2003.

Para compor este artigo, realizaram-se pesquisas em artigos de autores com contribuições relevantes no campo da CI, tais como Rafael Capurro, Birger Hjørland, Michael K. Buckland, Tefko Saracevic, Haroldo Borko, Carlos Alberto Ávila Araújo, entre outros, no intuito de compreender os conceitos de Informação e Ciência da Informação. Posteriormente, contextualiza-se a gestão de dados científicos, para a partir disso, buscar verificar dentre os paradigmas da CI propostos por Capurro, a qual paradigma converge a gestão de dados atualmente.

Essa aproximação teórica se fará por meio da análise do diagrama do ciclo de gestão de dados científicos, utilizado no Repositório Científico de Acesso Aberto de Portugal (RCAAP). Esse modelo foi selecionado, pois consolida três modelos conceituais da gestão de dados (Australiano, Americano e Sul-Africano), além de distinguir as fases da gestão de dados em planejamento, produção e disseminação. (CARVALHO; FURTADO; PRÍNCIPE, 2017).

Para abordar essa temática, a estrutura contará com breve revisão conceitual nos temas: Ciência da Informação, paradigmas da CI, gestão de dados científicos, ciclo de gestão dos dados propostos pelo RCAAP, e finalmente, a gestão de dados no contexto dos paradigmas de Rafael Capurro.

O levantamento da literatura pertinente foi realizado em periódicos de acesso aberto disponíveis na internet, em que os temas abordados estivessem atrelados a área de CI, além de observar a relevância dos autores nos temas pesquisados.

2 CIÊNCIA DA INFORMAÇÃO

Como o estudo aborda a temática de gestão de dados científicos no campo da ciência da informação, é relevante trazer algumas definições do que trata a CI.

Becker (1976 apud SARACEVIC, 1996, p. 47), “definiu a CI como o estudo do modo pelo qual as pessoas criam, usam e comunicam informações”.

De acordo com Borko (1968, p.05), a ciência da informação é:

uma ciência interdisciplinar que investiga as propriedades e comportamento da informação, as forças que governam os fluxos e os usos da informação, e as técnicas, tanto manual quanto mecânica, de processamento da informação, visando sua armazenagem, recuperação, e disseminação ideal.

Apresentar relação com várias outras áreas e tratar de problemas informacionais complexos, de várias formas e em campos diferentes, além de dialogar com distintas áreas para tornar a informação mais acessível, são alguns dos aspectos que a caracterizam como interdisciplinar.

Para Saracevic (1996, p. 47), a Ciência da Informação “é um campo dedicado às questões científicas e a prática profissional voltadas para os problemas da efetiva comunicação do conhecimento e de seus registros entre os seres humanos, no contexto social, institucional ou individual do uso e das necessidades de informação.”

Neste cenário a CI é a área de estudo:

voltada às atividades de organização, representação e recuperação da informação. Dentre seus objetivos, possui a compreensão sobre como representar o conhecimento, quais conhecimentos podem ser representados e como cada área do saber é representada. Contudo,

conhecimento e informação são termos de difícil conceituação devido à amplitude semântica e às diversas perspectivas de análise, domínios e concepções de cada área. (SCHIESSL; SHINTAKU, 2012 apud OTTONICAR; SANTOS; MORAES, 2017, p. 632).

Na concepção de Capurro e Hjørland (2007, p.186), a “CI se ocupa com a geração, coleta, organização, interpretação, armazenamento, recuperação, disseminação, transformação e uso da informação, com ênfase particular, na aplicação de tecnologias modernas nestas áreas.”

Para Hawkins (2001 apud ZINS, 2011, p.162), a Ciência da Informação é:

um campo interdisciplinar preocupado com os conceitos teóricos e práticos, bem como os tecnológicos, as leis e a indústria que lida com a transferência do conhecimento, as fontes, a geração, a organização, a representação, o processamento, a distribuição, a comunicação e os usos da informação e com a comunicação entre usuários e seus comportamentos quando eles buscam satisfazer suas necessidades informacionais.

Segundo Neves (2006 p. 40), sua principal função “é produzir conhecimentos que contribuam para a solução de problemas relacionados à organização de sistemas de informação especializados na incorporação, sistematização, disseminação e recuperação da informação”.

Nos conceitos acima observa-se que a CI permeia os estudos de todo fluxo informacional, está atrelada as tecnologias de informação e comunicação, possui dentre outras características, a de ser interdisciplinar e não deve estar dissociada da prática e do contexto social a quem pretende atender.

Para Wersig e Nevellng (1975 apud SARACEVIC, 1996, p.43), a CI “desenvolveu-se historicamente porque os problemas informacionais modificaram completamente sua relevância para a sociedade ou, em suas palavras, ‘atualmente, transmitir o conhecimento para aqueles que dele necessitam é uma responsabilidade social, e essa responsabilidade social parece ser o verdadeiro fundamento da CI`.”

Este cenário informacional dinâmico requer novas abordagens por parte da CI, e no campo da gestão de dados científicos a CI tem terreno fértil para estudos teóricos e práticos, construção de soluções (serviços e produtos) com enfoque social, na coletividade (perspectiva social) e no avanço da ciência.

Para citar os paradigmas, inicialmente se faz necessário elucidar o conceito de paradigma. Segundo Capurro (2003, p.04), “paradigma é um modelo que nos permite ver uma coisa em analogia a outra”. Para Kuhn (2003 apud ALMEIDA, *et al*, 2007, p.19), “um paradigma pode ser considerado como um conjunto de realizações científicas reconhecidas por uma determinada comunidade, durante algum tempo, fornecendo problemas e soluções modelares que podem ser identificados no seu campo de atuação”.

Dentre os paradigmas da informação em CI, o primeiro autor a ser destacado é Saracevic (1996),

que identificou três grandes conceitos de informação na Ciência da Informação. O primeiro é o **sentido restrito**: informação consiste em sinais ou mensagens envolvendo pequeno ou nenhum processamento cognitivo (ou, então, tal processamento pode ser expresso em termos de algoritmos ou probabilidades). Informação é então a propriedade de uma mensagem, que pode ser estimada por uma probabilidade. O segundo é o **sentido amplo**: informação envolve diretamente processamento cognitivo e compreensão. Ela resulta da interação entre duas estruturas cognitivas, uma “mente” e um “texto” (num sentido amplo dessa palavra). Informação é o que afeta ou altera um estado de conhecimento, ou seja, para determinar algo como sendo informação é preciso ver o que o leitor entendeu de um texto ou documento. O terceiro é o **sentido ainda mais amplo**: informação existe em um contexto. Sua definição envolve não apenas as mensagens (sentido restrito) que são cognitivamente processadas (sentido amplo), mas também um contexto, uma situação específica, e uma ação, ou tarefa, no decurso da qual a informação é cognitivamente processada. Assim, informação envolve motivação e intencionalidade do indivíduo, mas sempre conectadas a um horizonte social, do qual fazem parte a cultura e as ações desempenhadas. (ARAÚJO, 2014, p.18).

O autor Buckland (1991, p.01-02), cita três paradigmas para o uso da palavra informação:

a) Informação-como-processo: Quando alguém é informado, aquilo que conhece é modificado. Nesse sentido “informação” é “o ato de informar...; comunicação do conhecimento ou “novidade” de algum fato ou ocorrência; a ação de falar ou o fato de ter falado sobre alguma coisa. **b) Informação-como-conhecimento**: “Informação” é também usado para denotar aquilo que é percebido na “informação-como-processo”: o “conhecimento comunicado referente a algum fato particular, assunto ou evento; aquilo que é transmitido, inteligência, notícias.” A noção de que informação é aquela que reduz a incerteza poderia ser entendida como um caso especial de “informação-como-conhecimento”. Às vezes informação aumenta a incerteza. **c) Informação-como-coisa**: O termo “informação” é também atribuído para objetos, assim como dados para

documentos, que são considerados como “informação”, porque são relacionados como sendo informativos, tendo a qualidade de conhecimento comunicado ou comunicação. (BUCKLAND, 1991, p.01-02).

Como base para este estudo, nos ateremos aos três paradigmas apontados por Rafael Capurro (2003): físico, cognitivo e social. Segundo Capurro (2003) a CI teria nascido sob a vigência de um paradigma físico.

[...]. Conforme tal visão, a informação é algo, um objeto **físico**, que um emissor transmite a um receptor. Um segundo modelo, o **cognitivo**, emergiu nos anos 1970. [...]. Tal modelo relaciona informação a conhecimento: algo é informacional na medida em que altera as estruturas de conhecimento do sujeito que se relaciona com dados ou documentos. Em anos mais recentes, estaria emergindo um **paradigma social**, voltado para a constituição social dos processos informacionais. A partir da crítica ao modelo anterior, que via o usuário como um ser isolado da realidade e apenas numa dimensão cognitiva, busca-se aqui reinseri-lo nos seus contextos concretos de vida e atuação, numa perspectiva claramente fenomenológica: ver os sujeitos como “ser no mundo”, tal como [...] as “comunidades de discurso” estudadas por Hjørland e Albrechtsen a partir de uma inspiração em Wittgenstein. Daí a famosa fórmula de Capurro, para quem não é a informação que é a matéria prima do conhecimento: antes, é apenas a existência de um conhecimento partilhado entre diferentes atores que faz com que algo seja reconhecido como “informação”. (ARAÚJO, 2014, p.20-21).

Capurro (2003) ao denominar os paradigmas menciona Hjørland (2003) e Albrechtsen (1995), autores que desenvolveram um paradigma social-epistemológico, denominado “análise de domínio (*domain analysis*), no qual o estudo de campos cognitivos está em relação direta com comunidades discursivas (*discourse communities*), ou seja, com distintos grupos sociais e de trabalho que constituem uma sociedade moderna”. (CAPURRO, 2003, p.10).

Segundo Araújo (2010, p.26), uma primeira referência ao paradigma social, “seria a proposta da epistemologia social de Shera (que serviu, aliás, de inspiração a Capurro para a denominação deste paradigma).”

Ainda nas décadas de 1960 e 1970, Shera propunha a criação de uma nova disciplina para estudar a inserção social do conhecimento humano, isto é, as interações entre os processos intelectuais (conhecimento produzido) e as atividades humanas e sociais. Trata-se, em sua visão, de uma disciplina que ‘deveria fornecer uma estrutura para a investigação eficiente de todo o complexo problema dos processos intelectuais das

sociedades – um estudo pelo qual a sociedade como um todo procura uma relação perceptiva para seu ambiente total'. (SHERA, 1977 apud ARAÚJO, 2010, p.26).

Em síntese, sobre os paradigmas desenvolvidos por Capurro (2003), podemos mencionar que a CI iniciou sob o **paradigma físico**, que foi “centrado em sistemas informatizados, onde o conceito de informação aproxima-se de um sentido estritamente técnico, uma informação mensurável que não necessariamente abarca significado semântico”. (ALMEIDA, *et al.*, 2007, p.19).

Posteriormente, no **paradigma cognitivo**:

considera os modelos mentais dos usuários, utilizando abordagens cognitivas - centradas no processo interpretativo do sujeito cognoscente, observando-se suas características fenomenológicas e individuais, valorizando assim tentativas de inclusão das dimensões semânticas e pragmáticas nos sistemas de recuperação da informação [...] a partir da análise de ‘como as informações são compreendidas pelos usuários’”. (ALMEIDA, *et al.*, 2007, p.23).

Já o **paradigma social**, modelo mais recente, tem como consequência prática “abandonar a busca de uma linguagem ideal para representar o conhecimento ou de um algoritmo ideal para modelar a recuperação da informação, como aspiram os paradigmas ‘físico e o cognitivo’, pois considera as possíveis perspectivas ou pontos de acessos distintos de acordo com o interesse do usuário ou comunidade”. (CAPURRO, 2003, p.09).

Cabe destacar que no paradigma social, os critérios e perspectivas de uso de determinadas informações são processos intersubjetivos, mas, que derivam do fazer sentido coletivo. “Esses critérios são utilizados também para julgar as informações (como boas ou ruins, exatas ou inexatas, completas ou incompletas)”. (ARAÚJO, 2010, p.27). O autor enfatiza a participação e construção coletiva do conhecimento, por meio de sucessivas interações, nas quais influenciemos e somos influenciados, ou seja, não estamos fora do processo e sim, somos parte integrante dele.

3 GESTÃO DE DADOS CIENTÍFICOS

Antes de tratar da gestão de dados, especificamente, cabe questionar: afinal o que são dados científicos? Conforme definição da Organização de Cooperação e de Desenvolvimento Económico - OCDE (2007 apud SALES; SAYÃO, 2012, p.122) **dados**

científicos ou dados de pesquisas, [...] são “registros de fatos usados como fontes primárias na investigação científica e que geralmente são aceitos na comunidade científica como necessários para a validação dos resultados da pesquisa”.

As ações que coletivamente permeiam o ciclo de vida dos dados de pesquisa vêm sendo chamadas de *gestão de dados de pesquisa*, e se consolidam, de maneira ideal, em consonância com padrões de ampla aceitação, condicionantes de domínios disciplinares, requisitos estabelecidos pelos pesquisadores e, quando existentes, políticas institucionais e diretrizes de alcance nacional e internacional. (SAYÃO; SALES, 2016, p.92).

Para Curty e Aventurier (2017, p.3-4), a gestão de dados científicos

sinaliza para uma iminente movimentação em torno de políticas e iniciativas que valorizem a publicação de dados e seu potencial de reuso. A partir da caracterização desse novo paradigma da publicação científica, e da discussão das diferentes abordagens para a publicação de dados, espera-se contribuir não apenas para um melhor esclarecimento conceitual acerca do tema, mas também com informações que dêem suporte a decisões por parte das instituições científicas, editores científicos, agências de fomento, mantenedores de repositórios e pesquisadores de modo geral.

Essa temática é abordada no campo da ciência da informação tendo em vista as próprias características dessa área, de ser interdisciplinar, ciência humana e social emergente, pós-moderna, estar ligada as tecnologias de informação, além de ser voltada a estudar problemas tanto do campo da “pesquisa científica quanto à prática profissional”. (SARACEVIC, 1996, p. 41).

De acordo com Aventurier e Alencar (2016, p.03),

compartilhar os dados de pesquisas é uma necessidade. Hoje, a sociedade enfrenta desafios científicos muito mais complexos que pressupõem a interdisciplinaridade, colaborações entre laboratórios de pesquisa, permuta de informações e de competências. Com isso, novas competências associadas ao trabalho com os dados de pesquisas estão surgindo para corroborar com o descobrimento de novos conhecimentos a partir desses volumes de dados.

A gestão de dados científicos abertos pode propiciar vários benefícios no campo acadêmico, dentre eles, permitir o reuso/compartilhamento de dados, reduzir custos e retrabalho, dar credibilidade as pesquisas na medida em que permite testar os resultados, ampliar a visibilidade dos autores e possibilitar novos estudos a partir de abordagens ou

novas interações entre estes dados. A gestão de dados é um tema emergente e como tal, deve ser estudado e ter suas práticas incorporadas no contexto da CI.

Para que o processo de gestão de dados ocorra existe um ciclo de ações envolvidas. Esse processo de gestão já conta com alguns modelos, desenvolvidos por países que adotam a gestão de dados em suas instituições de pesquisa. Por exemplo, “na Europa, o programa *Horizon 2020* especifica que os dados de pesquisas necessários para validar os resultados apresentados em publicações devem ser depositados concomitantemente aos artigos, de preferência em um repositório de dados”. (AVENTURIER; ALENCAR, 2016, p.08-09).

No Brasil ainda não existe uma política nacional estruturada, porém, algumas instituições e agências de fomento já solicitam aos pesquisadores um plano de gestão de dados, seguindo as iniciativas internacionais da área. Este movimento tende a se tornar mandatório pelas instituições financiadoras de pesquisas no país.

Para que o ciclo da gestão de dados possa ocorrer e os dados possam ser recuperados por máquinas ou humanos, sugere-se a adoção dos **Princípios FAIR** (do inglês *Findable, Accessible, Interoperable, and Reusable - FAIR*), traduzindo para o português: Localizável, Acessível, Interoperável e Reutilizável. Estes princípios estabelecem critérios mínimos de qualidade para que os dados sejam usados e citados corretamente. Nesses princípios, um dado é **localizável** quando apresenta um identificador único e persistente; é **acessível** quando oferece a possibilidade de ser obtido por máquinas e seres humanos, é **interoperável** se puder ser processado por máquinas, em formato de dados que utilizem vocabulários compartilhados e/ou ontologias, e se for semântica e sintaticamente acessível para máquinas. Para que os dados sejam **reutilizáveis**, devem ter metadados ricos e bem descritos, que possam “ser automaticamente (ou com o mínimo de esforço humano) ligados ou integrados”. (AVENTURIER, 2017, p.01).

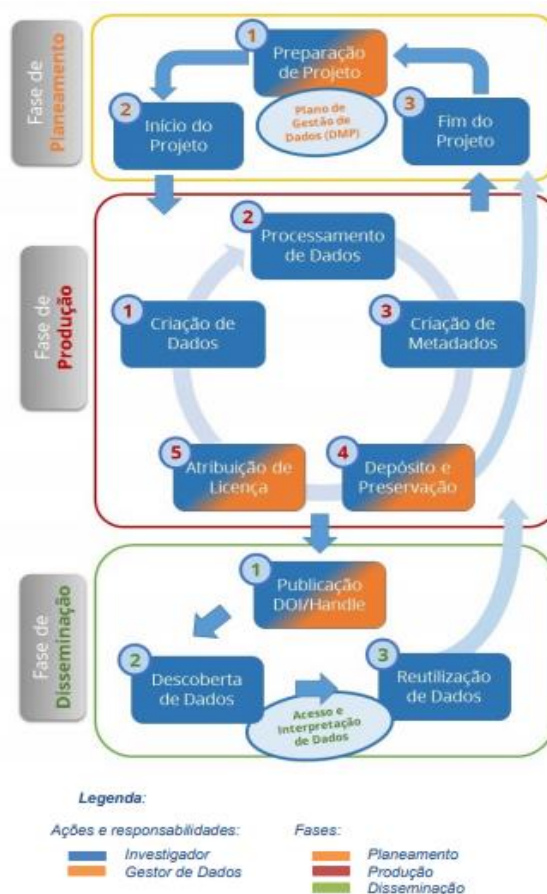
Segundo Aventurier (2017, p.01), é muito importante definir “critérios para avaliar a qualidade dos conjuntos de dados. Uma boa gestão de dados é também essencial para facilitar descobertas, inovação e reutilização para uma comunidade, depois do processo de publicação”.

Neste trabalho, optou-se por analisar o modelo do ciclo de gestão de dados adotado pelo Repositório Científico de Acesso Aberto de Portugal (RCAAP), por ser um *workflow* genérico, desenvolvido a partir da compilação dos modelos Australiano, Americano e Sul-

Africano, e por estar em uso em uma instituição que já pratica a gestão de dados, ou seja, um modelo com aderência às ações práticas.

Em meados de 2008, em Portugal, foi criado o Repositório Científico de Acesso Aberto de Portugal/RCAAP. A iniciativa teve como objetivo aumentar a visibilidade, acessibilidade e divulgação dos resultados da pesquisa científica no país. Em 2010, o RCAAP produziu o relatório *Open data Repositories: State of the Art*, que marca o início de suas intervenções no domínio da curadoria de dados de pesquisa, sua organização em repositórios e de seu acesso. O documento tinha o objetivo, entre outros, de informar e orientar sobre a criação de um projeto piloto de repositório de dados científicos, que se encontra em fase de desenvolvimento. (CAVALCANTI; SALES, 2017, p.94).

Imagem 01 - Diagrama do ciclo de gestão de dados de investigação do RCAAP



Fonte: (CARVALHO; FURTADO; PRÍNCIPE, 2017, p.15).

No diagrama, além de ilustrar as ações pertinentes ao ciclo de gestão de dados, categoriza em fases o processo (planejamento, produção e disseminação), e situa as ações e responsabilidades do pesquisador e do gestor de dados.

Posto este modelo, na sequência buscar-se-á identificar sua convergência com os paradigmas da CI (físico, cognitivo e social), propostos por Capurro (2003).

4 GESTÃO DE DADOS CIENTÍFICOS E OS PARADIGMAS

Neste cenário informacional dinâmico e complexo, onde novas relações e interações podem ser estabelecidas, este artigo buscou refletir sobre a gestão de dados científicos e os paradigmas da CI, conforme Capurro (2003).

Tendo como base o diagrama do ciclo de gestão de dados da RCAAP (Imagem 01), observa-se que apesar de destacar as fases distintas (planejamento, produção e disseminação), todo o processo de gestão de dados é interligado e busca propiciar ao final, o reuso dos dados, de forma que o processo possa ser cíclico.

A fase de **planejamento** envolve pensar sobre a criação dos dados (início e fim do projeto), elaboração do plano de gestão de dados, conforme os requisitos das agências de fomento. Na fase de **produção** é quando inicia o projeto em si e os dados são gerados e transformados, esta etapa abrange a criação dos dados, processamento, criação de metadados, preservação e atribuição de licenças, preparando para a publicação em repositórios. Nesta fase também ocorrem as atividades de curadoria e os dados passam do domínio restrito para o domínio público.

A fase de **disseminação** se dá após a publicação dos dados, com as etapas de divulgação de identificador persistente (DOI/*Handle*), potencialização da visibilidade dos dados para a descoberta, e finalmente, com a reutilização dos dados. “A reutilização de dados de investigação pressupõe a sua descoberta (uma vez no domínio público), o seu acesso (aberto ou não) e condições favoráveis à sua reutilização (regulada pela licença em vigor)” (CARVALHO; FURTADO; PRÍNCIPE, 2017, p.20). Envolve o acesso e reutilização dos dados de forma que o ciclo descrito possa reiniciar e transcorrer sem barreiras.

Dentro do **paradigma físico** não se observa esta tendência colaborativa, apesar da gestão de dados utilizar a tecnologia e sistemas de recuperação, seu objetivo não está focado nestes sistemas (tanto que propõem diferentes plataformas) e se utiliza delas como um meio para alcançar os objetivos. As tecnologias atuais fornecem facilidades no trato dos dados científicos, porém, o trabalho de curadoria na fase de produção e a criação dos metadados, indispensáveis para a descoberta e reutilização, deve ser pautado no contexto, dentro de áreas de domínio, com parcerias entre os pesquisadores envolvidos,

adotando critérios de qualidade, de forma que estes dados possam estar localizáveis, acessíveis, interoperáveis e reutilizáveis.

O **paradigma cognitivo** na gestão de dados estaria envolvido no momento em que as pessoas fazem uso da informação (por exemplo na fase de disseminação) no processo de apropriação, porém, essa informação disponibilizada já é resultado de interações sociais e será usada em um determinado contexto pelo indivíduo, logo, ela também é produto de construção social. Dessa forma, a gestão de dados não está diretamente ligada ao paradigma cognitivo. A partir do momento que os gestores se preocupam com a criação do plano de gestão de dados para o reuso (na fase de planejamento), a informação está sujeita a um contexto social visando sua disponibilidade para determinadas áreas de domínio. Nesse aspecto, já apresenta identificação maior com o modelo social.

Observa-se que os paradigmas migram de um modelo a outro, conforme as mudanças de contextos sociais, como uma forma de suprir novas demandas, porém, os modelos não são excludentes e sim complementares. Inicia com o modelo físico e a questão da informatização e uso de sistemas sem considerar a semântica e os usuários; passa pelo modelo cognitivo, que vê a informação como produto de um sujeito isolado, sem a interferência do contexto e das relações interpessoais. Após esse modelo, migra para o paradigma social da informação, que tem enfoque na construção coletiva e nas relações interpessoais, intersubjetivas, onde a informação faz sentido para determinados grupos e comunidades, é situacional e temporal, e está fortemente atrelada aos contextos sociais. (ARAÚJO, 2010).

Mais especificamente, de forma resumida, Capurro procura destacar as seguintes características essenciais: necessidade de interpretar a informação (hermenêutica); necessidade de contexto para permitir a interpretação da informação; informação como um conceito interdisciplinar (CAPURRO; HJØRLAND, 2003, p. 356); a informação tem significados diferentes para sujeitos diferentes, que têm diferentes interesses. Ao considerar tais características essenciais da informação, e não apenas a informação como coisa (materializada), é possível avaliar os aspectos pragmáticos, ou resultados e interesses sociais, envolvidos no uso, geração e disseminação da informação (CAPURRO; HJØRLAND, 2003, apud MATHEUS, 2005, p.162).

Dentre os paradigmas físico, cognitivo e social, e analisando o que se propõe a ser a gestão de dados, pode-se aferir que as ações pertinentes às fases de planejamento, produção e disseminação são convergentes ao **paradigma social**. O que pode ser

observado em Capurro (2003, p.10), ao considerar “as possíveis perspectivas ou pontos de acessos distintos de acordo com o interesse do usuário ou comunidade”.

Na fase de produção dos dados, além da criação e processamento, é vital a devida atenção aos metadados, de forma a facilitar a descoberta e o reuso, visto que sempre busca atender a um contexto, seja ele: acadêmico, temporal, cultural, histórico, econômico ou social. Portanto, a gestão de dados não se atém somente a disponibilizar os dados, ela envolve no processo de curadoria todo um contexto político, social e ético entre os pesquisadores, gestores e instituições, de forma que os dados sejam adequados para reuso, construindo assim, uma ciência mais aberta em prol do bem coletivo.

Segundo Rendón Rojas (2005 apud ARAÚJO, 2010, p.27),

na busca por uma abordagem realista-dialética da informação, recorre a Piaget para postular que o processo de conhecimento não é um fenômeno em que um sujeito com um “vazio” na mente busca algo para preencher esse vazio. O sujeito nunca é vazio, ele possui vários conhecimentos e também estruturas nas quais cada novo conhecimento se acomoda, não numa lógica cumulativa, mas num processo interativo, de alterar-se e ser alterado. Também a informação não é um “pacote” fechado que, apropriada pelo sujeito, ocupa um lugar na sua mente como se fosse uma peça de quebra-cabeças.

Cada indivíduo pode contribuir com seus saberes para esta construção coletiva, num processo de inter-relações contínuas e importantes para a construção de novos conhecimentos.

Nesse contexto, a participação de profissionais interdisciplinares é relevante na construção de repositórios de dados confiáveis e de qualidade (destacam-se os princípios *FAIR*), e em áreas de domínio diferentes. “Isso significa, em outras palavras, uma integração da perspectiva individualista e isolacionista do paradigma cognitivo dentro de um contexto social no qual diferentes comunidades desenvolvem seus critérios de seleção e relevância” (CAPURRO, 2003, p.10). Ressalta-se a importância do trabalho colaborativo entre os pesquisadores e gestores de dados, pois é possível observar estas interseções de atuação no processo, em todas as fases do ciclo de gestão de dados (desde o planejamento até a disseminação). Cada área de conhecimento apresenta suas abordagens, conceitos, metodologias, paradigmas, que podem ser discutidos/repensados, visando a solução de problemas mais complexos ou para a construção de novos saberes em conjunto.

A reutilização de dados visa o uso colaborativo, esta construção de saberes através do reuso de dados, converge para o paradigma social, que é o paradigma mais atual da CI. Isso não significa que os demais modelos não tenham pontos de conexão, porém, avalia-se que o paradigma mais aderente ao ciclo da gestão de dados analisado seja o paradigma social.

5 CONSIDERAÇÕES FINAIS

A gestão de dados científicos é um tema que ainda está construindo seus alicerces no país. Mas, já conta com algumas iniciativas práticas e teóricas, no intuito de inserir e ressaltar a importância da temática nas práticas de gestão em organizações que atuam com pesquisas, como uma forma de validar resultados científicos, permitir reuso de dados, dar visibilidade aos pesquisadores e estimular o avanço da ciência aberta no país.

Buscou-se no artigo analisar o ciclo da gestão de dados científicos conforme os paradigmas da CI (físico, cognitivo, social) propostos por Capurro (2003), todavia, ressalta-se que outras perspectivas podem ser analisadas, em diferentes áreas de domínio, neste estudo em específico foi na área de CI.

Observou-se maior convergência do ciclo de gestão de dados ao paradigma social, mas isso não significa que os demais modelos não tenham relevância, pois são construídos e evoluem com as demandas sociais. Cabe destacar, que foram apenas algumas reflexões acerca do tema, que não tiveram o intuito de demonstrar qual seria o melhor paradigma, mas somente levantar alguns aspectos da gestão de dados correlacionando-os aos paradigmas propostos por Rafael Capurro.

A CI se propõe a estudar como as pessoas geram, utilizam e transmitem informações; e a gestão de dados faz parte deste contexto. No ciclo de gestão de dados apresentado, ficou clara a integração entre as atividades que compõem as fases, desde o planejamento até a de disseminação. Esse fluxo, se orquestrado de forma ética e responsável pelos profissionais envolvidos, permite que os dados possam ser reutilizados de forma confiável e segura, com benefícios a todos os envolvidos. Benefícios estes, que se estendem para a sociedade com o acesso a dados abertos, que podem servir de base a novas pesquisas ou estudos, em áreas e contextos distintos dos previstos pelos criadores originais.

Como o tema da gestão de dados ainda está buscando construir suas bases teóricas, a CI também está se consolidando como uma ciência, ambos ainda deixam mais indagações do que respostas sobre as questões que surgem. Nossa intenção foi indicar novas perspectivas de estudos sem a pretensão de encontrar respostas definitivas, mas provocar novas reflexões no tema.

Neste contexto a CI pode dialogar com outras disciplinas ou até ciências, enfatizando seu cunho interdisciplinar, quem sabe até transpondo isso, e sendo transdisciplinar, ao construir novas possibilidades não só para a área, como também para outras disciplinas ou ciências.

REFERÊNCIAS

ALMEIDA, D. P. R. *et al.* Paradigmas contemporâneos da Ciência da Informação: a recuperação da informação como ponto focal. **Revista Eletrônica Informação & Cognição**, Marília, v. 6, n. 1, p. 16-27, 2007. Disponível em: <http://www2.marilia.unesp.br/revistas/index.php/reic/article/view/745>. Acesso em: 30 maio 2019.

ARAÚJO, Carlos Alberto Ávila. Estudos de usuários conforme o paradigma social da ciência da informação: desafios teóricos e práticos de pesquisa. **Informação & Informação**, [S. l.], v. 15, n. 2, p. 23-39, jul./dez. 2010. ISSN 1981-8920. Disponível em: <http://www.uel.br/revistas/uel/index.php/informacao/article/view/6485>. Acesso em: 01 out. 2018.

ARAÚJO, Carlos Alberto Ávila. O que é Ciência da Informação? **Informação & Informação**, [S.l.], v. 19, n. 1, p. 01-30, dez. 2014. ISSN 1981-8920. Disponível em: <http://www.uel.br/revistas/uel/index.php/informacao/article/view/15958>. Acesso em: 01 out. 2018.

AVENTURIER, Pascal. **Princípios FAIR**: critérios de qualidade para dados de pesquisa. 2017. Disponível em: <https://publicient.hypotheses.org/1456>. Acesso em: 20 set. 2018.

AVENTURIER, Pascal; ALENCAR, Maria Cleofas de. Os desafios de dados de pesquisa abertos. **Revista Eletrônica de Comunicação, Informação e Inovação em Saúde**, [S.l.], v. 10, n. 3, sep. 2016. ISSN 1981-6278. Disponível em: <https://www.reciis.icict.fiocruz.br/index.php/reciis/article/view/1069>. Acesso em: 02 out. 2018.

BORKO, H. Information Science: What is it? **American Documentation**, v.19, n.1, p.3-5, Jan. 1968.

CAPURRO, Rafael. Epistemologia e ciência da informação. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO - ENANCIB, 5., 2003. Belo Horizonte. **Anais...** Belo Horizonte: UFMG, 2003. Disponível em: http://www.capurro.de/enancib_p.htm. Acesso em: 25 set. 2018.

CAPURRO, Rafael; HJORLAND, Birger. O conceito de informação. **Perspect. Ciênc. Inf.**, Belo Horizonte, v. 12, n. 1, p. 148-207, Abr. 2007. Disponível em: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1413-99362007000100012&lng=en&nrm=iso. Acesso em: 15 set. 2018.

CARVALHO, José; FURTADO, Filipe; PRÍNCIPE, Pedro. WP4 – D25 – Kit sobre dados de investigação. **Repositório Científico de Acesso Aberto de Portugal**. Abr., 2017. Disponível em: <http://projecto.rcaap.pt/index.php/lang-pt/consultar-recursos-de-suporte/remository?func=fileinfo&id=395>. Acesso em: 27 set. 2018.

CAVALCANTI, M. T.; SALES, L. F. Gestão de dados de pesquisa: um panorama da atuação da união europeia. **BIBLOS - Revista do Instituto de Ciências Humanas e da Informação**, v. 31, n. 1, p. 73-98, jan./jun., 2017. Disponível em: <https://periodicos.furg.br/biblos/article/view/5789>. Acesso em: 29 set. 2018.

CURTY, Renata Gonçalves; AVENTURIER, Pascal. O paradigma da publicação de dados e suas diferentes abordagens. In: Encontro Nacional de pesquisa em ciência da informação, 18, ENANCIB, 2017. **Anais...** Out., 2017, Marília, Brasil. 2017. Disponível em: <https://hal.archives-ouvertes.fr/hal-01637987/>. Acesso em: 06 set. 2018.

MATHEUS, Renato Fabiano. Rafael Capurro e a filosofia da informação: abordagens, conceitos e metodologias de pesquisa para a Ciência da Informação. **Perspectivas em Ciência da Informação**, Belo Horizonte, v.10 n.2, p.140-165, jul./dez. 2005. ISSN 19815344. Disponível em: <http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/341>. Acesso em: 30 set. 2018.

NEVES, Dulce Amélia de B. Ciência da informação e cognição humana: uma abordagem do processamento da informação. **Ciência da Informação**, Brasília, v. 35, n. 1, p. 39-44, jan./abr. 2006. Disponível em: <http://www.scielo.br/pdf/ci/v35n1/v35n1a05.pdf>. Acesso em: 30 maio 2019.

OTTONICAR, Selma Leticia Capinzaiki; SANTOS, Beatriz Rosa Pinheiro dos; MORAES, Isabela Santana de. Aplicabilidade da Competência em Informação e da Organização do Conhecimento no processo de Gestão da Informação. **RDBCI: Revista Digital de Biblioteconomia e Ciência da Informação**, Campinas, SP, v. 15, n. 3, p. 629-646, jul. 2017. Disponível em: <https://periodicos.sbu.unicamp.br/ojs/index.php/rdbci/article/view/8649647>. Acesso em: 13 set. 2018.

PAVÃO, Caterina Groppo, *et al.* **Acesso aberto a dados de pesquisa no Brasil: repositórios brasileiros de dados de pesquisa: Relatório 2018**. Disponível em: <https://dadosdepesquisa.rnp.br/wp-content/uploads/2018/09/RDPBrasil-AcessoAbertoDadosPesquisaBrasil-RepositoryosBrasileiros.pdf>. Acesso em: 29 set. 2018.

SALES, L. F.; SAYÃO, Luana. F. Curadoria digital: um novo patamar para preservação de dados digitais de pesquisa. **Inf. & Soc.: Estudos, João Pessoa**, v. 22, n.3. p.179-191, set./dez. 2012. Disponível em: https://www.icict.fiocruz.br/sites/www.icict.fiocruz.br/files/Curadoria%20digital_Luis%20Fernando%20Sayao.pdf. Acesso em: 10 mar. 2018.

SALES, L. F.; SAYÃO, Luana. F. Algumas considerações sobre os repositórios digitais de dados de pesquisa. **Informação & Informação**, [S.l.], v. 21, n. 2, p. 90-115, dez. 2016. ISSN 1981-8920.

Disponível em: <http://www.uel.br/revistas/uel/index.php/informacao/article/view/27939>.
Acesso em: 27 set. 2018. doi:<http://dx.doi.org/10.5433/1981-8920.2016v21n2p90>.

SARACEVIC, T. Ciência da informação: origem, evolução e relações. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 1, n. 1, p. 41- 62, jan./jun. 1996. Disponível em: <http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/235>. Acesso em: 10 set. 2018.

ZINS, Chaim. Redefinindo a Ciência a informação: da “Ciência da informação” para a “ciência do conhecimento” **Inf. & Soc. Est.**, João Pessoa, v.21, n.3, p. 155-167, set./dez. 2011. Disponível em: <http://www.periodicos.ufpb.br/index.php/ies/article/view/11901>. Acesso em: 30 maio 2019.