

## XVIII ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2017

### GT-8 – Informação e Tecnologia

#### SNIIC: A PLATAFORMA DIGITAL DE CULTURA BRASILEIRA SOB A PERSPECTIVA DAS CONVERGÊNCIAS ENTRE O DESIGN E A CIÊNCIA DA INFORMAÇÃO

Anahi Rocha Silva (Universidade Estadual Paulista - UNESP)

Maria José Vicentini Jorente (Universidade Estadual Paulista - UNESP)

#### *SNIIC: DIGITAL PLATFORM OF BRAZILIAN CULTURE UNDER THE PERSPECTIVE OF CONVERGENCES BETWEEN DESIGN AND INFORMATION SCIENCE*

#### Modalidade da Apresentação: Comunicação Oral

**Resumo:** Esta comunicação traz resultados da pesquisa de mestrado defendida em um Programa de Pós-graduação em Ciência da Informação no ano de 2016. Esta pesquisa teve como objeto o Sistema Nacional de Informações e Indicadores Cultural (SNIIC), criado pelo governo federal com o objetivo de mapear e agregar, em base de dados única, informações referentes à cultura brasileira e que apresenta grandes desafios para sua implementação e execução, ante suas características de sistema complexo, híbrido e convergente. Nesse sentido nos propomos a analisar o SNIIC sob as perspectivas da Ciência da Informação e do Design da Informação, estabelecendo como objetivo geral identificar recursos e subsídios que ampliem as possibilidades de acesso, compartilhamento e apropriação da informação. A pesquisa é classificada como bibliográfica, exploratória, teórico-aplicada, pois visa estabelecimento de um produto/serviço para melhorar o sistema de busca e recuperação da informação. Os resultados da análise confirmaram o grande potencial da plataforma, mas revelaram algumas dificuldades, como o sistema de busca e recuperação da informação. Para a correção deste problema sugerimos a aplicação prática de um tesouro sobre cultura, convertido no modelo *Simple Knowledge Organization System (SKOS)* em padrão *Resource Description Framework (RDF)*, com vistas à organização, recuperação e interoperabilidade entre diferentes sistemas de informação e Web Semântica, permitindo que o usuário realize consultas utilizando seu vocabulário de comum uso, provendo assim uma maior acessibilidade às informações mantidas pelo SNIIC.

**Palavras-Chave:** Ciência da Informação. Design da Informação. Convergências de Linguagens. Sistema Nacional de Informações e Indicadores Culturais. Curadoria digital da Cultura.

**Abstract:** This communication brings results of the master's research defended in a Graduate Program in Information Science in the year 2016. The National System of Information and Cultural Indicators (SNIIC) is a digital platform created by the federal government as a tool for the National Culture Plan, with the aim to map and aggregate in a single database, information related to Brazilian culture. It is a dynamic environment information that meets complex system characteristics,

convergent, hybrid, and that presents great challenges for its implementation and enforcement. In this sense we propose to analyze the SNIIC from the perspectives of Information Science and Information Design establishing the general objective of identifying resources and subsidies that increase the possibilities of access, sharing and appropriation of information. The research is classified as literature, exploratory, theoretical and applied, since it starts from a literature survey in order to understand and contextualize the information systems and cultural indicators and thus outlining the problem; It is exploratory because it has the platform itself as an object for data collection, research and analysis; theoretical and applied, it aims to establish a product / service to improve the system of search and information retrieval. The test results confirmed the great potential of the platform, but had some problems, such as deficiencies in information retrieval system. In order to correct this problem we suggest the practical application of a culture thesaurus, converted into the Simple Knowledge Organization System (SKOS) model in the Resource Description Framework (RDF) model, with a view to the organization, retrieval and interoperability between different information systems and Semantic Web , Allowing the user to perform queries using their common vocabulary, thus providing greater accessibility to the information maintained by SNIIC.

**Keywords:** Information Science. Information Design. Languages of Convergences. National System of Information and Cultural Indicators. Digital Curation of Culture.

## 1 INTRODUÇÃO

O Projeto Padrões Web em Governo Eletrônico (e-PWG), Programa de Governo Eletrônico do Governo Federal (GOVERNO ELETRÔNICO, 2010), define as estratégias de divulgação, estrutura de informação, conteúdo disponibilizado, bem como estabelece o nível de interação com os sujeitos, que podem ser classificados como:

- **página:** é a unidade mínima de informação;
- **sítio:** é um conjunto de páginas contendo serviços e informações de um dado órgão, empresa ou departamento;
- **portal ou plataforma:** agrega conjunto de informações de órgãos e departamentos diferentes, oferece acesso direto a diversas aplicações, informações e serviços de outros sítios a ele agregado, unificados pelo desenho e pelo fluxo de interação.

Segundo a classificação acima, as chamadas “plataformas digitais de cultura” são ambientes informacionais que visam medir e processar informações de forma colaborativa.

Estas plataformas digitais correspondem a sistemas integrados de informações culturais que promovem mediações entre as esferas econômicas, culturais e políticas. São progressivamente utilizadas pelos governos para auxiliar na gestão da cultura com o fim de buscar transparência das ações governamentais e cumprimento de questões afetas ao direito à informação (CALCAGNO; LERMAN, 2007, p.101).

No Brasil, o Governo Federal propôs a elaboração do Sistema Nacional de Informações e Indicadores Culturais (SNIIC), uma plataforma que pode ser considerada híbrida, convergente e complexa, na medida em que é criada e disponibilizada pelo poder público e precisa da participação popular para sua consecução, reúne objetos e recursos digitais, com o objetivo de agregar, em base de dados única, informações referentes à cultura brasileira e realizar o mapeamento de manifestações artísticas e culturais, localizando-as geograficamente.

A formação deste repositório de dados e informações governamentais pressupõe a colaboração de sujeitos registrados, os quais fornecem informações que lhes são solicitadas em formulários eletrônicos (RAC), posteriormente armazenados nesse sistema para consulta e formação de estatísticas e indicadores.

O SNIIC é um sistema de informações de grande importância nas formas de interação, participação social e construção da memória coletiva e, apesar de já ter sido

implementado, percebe-se que há um grande volume de informações que ainda precisam ser inseridas, organizadas, classificadas, descritas, catalogadas e disponibilizadas.

Em razão do tema ser relativamente recente no Brasil, verificamos a escassez de estudos sob o prisma da Ciência da Informação e do Design da Informação, e que desperta uma ampla agenda de pesquisa, pois apresenta questões importantes a serem consideradas:

- quais as possibilidades de construção colaborativa que a plataforma digital de cultura pode oferecer para a promoção da diversidade cultural;
- qual seu papel na construção, preservação e disseminação da memória coletiva e do patrimônio cultural;
- se os documentos produzidos nesse contexto subsumem ao conceito de documento arquivístico eletrônico e assim podem ser considerados;
- a possibilidade de utilização de um tesouro de patrimônio cultural brasileiro no SNIIC, como instrumento para transferência efetiva da informação;
- quais as vantagens de utilização de um tesouro de patrimônio cultural baseado no conceito de metadados e RDF, adotado pelo Simple Knowledge Organization System (SKOS).

Todos estas questões foram trabalhados na dissertação de mestrado, estruturada em cinco capítulos, pensados e distribuídos conforme os objetivos específicos a serem alcançados. A dissertação foi realizada na linha de pesquisa Informação e Tecnologia da área da Ciência da Informação na Universidade Estadual Paulista, por entendermos que é tarefa da academia fomentar a reflexão e o debate sobre temas atuais que envolvam, sobretudo, a disseminação e a mediação da informação em ambientes informacionais digitais, por meio do uso dos recursos tecnológicos.

No caso desta comunicação oral, optamos por destacar “A importância do tesouro no Sistema Nacional de Informações e Indicadores Culturais (SNIIC)”, onde discutimos a importância da utilização de um tesouro de patrimônio cultural brasileiro no Mapeamento da Diversidade Cultural, convertido no modelo *Simple Knowledge Organization System (SKOS)* em padrão *Resource Description Framework (RDF)*, com vistas à organização, recuperação e interoperabilidade entre diferentes sistemas de informação e *Web Semântica* e apresentamos os resultados deste desenvolvimento e aplicação.

A pesquisa é classificada como bibliográfica, exploratória, teórico-aplicada, pois visa estabelecimento de um produto/serviço para melhorar o sistema de busca e recuperação da informação, cujo processo de execução será apresentado nas próximas seções, com o detalhamento de cada etapa.

## **2 SNIIC: A PLATAFORMA DIGITAL BRASILEIRA DE CULTURA**

A plataforma digital cultural SNIIC, ao lado das tradicionais bibliotecas, arquivos e museus, representa espaço e ambiente de registro projetado para guardar, preservar e disseminar a informação e a memória; trata-se de um repositório de dados e informações para estudos, consultas, criação de indicadores.

O SNIIC está disponível ao público em um URL próprio<sup>1</sup>, foi baseado na colaboração aberta para a composição de uma base de dados pública, elenca os conceitos da web 2.0 (web colaborativa), com as potencialidades da participação direta da sociedade civil através de interfaces típicas das mídias sociais somados aos da web 3.0 (a web semântica) na elaboração e implementação de arranjos de metadados integradores para recuperação da informação e gerenciamento de identidade (atribuição/autoria) e serviços de busca de dados georreferenciados empregados na elaboração de mapa cultural (cartografias colaborativas) (BRASIL, 2011).

Essa dinâmica interativa pode ser realizada em diversos graus de participação pelo cidadão, de acordo com a atividade realizada, desde sua inscrição no sistema, com o preenchimento do formulário RAC, ao compartilhamento de tipos variados de mídia (fotografias, audiovisuais, etc), relativas às experiências culturais vividas ou realizadas em sua comunidade local (CULTURA DIGITAL, 2012).

As opções de consulta relacionadas aos sujeitos inscritos no SNIIC são: através do número do cadastro SNIIC, ou de sua localização (Estado/Cidade) georreferenciado no mapa, pela identificação de parte do número do CPF, se pessoa física ou CNPJ, caso trate de pessoa jurídica, bem como o mantenedor, tipo de objeto cultural e a ação pretendida (no caso visualização).

---

<sup>1</sup> <http://sniic.cultura.gov.br/>

Figura 1: Interface de consulta de sujeito.



Fonte: SNIIC (BRASIL, 2011)

A partir desse resultado geral, o usuário pode refinar sua busca, elegendo qual dos resultados pretende visualizar. Assim, após a escolha, o sistema apresenta uma tela contendo detalhadamente as informações constantes no formulário de cadastro, que foram prestadas pelo usuário, onde destacamos na Figura 2 o *locus* da atividade cultural e sua descrição através da auto-declaração.

Figura 2: Interface de resultado à consulta de usuário.

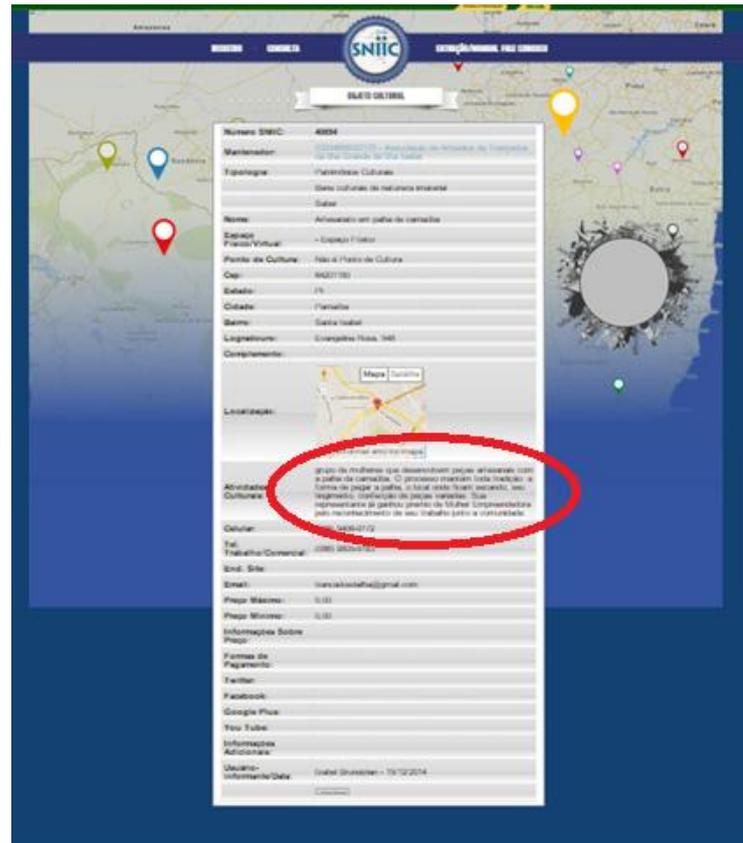
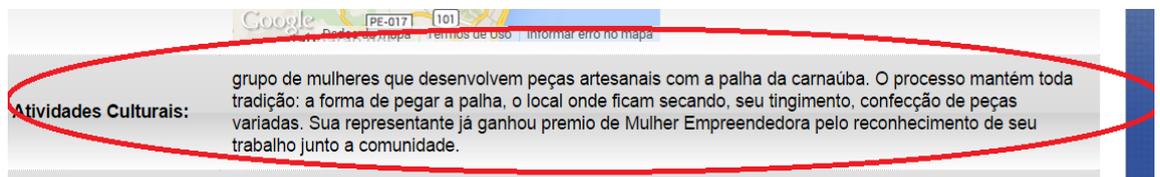


Figura 3: ítem destacado na Figura 2<sup>2</sup>.



Fonte: SNIIC (BRASIL, 2011)

Neste caso específico utilizado como exemplo do processo de busca e recuperação da informação no SNIIC, percebe-se que a informação descrita pelo usuário cadastrado localizado na cidade de Parnaíba, no Estado do Piauí, acerca de sua atividade cultural, é realizada em Linguagem Natural (LN)<sup>3</sup>. A partir desse texto descrito é possível identificar palavras-chaves: peças artesanais, palha de carnaúba, processo artesanal, tradição.

<sup>2</sup> <http://sniic.cultura.gov.br/index.php/usuarioInternetAction/consultaInformacoesView/40654>

<sup>3</sup> A linguagem natural (LN) é definida por Lancaster (1993, p. 200) como “Esta linguagem é pouco usada porque apresenta grande inconsistência devido inúmeras inconsistências causadas pela subjetividade. Para que o usuário venha obter uma maior precisão na sua busca a maior parte dos sistemas adota linguagens controladas com normas previamente estabelecidas. A linguagem controlada (LC), também denominada de vocabulário controlado, pode ser definida como um conjunto limitado de termos autorizados para uso na indexação e busca de documentos.

Por outro lado, além dessas informações revelarem o contexto de atividade vivenciada por uma pessoa inscrita no SNIIC, referidos termos destacados também possuem uma carga semântica e sintática com carga informacional de conceitos e definições próprias, que podem ser representadas na Web Semântica:

- Carnaúba é uma palmeira nativa do Nordeste do Brasil.
- Processo artesanal utilizando a palha da carnaúba.
- Processo artesanal que mantém toda a tradição local.
- Grupo de mulheres produzem artesanato com palha de carnaúba.

Dessa forma, a partir da Web Semântica, o computador consegue entender e fazer relações que carnaúba é uma palmeira nativa do nordeste e que com sua palha um grupo de mulheres da cidade de Parnaíba, no Estado do Piauí produzem peças de artesanato mantendo a tradição daquela comunidade local.

Ainda, tem-se que essas palavras-chaves serviriam como descritores para indexação, assumindo papel crucial na formulação correta da consulta com vistas à recuperação automática, bem como para própria formulação da política e dos indicadores culturais.

Inúmeros estudos sobre o uso da Linguagem Controlada e da Linguagem Natural na recuperação da informação têm se concentrado na utilização conjunta das duas linguagens na estratégia de busca, comprovando que o uso simultâneo dessas linguagens proporciona melhor desempenho nos resultados (LOPES, 2002).

Na prática, quem procura no SNIIC informações sobre determinada manifestação ou expressão cultural, ou mesmo aquele que insere informações na plataforma, o fazem a partir da linguagem natural com a qual está familiarizado (muitas vezes utilizando-se de termos locais, regionais ou coloquialismos) e em razão disso, têm grande chance de não obter êxito em sua busca.

Isso se deve ao fato do SNIIC não contar com uma política clara de tratamento da informação, nem instrumentos de controle de vocabulário, instrumentos básicos para a organização da informação; Além disso, também não oferece o usuário outros meios internos de busca e recuperação da informação.

## **2.1 TESAURO EM AMBIENTE DIGITAL**

O conceito tradicional de tesouro indica-o como linguagem de representação com a finalidade de homogeneizar terminologias e controlar vocabulários, conforme a norma norte americana ANSI/NISO Z39.19-2003:

“Vocabulários controlados organizados em uma ordem conhecida em que as relações de equivalência, homográficas, hierárquicas e associativas entre os termos são claramente exibidas e identificadas por indicadores padronizados de relacionamentos” (ANSI/NISO Z39.19-2003).

Apesar do tesouro ser uma linguagem documental que usa o controle de vocabulário para resolver problemas de ambiguidade da linguagem natural, a literatura mostra que a função do tesouro modificou-se ao longo do tempo (SERGE, 1997; SPARCK JONES, 1997) podendo ser utilizado como apoio:

- 1) para classificação de documentos, na caracterização de temas e categorização de conceitos;
- 2) à produção e à tradução de textos, principalmente na seleção de vocabulário;
- 3) à comunicação e aprendizado, na geração da estrutura conceitual;
- 4) e base conceitual para projetos, na produção do contexto conceitual;
- 5) à tomada de decisão, na classificação de assuntos;
- 6) à sumarização de textos, na identificação e associação dos principais conceitos desenvolvidos;
- 7) nos sistemas de recuperação de informação;
- 8) nos sistemas de navegação e rotulagem.

Assim, percebe-se que a função de tesouro em ambientes digitais vai além de uma mudança de suporte impresso para o digital, corresponde a uma mudança de paradigma com verdadeira redefinição dos princípios conceituais (Arano e Codina 2004).

De acordo com a norma ANSI/NISO Z39.19-2003 os tesouros em meio digital podem desempenhar 4 funções, sendo que uma quinta função foi incorporada pela revisão ANSI/NISO Z39.19-200X (MURAKAMI, 2005):

1. **Tradução:** promove um modo para traduzir a linguagem natural dos autores, indexadores e usuários para um vocabulário controlado usado para indexação e recuperação;
2. **Consistência:** promove consistência na designação de termos de indexação;
3. **Indicação de Relacionamentos:** indica relacionamentos semânticos entre termos;
4. **Recuperação:** serve como uma ajuda na busca e recuperação de documentos;

5. **Navegação:** provê hierarquias claras e consistentes em um sistema de navegação para ajudar usuários a localizar objetos de conteúdo desejados.

Referidas modificações no uso dos tesouros ocorreram à medida do desenvolvimento das tecnologias da informação e comunicação, e à nova dinâmica de gestão de informações na Web, cujas relações semânticas existentes tornam-se essenciais para o aprimoramento da Web Semântica (ARANO, 2005; RAMALHO, VIDOTTI, FUJITA, 2008; PASTOR SÁNCHEZ *et al*, 2009; ALVITE DÍEZ *et al*, 2010; PEÑA *et al*, 2013, p.110).

Shiri *et al* fizeram uma listagem dos tipos de tecnologias de representação dos tesouros presentes na Web (SHIRI, CHOWDHURY, CRAWFORD, 2002) :

1. Tesouro em um formato de texto estático simples;
2. Tesouro no formato HTML mas ainda estático, sem o uso efetivo de hyperlinks;
3. Tesouro no formato HTML com hyperlinks completamente navegáveis;
4. Tesouro com interface gráfica e visual avançada;
5. Tesouro em formato XML.

Além dessas, a representação de tesouros em *Resource Description Framework* (RDF)<sup>4</sup> proposto pela World Wide Web Consortium (W3C)<sup>5</sup> visa e promover interoperabilidade entre aplicações, trocando informações processáveis por máquinas na Web.

Apesar dessa recomendação da W3C, ainda a maioria dos tesouros são representados em bases de dados como linguagens de marcação tipo XML (*Extensible Markup Language*) e muito raramente são publicados em RDF, problema que pode ser resolvido através da aplicação de métodos para conversão de seu formato nativo para RDF (Miles *et al*. 2006).

Em razão da temática do objeto desta pesquisa versar sobre a cultura brasileira, pesquisamos tesouros como instrumento terminológico para realizar o controle de vocabulário. Para tanto, elegemos o Tesouro de Folclore e Cultura Popular Brasileira do Centro Nacional de Folclore e Cultura Popular do Instituto do Patrimônio Histórico e Artístico

---

<sup>4</sup> O *Resource Description Framework* (RDF) não é linguagem, é um modelo, um padrão para descrição de metadados, sendo possível descrever fatos sobre recursos da Web e promover interoperabilidade entre aplicações, trocando informações processáveis por máquinas na Web, recomendado pela W3C (MENDEZ, 1999) .

<sup>5</sup> o *World Wide Web Consortium* (W3C) é um consórcio internacional para difundir recomendações e tecnologias com padrões universais com vistas à Web Semântica, a fim de tornar todos os dados da Web legíveis tanto pelos humanos, quanto por máquinas. Dentre as recomendações da Web Semântica está o RDF.

Nacional (CNFCP, 2004), o qual consiste em um instrumento terminológico de indexação com classificação e conceituação dos termos associados à arte popular, ao indivíduo produtor, o produto e os processos empregados. Pode-se dizer que este tesouro possui um caráter inédito nas áreas de conhecimento denominadas folclore e cultura popular brasileira (MENDONÇA, 2012).

Figura 4: Tesouro de Folclore e Cultura Popular Brasileira do IPHAN

Tesouro de Folclore e Cultura Popular Brasileira

apresentação introdução parte sistemática parte alfabética fontes consultadas créditos

Índice  
A B C D E F G  
H I J K L M  
N O P Q R S T  
U V W X Y Z

[Busca nos Acervos Digitais](#)  
[Busca na Internet](#)

### Baião (dança)

Dança de pares em pequenos círculos com pares solistas ao centro. É dança caracteristicamente ligeira, com movimentos improvisados e ágeis, sapateados, palmas, giros, requebros, além do volteado e da "roda do galo"; às vezes há umbigada simulada. A música é feita com rabeca ou viola, pandeiro, triângulo. Dança obrigatória nos forrós nordestinos e que integra o bumba-meu-boi da Paraíba, do Rio Grande do Norte e de Pernambuco como solo musical dos personagens Mateus, Berico e Fidélis.

**Usado por**

- Baião (dança)

**Termo Genérico**

- Dança

**Termos Associados**

- Baião (atividade musical)
- Bumba-meu-boi
- Dança de roda
- Dança de umbigada
- Dança folclórica
- Forró
- Milindô

Fonte: Tesouro de Folclore e Cultura Popular Brasileira

Nossa proposta é que esse tesouro digital seja utilizado experimentalmente no SNIIC como parte de um *design* conceitual e estruturador desse ambiente digital como formas de representação, gestão uso e integração de recursos multimídias, complementando de forma mais eficaz e eficiente a consulta e a recuperação da informação.

## 2.2 SIMPLE KNOWLEDGE ORGANIZATION SISTEM (SKOS)

Para a consecução do tesouro, além das normas internacionais, que servem como diretrizes para sua elaboração, formatação e manutenção (ISO 2788 de 1986; ISO 5964 de 1985; ANSI/NISO Z39.19-2003 e ISO 25964-1 de 2011), o uso do padrão *Simple Knowledge*

*Organization System* (SKOS) passou também a ser recomendado pelo W3C a partir de 2009 (MOREAL et al 2010).

O SKOS foi desenvolvido pelo Grupo de Pesquisa em Tecnologia da Informação na Universidade de Murcia e da Cátedra UNESCO em Gestão da Informação nas Organizações. É um modelo que deve ser aplicado na construção de esquemas de conceito, fornece uma linguagem de modelagem conceitual simples e intuitiva, baseada em RDF, fazendo com que sejam lidos por agentes inteligentes e interoperáveis entre diferentes tipos de sistemas de organização do conhecimento (PÉREZ AGÜERA, 2004), visando a Web Semântica (BAKER et al, 2013).

Assim, o SKOS possibilita publicar linguagens de indexação através de um mecanismo (triplos - RDF) para representar relações entre diferentes esquemas conceituais, visando a recuperação de informação documental, o que possibilita migrar para a web semântica a baixo custo (MOREIRO GONZALEZ, 2011, p. 45)

O modelo SKOS relaciona os nós que representam o sujeito, predicado ou propriedade, e objeto, e triplos de formulário, ou seja, conecta conceitos seguindo a premissa de que o significado é expresso em RDF (LARA, 2013).

Segundo Moreira Gonzalez, a finalidade do SKOS está em representar a realidade dos termos incluídos em um determinado vocabulário controlado, respeitando a finalidade e o uso para o qual foi concebido (MOREIRO GONZALEZ, 2011, p. 46), aumentando o valor dos modelos configurados pelos tesouros e a capacidade de estabelecer relações entre unidades léxicas ou etiquetas, entre esquemas de conceitos e agrupá-los em coleções (PASTOR et al, 2009, p.12).

Isso se deve ao fato dos conceitos estarem ligados entre si por hierarquias, associações e agregados em esquemas conceituais que podem ser mapeados para os termos de outros regimes (LARA, 2013), propiciando padrões de interoperabilidade entre aplicações que são legíveis por máquinas e facilmente recuperar a informação.

No mundo, várias iniciativas, incluindo alguns vocabulários importantes, foram migrados para o formato SKOS e estão disponíveis no domínio público, a exemplo do EuroVoc, AGROVOC e GEMET. A “skosificação” já se faz presente em outros países, com base nisso, sugere-se sua replicação no SNIIC para uma melhor interatividade do sistema, bem como para a sua eficaz interoperabilidade entre outras aplicações e sistemas.

### 2. 3 DA TEORIA À PRÁTICA: APLICAÇÃO DO TESAURO NA PLATAFORMA SNIIC

A implementação da teoria descrita nesta comunicação envolveu uma equipe multidisciplinar para sua consecução, formada pelo grupo de estudo sob orientação do professor e pesquisador Ricardo da Silva Torres, do Instituto de Computação da Universidade Estadual de Campinas (UNICAMP), discentes Alberto Arruda de Oliveira, Karina Olga Maizman Bogdan e Pedro Ribeiro Mendes Júnior.

Este grupo desenvolveu um sistema de expansão de consultas para a plataforma SNIIC, para minimizar os casos de consultas mal sucedidas, permitindo que sujeito informacional realize consultas utilizando seu vocabulário natural, provendo assim um maior acesso às informações mantidas pelo SNIIC. O processo de execução do trabalho foi dividido em etapas nas quais se deu:

1. o processamento de consultas no TFCPB, subdividida em: (a) obtenção do tesauro, (b) representação no modelo RDF/SKOS e (c) elaboração e expansão de consultas utilizando o novo modelo do tesauro baseado no SKOS;
2. o processamento de consultas na base de dados do SNIIC, que consiste essencialmente no processamento das consultas, expandidas e não expandidas, diretamente na base do SNIIC.

O sistema de processamento de consultas desenvolvido, cujo código-fonte está disponível para *download*<sup>6</sup>, abrangeu duas funcionalidades principais: (1) Interface de expansão de consultas; (2) e interface de consultas (application programming interface - API) no SNIIC com termos expandidos; A primeira funcionalidade é provida por uma interface web que permite ao usuário a inserção de múltiplos termos que deseja expandir e selecionar os termos expandidos resultantes que julgar mais interessantes para busca no SNIIC.

O internauta pode entrar com um ou mais termos que deseja expandir, se deseja expandilos ou não e qual a profundidade de expansão desejada. Os termos devem ser inseridos separados por espaço ou unidos por símbolos "+", "," e ";".

Caso o termo desejado seja composto de diversas palavras, como, por exemplo, "Teatro de boneco", este deve estar entre aspas duplas. Se o botão de expansão de

---

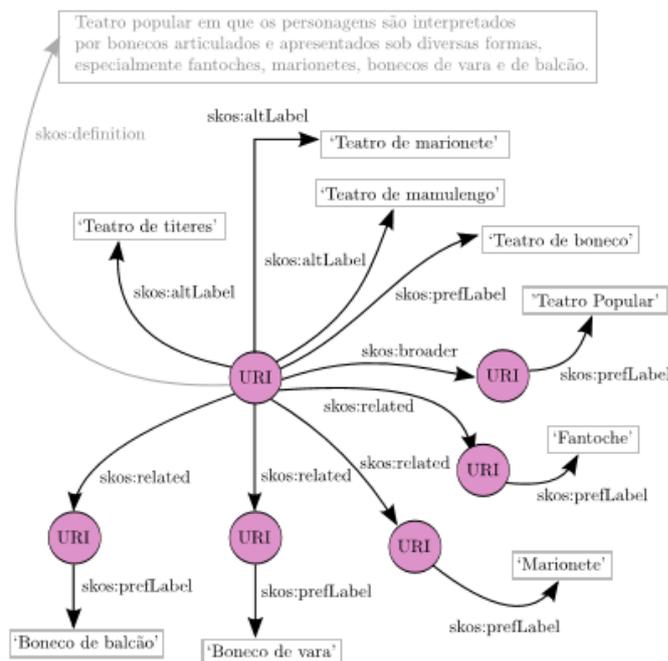
<sup>6</sup> Disponível em: <https://bitbucket.org/pedrorjunior/mo410-query-expansion/wiki/Home>

consultas for marcado, o sistema buscará no TFCPB representado em RDF os termos inseridos, listando como resultado os termos sinônimos, genéricos, específicos e associados.

A partir do tesouro anteriormente definido foi construído o grafo de representação considerando a definição de tripla- recurso, propriedade e valor- onde:

- **recurso:** representado pela URI do termo;
- **propriedade:** relacionamentos `prefLabel`, `altLabel`, `broader`, `narrower` ou `related`;
- **valor:** representado por um termo literal para as propriedades `prefLabel` (o termo) e `altLabel` (sinônimo do termo), ou por uma outra URI para as propriedades `broader`, `narrower` e `related`, que indicam o relacionamento entre conceitos.

Figura 5: Exemplo de mapeamento do termo “Teatro de boneco”



Fonte: Relatório final dos pesquisadores da Unicamp

Por fim, no campo *Search Depth*, o sujeito informacional indica a profundidade máxima de expansão, que só se aplica para a expansão de termos associados. Esse valor, um inteiro entre 1 e 4, indica se para os termos resultantes da busca original, novas buscas deverão ser realizadas.

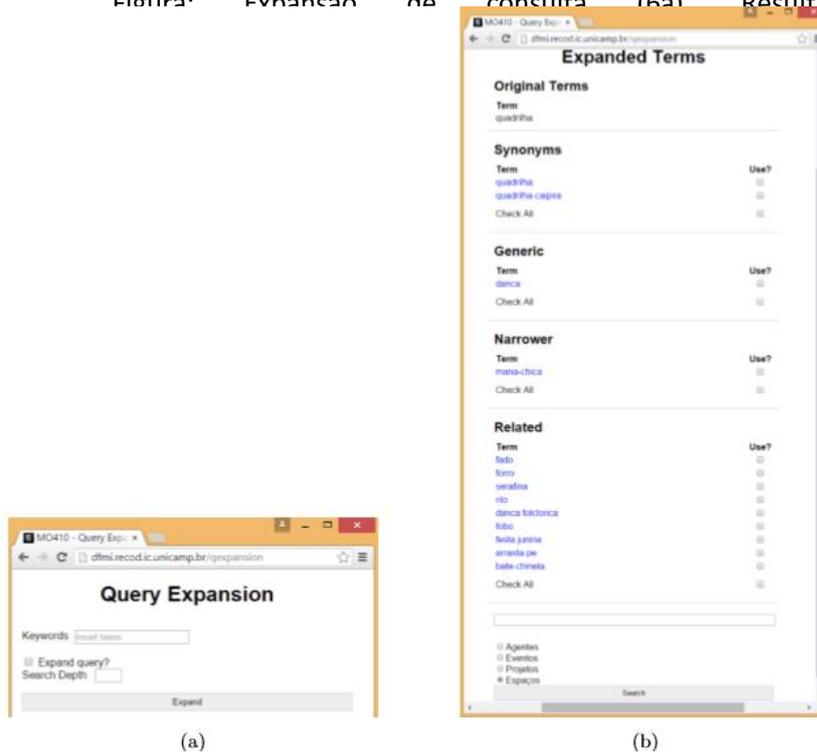
A profundidade 1, indica que apenas os termos originais, entrados pelo usuário, serão expandidos. Uma profundidade de 2, por exemplo, indica que para cada termo associado aos termos entrados pelo usuário, o sistema também realizará a expansão para eles. A profundidade é limitada em 4 pois, além de números maiores resultarem em uma expansão em termos que muitas vezes fogem do escopo do esperado pelo usuário, o processo de busca também torna-se muito lento.

Uma vez buscados os termos, o sistema avança para a tela de resultados da expansão, apresentada na Figura 6a. A tela é separada de acordo com a categorização dos resultados em termos sinônimos, genéricos, específicos e associados.

Cada um dos termos resultantes possui links para a página com sua descrição no site do TFCPB e uma caixa para marcar se o termo será utilizado na busca na base de dados do SNIIC. Permitindo ao internauta quais dos termos expandidos ele deseja utilizar na busca, o sistema torna possível uma melhor análise do impacto da expansão de consultas. É importante citar que os termos “entrados” pelo sujeito (que foram expandidos) aparecem na listagem de sinônimos, e devem ser marcados para serem adicionados à consulta final.

Os termos marcados aparecem unidos por “+” na caixa de texto mostrada abaixo da listagem de termos resultantes, podendo ser copiados para busca em outras plataformas. O sujeito então seleciona em qual das bases do SNIIC gostaria de realizar a busca, cujas opções são: Agentes, Eventos, Projetos e Espaços culturais. Clicando no botão *Search*, a consulta é realizada no SNIIC e o sujeito é direcionado para a página de resultados (ver Figura 6b). Se o internauta não selecionou nenhum termo para busca, o sistema volta para a tela inicial.

Figura: Expansão de consulta (6a) Resultados da consulta (6b).



Fonte: Relatório final dos pesquisadores da Unicamp

Os resultados apresentados na página de resultados (Figura 6 b) os termos originais são mostrados destacados em negrito e azul e os termos expandidos em negrito. Uma das limitações da API de busca do SNIIC é não considerar os termos buscados como partes de palavras existentes nos campos buscados. O código-fonte do sistema desenvolvido está disponível para *download* e as instruções para uso encontra-se acessível pelo mesmo link<sup>7</sup>.

### 3 CONSIDERAÇÕES FINAIS

Organizar, sistematizar e difundir informações objetivas sobre cultura, não é uma atividade simples ou auto-descritiva; representa grande desafio para a Ciência da Informação (CI). Torna-se, portanto, indispensável o emprego de novas abordagens tecnológicas de representação e recuperação de recursos informacionais (RAMALHO, VIDOTTI, FUJITA, 2007).

O SNIIC já continha um grande volume de informações culturais que necessitava ser processado e estruturado por meio de um *Design* de Informação apropriado para que os

<sup>7</sup> Disponível em: <https://bitbucket.org/pedrorrmjunior/mo410-query-expansion/wiki>.

aspectos inerentes às informações fossem tratados e recuperados. Assim, idealizamos o uso de um tesouro de patrimônio e expressões culturais para auxiliar na organização de informações no ambiente digital.

No primeiro momento o tesouro serviria para controle terminológico na indexação e disponibilização para o cidadão, leigo ou especializado e no momento da interação, converteria a linguagem natural em uma linguagem controlada. A seguir, a utilização dos padrões SKOS em RDF possibilitaria a criação de ligações e permitiria a melhoria da interoperabilidade entre outros sistemas de informação.

A utilização do tesouro também contribuiria para a construção de ontologias, possibilitando níveis de relacionamentos avançados que podem ser processados de forma automatizada por computadores (RAMALHO, VIDOTTI, FUJITA, 2007; BAÑOSMORENO et al, 2013).

No entanto, pesquisas sobre conhecimento tecnológico representam um grande desafio, pois de um lado estão os pesquisadores diretamente envolvidos com o desenvolvimento de sistemas e produtos e, de outro, os que são afetados pelas incorporações desses aperfeiçoamentos e avanços e finalmente, e aqueles que pesquisam e estudam esses fenômenos.

Esse desafio promovido pelas mudanças e alterações tecnológicas foi experimentado durante o desenvolvimento dessa pesquisa de mestrado, uma vez que a plataforma SNIIC, objeto de nosso estudo, sofreu profundas alterações em sua configuração inicial e isso nos leva a traçar algumas comparações e refletir sobre o assunto sob os pontos de vista da Ciência da Informação e do Design da Informação.

O sistema criado para solucionar problemas na recuperação da informação no SNIIC foi afetado pelas grandes alterações em sua interface gráfica e em suas funcionalidades advindas no fim de outubro de 2015. Isso porque a parte do sistema que havia sido implementado e que direcionava a consulta diretamente para a plataforma SNIIC, foi inutilizada e essa mudança afetou a disponibilidade dos dados referentes aos eventos culturais, o que impossibilitou testes do API.

Embora a plataforma SNIIC da maneira que foi projetada e apresentada primeiramente, não exista mais, nosso estudo e as sugestões desenvolvidas a partir do Laboratório multidisciplinar com o docente Ricardo Torres e seus discentes para a

incorporação de um Tesouro no ambiente, permanecem válidas e aplicáveis de forma genérica em outras plataformas digitais. Referidas mudanças intensificam ainda mais a necessidade de continuidade das pesquisas sobre a plataforma, sobretudo questões afetas às áreas da CI e do Design da Informação, que podem oferecer subsídios na construção de soluções.

Observou-se que estudar aspectos relacionados ao desenvolvimento e implementação de soluções na busca e recuperação da informação na plataforma SNIIC configurou-se em trabalho metodológico complexo e de difícil aplicação que na prática, extrapolava a área de conhecimento da Ciência da Informação, e, portanto, foi preciso a formação de equipe multidisciplinar.

Concluimos que a convergência entre a Ciência da Informação e o Design da Informação na plataforma SNIIC se dá em diversas camadas e diferentes graus de aplicações, tanto teórico como prático. Preferimos explorar a dimensão da interface de apresentação do SNIIC sob a perspectiva das duas áreas, e cuja pesquisa se direcionou para a necessidade de atuação pós-custodial do profissional da informação, por meio da curadoria digital e contextos multirelacionais de criação e uso da informação pelos sujeitos.

## REFERÊNCIAS

ALVITE DÍEZ, María Luisa *Beatriz*; PÉREZ-LEÓN, Mercedes; MARTÍNEZ-GONZÁLEZ, Dámaso-Javier *Vicente Blanco*. 2010. Propuesta de representación del tesouro EuroVoc en SKOS para su integración en sistemas de información jurídica. **Scire: Representación y Organización del Conocimiento**, volume 16, número 2, 2010. P. 47-51. Disponível em: <<http://ibersid.eu/ojs/index.php/scire/article/view/4015>>. Acesso em: jul.2015.

ARANO, Silvia. Los tesauros y las ontologías en la biblioteconomía y la documentación. *Thesauri and ontologies in librarianship and documentation*. **Hipertext.net**, Volume 14, número 3, 2005. Disponível em: <[http://eprints.rclis.org/8972/1/Los\\_tesauros\\_y\\_las\\_ontolog%C3%ADas\\_en\\_la\\_Biblioteconom%C3%ADa\\_y\\_la\\_Documentaci%C3%B3n.pdf](http://eprints.rclis.org/8972/1/Los_tesauros_y_las_ontolog%C3%ADas_en_la_Biblioteconom%C3%ADa_y_la_Documentaci%C3%B3n.pdf)><<http://www.hipertext.net/Web/pag260.htm>>. Acesso em: jun.2015.

BAKER, Thomas; BECHHOFER, Sean; ISAAC, Antoine; MILES, Alistair; SCHREIBER, Guss; SUMMERS, Ed. Key choices in the design os Simple Knowledge System (SKOS). **Web Semantics: Science, Services and Agents on the World Wide Web.**, Volume 20, Maio, 2013, P. 35-49. Disponível em: <<http://arxiv.org/pdf/1302.1224.pdf>>. Acesso em: jun.2015.

BAÑOS-MORENO, María-José; PASTOR-SÁNCHEZ, Juan-Antonio; MARTÍNEZ-BÉJAR, Rodrigo. Propuesta de actualización de macro-tesauros a partir de noticias de divulgación científico-tecnológica. 2013 . In: I Congresso ISKO España y Portugal / XI Congreso ISKO España, Porto (Portugal), 7-9 November 2013. Disponível em: <<http://eprints.rclis.org/20684/>>. Acesso em: jun.2015

BOCCATO, Vera Regina. CASARI; Ramalho, Rogério Aparecido; FUJITA, Mariângela Spotti Lopes. A contribuição dos tesauros na construção de ontologias como instrumento de organização e recuperação da informação em ambientes digitais. **Ibersid: revista de sistemas de información y documentación, IBERSID. (2008)**. Zaragoza, volume 2, 2008. Disponível em: <<http://ibersid.eu/ojs/index.php/ibersid/article/view/2235> >. Acesso em: jul.2015

BRASIL. MINISTÉRIO DA CULTURA. Sistema Nacional de Cultura. Guia de orientações para os municípios. Brasília, DF, 2011. Disponível em: <[http://www.cultura.gov.br/documents/10907/963783/cartilha\\_web.pdf/8cbf3dae-0baf-4a30-88af-231bd3c5cd6e](http://www.cultura.gov.br/documents/10907/963783/cartilha_web.pdf/8cbf3dae-0baf-4a30-88af-231bd3c5cd6e)>. Acesso em: jul. 2015.

\_\_\_\_\_. Sistema Nacional de Informações e Indicadores Culturais (**SNIIC**). Brasília, DF, 2011. Disponível em:< <http://sniic.cultura.gov.br/>>. Acesso em: jul. 2015.

\_\_\_\_\_. Metas do Plano Nacional de Cultura. Brasília, DF, 2011. Disponível em: <[http://www.cultura.gov.br/documents/10883/11294/METAS\\_PNC\\_final.pdf/3dc312fb-5a42-4b9b-934b-b3ea7c238db2](http://www.cultura.gov.br/documents/10883/11294/METAS_PNC_final.pdf/3dc312fb-5a42-4b9b-934b-b3ea7c238db2)>. Acesso em: jul. 2015.

CENTRO NACIONAL de FOLCLORE e CULTURA POPULAR. **Tesouro de Folclore e Cultura popular**. Rio de Janeiro: **CNFCP/IPHAN/MinC**; Brasília: Unesco, 2004.

CULTURA DIGITAL. SNIIC: versão 1.0 do Registro Aberto da Cultura (RAC) e Nuvem da Cultura. Disponível em: <<http://culturadigital.br/sniic/>>. Acesso em: mai. 2015.

IPHAN. **Tesouro de Folclore e Cultura Popular Brasileira**. Edição multimídia. Rio de Janeiro: 2006. Disponível em: <[http://www.cnfcp.gov.br/interna.php?ID\\_Secao=30](http://www.cnfcp.gov.br/interna.php?ID_Secao=30)> Acesso em: jul. 2015.

LOPES, Ilza Leite. Uso das linguagens controlada e natural em bases de dados: revisão da literatura. Ci. Inf., Brasília, v. 31, n. 1, p. 41-52, jan./abr. 2002. Disponível em: <<http://www.scielo.br/pdf/ci/v31n1/a05v31n1.pdf>>. Acesso em: jul. 2015.

Governo Eletrônico do Brasil. Programa de Governo Eletrônico Brasileiro - Sítio Oficial. **Padrões Web em Governo Eletrônico e-PWG - Cartilha de Redação Web (Webwriting)**. Versão 1.0 - Abril 2010 Disponível em: <<http://epwg.governoeletronico.gov.br/cartilha-redacao#padroes>>. Acesso em: set. 2017.

MÉNDEZ, E. RDF: un modelo de metadatos flexible para las bibliotecas digitales del próximo milenio. In: JORNADES CATALANES DE DOCUMENTACIÓ, 7., 1999, Barcelona. Actas...

Barcelona: E-LIS, 1999. Disponível em: <<http://eprints.rclis.org/12694/>>. Acesso em: 6 fev. 2016.

MENDONÇA, Elizabete. Narrativa sobre arte popular: estudo de caso sobre tesouro e exposições permanentes elaboradas pelo Centro Nacional de Folclore e Cultura Popular. *Museologia & Interdisciplinaridade*. Revista do Programa de Pós-Graduação em Ciência da Informação da Universidade de Brasília Vol.1, nº1, jan/jul de 2012. Disponível em: <<http://periodicos.unb.br/index.php/museologia/article/view/6845/5517>>. Acesso em: jul. 2015.

MILES, Alister; Van Assem, Mark; Malaise, Veronica; Schreiber, Guss. (2006). A method to convert thesauri to SKOS. In Proc. of the 3rd Annual European Semantic Web Conf., pp. 95–109. Disponível em: <<http://www.cs.vu.nl/~guus/papers/Assem06b.pdf>>. Acesso em: jul. 2015.

MONREAL CONCHA SOLER; Leiva, ISIDORO GIL. Posibilidades y límites de los tesauros frente a otros sistemas de organización del conocimiento: folksonomías, taxonomías y ontologías. **Revista . Interamericana. de Bibliotecología**. Medellín (Colombia), Volume. 33, número 2, julho-dezembro de 2010. Disponível em: <<http://www.scielo.org.co/pdf/rib/v33n2/v33n2a4.pdf>>. Acesso em: jul.2015

MURAKAMI, Tiago Rodrigo Marçal. **Tesauros e a World Wide Web**. São Paulo, 2005. Monografia (Trabalho de Conclusão de Curso) – Curso de Biblioteconomia e Documentação, Escola de Comunicação e Artes, Universidade de São Paulo. Disponível em: <<http://eprints.rclis.org/9863/1/murakami-tesauros.pdf>>. Acesso em: jul.2015

PASTOR-SANCHEZ, J.-A., MARTÍNEZ MENDEZ, F. J., & RODRÍGUEZ-MUÑOZ, J. V. Advantages of thesaurus representation using the Simple Knowledge Organization System (SKOS) compared with proposed alternatives. **Information Research**. SUÉCIA, VOL 14 Nº. 4, DECEMBER 2009. volume 14, número 4, dezembro 2009. Disponível em: <<http://www.informationr.net/ir/14-4/paper422.html>>. Acesso em: jul. 2015

PEÑA, CATALINA NAUM; MORÁN, ARIEL ANTONIO; MORALES. Consideraciones para construir un tesouro sobre leyes político-electorales usando tecnología digital. **Complexidade e organização do conhecimento: Desafios de nosso século**. Vera Dodebei & José Augusto Chaves Guimarães (Coords.), pp. 110-131, ISKO: FUNDAPE 2013. Disponível em: <<http://isko-brasil.org.br/wp-content/uploads/2013/02/Estudos-avan%C3%A7ados-2.pdf>>. Acesso em: jul. 2015.

PÉREZ AGÜERA, José Ramón. “Automatización de tesauros y su utilización en la web semántica”. *BiD: textos universitaris de biblioteconomia i documentació*, 2004, n. 13. Disponível em: <<http://bid.ub.edu/13perez2.htm>>. Acesso em: jul. 2015.

RAMALHO, ROGÉRIO APARECIDO SÁ; VIDOTTI, SILVANA APARECIDA BORSETTI GREGORIO; FUJITA, MARIÂNGELA SPOTI LOPES. Web semântica: uma investigação sob o olhar da Ciência

da Informação. **DataGramaZero – Revista. Ciência da Informação**. Inf., volume .8, número. 6, 2007. Disponível em: < [http://www.dgz.org.br/dez07/Art\\_04.htm](http://www.dgz.org.br/dez07/Art_04.htm)>. Acesso em jul. 2015.

SERGE. Cacaly. Dictionnaire de l'information sous la. dir. de Cacaly. Paris: Nathan. 2004. p 243

SHIRI, ALI; CHOWDHURY, GOBINDA; REVIE, CRAWFORD W. **Thesaurus-assisted search term selection and query expansion: a review of user-centred studies**. **Knowledge Organization**, Department of Computer and Information Sciences, University of Strathclyde, Glasgow, UK. pp. 1-19. Disponível em: <[https://www.researchgate.net/publication/230626824\\_Thesaurus-assisted\\_search\\_term\\_selection\\_and\\_query\\_expansion\\_A\\_review\\_of\\_user-centred\\_studies](https://www.researchgate.net/publication/230626824_Thesaurus-assisted_search_term_selection_and_query_expansion_A_review_of_user-centred_studies)>. Acesso em jul. 2015.

SOUZA, Rosali Fernades de. Thesaurus como linguagem de representação da informação. **Documentação em Museus /Museu de Astronomia e Ciências Afins**. Organização: Marcus Granato, Claudia Penha dos Santos e Maria Lucia N. M. Loureiro. Rio de Janeiro: MAST, 2008.

SPARCK JONES, K. E WILLET, P. (org.). Readings in Information Retrieval. San Francisco: Morgan Kaufmann Publishers, 1997. 589 p.

UNESCO. Convenção sobre a proteção e a promoção da diversidade das expressões culturais. Brasília: UNESCO, 2006.