

RELATÓRIO TÉCNICO

Relatório de integração entre a Biblioteca Digital e
o Thesaurus da Biblioteca do Ministério da Mulher,
Família e Direitos Humanos

GOVERNO FEDERAL

MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E INOVAÇÕES

Marcos Cesar Pontes

Ministro da Ciência, Tecnologia e Inovações

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

RELATÓRIO TÉCNICO

Relatório de integração entre a Biblioteca Digital e o Thesaurus da Biblioteca do Ministério da Mulher, Família e Direitos Humanos



Coordenação de Articulação, Geração
e Aplicação de Tecnologia (CoTec)

Brasília
2020

EQUIPE TÉCNICA

Diretora do Instituto Brasileiro de Informação em Ciência e Tecnologia

Cecília Leite Oliveira

Coordenador-Geral de Tecnologias de Informação e Informática – CGTI

Tiago Emmanuel Nunes Braga

Coordenador do Projeto

Milton Shintaku

Autores

Milton Shintaku

Frederico Oliveira

Lucas Angelo da Silveira

Ítalo Barbosa Brasileiro

Guilherme Enéas Vaz Silva

Raíssa da Veiga de Meneses

Este Relatório de Técnico 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 e Inovações.

É permitida a reprodução deste texto e dos dados nele contidos, desde que citada a fonte. Reproduções para fins comerciais são proibidas.

SUMÁRIO

1. INTRODUÇÃO	6
2. OBJETIVOS	6
2.1 Objetivo Geral	6
2.2 Objetivos Específicos	6
3. RESULTADOS	6
3.1 Catalogação no DSpace com uso de tesouros do Tematres	7
3.2 Alteração no Tematres e DSpace para a Integração	11
3.3 Alteração no DSpace para a Integração	11
3.4 Alteração no Tematres para a integração	17
4. CONSIDERAÇÕES FINAIS	18

1. INTRODUÇÃO

A internet fomentou o uso de documentação no formato digital, com um grande aumento na chamada web 2.0, em que usuários passaram a ser produtores de informação. Nesse contexto, várias iniciativas possibilitaram que se disponibilizassem documentação digital em formato de texto completo, independentemente do tamanho, na medida em que certas limitações estão apenas na forma de conexão ou velocidade de transmissão.

O uso cada vez maior da internet e suas funcionalidades afetaram todas as atividades humanas, inclusive as de governo, tanto que no Brasil surgiu o chamado e-GOV, ofertando serviços e informações. Outro ponto a ser destacado foi a publicação da Lei de Acesso à Informação (LAI), Lei nº 12.527 de 18 de novembro de 2011, que promove a transparência nas atividades administrativas do governo. Assim, as Bibliotecas Digitais de Governo (BDG) podem ser consideradas como ferramentas alinhadas ao e-GOV na oferta de acesso à documentação digital, atendendo à LAI.

A implementação de uma BDG impõe desafios que transcendem a tecnologia, requerendo estudos voltados à organização e representação da informação, a fim de facilitar a recuperação dos documentos pertencentes ao acervo. Assim, definir a estrutura da BDGI organiza o acervo pelas áreas de conhecimento mantidas pela biblioteca, da mesma forma que as estantes e as prateleiras a fazem em bibliotecas físicas. A representação fica a cargo dos metadados escolhidos para descrever os documentos, podendo ser descritivos, como título e nome do autor, técnicos, como formato e tipo do arquivo, ou mesmo voltados aos direitos autorais e licenciamento.

O maior desafio está na qualidade das informações inseridas nos metadados, visto que a riqueza das línguas permite que se descreva um documento de várias maneiras. Assim, são criados vocabulários controlados para alguns campos de metadados, principalmente para aqueles que são sensíveis à recuperação ou possuem poucas opções de preenchimento. Assim, idioma, tipo e formato de arquivo e as datas podem ser padronizados a fim de facilitar a indexação.

Entretanto, o maior desafio está no campo “assunto”, que pode apresentar uma infinidade de possibilidades. Para tanto, bibliotecas geralmente desenvolvem os chamados bancos de autoridades, com os termos autorizados para descrever os assuntos ou, mesmo, nome de autores. Assim, podem-se padronizar os termos inseridos nos assunto por meio de um compêndio de termos autorizados que, geralmente, fazem parte do sistema de gestão de biblioteca.

No caso de ferramentas para bibliotecas digitais, como o DSpace, esses bancos de autoridade não existem. Por isso, no projeto da Biblioteca do MMFDH optou-se por utilizar um sistema de tesouros (Tematres) para montar o banco de autoridades, integrando-o ao sistema de biblioteca digital (DSpace). Com consequência, pode-se utilizar um Tesauros (conjunto selecionado de termos organizados e descritos) como base de autoridade para uma biblioteca digital.

2. OBJETIVOS

2.1 Objetivo Geral

Apresentar a integração entre os sistemas de Biblioteca Digital (DSpace) e o de Tesouros (Tematres).

2.2 Objetivos Específicos

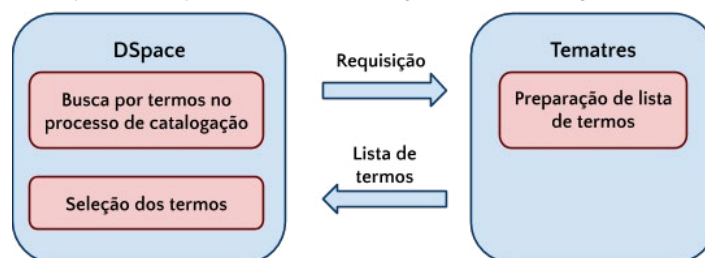
- Apresentar o modelo conceitual da integração entre os dois sistemas;
- Apresentar como a integração se realiza;
- Apresentar as alterações necessárias em cada um dos sistemas.

3. RESULTADOS

A integração entre o Sistema de Tesouros (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 tesouros de forma automática, diretamente na biblioteca digital, conforme esquema apresentado na Figura 1.

Figura 1 – Esquema do processo de catalogação com integração ao Tematres



Fonte: Elaboração dos autores (2020).

Nesse sentido, os dois sistemas se complementam, na medida em que o Tematres oferta termos autorizados para o uso no DSpace, normalizando as informações inseridas em campos como palavras-chave ou assunto. Logo, o uso da integração é vista apenas no DSpace, ao longo do 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 use 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.

3.1 Catalogação no DSpace com uso de tesouros do Tematres

Com o integrador implantado, todos os termos inseridos no TemaTres estarão disponíveis para serem utilizados no DSpace no processo de depósito, no momento da catalogação de um item na Biblioteca Digital, 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 utilização do integrador se 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 2. 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 2 – Campo palavra-chave

Abstract

Inserir as informações relevantes que não serão adicionadas em outros campos.

Notas

Inserir palavras-chaves que correspondam ao assunto/tema da obra.

Palavra-Chave *

Publicação Relacionada

Inserir o nome da pessoa/gerenciadora do Direitos autorais.

Detentor dos Direitos Autorais *



Fonte: Elaboração dos autores (2020).

Como mostra a Figura 3, a presença do integrador pode ser notada pela existência da lupa ao lado do campo de palavras-chave existente 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 3 - Busca de termo

Insira palavras-chaves que correspondam ao assunto/tema da obra.

Palavra-Chave *

instituição  Digitar termo da busca  + Adicionar mais

Fonte: Elaboração dos autores (2020).

A resposta surge de uma janela PopUp, como mostra a Figura 4. A palavra inserida no campo será buscada no Tematres. Como mostra o exemplo, a busca deu-se pela palavra "instituição", assim, na base de termos do Tematres havia apenas o termo "Instituição de longa permanência". Como consequência, o usuário pode utilizar o termo sugerido ou manter o termo pesquisado, mesmo que não seja autorizado.

Figura 4 - Resultados da busca

Biblioteca Digital: DSpace Valor Lookup - Google Chrome

bibliotecadigital.mdh.gov.br/jspui/tools/lookup.jsp?field=dc_subject_keyword&formID...

Vocabulário Controlado pelo Software Tematres

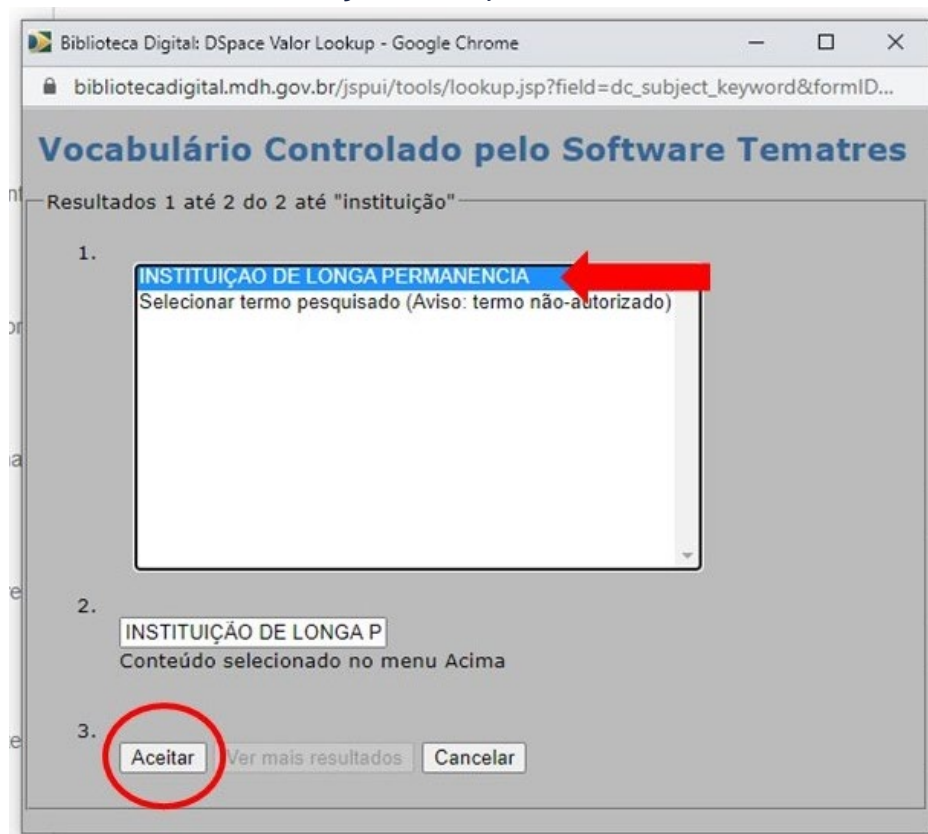
Resultados 1 até 2 do 2 até "instituição"

1. INSTITUIÇÃO DE LONGA PERMANÊNCIA
Selecionar termo pesquisado (Aviso: termo não-autorizado)
2.
Conteúdo selecionado no menu Acima
3.

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 5. Assim, para utilizar o termo selecionado (autorizado ou não) como palavra-chave na catalogação, basta clicar sobre o termo desejado e então no botão **Aceitar**.

Figura 5 - Seleção do termo



Fonte: Elaboração dos autores (2020).

Ao clicar em Aceitar, a janela do Tesouro fechará automaticamente e o termo selecionado aparecerá no campo do formulário de entrada (figura 6). Com isso, o termo autorizado preenche 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 nova busca.

Figura 6 - 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 7, clicando em *Excluir esta entrada*. Assim, pode-se reiniciar o processo e buscar novos termos. Na grande maioria dos casos, o campo das palavras-chave ou assunto é obrigatório, por customização do formulário de entrada, não podendo ser deixado em branco.

Figura 7 - Excluir entrada

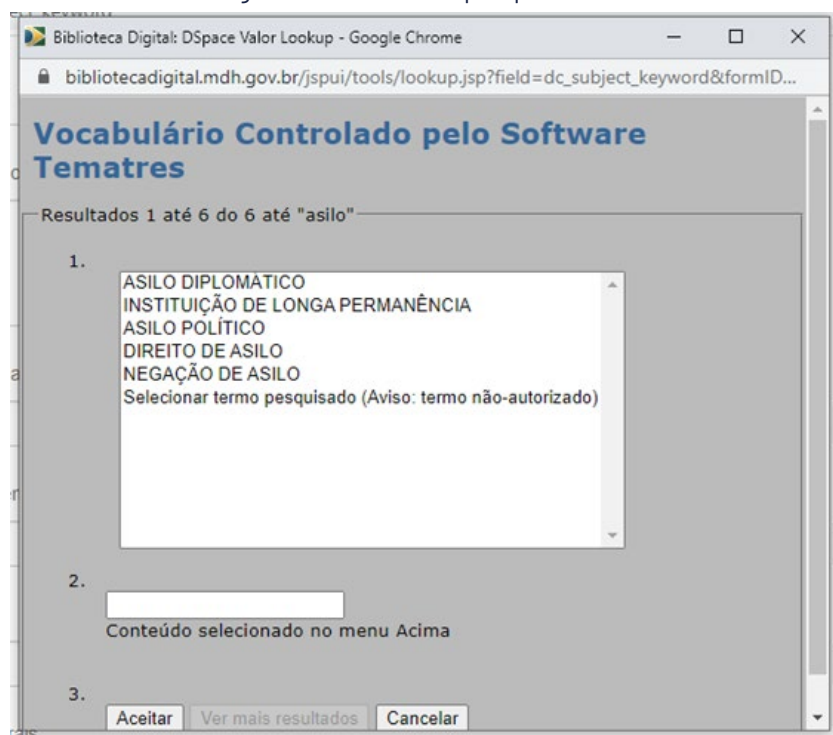


Fonte: Elaboração dos autores (2020).

O DSpace permite que sejam inseridos **Termos não autorizados** no processo de 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 8, a palavra pesquisada é *ASILO*.

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

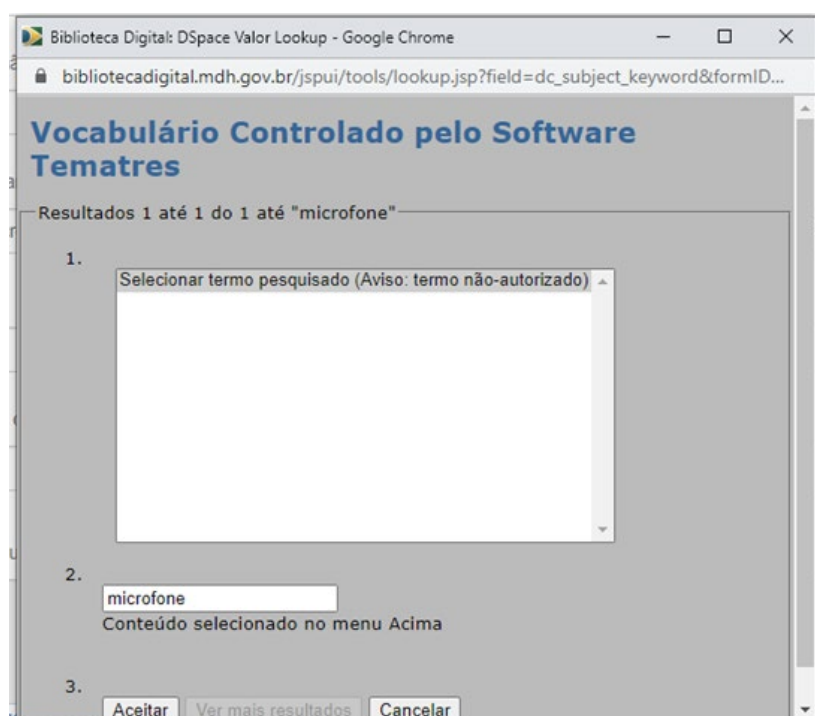
Figura 8 – Resultados pesquisa ASILO



Fonte: Elaboração dos autores (2020).

Quando uma busca não obtiver registro, a tela aparecerá de acordo com apresentado na Figura 9. Quando isso acontecer, devem ser feitas outras buscas no sentido de encontrar outras palavras-chave que representem aquele documento ou, então, utilizar um termo não autorizado, clicando na opção "Selecionar o termo pesquisado".

Figura 9 – Nenhum resultado



Fonte: Elaboração dos autores (2020).

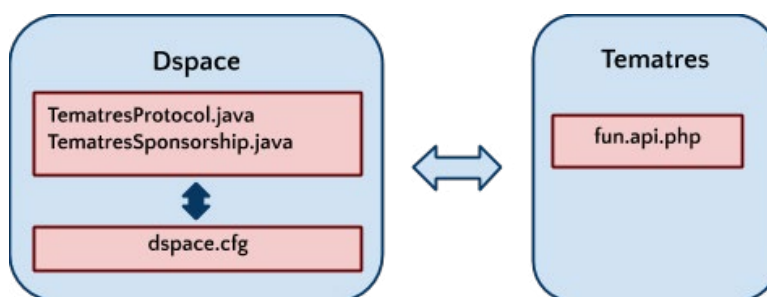
A integração entre o Tematres e o DSpace visa manter o controle na inserção de palavras-chaves utilizadas na catalogação de novos itens. Assim, mantém o Tematres como um gestor de termos autorizados. Nesse sentido, tem-se um sistema apropriado para a gestão de tesouros, que pode ser utilizado como banco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

3.2 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. Nesse 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 10 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 no Tematres apenas ajusta o existente. Com isso, o DSpace pode requisitar informações do Tematres e obter respostas.

Figura 10 – 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 que os testes com java 8, tanto em ambiente de homologação quanto de produção, optem por trabalhar com java oracle. Note que a integração é de mão única, de modo que o DSpace requisita informações, mas não possibilita que o Tematres possa requisitar informações do DSpace. Por isso, a maior alteração efetuada no DSpace foi a criação dos protocolos de comunicação.

3.3 Alteração no DSpace para a Integração

A maior parte das alterações são efetuadas no DSpace, com a criação de protocolos e sua ativação através do arquivo de configuração principal dspace.cfg. Nesse projeto utilizou-se a versão 6, mais precisamente a última versão estável (6.3), que pode ser baixada no site o Git¹ do DSpace, optando pelas versões *src-release*. Assim, ao descompactar por convenção, será 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**.

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

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.

¹ Disponível em: <https://github.com/DSpace/DSpace/releases/tag/dspace-6.3>

```

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)
{
    HttpClient hc = new DefaultHttpClient();
    String srUrl = url + "?" + URLEncodedUtils.format(args, "UTF8");
    HttpGet get = new HttpGet(srUrl);
    log.debug("Trying Tematres Query, URL="+srUrl);
    try
    {
        HttpResponse response = hc.execute(get);
        if (response.getStatusLine().getStatusCode() == 200)
        {
            SAXParserFactory spf = SAXParserFactory.newInstance();
            SAXParser sp = spf.newSAXParser();
            XMLReader xr = sp.getXMLReader();
            TematresHandler handler = new TematresHandler(result, label, authority);
            xr.setFeature("http://xml.org/sax/features/namespaces", true);
            xr.setContentHandler(handler);
            xr.setErrorHandler(handler);
            xr.parse(new InputSource(response.getEntity().getContent()));
            int confidence;
            if (handler.total == 0)
            {
                confidence = Choices.CF_NOTFOUND;
            }
            else if (handler.total == 1)
            {
                confidence = Choices.CF_UNCERTAIN;
            }
            else
            {
                confidence = Choices.CF_AMBIGUOUS;
            }
            return new Choices(handler.result, start, handler.total, confidence, false);
        }
    }
    catch (IOException e)
    {
        log.error("Tematres query failed: ", e);
        return null;
    }
    catch (ParserConfigurationException e)
    {
        log.warn("Failed parsing Tematres result: ", e);
        return null;
    }
    catch (SAXException e)
    {
        log.warn("Failed parsing Tematres result: ", e);
        return null;
    }
}

```

```

    finally
    {
        get.releaseConnection();
    }
    return null;
}

private static class TematresHandler
    extends DefaultHandler
{
    private Choice result[] = null;
    int rindex = 0;
    int total = 0;
    private String resultElement = null;
    private String labelElement = null;
    private String authorityElement = null;
    protected String textValue = null;
    public TematresHandler(String result, String label, String authority)
    {
        super();
        resultElement = result;
        labelElement = label;
        authorityElement = authority;
    }

    @Override
    public void characters(char[] ch, int start, int length)
        throws SAXException
    {
        String newValue = new String(ch, start, length);
        if (newValue.length() > 0)
        {
            if (textValue == null)
            {
                textValue = newValue;
            }
            else
            {
                textValue += newValue;
            }
        }
    }

    @Override
    public void endElement(String namespaceURI, String localName, String qName)
        throws SAXException
    {
        if (localName.equals("cant_result"))
        {
            String stotal = textValue.trim();
            if (stotal.length() > 0)
            {
                total = Integer.parseInt(stotal);
                result = new Choice[total];
                if (total > 0)

```

```

        {
            result[0] = new Choice();
            log.debug("Got "+total+" records in results.");
        }
    }
}
else if (localName.equals(resultElement))
{
    // after start of result element, get next hit ready
    if (++rindex < result.length)
    {
        result[rindex] = new Choice();
    }
}
else if (localName.equals(labelElement) && textValue != null)
{
    // plug in label value
    result[rindex].value = textValue.trim();
    result[rindex].label = result[rindex].value;
}
else if (authorityElement != null && localName.equals(authorityElement) && textValue != null)
{
    result[rindex].authority = textValue.trim();
}
else if (localName.equals("message") && textValue != null)
{
    log.warn("Tematres response error message: " + textValue.trim());
}
}

@Override
public void startElement(String namespaceURI, String localName,
                        String qName, Attributes atts)
    throws SAXException
{
    textValue = null;
}

@Override
public void error(SAXParseException exception)
    throws SAXException
{
    throw new SAXException(exception);
}

@Override
public void fatalError(SAXParseException exception)
    throws SAXException
{
    throw new SAXException(exception);
}
}
}

```

Após a criação do `TematresProtocol.java`, será preciso criar o arquivo **`TematresSponsorship.java`**, o qual será responsável pelo processamento das requisições no DSpace. Para tanto, execute os seguintes comandos:

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

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

```
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);
    }
}
```


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

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

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

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

Na sequência, adicionar o seguinte trecho após a linha 1429, para adicionar a opção do uso do Tematres.

```
org.dspace.content.authority.TematresSponsorship = TematresSponsorship
```

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

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

```
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
```

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, como visto no arquivo de configuração. Isto feito, deve-se adicionar o do Tematres, seguindo o padrão utilizado na ferramenta, facilitando ajustes em caso de atualização de versão do DSpace.

3.4 Alteração no Tematres para a integração

As alterações no Tematres são menores, visto que ele já possui protocolos de comunicação que serão utilizados na integração (services.php e fun.api.php). Entretanto, pela necessidade apresentada no projeto, é necessária a quantidade de registros em cada requisição efetuada pelo DSpace. Assim, tomando como base que o Tematres esteja instalado em [/var/www/html/TemaTres](#), acessar o arquivo [common/include/fun.api.php](#) e na linha 1055 adicionar o trecho de código abaixo:

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

Em seguida, deve-se salvar e verificar se a quantidade de termos está sendo renderizada no navegador e a resposta à seguinte requisição:

<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 expõe quantos termos da base iniciam com a letra “a” nas primeiras linhas da resposta, como mostrado abaixo:

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

Observações: IPTemaTres faz referência ao url no qual está sendo acessado o TemaTres. O “X” no xml, apresentado no navegador, é apenas um exemplo de como será devolvido um número referenciando a quantidade de termos iniciados com a vogal “a” no tesouro.

4. CONSIDERAÇÕES FINAIS

A integração entre sistemas especializados tem sido uma constante na formação de ecossistemas de informação, nos quais o intercâmbio de informação é a tônica. Desse modo, 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 a 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. Tal 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-chaves de um novo documento a ser inserido no repositório estará sujeito à normalização imposta pelos termos inseridos no Tematres. A redução para um mini mundo de palavras-chaves controlado possibilita ao repositório, por meio do Tematres, o direcionamento de indexador de conteúdo ao campo de pesquisa da instituição, que será de grande ajuda pensando na recuperação da informação posteriormente.

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 ou órgãos de governo, 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 conteúdos dos campos de descrição dos itens, abrindo inúmeras possibilidades de criar bancos de autoridades com o Tematres a serem integrados ao DSpace.

SAS - Quadra 05 - Lote 06 -
Bloco H - Sobreloja
Cep: 70070-912 - Brasília / DF

Telefone: +55 61 3217 6213
E-mail: shintaku@ibict.br



MINISTÉRIO DA
CIÊNCIA, TECNOLOGIA
E INOVAÇÕES



PÁTRIA AMADA
BRASIL
GOVERNO FEDERAL