GUIA

para a
integração entre

DSPACE E TEMATRES

Lucas Angelo Silveira
Rafael Fernandez Gomes
Raíssa da Veiga de Menêses
Milton Shintaku

ibict

MINISTÉRIO DA MULHER, DA FAMÍLIA E DOS DIREITOS HUMANOS

Secretaria Nacional da Juventude

MINISTÉRIO DA CIENCIA, TECNOLOGIA E INOVAÇÕES

Instituto Brasileiro de Informação em Ciência e Tecnologia

GUIA PARA A INTEGRAÇÃO ENTRE DSPACE E TEMATRES

Brasília

2020

PRESIDÊNCIA DA REPÚBLICA

Jair Messias Bolsonaro
Presidente da República

Hamilton Mourão
Vice-Presidente da República

MINISTÉRIO DA MULHER, DA FAMÍLIA E DOS DIREITOS HUMANOS

Damares Regina Alves
Ministra da Mulher, da Família e dos Direitos Humanos

Jailton Almeida do Nascimento
Chefe de Gabinete

Christiane de Oliveira Souza Zanzoni
Coordenação-Geral do Gabinete

Tatiana Barbosa de Alvarenga
Secretária Executiva

Cristiane Rodrigues Britto
Secretária Nacional de Política para as Mulheres

Angela Vidal Gandra da Silva Martins
Secretária Nacional da Família

Mauricio Jose Silva Cunha
Secretário Nacional dos Direitos da Criança e do Adolescente

Emily Rayanne Silva Coelho
Secretária Nacional da Juventude

Alexandre Magno Fernandes Moreira
Secretaria Nacional de Proteção Global

Sandra Terena
Secretaria Nacional de Políticas de Promoção da Igualdade Racial

Priscilla Roberta Gaspar de Oliveira
Secretaria Nacional dos Direitos da Pessoa com Deficiência

Antonio Fernandes Toninho Costa
Secretaria Nacional de Promoção e Defesa dos Direitos da Pessoa Idosa

MINISTÉRIO DA CIÊNCIA, TECNOLOGIA, INOVAÇÕES E COMUNICAÇÕES

Marcos Cesar Pontes
Ministro da Ciência, Tecnologia, Inovações e Comunicações

INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA

Cecília Leite Oliveira
Diretora

Reginaldo de Araújo Silva
Coordenação de Administração – COADM

Gustavo Saldanha
Coordenação de Ensino e Pesquisa, Ciência e Tecnologia da Informação – COEPPE

José Luis dos Santos Nascimento
Coordenação de Planejamento, Acompanhamento e Avaliação – COPAV

Anderson Itaborahy
Coordenador-Geral de Pesquisa e Desenvolvimento de Novos Produtos - CGNP

Bianca Amaro de Melo
Coordenadora-Geral de Pesquisa e Manutenção de Produtos Consolidados - CGPC

Tiago Emmanuel Nunes Braga
Coordenador-Geral de Tecnologias de Informação e Informática – CGTI

Milton Shintaku
Coordenador de Articulação, Geração e Aplicação de Tecnologia – COTEC

MINISTÉRIO DA MULHER, DA FAMÍLIA E DOS DIREITOS HUMANOS

Secretaria Nacional da Juventude

MINISTÉRIO DA CIENCIA, TECNOLOGIA E INOVAÇÕES

Instituto Brasileiro de Informação em Ciência e Tecnologia

GUIA PARA A INTEGRAÇÃO ENTRE DSPACE E TEMATRES

Lucas Angelo Silveira
Rafael Fernandez Gomes
Raíssa da Veiga de Menêses
Milton Shintaku

Brasília
2020

© 2020 Instituto Brasileiro de Informação em Ciência e Tecnologia
Esta obra é licenciada sob uma licença Creative Commons - Atribuição CC BY 4.0, sendo permitida a reprodução parcial ou total desde que mencionada a fonte.



Autores
Lucas Ângelo Silveira
Rafael Fernandez Gomes
Raíssa da Veiga de Menêses
Milton Shintaku

Revisor de texto
Flávia Sousa dos Santos

Diagramação e projeto gráfico
Mariana Lozzi Teixeira

Normalização
Ingrid Torres Schiessl

Impresso no Brasil

Guia para a Integração do DSpace com o TemaTres / Instituto Brasileiro de Informação em Ciência e Tecnologia e Secretaria Nacional da Juventude. -- Brasília: Ibict, 2018.

p. 42.

1. Interoperabilidade de sistemas. 2. Sistemas de informação. 3. eMag. 4.. I. Instituto Brasileiro de Informação em Ciência e Tecnologia. II. Secretaria Nacional da Juventude. III. Título

DD 025.47
CDU 027.43:004

Ficha catalográfica elaborada por Ingrid Torres Schiessl CRB1/3084

Esta obra é um produto do Projeto de pesquisa e desenvolvimento de soluções tecnológicas integradas para a gestão do conhecimento no âmbito da Secretaria Nacional de Juventude da Secretaria Geral da Presidência da República (SNJ/SG/PR), nas temáticas de Juventude e Participação Social - Plano de Trabalho Adicional - SNJ/MMFDH.Ref. SNJ - Processo SEI nº 00135.202553/2017-26

Ref. IBICT 1245/2015 - Processo SEI
Ref. FUNDEP 23506

As opiniões emitidas nesta publicação são de exclusiva e inteira responsabilidade dos autores, não exprimindo, necessariamente, o ponto de vista do Instituto Brasileiro de Informação em Ciência e Tecnologia ou do Ministério da Ciência, Tecnologia, Inovações e Comunicações.

SUMÁRIO

APRESENTAÇÃO 7

1 INTRODUÇÃO 10

2 DSPACE 13

3 TEMATRES 16

4 INTEGRAÇÃO DSPACE TEMATRES 19

4.1. CATALOGAÇÃO NO DSPACE COM USO DE TESAUROS DO TEMATRES 20

4.2 ALTERAÇÃO NO TEMATRES E DSPACE PARA A INTEGRAÇÃO 27

4.2.1 ALTERAÇÃO NO DSPACE PARA A INTEGRAÇÃO 28

4.2.2 ALTERAÇÃO NO TEMATRES PARA A INTEGRAÇÃO 37

5 CONSIDERAÇÕES FINAIS 40

APRESENTAÇÃO

Este guia apresenta a integração entre a Biblioteca Digital implementada com o DSpace e um Tesauros implementado com o TemaTres, resultante de estudos desenvolvidos no âmbito do projeto firmado entre a Secretaria Nacional da Juventude (SNJ), vinculado ao Ministério da Mulher, Família e Direitos Humanos (MMFDH), e o Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict), Unidade de Pesquisa do Ministério da Ciência, Tecnologia e Inovação (MCTI). Com isso, busca apresentar resultados de estudos aplicados às tecnologias utilizadas no projeto.

O projeto de pesquisa visa implementar um ecossistema de informação composto por vários sistemas independentes, mas complementares em suas funções. Para o presente guia destacam-se os sistemas de informação formados pela Biblioteca Digital e Tesauros, ambos voltados à Biblioteca do MMFDH. Assim sendo, o guia trata da integração do DSpace com o TemaTres, além de descrever como os dois sistemas podem ser integrados, compartilhando dados.

Cabe destacar a necessidade e os desafios enfrentados na gestão da informação pelas bibliotecas, com o uso cada vez mais frequente do formato digital, principalmente pelos órgãos de governo. Bibliotecas digitais têm sido utilizadas para a disponibilização de acesso de memória técnica de órgãos de governo,

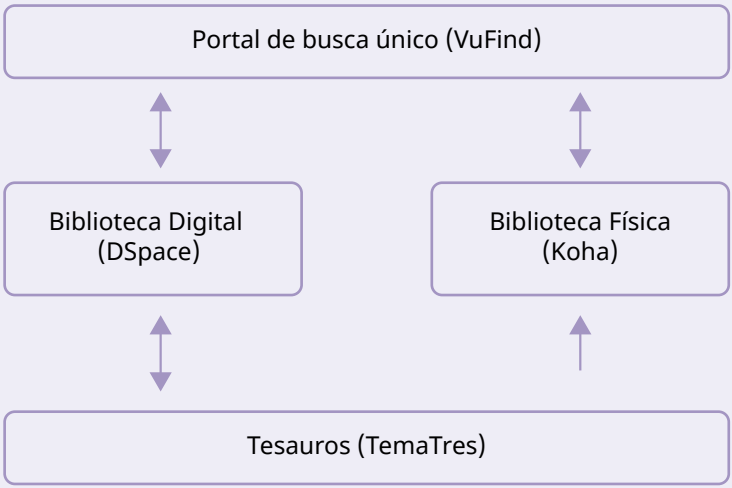
mas requerem esforços quanto à qualidade dos metadados, principalmente em relação aos assuntos tratados pelos documentos. Em alguns casos, Tesauros têm sido criados para criar listas de termos autorizados, a fim de facilitar o processo de indexação de memória técnica. Dessa maneira, visa integrar o Tesauros à biblioteca digital e possibilitar o uso dos termos mantidos no tesauros nos processos de catalogação e recuperação de documentos da Biblioteca Digital.

INTRODUÇÃO

1 INTRODUÇÃO

A Biblioteca do Ministério da Mulher, Família e Direitos Humanos gerencia quatro sistemas de informação distintos, o Sistema de Gestão de Acervo Físico, Biblioteca Digital, Tesauros e Portal de Busca Único. Assim, o ecossistema de informação mantido pela Biblioteca MMFDH (Figura 1) compõe-se de ferramentas livres, como o Koha, DSpace, TemaTres de VuFind. Todos os sistemas são independentes e podem ser acessados diretamente, mas possuem certa integração de dados.

Figura 1 - Estrutura de funcionamento do sistema de informação da Biblioteca MMFDH



Fonte: Elaboração dos autores (2020).

Usuários podem recuperar conteúdos digitais e físicos por meio do Portal de Busca Único, implementado com o sistema

de descoberta e entrega VuFind. Caso o registro for de um livro físico virá do Sistema de Biblioteca implementado com o Koha. Se for de um conteúdo digital, o registro virá da Biblioteca Digital, implementada com o DSpace. Tanto o Koha quanto o DSpace utilizam os termos mantidos pelo Tesauros para fazer a indexação.

No caso do sistema de biblioteca física não há integração, mas um processo de exportação dos termos do tesauros e uma importação no banco de autoridades de assunto do Koha. Mesmo que seja um processo manual, as tecnologias envolvidas – Koha e TemaTres – não possibilitam a integração fácil. Estudos mostraram impactos significativos relacionados ao processo.

O DSpace e o TemaTres, por outro lado, têm aspectos impactantes, mas a integração apresenta maiores vantagens. Geralmente os itens físicos passaram por processo editorial e, na maioria dos casos, já oferecem facilidades para a catalogação e indexação. Memórias técnicas, por serem itens únicos, apresentam maior desafio para esses processos. Assim, o protocolo de catalogação e indexação de memória técnica em biblioteca digital requer ajuda de vocabulários controlados com termos autorizados, a fim de oferecer maior qualidade aos metadados, facilitando o processo de recuperação.

2

DSPACE

2 DSPACE

O DSpace é um software livre, de código aberto, desenvolvido originalmente para a Biblioteca do Massachusetts Institute of Technology (MIT) pela Hewlett Packard (HP). Posteriormente foi aberto a todos que desejavam utilizar a ferramenta, de modo que a gestão foi transferida para a organização sem fins lucrativos DuraSpace. Como consequência, desde o seu lançamento em 2002, tornou-se a ferramenta livre mais utilizada no mundo para a criação de Repositórios Digitais (RD) e Bibliotecas Digitais (BD).

No Brasil, as universidades preferem o termo RD, por disponibilizar em grande parte documentação científica que passa por processo editorial. Os órgãos de governo, por sua vez, utilizam a denominação BD por disponibilizar memória técnica, sendo a primeira fonte. As diferenças de denominação encontram fundamentação nos primeiros sistemas de informação em que BD eram utilizados para disponibilizar literatura cinzenta que não passava por processo editorial e RD disponibiliza cópias de literatura científica.

O DSpace é uma ferramenta altamente customizável, que possibilita a criação de diversos tipos de BD e pode apresentar temáticas, multitemáticas, manter acervos de único autor ou de muitos. De igual modo, pode ser usada para criar repositórios digitais, institucionais, temáticos e outros. Independentemente se for biblioteca ou repositório, o DSpace oferece funcionalidades para tal e recuperação de itens de informação, de maneira que a qualidade dos metadados é fator imprescindível.

A alimentação no DSpace é feita por meio de processo simples, durante o qual a etapa de catalogação é feita por meio de formulário, que pode ser customizado no sentido de adaptar melhor às necessidades da BD ou RI. O formulário pode ser composto por vários tipos de campos, dentre os quais apenas o campo título seria obrigatório, mesmo que isso pudesse ser ajustado. Título refere-se, aqui, ao nome do item e à sua identificação, geralmente por meio de um campo livre.

O DSpace permite que os campos de metadados possam utilizar vários tipos de vocabulários controlados no intuito de normalizar alguns conteúdos. Esses campos normalizados podem tornar-se facetas para facilitar a recuperação. Assim, tipo de documento, idioma, formato do arquivo, data etc. são facilmente implementados com listas simples, mas assunto ou palavras-chave apresentam maior desafio. Em todo caso, de forma padrão pode-se implementar taxonomias estáticas para esses campos.

Assim, todas as funcionalidades oferecidas pelo DSpace estão voltadas à gestão de um acervo digital, nos processos de depósito e recuperação dos itens de informação. Dessa maneira, gestão de usuários e permissões são efetuados para dar acesso aos recursos e itens de informação mantidos pelo DSpace.

3

TEMATRES

3 TEMATRES

O TemaTres é uma ferramenta simples, mas muito flexível para criação e gestão de vocabulários controlados como taxonomias e tesauros. Tal ferramenta é mantida pelo Diego Ferreyra, da Argentina, e tem como slogan "gerencie, publique, compartilhe e re-use", "O software livre para gerenciar representações formais de conhecimento". Assim, oferece funcionalidades para adequação e disseminação de vocabulários controlados.

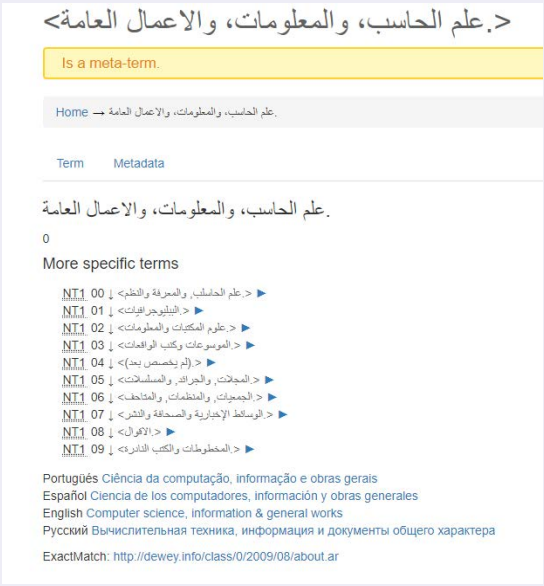
Mesmo sendo uma ferramenta livre criada na Argentina, o TemaTres é utilizado em vários países, levando-se em conta sua simplicidade e robustez. Este é um caso de sucesso incomum, na medida em que grande parte dos softwares livres são oriundos de países desenvolvidos como os Estados Unidos, muito provavelmente pela sua influência e forma de atuação. Atualmente, o TemaTres é utilizado por mais de 450 vocabulários controlados, disponíveis na Internet.

O TemaTres possibilita a criação de vários tipos de vocabulários controlados, dos mais simples, como os glossários e taxonomias, a estruturas mais complexas, como tipos de ontologias. O software permite registrar termos e suas definições em uma estrutura hierárquica, a fim de criar vocabulários controlados. Da mesma forma, possibilita recuperar tais termos por meio de ferramenta de busca ou navegação.

Ademais, o TemaTres possibilita vocabulários controlados em mais de um idioma, no intuito de possibilitar relacionamento entre vocabulários. Um exemplo do uso do TemaTres em idio-

ma que usa alfabeto diferente e registra outros idiomas pode ser visto na Figura 2, com um vocabulário controlado em árabe e com traduções para outros idiomas.[1]

Figura 2 - TemaTres em outro idioma



Fonte: Captura de tela (2020)

Assim, o registro dos termos no TemaTres pode seguir padrões internacionais, como o *Simple Knowledge Organization System* (SKOS), *Dublin Core* e outros. Com isso, possibilita o intercâmbio de dados com outros sistemas, visto que pode trocar dados em vários formatos e padrões. Atualmente já existe um plugin para a integração do TemaTres com *o Open Journal System* (OJS).

1 Disponível em: https://vocabularyserver.com/ddc/ar/index.php?tema=112&/aalm-lhsb-olmaalomt-olaaml-laam.

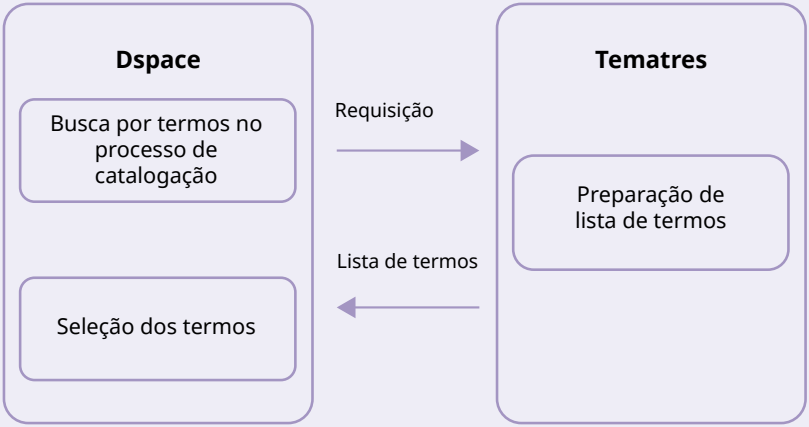
4

INTEGRAÇÃO DSPACE TEMATRES

4 INTEGRAÇÃO DSPACE TEMATRES

A integração entre o Sistema de Tesauros (TemaTres) e a Biblioteca Digital (DSpace) tem relação com a normalização dos termos inseridos no campo de assunto, por meio de um vocabulário com os termos autorizados. Assim, no processo de catalogação de itens – principalmente os pertencentes à memória técnica –, o usuário poderá selecionar termos pertencentes ao sistema de tesauros de forma automática, diretamente na biblioteca digital, conforme esquema apresentado na Figura 3.

Figura 3 - Estrutura de funcionamento da integração do DSpace com TemaTres



Fonte: Elaboração dos autores (2020).

Nesse sentido, os dois sistemas se complementam, na medida em que o TemaTres oferta os termos autorizados para o uso

no DSpace, normalizando as informações inseridas em campos como palavras-chave ou assunto. Assim, o uso da integração é vista apenas no DSpace, no processo de catalogação, mesmo que tenham sido efetuadas alterações nos dois sistemas.

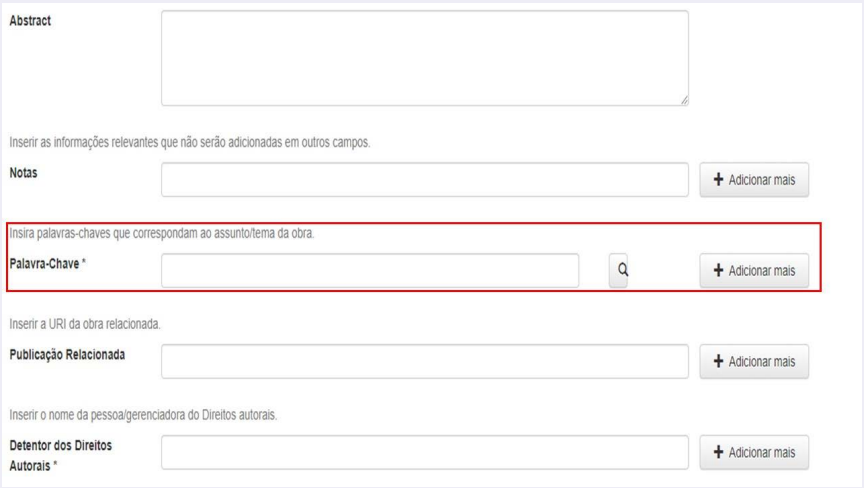
A inserção de termos autorizados apresenta impactos positivos em vários pontos, como na melhoria do processo de busca e descoberta, além de assegurar que a Biblioteca Digital usa apenas os termos corretos na descrição dos itens. O processo do uso de base de autoridades é comum em sistemas de bibliotecas tradicionais de objetos físicos, mas ainda com menor uso em bibliotecas digitais.

4.1. CATALOGAÇÃO NO DSPACE COM USO DE TESAUROS DO TEMATRES

Com o integrador implantado, todos os termos inseridos no TemaTres estarão disponíveis para serem utilizados no DSpace durante o processo de depósito, no momento da catalogação de um item na Biblioteca Digital e no preenchimento de campos como assunto ou palavras-chave. Para tanto, o DSpace tem que estar configurado no formulário de entrada, com os campos a serem integrados com o TemaTres.

A forma pela qual se deve utilizar o integrador inicia com um novo depósito, mais especificamente no processo de catalogação, com o preenchimento do formulário de entrada, no campo de palavras-chave, como mostra a Figura 4. O campo de palavras-chave ou assunto geralmente é colocado como obrigatório por muitas bibliotecas digitais, mas isso depende da configuração do formulário de entrada.

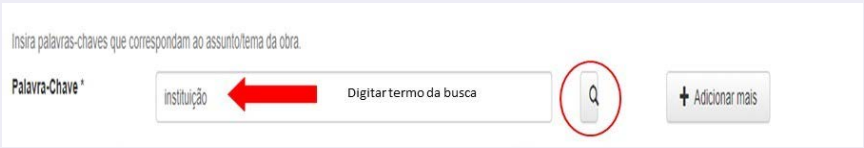
Figura 4 - Campo palavra-chave no DSpace



Fonte: Elaboração dos autores (2020).

Como mostra a Figura 5, uma a presença do integrador pode ser notada pela existência da lupa ao lado do campo de palavras-chave, inexistente nos formulários padrão. Assim, basta preencher o campo e clicar na lupa para buscar o termo inserido no TemaTres. No exemplo, utilizou-se a palavra “instituição”, para ver se existia na base dos termos autorizados contidos no tesauros.

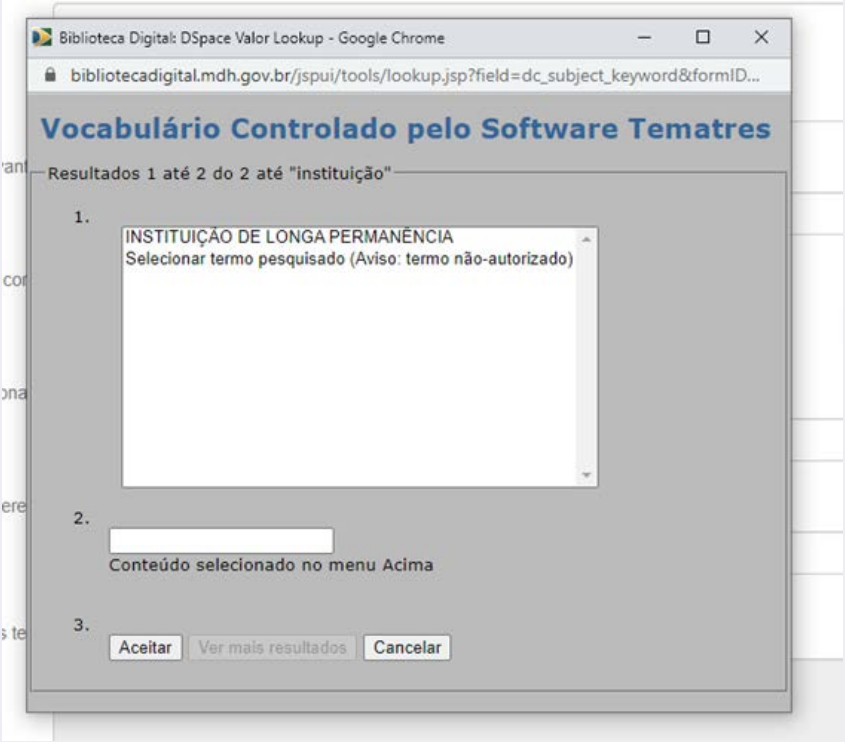
Figura 5 - Botão para buscar termos no DSpace



Fonte: Elaboração dos autores (2020).

A resposta vem em forma de uma janela PopUp, como mostra a Figura 6. A palavra inserida no campo será buscada no TemaTres. Como mostra o exemplo, a busca deu-se pela palavra "instituição" e, assim, na base de termos do TemaTres, havia apenas o termo "Instituição de longa permanência". Logo, o usuário pode utilizar o termo sugerido ou manter o termo pesquisado, mesmo que não seja autorizado.

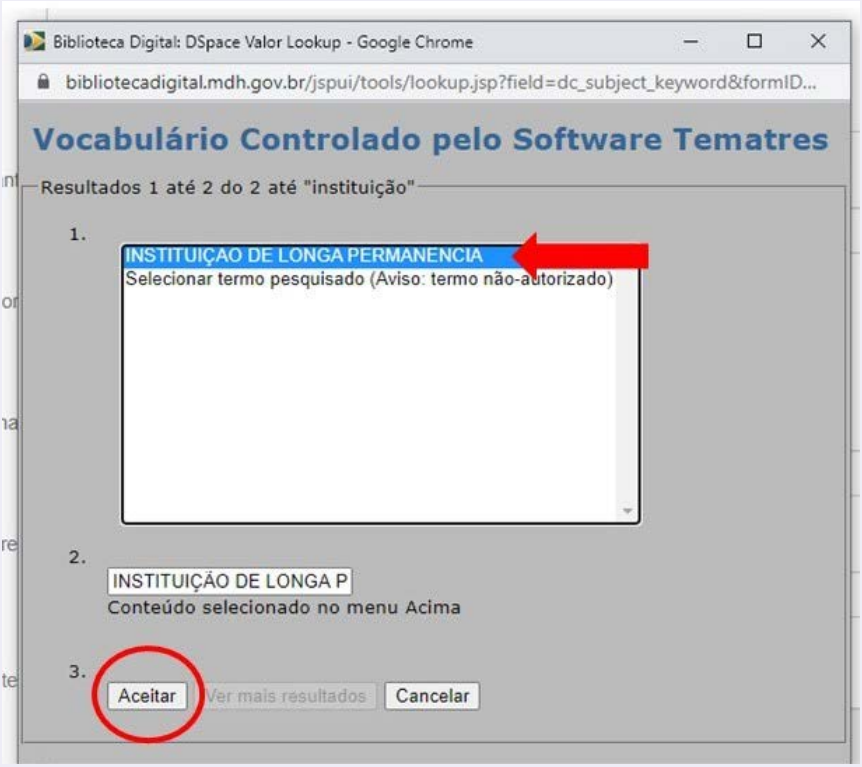
Figura 6 - Resultado da busca no TemaTres, via PopUp do DSpace



Fonte: Elaboração dos autores (2020).

Ao clicar no termo desejado, autorizado ou não, o termo vai para o campo 2, a fim de que possa ser validado pelo usuário, como mostra a Figura 7. Assim, para utilizar o termo selecionado (autorizado ou não) como palavra-chave na catalogação, basta clicar sobre o termo desejado e, depois, no botão Aceitar.

Figura 7 - Seleção do terno no TemaTres



Fonte: Elaboração dos autores (2020).

Ao clicar em Aceitar, a janela do Tesauro fechará automaticamente, e o termo selecionado aparecerá no campo do formulário de entrada (Figura 5). Com isso, o termo autorizado preen-

che o campo. Para adicionar outras palavras-chave na mesma catalogação, basta clicar em Adicionar mais e seguir os mesmos passos explicados. Caso queira desistir do termo, basta limpar o campo e iniciar uma nova busca.

Figura 8 - Botão “Adicionar mais palavras-chave”



Fonte: Elaboração dos autores (2020).

Caso adicione o termo e depois queira mudar, pode-se excluir o termo inserido, como mostra a Figura 6, clicando em Excluir esta entrada. Assim, pode-se reiniciar o processo e buscar novos termos. Na grande maioria dos casos, o campo palavras-chave ou assunto é obrigatório, por customização do formulário de entrada, não podendo ser deixado em branco.

Figura 9 - Excluir entrada de assunto

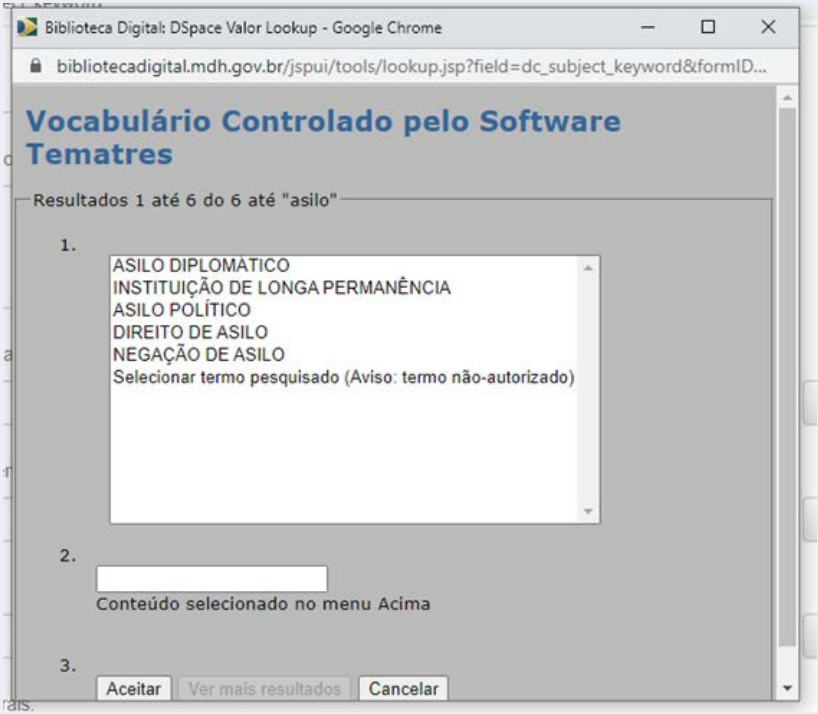


Fonte: Elaboração dos autores (2020).

O DSpace permite que sejam inseridos **Termos não autorizados** no processo e catalogação, ou seja, termos que não estão presentes no TemaTres. Assim, se oferece maior liberdade aos catalogadores na seleção e escolha dos melhores termos. No exemplo apresentado na Figura 7, a palavra pesquisada foi

ASILO. Cabe notar, no exemplo, que o termo INSTITUIÇÃO DE LONGA PERMANÊNCIA não possui a palavra asilo, diferente dos demais resultados. Isso acontece porque o termo Asilo para idosos é o **Termo não-autorizado** para INSTITUIÇÃO DE LONGA PERMANÊNCIA.

Figura 10 - Resultados pesquisa ASILO

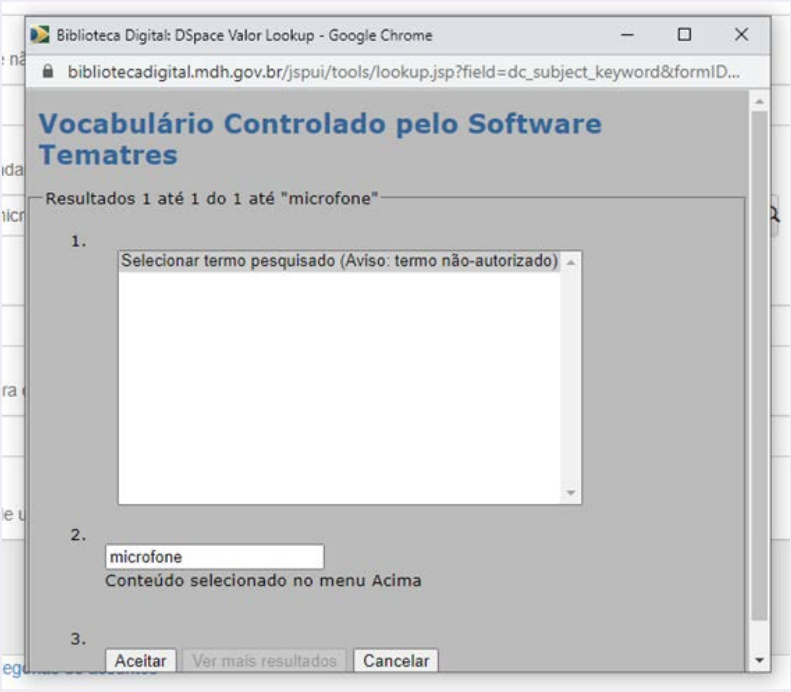


Fonte: Elaboração dos autores (2020).

Quando uma busca não resultar em nenhum registro, a tela aparecerá de acordo com apresentado na Figura 8. Quando isso acontecer, devem ser feitas outras buscas, a fim de que se possam encontrar outras palavras-chave que representem

aquele documento, ou então utilizar um termo não autorizado, clicando na opção “Selecionar o termo pesquisado”.

Figura 11 - Nenhum resultado ao buscar um termo que não exista



Fonte: Elaboração dos autores (2020).

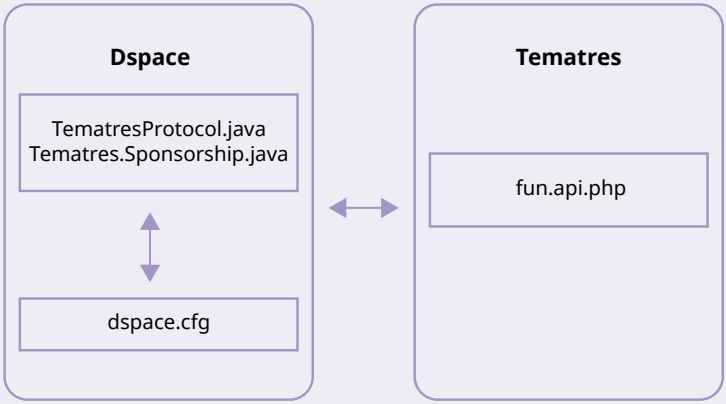
A integração entre o TemaTres e o DSpace visa manter o controle na inserção de palavras-chaves utilizadas na catalogação de novos itens. Assim, mantém o TemaTres como um gestor de termos autorizados, tendo-se um sistema apropriado para a gestão de tesauros, o qual pode ser utilizado como banco de autoridade pela Biblioteca Digital, o que constitui uma forma prática de integração do TemaTres com o DSpace.

4.2 ALTERAÇÃO NO TEMATRES E DSPACE PARA A INTEGRAÇÃO

A integração entre ferramentas requer a criação de meios que permitam uma comunicação contínua ou por demanda, executadas principalmente por meio de protocolos de comunicação. No caso, é preciso adicionar um protocolo de comunicação no DSpace e ajustar outro no TemaTres, com vistas a atender à necessidade de comunicação entre os softwares.

De forma simplificada, a Figura 9 apresenta a integração entre as duas ferramentas. Da parte do DSpace, é preciso criar todos os protocolos de comunicação a serem utilizados para a integração, enquanto o TemaTres apenas ajusta o existente. Com isso, torna-se possível ao DSpace a requisição de informações do TemaTres e obter respostas.

Figura 12 - Estrutura do conjunto de classes envolvidas



Fonte: Elaboração dos autores (2020).

A integração foi efetuada com o DSpace na versão 6 e o TemaTres na versão 3. Para tal procedimento, será necessário intervenção de um informático, e os testes com java 8 – tanto em ambiente de homologação quanto de produção – devem optar por trabalhar com java oracle. Note que a integração é de mão única e o DSpace requisita informações, mas não possibilita que o TemaTres possa requisitar informações do DSpace. Por isso, a maior alteração do DSpace foi a criação dos protocolos de comunicação.

4.2.1 ALTERAÇÃO NO DSPACE PARA A INTEGRAÇÃO

A maior parte das alterações são efetuadas no DSpace, com a criação de protocolos e sua ativação por meio do arquivo de configuração principal **dspace.cfg**. Nesse projeto, utilizou-se a versão 6, mais precisamente a última versão estável (6.3), que podem ser baixadas no site o Git[2] do DSpace, optando pelas versões **src-release**. Assim, ao se descompactar por convenção, deverá ser denotado o DSpace-6.3.

Inicialmente, é preciso criar o protocolo responsável por controlar as requisições entre DSpace e TemaTres (**TemaTresProtocol.java**) no diretório que agrega os outros programas voltados ao controle de autoridade, a fim de manter o padrão do software. Os comandos apresentados seguem o padrão do sistema operacional **debian/ubuntu**.

2 Disponível em: https://github.com/DSpace/DSpace/releases/tag/dspace-6.3.

Quadro 1 – Comando para criação de protocolos e ativação do arquivo de configuração principal

```
$ cd DSpace-6.3/dspace-api/src/main/java/org/dspace/content/authority

$ echo > TematresProtocol.java
```

Fonte: Elaboração dos autores (2020).

Depois de criado o arquivo **TemaTresProtocol.java**, utilize um editor de preferência e adicione o código do protocolo, como apresentado a seguir. O código, em linguagem java, implementa a comunicação entre os sistemas.

Quadro 2 – Comando para editar de preferência e adicionar o código do protocolo

```
package org.dspace.content.authority;


import java.io.IOException;
import java.util.List;
import javax.xml.parsers.SAXParser;
import javax.xml.parsers.SAXParserFactory;
import javax.xml.parsers.ParserConfigurationException;
import org.xml.sax.XMLReader;
import org.xml.sax.InputSource;
import org.xml.sax.helpers.DefaultHandler;
import org.xml.sax.Attributes;
import org.xml.sax.SAXException;
import org.xml.sax.SAXParseException;
import org.apache.log4j.Logger;
import org.dspace.content.Collection;
import org.dspace.core.ConfigurationManager;
import org.apache.http.HttpResponse;
import org.apache.http.client.HttpClient;
import org.apache.http.client.methods.HttpGet;
import org.apache.http.client.utils.URLEncodedUtils;
import org.apache.http.impl.client.DefaultHttpClient;
import org.apache.http.message.BasicNameValuePair;


public abstract class TemaTresProtocol implements ChoiceAuthority
```

```
    private static Logger log = Logger.getLogger(TemaTresProtocol.class);


    private static String url = null;


    public TemaTresProtocol()
    {
        if (url == null)
        {
            url = ConfigurationManager.getProperty(“TemaTres.url”);
            if (url == null)
            {
                throw new IllegalStateException(“Missing DSpace configuration keys for
TemaTres Query”);
            }
        }
    }


    public abstract Choices getMatches(String text, Collection collection, int start, int limit,
String locale);


    @Override
    public Choices getBestMatch(String field, String text, Collection collection, String locale)
    {
        return getMatches(field, text, collection, 0, 2, locale);
    }
```

```
@Override
public String getLabel(String field, String key, String locale)
{
   return key;
}

protected Choices query(String result, String label, String authority,
                        List<BasicNameValuePair> args, int start, int limit)
```

Fonte: Elaboração dos autores (2020).

Após a criação do arquivo **TemaTresProtocol.java** será preciso criar o arquivo **TemaTresSponsorship.java**, que será responsável pelo processamento das requisições no DSpace. Para tanto, execute os seguintes comandos:

Quadro 3 - Criação do arquivo TemaTresSponsorship.java, o qual será responsável pelo processamento das requisições no DSpace

```
cd DSpace-6.3/dspace-api/src/main/java/org/dspace/content/authority
$ echo > TemaTresSponsorship.java
```

Fonte: Elaboração dos autores (2020).

Utilizando um editor de preferência, adicione o seguinte código java:

Quadro 4 - Código java para criar o arquivo

```
package org.dspace.content.authority;

import java.util.ArrayList;

import java.util.List;

import org.dspace.content.Collection;

import org.apache.http.message.BasicNameValuePair;

public class TemaTresSponsorship extends TemaTresProtocol{

   private static final String RESULT = “term”;

   private static final String LABEL = “string”;

   private static final String AUTHORITY = “term_id”;


   public TemaTresSponsorship()

   {

      super();

   }


   @Override

   public Choices getMatches(String text, Collection collection, int start, int limit, String
locale)

   {

      if (text == null || text.trim().length() == 0)

      {

         return new Choices(true);

      }

      List<BasicNameValuePair> args = new ArrayList<BasicNameValuePair>();

      args.add(new BasicNameValuePair(“arg”, text));

      args.add(new BasicNameValuePair(“task”,”search”));
```

```
        Choices result = query(RESULT, LABEL, AUTHORITY, args, start, limit);
        if (result == null)
        {
            result =  new Choices(true);
        }
        return result;
    }

    @Override
    public Choices getMatches(String field, String text, Collection collection, int start, int
limit, String locale) {
        return getMatches(text, collection, start, limit, locale);
    }
}
```

Fonte: Elaboração dos autores (2020).

Após a criação dos arquivos java que processam a comunicação com o TemaTres, é necessário configurar o **dspace.cfg** para utilizar o serviço.

Quadro 5 - Configuração no dspace.cfg

```
$ cd DSpace-6.3/dspace/target/dspace-installer/config}
```

Fonte: Elaboração dos autores (2020).

Então, no arquivo **dspace.cfg**, deve-se descomentar a linha 1429, que ativa o uso de controle de autoridade.

Quadro 6 - Descomentar a linha 1429 que ativa o uso de controle de autoridade

```
plugin.named.org.dspace.content.authority.ChoiceAuthority = \
```

Fonte: Elaboração dos autores (2020).

Na sequência, deve-se adicionar o seguinte trecho após a linha 1429, para adicionar a opção do uso do TemaTres.

Quadro 7 -Trecho para adicionar a opção do uso do TemaTres

```
org.dspace.content.authority.TemaTresSponsorship = TemaTresSponsorship
```

Fonte: Elaboração dos autores (2020).

Por fim, tem-se a seção de controle de autoridade, com alguns PlugIns comentados, como o do *Library of Congress* (LCNameAuthority), mas com o TemaTres ativo.

Quadro 8 - TemaTres ativo

```
##### Authority Control Settings #####
plugin.named.org.dspace.content.authority.ChoiceAuthority = \
 org.dspace.content.authority.SampleAuthority = Sample, \
 org.dspace.content.authority.LCNameAuthority = LCNameAuthority, \
 org.dspace.content.authority.SHERPARoMEOPublisher = SRPublisher, \
 org.dspace.content.authority.SHERPARoMEOJournalTitle = SRJournalTitle, \
 org.dspace.content.authority.TemaTresSponsorship = TemaTresSponsorship
```

Fonte: Elaboração dos autores (2020).

Após finalizar as customizações efetuadas no TemaTres, é necessário configurar o **dspace.cfg** com o propósito de o DSpace consultar, de forma assíncrona, a base de dados do TemaTres. Por questão de padrão, após as linhas que fazem referência ao plugin TemaTres, adicione as seguintes linhas:

Quadro 9 - Comando para consulta assíncrona

```
TemaTres.url = https://IPTemaTres/vocab/services.php
choices.plugin.dc.subject.keyword = TemaTresSponsorship
choices.presentation.dc.subject.keyword = lookup
authority.controlled.dc.subject.keyword = true
```

Fonte: Elaboração dos autores (2020).

Observação: “**dc.subject.keyword**” está informando o metadado no input-forms.xml, que mapeia as palavras-chaves.

O DSpace já possui alguns PlugIns para implementação de controle de autoridade, com o uso de vocabulários controlados, conforme visto no arquivo de configuração. Com isso, adiciona-se o TemaTres, seguindo o padrão utilizado na ferramenta, facilitando ajustes em caso de atualização de versão do DSpace.

4.2.2 ALTERAÇÃO NO TEMATRES PARA A INTEGRAÇÃO

As alterações no TemaTres são menores, visto que ele já possui protocolos de comunicação que serão utilizados nesta integração (services.php e fun.api.php). Entretanto, pela necessidade apresentada no projeto, se requer a quantidade de registros em cada requisição efetuada pelo DSpace. Assim, considera-se que o TemaTres esteja instalado em **/var/www/html/TemaTres**. Ao acessar o arquivo **common/include/fun.api.php**, na linha 1055, deve-se adicionar o trecho de código abaixo:

Quadro 10 - Comando para alterar o TemaTres de forma a possibilitar a integração

```
$arrayResume[“cant_result”] = count($response[“result”]);
```

Fonte: Elaboração dos autores (2020).

Em seguida, deve-se salvar e verificar se os termos estão sendo renderizados no navegador. Nesses casos, a resposta à requisição deve ser a seguinte: **https://IPTemaTres/vocab/services.php?task=letter&arg=a**. Note que o argumento de busca é a letra "a", indicada pela "task=letter" e "arg=a". Como resposta, o sistema responde quantos termos da base iniciam com a letra "a" nas primeiras linhas, segundo é mostrado abaixo:

Quadro 11 - Resposta do sistema ao teste de renderização

```
<vocabularyservices>
<cant_result>X</cant_result>
```

Fonte: Elaboração dos autores (2020).

Observações: *IPTemaTres faz referência ao url em que se está acessando o TemaTres, e "X" no xml, apresentado no navegador, é apenas um exemplo que será devolvido, ou seja, um número referenciando a quantidade de termos iniciados com a vogal "a" no tesauro.*

5

CONSIDERAÇÕES FINAIS

5 CONSIDERAÇÕES FINAIS

A integração entres sistemas especializados tem sido uma constante na formação de ecossistemas de informação, para os quais o intercâmbio de informação é a tônica. Assim, cada sistema oferta serviços de informações a outras, como se fossem módulos em uma grande plataforma informacional.

O DSpace é um software que já vem preparado para se integrar com outras ferramentas por meio da oferta de estrutura voltada à integração, que responde a protocolos de comunicação. Assim, já disponibiliza algumas possibilidades de integração e torna fácil o desenvolvimento de outros. A estrutura já disponibiliza parte básica do processamento, bastando o desenvolvimento da parte específica.

Neste trabalho, integrou-se o DSpace com o software TemaTres, cujo serviço disponibilizado baseia-se na criação e gestão de vocabulários controlados. Como estudo de caso, o integrador proporciona a normalização do conteúdo do campo de palavras-chave ou assunto no formulário de entrada, ligado ao metadado conhecido como **subject.keyword** no padrão *Dublin Core.* Dessa forma, todo preenchimento relacionado a palavras-chave de um novo documento a ser inserido no repositório estará sujeito a normalização imposta pelos termos inseridos no TemaTres. A redução para um mini mundo de palavras-chave controlado possibilita ao repositório, por meio do TemaTres, o direcionamento de indexador de conteúdo ao campo de pesquisa da instituição, o qual será de grande ajuda quando se pensa posteriormente na recuperação da informação.

Vale ressaltar que o integrador pode ser utilizado para qualquer metadado utilizado no formulário de entrada, sendo unicamente necessário configurá-lo no **dspace.cfg**. Assim, o TemaTres pode ser gerenciador de outras formas de representação da informação, além de assunto, podendo gerenciar nome de instituição, órgãos de governo e linhas de pesquisa, bem como outras informações que podem ser estruturadas como taxonomias.

Por fim, a integração entre o DSpace e o TemaTres traz significativas vantagens à representação e recuperação da informação inseridas no DSpace, visto que normaliza o conteúdo dos campos de descrição dos itens, abrindo inúmeras possibilidades de criar bancos de autoridades com o TemaTres a serem integrados ao DSpace.

