

A GÊNESE DO REPOSITÓRIO FILATÉLICO BRASILEIRO: uma experiência interdisciplinar nas Humanidades Digitais

artigo de revisão

Diego Salcedo*

Vinícius Cabral Accioly Bezerra**

RESUMO

Apresenta o projeto em andamento intitulado “Repositório Filatélico Brasileiro – REFIBRA”. Explica as ações de curadoria digital realizadas pelo Grupo de Pesquisa IMAGO vinculado ao Departamento de Ciência da Informação da Universidade Federal de Pernambuco. De forma inédita no país, o projeto tem como objetivo desenvolver um repositório digital do conjunto de documentos filatélicos brasileiros produzidos desde 1843. O artigo indica aplicações de tecnologias computacionais articuladas à interdisciplinaridade das humanidades digitais. Como metodologia faz a revisão bibliográfica desde alguns elementos propostos na literatura científica especializada da Biblioteconomia, Ciência da Informação, Computação e Filatelia. Aponta os critérios utilizados até então para a escolha da tecnologia computacional mais eficaz e eficiente. Ilustra uma proposta preliminar de aplicação web semântica para o tratamento informacional de 32 selos postais comemorativos brasileiros. Utiliza o padrão de metadados Dublin Core, com necessárias adequações, o sistema de visualização *Linked Data* e o *Resource Description Framework*. Os resultados obtidos e divulgados neste artigo colaboram com o avanço da pesquisa e são evidência de que o campo das humanidades digitais articula-se com outras áreas, enaltecendo sua faceta interdisciplinar, mas, também aprendendo como certa articulação de tecnologias computacionais com demandas informacionais podem resultar em modelos de alta qualidade.

Palavras-chave: Humanidade Digital. Memória. Repositório. Selo Postal. Web Semântica.

* Doutor em Comunicação pela Universidade Federal de Pernambuco, Brasil. Professor da Universidade Federal de Pernambuco, Brasil. Docente permanente do Programa de Pós-Graduação em Ciência da Informação da Universidade Federal de Pernambuco, Brasil. E-mail: salcedo.da@gmail.com.

** Graduado em Sistemas de Informação pelas Faculdades integradas Barros Melo, Brasil. Mestrando no Programa de Pós-Graduação em Ciência da Informação da Universidade Federal de Pernambuco, Brasil. E-mail: vviniciuscabral@gmail.com.

I INTRODUÇÃO

O projeto de desenvolvimento do Repositório Filatélico Brasileiro (REFIBRA), no qual serão inseridos a Bibliografia Filatélica Brasileira (BIFIBRA) e o Dicionário Filatélico Brasileiro (DIFIBRA), todos eles produtos inéditos no país, e que está devidamente homologado pelas respectivas instâncias da Universidade

Federal de Pernambuco (UFPE), bem como cadastrado no grupo de pesquisa do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), articula atividades de pesquisa, ensino, extensão e cultura coordenadas pelo Grupo de CoPesquisa Imago e Humanidades Digitais, em parceria com o Departamento de Ciência da Informação (DCI) e o Programa de Pós-Graduação em Ciência da Informação (PPGCI), da UFPE, desde 2014.

Em maior escala, o REFIBRA contribui com um conjunto de princípios que norteiam os programas estratégicos de uma rede nacional de instituições comprometidas com políticas de preservação e digitalização de acervos memoriais brasileiros. O REFIBRA, assim, tem se constituído como o ambiente em que está acontecendo um processo de curadoria digital do patrimônio memorial filatélico brasileiro que, por sua vez, tem sido produzido ininterruptamente desde 1843. Esse patrimônio é constituído por diversos e distintos documentos filatélicos: selos, inteiros, blocos, editais, catálogos, livros, periódicos, boletins, folhetos etc. Vale lembrar que o Brasil, em 1843, foi o terceiro país a emitir e utilizar o selo postal para franquear correspondências, de fato, um pequeno embaixador de papel (ALTMAN, 1991; SALCEDO, 2010).

O projeto corrobora com as ações de preservação da memória digital recomendadas pela Organização das Nações Unidas para a Educação, a Ciência e a Cultura (UNESCO), na sua Declaração de Vancouver (2012). Logo, insere o Brasil como um dos pioneiros no cenário internacional, particularmente, no campo da Filatelia, da História e da Memória Postal. Na esfera nacional, o projeto tem desenvolvido um ambiente digital que contribuirá para a preservação, conservação e difusão da memória, da cultura e da identidade nacional em consonância, também, com o Plano Nacional de Cultura - PNC, do Ministério da Cultura do Brasil (2010).

Nesse contexto, o artigo apresenta o desenvolvimento da etapa inicial do REFIBRA, como ocorreu sua concepção e quais as tecnologias computacionais utilizadas. Como metodologia faz a revisão bibliográfica desde alguns elementos propostos na literatura científica especializada da Biblioteconomia, Ciência da Informação, Computação e Filatelia, caracterizando o que pensamos constituir certo desafio interdisciplinar. Ilustra uma proposta preliminar de aplicação web semântica para o tratamento informacional de 32 selos postais comemorativos brasileiros e utiliza o padrão de metadados Dublin Core, com necessárias adequações, o sistema de visualização *Linked Data* e o *Resource Description Framework*.

O artigo está estruturado da seguinte maneira: após esta introdução segue a seção 2, que apresenta a importância do selo postal e sua

relação com a memória nacional. Em seguida, a seção 3 apresenta o que há de se conceituar, entender e aplicar para que possa trabalhar com recursos na web semântica, nesse caso, os selos postais. A seção 4 apresenta como foi desenvolvido a fase inicial o REFIBRA com a aplicação das tecnologias apresentadas na seção anterior. Por fim, a seção 5 apresenta as conclusões e as proposições de continuidade do projeto.

2 O SELO POSTAL COMEMORATIVO BRASILEIRO COMO INTERFACE DA MEMÓRIA

O que dizer de tão rica, lúdica e poderosa interface de informação? Esse pequeno pedaço de papel, indiferente às diversas formas como se apresenta e aos suportes aos quais é agregado, elimina distâncias, preserva com criatividade uma possível história da humanidade. Resgata na forma de documento, pessoas, fatos, eventos, processos e o tempo, de forma geral, funcionando como um elo entre o humano, sua história e o conhecimento político, econômico, social e cultural. Assim, conforme sugere Salcedo (2010, p. 73),

o selo postal oferece a oportunidade para que possamos, se olharmos atentamente, perceber as transformações pelas quais temos passado, como conduzimos o desenvolvimento tecnológico, como nos distanciamos ou aproximamos do Outro, como lidamos com as diferenças e as semelhanças, como continuamos contando a nossa própria história e a da Natureza, como dizemos ou silenciemos nossos discursos e como os Estados ramificam os seus.

O selo postal, enquanto interface documental tem características marcantes, entre as quais, ser aproveitado tanto para enviar mensagens ao grande público como educar novos leitores de mundo. Isso é possível porque sua estrutura física enquanto objeto material é de fácil manuseio. Aos selos poder-se-ia aplicar a síntese feita por Abraham Moles (2004) para o cartaz. Eles têm a função de informar (qual a comemoração, o feito, o objeto da emissão) e de educar (dando a conhecer valores ou temas que escolheram).

Os selos postais servem, por exemplo, para reconhecer e legitimar uma nação, estão entre os produtos do Estado que possuem maior visibilidade interna e externamente. Enquanto a moeda é produzida conceitualmente para o uso interno de uma unidade política, o selo postal, mais que isso, assume um papel diplomático ao ser utilizado, também, além das fronteiras de sua unidade política emissora. Selos postais são produtos do Estado emissor, difundem universalmente o simbolismo público e as tradições nacionais tanto dentro de suas fronteiras, quanto ultramar.

De acordo com as especificações de texto e imagem impressos os selos postais podem ser categorizados como uma mídia, um meio de comunicação. Além disso, a mensagem impressa no selo é intencional, detentora de mensagens subliminares que, em certa medida, passam despercebidas ao leitor comum, ou ao grande público, que apenas o identifica como taxa devida ao Correio para o envio de missivas postais.

Ora, os selos constituem um excelente meio de propaganda e, assim sendo, justo é que os governos se esforcem de representar fatos, episódios e vultos do nosso país. Os selos postais tornaram-se úteis artefatos ideológicos e culturais para os governos no sentido de promover certa imagem em casa e fora dela.

Interessam, então, as mudanças ocorridas nos períodos imperial e republicano brasileiros, particularmente, a partir de 1843, posto que refletiram diretamente na produção e circulação de selos postais, seus desenhos, suas tarifas e seus tipos. Ao mesmo tempo em que os próprios selos postais, de alguma forma, integravam um acervo documental (coleções) que materializava aquelas mudanças.

Ao ser traçado um panorama das linguagens visuais dos selos postais imperiais brasileiros é possível perceber um interesse em arrecadar receita por meio do erário elitizado, impondo um limite intransponível tanto em torno do sistema de comunicação quanto no do colecionismo às pessoas de menor poder aquisitivo. A popularização dessa linguagem diz respeito não só ao meio, propriamente dito, ao regime de informação em que se constitui, mas, também, às pessoas que o formam e seus parâmetros sócio-culturais.

Não é adequado separar esses pequenos textos do quadro social em que foram produzidos

e nos quais circularam. No entanto, também não é adequado analisar o selo postal sem considerar a interface documental iconográfica como registro produzido pelo observador; o documento visual como registro ou parte do observável, na sociedade observada; e, finalmente, a interação entre observador e observado.

Por volta de 1900, no Brasil, existia o que pode ser denominado de “dialética do colecionismo filatélico”. Ela se baseia, principalmente, nas conexões entre o colecionismo individual e o coletivo. Parte fundamental desse movimento é o papel da instituição produtora dos objetos colecionáveis, no caso dos selos postais, os Correios. Assim, é possível perceber que, naquele período e que dura até hoje, teve início um movimento de retroalimentação, entre os Correios, e os colecionadores, incluídos os comerciantes filatélicos. Desde então, os Correios passaram, a produzir uma série de produtos para atrair a atenção dos colecionadores e comerciantes filatélicos. É nesse contexto que emerge, dentre variados tipos de documentos filatélicos, o selo postal do tipo comemorativo, objeto específico do que trata este artigo e que serviu como objeto para o protótipo do REFIBRA, tema tratado nas seções seguintes.

3 WEB SEMÂNTICA E HUMANIDADES DIGITAIS: O DESAFIO INTERDISCIPLINAR

As tecnologias de informação e comunicação têm aumentado a facilidade de distribuição de informação através da WEB, tendo o volume de documentos e *links* navegáveis disponíveis, para qualquer pessoa que tenha acesso à Internet, crescido muito mais se comparado ao começo da década. Todo esse volume e diversidade de informações têm criado dificuldade para que o usuário encontre conteúdos relevantes para suas necessidades informacionais. Como abordado por Souza e Alvarenga (2004), apesar da WEB ter sido projetada para facilitar a troca de informações, seu crescimento se deu de forma rápida e caótica, resultando em um enorme repositório de documentos recuperáveis. Esse referido repositório é conhecido atualmente como Web Sintática, uma vez que tratada as informações

de maneira a comparar estrutura de padrões de texto idênticos para recuperação da informação (MARCONDES, 2012) não levando em consideração o contexto e as ligações de inferência dos documentos recuperados.

Nesse cenário, os resultados de pesquisas na WEB feitos por meio dos motores de busca apresentam como resultado uma listagem de documentos que muitas vezes contêm os termos pesquisados pelo usuário, mas que nada tem de relevante com o contexto da sua necessidade informacional, assim as consultas apresentam problemas de atinência com baixa revocação e precisão (SOUZA; ALVARENGA, 2004). A nova abordagem semântica da WEB proporciona uma outra forma de se pensar a respeito de informação disponível na Internet, como os conhecimentos disponíveis podem ser melhor recuperados e indexados, ou seja, essa nova arquitetura de dados tem influência direta no dia a dia do profissional da Ciência da Informação e de áreas correlatas.

Dessa forma, Souza e Alvarenga (2004) acreditam que a Web Semântica e a Ciência da Informação são campos de pesquisa e desenvolvimento que convergem para um mesmo ponto, se complementam. Os autores observam possíveis melhorias nas atuais atividades do profissional da informação e do tratamento do seu objeto de trabalho: melhoria nos motores de busca, novas construção de interfaces para o usuário, construção automática de tesouros e vocabulários controlados, melhor indexação automática de documentos, melhor gestão do conhecimento organizacional e uma gestão da informação estratégica mais eficaz. Nesse contexto, a web semântica também influencia diretamente na forma como as bibliotecas, os museus e os arquivos podem trabalhar e disponibilizar seus catálogos, além de promover uma interoperabilidade entre eles, como proposto por Marcondes (2012). Outros autores, como Serra e Segundo (2017) sugerem exemplos práticos de como as bibliotecas atuais, físicas e digitais, podem converter seus catálogos, por exemplo em MARC (sigla para *Machine Readable Cataloging* - catalogação legível por computador), para um modelo de dados RDF (sigla para *Resource Description Framework* - uma conjunto de especificações da *World Wide Web Consortium* - W3C, originalmente planejada como um modelo de dados para metadados),

com o intuito de atender exigências de princípios do *Linked Data* (dados ligados entre si com a finalidade de publicar e estruturar dados na WEB).

Essa diversidade de agentes interconectados em rede (humano-humano, humano-máquina e máquina-máquina) foi observada por Luciano Floridi (2002) ao apresentar, num positivo e pertinente debate a respeito da ética da informação contemporânea, o conceito de infosfera, o qual consiste em um ecossistema informacional constituído por entidades informacionais chamadas *Inforgs*, não necessariamente humanas, que tem como principal função colaborar com a dinâmica da informação num ecoambiente computacional.

De fato, há muita pesquisa e desenvolvimento em andamento com relação a projetos da Web Semântica e *Linked Data* existindo um grande número de vocabulários controlados e padrões de dados propostos, por exemplo como o *Bibframe* abordado por Ramalho (2016a; 2016b). No entanto, o escopo do desenvolvimento do REFIBRA tratado neste texto é o estudo de um modelo de dados para um repositório do patrimônio filatélico nacional de acesso aberto, baseado em tecnologia web semântica com princípios do *Linked Data*, conectado e interoperável com outras bases de dados abertas (*machine to machine*), de maneira que o conteúdo esteja acessível ao público potencial e que, como pesquisa em andamento, está caracterizada pelo que pensamos constituir certo desafio interdisciplinar para as áreas envolvidas.

Por essa razão foi escolhido o padrão de metadados Dublin Core para descrever os itens catalogados até então. Tal padrão possui um conjunto de quinze elementos extensíveis para descrever recursos eletrônicos na WEB apresentando características como: simplicidade, interoperabilidade semântica, extensibilidade, flexibilidade e certo consenso internacional, uma vez que mais de vinte países aceitam esse modelo como um metadado usável para buscas em meio eletrônico (GRÁCIO, 2002). Como proposta para atualizar a forma de como os documentos são tratados e recuperado na web, Berners-Lee (2001) propôs uma forma de pensar a WEB de maneira que tanto as máquinas, quanto as pessoas pudessem entender melhor o contexto que o documento está inserido,

fazendo inferências sobre eles (a web semântica). Ducharme (2013, p. 35, tradução nossa) define a web semântica “um grupo de padrões e melhores práticas para compartilhamento de dados e seus elementos semânticos na WEB para uso aplicado” sendo esse “conjunto de padrões” tecnologias que definem uma estrutura necessária para que as informações sejam disponibilizadas de uma forma interoperável, tais padrões, como o RDF e SPARQL (linguagem para recuperação/consultas de dados em formato RDF) são mantidos pelo W3C.

Concomitantemente com a idéia de dados inseridos em uma WEB que não trabalha mais apenas com comparação de texto, mas sim com a inferência semântica da informação, emergiu a idéia de *Linked Data* que são princípios e boas práticas que tratam qualquer informação como um recurso representado por *Uniform Resource Identifier* (URI) e potencialmente ligado à outros recursos, formando uma teia de interligações: “Berners-Lee criou o *Linked Data* como melhor prática para compartilhar dados por meio da infraestrutura WEB, assim, oferecendo excelentes guias para o desenvolvimento de uma infraestrutura de web semântica (DUCHARME, 2013, p. 36, tradução nossa).

Atualmente a proposta da web semântica se encontra em pleno desenvolvimento tendo grandes expoentes incentivadores como a Dbpedia que é um projeto que extrai informação da Wikipedia para ser estruturada de maneira semântica e disponibilizada no modelo necessário para que agentes inteligentes consigam usar essa informação. Porém o processo de mudança da web sintática para o novo modelo semântico é demorado, uma vez que depende de uma estrutura de colaboração universal de diversos interesses para disponibilização de informação interconectada (SANTOS e CARVALHO, 2007, p. 05).

Por sua vez, o RDF é um modelo de representação de dados largamente utilizado para troca de informações semanticamente relacionadas por sistemas na web (PEREIRA et al, 2016). Tal modelo é baseado na idéia de que todo item pode ser descrito e relacionado por meio de um conceito de triplas o qual o *subject* é o recurso em questão, o *predicate* é a propriedade que tem relação com o *object* é o valor da propriedade (LIMA; CARVALHO, 2005). Com a incorporação do valor de URIs tanto para o *subject* como o *object*, o RDF permite

a criação de um mapeamento relacional entre os itens, de modo a permitir uma navegabilidade e relacionamento compreensível por humanos e máquinas (COYLE, 2017).

No entanto os problemas começam a surgir quando é preciso pensar na forma de armazenamento de documentos em formato RDF, em como persistir os itens mapeados em meio digital. Inicialmente é possível pensar numa forma de armazenamento feito num banco de dados relacional, porém ainda assim é preciso pensar na melhor forma de armazenamento de dados e seus relacionamentos, pois eles podem ser muito complexos quando o usuário precisar recuperar itens relacionados com outros itens de natureza diferentes. Nesse cenário existe a proposta do banco de dados relacional no modelo *triple-store*, que consiste na criação de uma única tabela contendo apenas três colunas (*subject*, *predicate* e *object*), em que qualquer tipo de item pode ser mapeado de modo simples.

Apesar de existirem críticas relacionadas a degradação de desempenho dessa abordagem, como estudado por Perreira (2016) e Bayer (2014), ela é utilizada em diversos projetos como no caso *Jena TDB* (<https://jena.apache.org>) que é um repositório que faz parte do framework do projeto Jena. Esse projeto permite o armazenamento de informações no modelo de dados RDF, além de permitir consultas complexas usando a ferramenta Jena Fuseki, a qual fornece uma interface para utilização da tecnologia SPARQL. Por sua vez, esta é uma linguagem de consulta eficiente para recuperação de dados em banco de dados RDF e recomendada pela W3C.

Pelo fato do repositório *Jena TDB* ser um projeto de uso irrestrito de código e de fácil distribuição, optou-se por utilizar esse *framework* no desenvolvimento inicial do REFIBRA, uma vez que ele atende as necessidades iniciais do repositório e as suas finalidades, assunto apresentado e debatido na próxima seção.

4 DISCUSSÃO SOBRE O DESENVOLVIMENTO DO REFIBRA

A WEB alterou radicalmente a forma como produzimos e utilizamos dados e informação reduzindo, para bilhões de pessoas, o seu acesso. Os links de hipertexto permitem que os usuários

surfem ou naveguem no ciberespaço por meio de navegadores, mas, também, que recuperem e indexem documentos digitais recuperados pelos motores de busca. Como sugeriu Heery (2004, p. 263), “a identificação de recursos, elementos de metadados termos de vocabulários controlados é fundamental para estabelecer a infraestrutura da web semântica”. Algumas iniciativas que utilizam web semântica vão ganhando reconhecimento, financiamento e auxiliam na materialização do discurso científico que defende essas iniciativas como parte de um campo intitulado Humanidades Digitais.

No contexto do trabalho em andamento esse campo das Humanidades Digitais é amplo conceitualmente: para alguns tem o sentido de uma disciplina em construção, para outros parece ser um conjunto de práticas que elenca a inserção das humanidades tanto no contexto da cultura digital quanto no debate científico do paradigma da pós-custódia. Nesse sentido

Spiro (2011, tradução nossa) sugere valores que norteiam essas práticas, quais sejam: “abertura, colaboração, colegialidade, diversidade e experimentação”.

Articula-se com a Ciência da Informação num universo amplo e complexo, tanto a partir dos debates teóricos, quanto sob uma perspectiva das práticas de organização, seleção, disseminação e curadoria digital. De fato, o debate epistemológico na Ciência da Informação não pode prescindir das práticas das Humanidades Digitais e os constantes desafios demandados pela contemporânea cultura digital e sua forte característica interdisciplinar. Nesse contexto, para o desenvolvimento do REFIBRA, foi adotado o padrão de metadados Dublin Core justificada pela sua aceitação internacional e capacidade de extensão.

Para ilustrar o que foi exposto até este momento foi utilizado um selo postal comemorativo brasileiro (Figura 1).

Figura 1: selo postal brasileiro utilizado como exemplo neste artigo



Fonte: coleção particular dos autores

A partir da análise desse selo postal comemorativo foi definido um padrão mínimo, não exaustivo, de metadados apresentado na lista 1:

Lista 1 - Relação dos metadados em Língua Inglesa e seus respectivos atributos em Língua Portuguesa

Title: Atributo referente ao nome dado ao título do recurso, podendo o ser

o nome oficial do selo postal ou batizado no momento do cadastro.

Subject: Atributo envolve todos os assuntos que o selo postal trata, levando em consideração a imagem impressa no item e todo seu contexto relevante.

Relation: Atributo referente a qualquer recurso que tenha relação com o selo postal e que não faça parte do assunto principal tratado pelo item.

Description: Atributo referente a descrição extensiva e mais detalhada do contexto que o selo postal está envolvido, nesse atributo é envolvido a bagagem cultural e cognitiva do catalogado.

Creator: Atributo referente entidade responsável pela elaboração do conteúdo do recurso. No presente artigo, todos os selos são de responsabilidade da Comissão filatélica nacional.

Publisher: Atributo referente a instituição responsável pela difusão do selo postal. No presente artigo, todos os selos são de responsabilidade dos Correios.

Contributor: Ao artista que fez a imagem estampada no selo postal, seja ela um desenho, uma foto ou qualquer outra arte expressa.

Date: Atributo referente ao ano de publicação do selo postal.

Identifier: Atributo referente a identificação não ambígua selo postal.

Language: Atributo referente ao idioma do conteúdo do selo postal.

Coverage: Atributo referente a região a qual o selo postal pôde ser utilizado.

Paper Type: Atributo referente tipo do papel utilizado no selo postal.

Color Type: Atributo referente tipo de coloração do selo postal.

Filigree: Atributo referente ao tipo de filigrana utilizado no selo postal.

Print Process: Atributo referente ao processo de impressão utilizado para emitir o selo postal.

Print Location: Atributo referente ao local de impressão do selo postal.

Stamp Dimension: Atributo referente às dimensões do selo postal.

Stamps Issued: Atributo referente ao número de selos emitidos.

Stamp Series: Atributo referente ao nome da série que o selo postal faz parte.

Face Value: Atributo referente ao valor facial estampado no selo postal.

Perforation: Atributo referente dentição do selo postal

Por sua vez, na lista 2, a seguir, os respectivos atributos estão indicados, alguns em

formatos de link, considerada a leitura analítica do selo:

Lista 2 - Metadados em Língua Inglesa e seus atributos em Língua Portuguesa (Links)

Title: Selo Comemorativo do Centenário de Gilberto Freire.

Subject: <http://dbpedia.org/resource/Gilberto_Freyre>

Relation: <<http://dbpedia.org/resource/Sociology>>
<<http://dbpedia.org/resource/Centenarian>>
<<http://dbpedia.org/resource/Anthropology>>
<<http://dbpedia.org/resource/History>>
<<http://dbpedia.org/resource/Pernambuco>>

Description: Centenário do Nascimento de Gilberto Freyre

Creator: Comissão Filatélica Nacional do Brasil

Publisher: <<http://dbpedia.org/resource/Correios>>

Contributor: Márcio Rocha

Date: 24.03.2000

Identifier: S002

Language: <http://dbpedia.org/resource/Portuguese_language>

Coverage: <<http://dbpedia.org/resource/Brasil>>

Paper Type: <http://dbpedia.org/resource/Coated_paper>

Color Type: <<http://dbpedia.org/resource/Polychrome>>

Filigra: Não possui.

Print Process: <http://dbpedia.org/resource/Offset_printing>

Print Location: http://dbpedia.org/resource/Casa_da_Moeda_do_Brasil

Stamp Dimension: 55 x 22mm

Stamps Issued: 2.160.000

Stamp Series: Não possui.

Face Value: R\$ 0,36

Perforation: 11.5 x 12 mm

Determinado o padrão de metadados para esse selo postal, destarte um protótipo para os demais selos postais constituintes do REFIBRA, foi possível fazer catalogá-los numa base de dados RDF. Para tal, como indicado na Figura 2, foi utilizado o *Jena Fueki* para a inserção das triplas RDF utilizando os metadados propostos. Vale ressaltar que estamos ilustrando, unicamente, os atributos *title* (<http://purl.org/>

dc/elements/1.1/title) e *print location* (<http://metadadosrefibra/printLocation>). O valor desses *predicates* tem como seus respectivos *objects* (Title

= Selo Comemorativo do Centenário de Gilberto Freire) e (*print location* = Casa da moeda do Brasil).

Figura 2: inserindo os primeiros metadados

```

1 ▾ INSERT DATA{
2   <http://refibra/resource/S001>
3     <http://purl.org/dc/elements/1.1/title>
4       "Selo comemorativo do Centenário do Nascimento de Gilberto Freyre" .
5   <http://refibra/resource/S001>
6     <http://metadadosrefibra/printLocation>
7       <http://dbpedia.org/page/Casa_da_Moeda_do_Brasil>.
8 }

```

Fonte: os autores (2018)

Observe como na figura 2, entre as linhas de código 1 e 8, estão os comandos em SPARQL para inserir triplas RDF no banco de dados *Jena TDB*. A primeira tripla inserida pode ser identificada entre as linhas 2 e 4. Na primeira linha citada está o valor "<http://refibra/resource/S001>", para representar o *subject* tratado; segue com "<http://purl.org/dc/elements/1.1/title>", que representa o atributo *title* do Dublin Core e do padrão de metadados proposto e, por fim, o valor "Selo comemorativo do Centenário do Nascimento de Gilberto Freyre", para representar o *object*. Nas linhas 5 a 8 é inserida outra tripla, na qual fica evidente que o valor do *object* do atributo "<http://metadadosrefibra/printLocation>" é outro item semântico disponível em outra base de dados: a

Dbpedia, da Wikipedia: (<http://dbpedia.org/page/Casa_da_Moeda_do_Brasil>).

O que está exposto na Figura 2 denota a maneira potencial de utilizar um padrão de metadados existente dentro de uma lógica de triplas RDF, bem como a forma de empregar o Linked Data, uma vez que esse selo postal inserido na base de dados é um link semântico que direciona a máquina e/ou o usuário para uma outra base de dados, para um outro local de conexão em rede e compartilhamento de informação. Uma vez executados os comandos apresentados na Figura 2 o *Jena Fuseki* armazenará as triplas RDF em um banco de dados, a partir do qual o usuário poderá recuperar essas informações no futuro, nessa mesma interface, como apresentado na Figura 3 abaixo:

Figura 3: seleção de triplas inseridas

```

1 select ?subject ?predicate ?object
2 ▾ where {
3   ?subject ?predicate ?object
4 }

```

Fonte: os autores (2018)

Com a execução do comando apresentado na Figura 3 o *Jena Fuseki* apresenta o resultado da consulta como uma tabela de três colunas

contendo o valor do *subject*, *predicate* e *object*, respectivamente, como apresentado na Figura 4, a seguir:

Figura 4: resultado da consulta

	subject	predicate	object
1	< http://refibra/resource/S001 >	< http://purl.org/dc/elements/1.1/title >	"Selo comemorativo do Centenário do Nascimento de Gilberto Freyre"
2	< http://refibra/resource/S001 >	< http://metadadorefibra/printLocation >	< http://dbpedia.org/page/Casa_da_Moeda_do_Brasil >

Fonte: Os autores (2018)

Como pode ser percebido ao analisar a Figura 4, o retorno da consulta realizada são as triplas RDF que foram cadastradas anteriormente. Isso significa que as tecnologias computacionais desenvolvidas e aplicadas no REFIBRA resultaram satisfatórias em relação ao objetivo da pesquisa em andamento. O REFIBRA, então, está emergindo como um ciberambiente baseado em tecnologia web semântica a partir do qual dezenas de tipos de documentos filatélicos convergirão, não, apenas entre si, mas com hiperlinks na rede mundial de computadores caracterizando, assim, um modelo de aplicação na articulação aplicada entre a Ciência da Informação e as Humanidades Digitais, como sugere (KOLTAY, 2016, tradução nossa) quando aproxima processos investigativos no campo da Ciência da Informação, por meio de sua imanente natureza interdisciplinar, com práticas que utilizam procedimentos e instrumentos tecnológicos de última geração no campo das Humanidades Digitais.

5 CONSIDERAÇÕES FINAIS

O projeto REFIBRA contribui para uma nova abordagem de se trabalhar itens informacionais na web pautada, basicamente, com aplicações de tecnologias computacionais articuladas ao campo interdisciplinar das Humanidades Digitais. Parte do que tem feito o Grupo de CoPesquisa IMAGO da UFPE é, justamente, estimular que estudantes e professores de cursos como Biblioteconomia, Gestão da Informação, Comunicação, Design,

Museologia e Filosofia sedimentem novas formas de conhecimento por meio de modelos de produção e uso de informação em ambientes ou espaços digitais. Os passos que estão sendo dados, a partir dos selos postais, não exclui outros documentos, sejam eles de natureza verbal ou verbovisual (Salcedo, 2010).

O crédito que, como o nome do grupo indica, damos ao trabalho colaborativo e multidisciplinar tem sido salutar e positivo no desenvolvimento de dispositivos que constituem a estrutura desse ambiente digital, bem como na experiência de leituras, debates e convivência das pessoas que ali participam. O REFIBRA, então, sem que pudéssemos planejar tem se constituído num dispositivo digital que oferecerá acesso irrestrito e aberto, por meio de Linked Data, ao conhecimento na rede mundial de computadores. De fato, uma experiência interdisciplinar que articula saberes diversos da Ciência da Informação e das Humanidades Digitais.

Nesse sentido temos um papel social de mediadores de informação, apesar de isso estar articulado com ensino, pesquisa e extensão na esfera acadêmico-científica (SALCEDO e CRUZ, 2017) e o REFIBRA uma função de mediação cultural (SALCEDO, PESSOA E SILVA, 2017), na medida em que serve como plataforma digital de patrimônio filatélico brasileiro conectado em rede com diversas e distintas bases de dados no mundo.

Outrossim, tanto o conjunto de dados abertos, quanto as informações contidas nos documentos podem servir como amostra para futuras análises que busquem compreender de forma crítica como se dão as interações sócio-

culturais de produção, circulação e disseminação do conhecimento. Nesse contexto contemporâneo é salutar perceber que o REFIBRA, em que pese o desafio tecnológico que teremos adiante, articula três dimensões das interrelações supracitadas, a saber: interatividade, hipertextualidade e transmediação (MIRANDA e SIMEÃO, 2005; COYLE, 2017). O trabalho prosseguirá com o processamento de 32 selos postais comemorativos relacionados a Pernambuco, depois 3500 selos postais do tipo comemorativo brasileiro, em seguida mais de 5000 selos postais diversos e distintos e, por fim, a documentação filatélica nacional.

Por fim, vale apontar para duas dimensões que, por meio do desenvolvimento do REFIBRA, indicam o futuro dos campos científico supracitados. Uma dimensão diz respeito ao que chamaremos de solução técnica e a outra de solução teórica. Ambas estão articuladas e são mais complementares do que suplementares.

As duas dimensões aludem ao futuro aplicado e teórico dos campos supracitados neste artigo, no sentido de que algumas terminologias estão ganhando força e terreno no cenário internacional e nacional, a saber: curadoria digital, ciências da memória, estudos das interfaces, web semântica e, por fim, aquela terminologia que hoje, em vários lugares, está alcançando status de ciência: humanidades digitais.

Considerando a curadoria digital como um processo de interação e ciclo vital entre dados, informações, documentos, pessoas e instituições, desde a produção, passando pela distribuição e, atualmente, seu consumo e reutilização, é possível articulá-la com as duas dimensões supracitadas num contemporâneo que designa certo momento histórico idiossincrático e que, em certa medida, traz para um conflituoso debate alguns paradigmas consolidados nas áreas emergentes nos séculos XIX e XX, criando as condições de pesquisas futuras.

Artigo recebido em 16/02/2018 e aceito para publicação em 15/10/2018

THE GENESIS OF THE BRAZILIAN PHILATELIC REPOSITORY: an interdisciplinary experience in the Digital Humanities

ABSTRACT

It presents the project in progress entitled "Brazilian Philatelic Repository - REFIBRA". Explains the actions of digital curation carried out by the IMAGO Research Group linked to the Information Science Department of the Federal University of Pernambuco. The project aims to develop a digital repository of the Brazilian philatelic documents produced since 1843. The article indicates applications of computational technologies articulated to the interdisciplinarity of the digital humanities. As a methodology, the bibliographical review makes use of some elements proposed in the specialized scientific literature of Library Studies, Information Science, Computing and Philately. It points out the criteria used until now to choose the most efficient and effective computational technology. It illustrates a preliminary proposal of semantic web application for the informational treatment of 32 commemorative postage stamps. It uses the Dublin Core metadata standard, with necessary adaptations, the Linked Data visualization system and the Resource Description Framework. The results obtained and published in this article collaborate with the advancement of the research and are evidence that the field of Digital Humanities articulate its self-praising its interdisciplinary facet, but also learning about how a certain level of articulation of computational technologies with informational demands may result in high quality models.

Key-words:

Digital Humanities. Memory. Postage Stamp. Repository. Web Semantics.

REFERÊNCIAS

- ALTMAN, D. **Paper ambassadors: the politics on stamps**. London: Angus and Robertson: 1991.
- BAYER, F. R.; NESE, L. L.; SCHROEDER, Rebeca. **ntSQL: um conversor de documentos RDF para SQL**. Departamento de Ciências da Computação - Universidade do Estado de Santa Catarina. Joinville - SC - Brasil, 2014.
- BERNERS-LEE, T., LASSILA, O.; HENDLER, J. The semantic web. **Scientific America**, Maio 2001.
- DUCHARME, B. **Learning SPARQL: Querying and Updating with SPARQL 1.1**. 2. ed. Sebastopol (usa): O'reilly, 2013.
- CONEGLIAN, C. S.; SANTARÉM SEGUNDO, J. E. Europeia no Linked Open Data: conceitos de Web Semântica na dimensão aplicada das Humanidades Digitais. **Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação**, v. 22, n.48, p.88-99, jan./abr., 2017.
- COYLE, K. **Two FRBRs, Many Relationships**. 17 de maio 2017. Disponível em: <<http://kcoyle.blogspot.com.br/2017/05>>. Acesso em: 29 de maio 2017.
- FLORIDI, L. **The fourth revolution: how the infosphere is reshaping human reality**. Grate Britain: Oxford Press, 2014.
- GRÁCIO, J. C. A. **Metadados para a descrição de recursos da Internet: o padrão Dublin Core, aplicações e a questão da interoperabilidade**. 2002.127 f. Dissertação (Mestrado em Ciência da Informação) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2002
- HEERY, R. Metadata futures: steps toward semantic interoperability. IN: HILLMANN, D. I.; WESTBROOKS, E. L. (Eds.). **Metadata in practice**. Chicago: American Library Association, 2004. p. 257-271.
- KOLTAY, T. "Library and information science and the digital humanities: perceived and real strengths and weaknesses". **Journal of Documentation**, v.72, n.4, p.781-792, 2016. Disponível em: <<https://bit.ly/2QLvZ60>>. Acesso em: 2 ago. 2018.
- LIMA, J. C. de; CARVALHO, Cedric L. de. **Resource Description Framework (RDF)**. Goiania: Instituto de Informática Universidade Federal de Goiás, 2005.
- MARCONDES, C. H. "Linked data" - dados interligados - e interoperabilidade entre arquivos, bibliotecas e museus na web. **Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação**, v. 17, n. 34, p.171-192, maio/ago., 2012.
- MIRANDA, A.; SIMEÃO, E. (Orgs.). **Informação e tecnologia: conceitos e recortes**. Brasília: UNB, 2005.
- MOLES, A. **O cartaz**. 2. dd. São Paulo: Perspectiva, 2004.
- PEREIRA, W.; HEINRICH, T.; SCHROEDER, R. **Avaliação de Desempenho de Sistemas Relacionais para Armazenamento de dados RDF**. Departamento de Ciências da Computação - Universidade do Estado de Santa Catarina. Joinville - SC - Brasil, 2016.
- RAMALHO, R. A. Sá. Bibframe: modelo de dados interligados para bibliotecas. **Informação & Informação**, Londrina, v. 21, n. 2, p.292-306, 2016a.
- RAMALHO, R. A. Sá. Análise dos modelos de dados SKOS e BIBFRAME: novas perspectivas de representação na era dos dados interligados. In: ENCONTRO NACIONAL DE PESQUISA EM PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO, 17, Salvador, BA. **Anais...**, 2016b.
- SALCEDO, D. A. **A ciência nos selos postais comemorativos brasileiros: 1900-2000**. Recife: EDUFPE, 2010.

SALCEDO, D. A.; CRUZ, M. Biblioteconomia, Ciência e Filosofia: um debate necessário sobre teoria e prática no campo acadêmico-científico. **Inf. & Soc.:Est.**, João Pessoa, v.27, n.1, p. 47-58, jan./abr. 2017

SALCEDO, D. A.; PESSOA E SILVA, J. R. Disseminação da informação: o papel do bibliotecário-mediador. **Revista ACB: Biblioteconomia em Santa Catarina**, Florianópolis, v. 22, n. 1, p. 23-30, dez./mar., 2017.

SOUZA, R. R.; ALVARENGA, L. A Web Semântica e suas contribuições para a Ciência da Informação. **Ciência da Informação**, Brasília, v. 33, n. 1, p. 132-141, jan./abr. 2004.

SANTOS, F. C.; CARVALHO, C. L. de. **Aplicações de Suporte à Web Semântica**. Goiania: -, 2007. 25 p.

SERRA, L. G.; SANTARÉM SEGUNDO, J. E. O catálogo da biblioteca e o linked data. **Em Questão**, Porto Alegre, v. 23, n. 2, p. 167-185, maio/ago. 2017.

SPIRO, L. "Getting Started in Digital Humanities". **Journal of Digital Humanities**, v.1, n.1, 2011. Disponível em: <<https://bit.ly/1zSHIJB>>. Acesso em: 2 ago. 2018.

W3C SPARQL Query Language for RDF. Disponível em: <<https://www.w3.org/TR/rdf-sparql-query/>>. Acesso em: maio de 2017.