

UNIVERSIDADE ESTADUAL PAULISTA “JÚLIO DE MESQUITA FILHO”
Campus de Marília – SP
Faculdade de Filosofia e Ciências
Programa de Pós-Graduação em Ciência da Informação

LISANDRO ROGÉRIO MODESTO

**Representação e Persistência para acesso
a Recursos Informacionais Digitais
gerados dinamicamente em sítios oficiais
do Governo Federal**

Marília
2013

UNIVERSIDADE ESTADUAL PAULISTA “JÚLIO DE MESQUITA FILHO”
Campus de Marília, São Paulo
Faculdade de Filosofia e Ciências
Programa de Pós-Graduação em Ciência da Informação

LISANDRO ROGÉRIO MODESTO

Representação e Persistência para acesso a Recursos Informacionais Digitais gerados dinamicamente em sítios oficiais do Governo Federal

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Ciência da Informação da Faculdade de Filosofia e Ciências - Universidade Estadual Paulista “Júlio de Mesquita Filho” – UNESP – campus de Marília, São Paulo, como requisito parcial para a obtenção do título de Doutor em Ciência da Informação.

Linha de Pesquisa: Informação e Tecnologia.

Orientador: Prof. Dr. Edberto Farneda.

Coorientador: Prof. Dr. Ricardo César Gonçalves Santana.

Marília

2013

FICHA CATALOGRÁFICA

Modesto, Lisandro Rogério

Representação e Persistência para acesso a Recursos
Informacionais Digitais gerados dinamicamente em sítios oficiais do
Governo Federal / Lisandro Rogério Modesto, 2013.
103f.; 30 cm.

Tese (Doutorado em Ciência da Informação) – Faculdade de Filosofia
e Ciências, UNIVERSIDADE ESTADUAL PAULISTA, 2013.

Orientador: Prof. Dr. Edberto Ferneda.

Coorientador: Prof. Dr. Ricardo César Gonçalves Santana.

1. Dados abertos. 2. Metadados. 3. Recursos Informacionais
Digitais. I. Lisandro Modesto. II. Edberto Ferneda. III. Ricardo César
Gonçalves Sant'Ana. IV. Título.

LISANDRO ROGÉRIO MODESTO

**Representação e Persistência para acesso a Recursos
Informacionais Digitais gerados dinamicamente em sítios
oficiais do Governo Federal**

Banca Examinadora:

Prof. Dr. Edberto Ferneda
Orientador
Universidade Estadual Paulista – UNESP

Prof^a. Dra. Zaira Regina Zafalon
Universidade Federal de São Carlos - UFSCAR

Prof. Dr. Elvis Fusco
Fundação de Ensino Eurípides Soares da Rocha - UNIVEM

Prof^a. Dra. Maria José Vicentini Jorente
Universidade Estadual Paulista – UNESP

Prof^a. Dra. Rachel Cristina Vesu Alves
Universidade Estadual Paulista – UNESP

Marília, 2013.

Dedico esse trabalho a minha esposa Juliana e a meu filho Lorenzo que serviram de inspiração nos momentos de dúvida e insegurança e sem os quais essa caminhada jamais teria sido concluída.

Agradecimentos

Agradeço primeiramente a Deus que, com sua sabedoria e bondade divina, iluminou-me em todos os meus aprendizados.

À minha esposa Juliana, pelo carinho, amor, compreensão e dedicação que sempre me concedeu no decorrer dos anos de doutoramento, e que me presenteou com a maior dádiva da vida, a paternidade.

Ao meu filho Lorenzo, que com sua chegada me fez enxergar o mundo de forma muito mais simples e me mostrou a felicidade plena. A cada sorriso seu recarrega minhas energias para enfrentar as dificuldades da vida.

À minha família, em especial aos meus pais, Vicente e Aparecida, que sempre me ensinaram a honestidade e a persistência na busca de meus ideais e me apoiaram e encorajaram, principalmente nas horas em que mais precisei.

Ao meu orientador e amigo, Prof. Dr. Edberto Ferneda, que com sua experiência e serenidade, sempre me orientou em qual caminho deveria prosseguir, fazendo isso com segurança e transmitindo a mim a confiança necessária durante toda a elaboração desta tese. Sua sabedoria me fez enxergar o caminho certo e me fez atingir este objetivo.

Ao meu coorientador e amigo, Prof. Dr. Ricardo César Gonçalves Sant'Ana, que muito contribuiu para a formação do contexto deste trabalho e pelas palavras de força e direcionamento.

À FAP – Faculdade de Apucarana, que me apoiou financeira e profissionalmente para que eu pudesse vencer mais essa etapa de minha vida profissional.

À amiga Rosimeiri Darc Cardoso, diretora geral da FAP, pelo apoio nos momentos difíceis e no encorajamento para ingressar neste programa de doutorado.

Aos amigos Flávio Navarro e Déverson Rando que contribuíram na concepção de idéias e desenvolvimento de atividades relacionadas à Tese.

Enfim, agradeço a todas as pessoas que de uma forma ou de outra, contribuíram para que eu conseguisse concluir mais esta etapa de minha vida.

MODESTO, Lisandro Rogério. **Representação e Persistência para acesso a Recursos Informativos Digitais gerados dinamicamente em sítios oficiais do Governo Federal**. Tese (Doutorado) – Programa de Pós-Graduação em Ciência da Informação – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista – UNESP, Marília, 2013.

RESUMO

A Lei de Acesso à Informação, sancionada em 18 de novembro de 2011 pela Presidenta da República em Exercício Dilma Rousseff, revogou a lei número 11.111/1995 a partir de 16 de maio de 2012. Esta lei regulamenta o direito constitucional de acesso às informações públicas pelos cidadãos brasileiros, sendo aplicada aos 3 poderes do Governo Federal, Governo Estadual e Governo Municipal. Com esta lei, os sítios oficiais do governo foram obrigados a deixarem disponíveis informações em seus portais. Embora esta lei imponha uma disponibilização de informações de interesse público, ela não define a forma como tais informações devam estar disponibilizadas. Mesmo com a Lei de Acesso à Informação, os sítios oficiais não disponibilizam as informações de forma que o cidadão possa entendê-la. Para resolver esta situação, o ideal seria que as informações fossem disponibilizadas em formato aberto, permitindo análise e utilização. Este trabalho está restrito aos sítios oficiais do Governo Federal podendo extrapolar outros contextos, que deveriam disponibilizar informações atendendo ao que rege a Lei de Acesso à Informação. Da forma como as informações estão disponibilizadas atualmente não é possível acessar novamente resultados de pesquisas com dados gerados dinamicamente a partir de parâmetros informados ou selecionados em sítios oficiais do Governo Federal. A partir da persistência / materialização dos resultados gerados dinamicamente e com metadados agregados é possível garantir o reuso de informações dos sítios oficiais do Governo Federal. O objetivo deste trabalho é propor um modelo de representação de Recursos Informativos Digitais gerados dinamicamente, disponibilizados pelo Governo Federal em seus sítios oficiais, usando metadados para descrever o acesso e persistir tais recursos a fim de possibilitar sua reutilização e garantir novas formas de recuperação. Para permitir a persistência dos Recursos Informativos Digitais gerados dinamicamente foi desenvolvida a ferramenta MDDIR Tool para descrever o acesso aos Recursos Informativos Digitais gerados dinamicamente, usando o HTTPTRACK Website Copier, que permite baixar um sítio para um diretório local, construindo recursivamente todos os diretórios, incluindo imagens e outros arquivos do servidor para o seu computador. Desta forma, o cidadão poderá localizar o ponto exato em que é gerada uma tabela de resultado de uma determinada consulta sem passar pela estrutura (caminho) que hoje é hierarquizado. Pretende-se com isso facilitar o acesso a dados abertos, permitindo, assim, que tais dados sejam passíveis de serem consolidados, organizados e interpretados, gerando novos conhecimentos e informações fáceis de serem interpretadas e entendidas pela população em geral, possibilitando o registro histórico e social dos resultados.

Palavras-chave: Dados abertos. Metadados. Recursos Informativos Digitais.

MODESTO, Lisandro Rogério. **Representation and Persistency for access to Digital Resources Informational dynamically generated on the official sites of the Federal Government.** Thesis (Ph.D.) - Graduate Program in Information Science - Faculty of Philosophy and Sciences, Universidade Estadual Paulista - UNESP, Marília, 2013.

ABSTRACT

The Access to Information Act, enacted on November 18, 2011 by the Acting President of the Republic Dilma Rousseff, repealed the law number 11.111/1995 from May 16, 2012. This law regulates the constitutional right of access to public information by Brazilian citizens, being applied to 3 powers of the Federal Government, State Government and Local Government. With this law, the official government sites were forced to leave information available on their portals. Although this law imposes an availability of information of public interest, it does not define how such information should be made available. Even with the Law on Access to Information, the official websites do not provide the information so that citizens can understand it. To resolve this situation, the ideal would be that the information be made available in an open format, allowing analysis and use. This work is restricted to official sites of the Federal Government can extrapolate other contexts, which should make information available given the governing Law on Access to Information. How the information is provided is not currently possible to access research results with data dynamically generated from input or selected on the official sites of the Federal Government parameters again. From the persistence / materialization of the results and dynamically aggregates generated metadata can ensure the reuse of information from official websites of the Federal Government. The objective of this work is to propose a model of representation of Digital Informational Resources dynamically generated, provided by the Federal Government on their official websites, using metadata to describe and access resources such persists to enable reuse and secure new forms of recovery. To allow the persistence of Digital Informational Resources dynamically generated was developed MDDIR Tool to describe access to Digital Resources Informational dynamically generated using HTTrack Website Copier, a site that lets you download to a local directory, building recursively all directories, including images and other files from the server to your computer. Thus, citizens can locate the exact point where it is generated a table of results for a given query without passing by the structure (path) which is now hierarchical. The aim of this easy access to open data, thus allowing such data are likely to be consolidated, organized and interpreted, generating new knowledge and information easy to be interpreted and understood by the general population, enabling the historical record and social results.

Keywords: Open data. Metadata. Digital Informational Resources.

Lista de Ilustrações

Figura 1 - Modelo de referência para visualização.....	63
Figura 2 - Representação na técnica Bifocal Display.	64
Figura 3 - Perspective Wall.	64
Figura 4 - Cone Tree.	65
Figura 5 - Hyperbolic Tree.....	66
Figura 6 - TreeMap.....	66
Figura 7 - Ícone de Acesso à Informação.....	70
Figura 8 - Página de despesas do Ministério da Educação.	70
Figura 9 - Modelo entidade-relacionamento (MER) do Sistemas Web MDDIR Tool.	83
Figura 10 - Tela inicial do sistema MDDIR Tool	85
Figura 11 - Tela de menu do sistema MDDIR Tool	86
Figura 12 - Tela de cadastro de Página Menu do sistema MDDIR Tool	87
Figura 13 - Tela de cadastro de Página Resultado do sistema MDDIR Tool	88
Figura 14 - Tela de consulta por parâmetro do sistema MDDIR Tool	89
Figura 15 - Tela de consulta por Palavras-chave do sistema MDDIR Tool	90
Figura 16 - Ícone para representar uma Página Menu.....	91
Figura 17 - Ícone para representar uma Página Resultado.....	91
Figura 18 - Tela de resultado de consultas do sistema MDDIR Tool	91
Figura 19 - Tela de apresentação da Página Menu e Páginas Resultado	92
Figura 20 - Tela de apresentação completa do recurso descrito.....	93
Figura 21 - Tela de detalhes do sistema MDDIR Tool.....	94
Figura 22 - Arquivo XML gerado pelo sistema MDDIR Tool.....	95
Figura 23 - Arquivo CSV gerado pelo sistema MDDIR Tool.....	95

Lista de Quadros

Quadro 1 - Descrição de recurso utilizando o padrão de metadados Dublin Core....	49
Quadro 2 - Descrição de recurso utilizando o padrão de metadados e-PMG	54
Quadro 3 - Procedimentos necessários para a recuperação de um determinado conjunto de dados	69
Quadro 4 - Sítios oficiais analisados	71
Quadro 5 - Descrição do acesso ao sítio do Ministério do Trabalho e Emprego no Padrão de Metadados Dublin Core.....	73
Quadro 6 - Descrição do acesso ao sítio do Ministério do Trabalho e Emprego no Padrão de Metadados e-PMG	74
Quadro 7 - Representação de um recurso informacional digital gerado dinamicamente	79

SUMÁRIO

1	INTRODUÇÃO.....	12
1.1	DELIMITAÇÃO DO TEMA	16
1.2	PROBLEMA DE PESQUISA	16
1.3	OBJETIVOS	17
1.3.1	Objetivo Geral	17
1.3.2	Objetivos Específicos	17
1.4	JUSTIFICATIVA.....	17
1.5	HIPÓTESE	18
1.6	TESE	19
1.7	PROCEDIMENTOS METODOLÓGICOS	19
1.8	ORGANIZAÇÃO DO TRABALHO.....	19
2	ACESSO A DADOS DA GESTÃO PÚBLICA	21
2.1	ASPECTOS LEGAIS DO ACESSO A DADOS GOVERNAMENTAIS	22
2.2	LEIS E PRINCÍPIOS DE DADO ABERTO	27
2.3	LEI DE ACESSO À INFORMAÇÃO (LAI)	29
2.4	ESTRUTURA DOS SÍTIOS OFICIAIS DO GOVERNO FEDERAL	32
2.5	RESUMO E DISCUSSÃO	34
3	METADADOS	37
3.1	REPRESENTAÇÃO DE RECURSOS INFORMACIONAIS DIGITAIS	38
3.2	PADRÕES DE METADADOS	39
3.2.1	Dublin Core	40
3.2.2	e-PMG.....	50
3.3	RESUMO E DISCUSSÃO	55
4	RECUPERAÇÃO E VISUALIZAÇÃO DE INFORMAÇÃO	57
4.1	QUERYING	58
4.2	BROWSING.....	61
4.3	VISUALIZAÇÃO DE INFORMAÇÃO	62
4.4	RESUMO E DISCUSSÃO	67

5	PROPOSTA DE METADADOS PARA DESCRIÇÃO DO ACESSO AOS RECURSOS INFORMACIONAIS DIGITAIS GERADOS DINAMICAMENTE	68
5.1	PROPOSTA DE METADADOS	69
5.2	RESUMO E DISCUSSÃO	80
6	SISTEMA WEB PARA DESCRIÇÃO DE RECURSOS INFORMACIONAIS DIGITAIS GERADOS DINAMICAMENTE	82
6.1	MODELO DE DADOS LÓGICO DO SISTEMA WEB MDDIR TOOL	82
6.2	FUNCIONALIDADES DO SISTEMA WEB MDDIR TOOL	84
6.3	RESUMO E DISCUSSÃO	96
7	CONSIDERAÇÕES FINAIS.....	97
	REFERÊNCIAS.....	99

1 INTRODUÇÃO

Silva e Lopes (2011) afirmam que todos os períodos da humanidade tiveram características e peculiaridades que lhes deram uma identidade própria. No século XX, grandes transformações provocaram mudanças em todos os setores da sociedade. A sociedade que emergiu nesse período recebeu diferentes denominações: sociedade pós-industrial (BELL, 1977), sociedade pós-moderna (LYOTARD, 1990), sociedade pós-capitalista (DRUCKER, 1994), sociedade em rede (CASTELLS, 1999), sociedade da modernidade líquida (BAUMAN, 2001), sociedade da informação (MATTELART, 2002) são algumas das denominações cunhadas pelos estudiosos desse período.

Coutinho (2011) afirma que um dos primeiros autores a referir-se ao conceito de Sociedade da Informação foi o economista Fritz Machlup, em seu livro publicado em 1962, *The Production and Distribution of Knowledge in the United States*. No entanto, o desenvolvimento do conceito deve-se a Peter Drucker que, em 1966, no bestseller *The Age of Discontinuity*, menciona pela primeira vez uma sociedade pós-industrial em que o poder da economia – que, segundo o autor, teria evoluído da agricultura para a indústria e desta para os serviços - estava agora assente num novo bem precioso: a Informação.

Referindo-se à Sociedade da Informação, John Naisbitt, afirma que atualmente a fonte de poder não é o dinheiro nas mãos de poucos, mas informação nas mãos de muitos.

Castells (1999, p. 46) estabelece diferenças entre os conceitos que definem a Sociedade da Informação e Sociedade Informacional, considerando que a principal característica da primeira é ter como elemento essencial a informação, enquanto a segunda apresenta-se como “[...] uma forma específica de organização social em que a geração, o processamento e a transmissão de informações tornam-se as fontes fundamentais de produção e poder devido às novas condições tecnológicas surgidas [...]”.

A chamada Sociedade de Informação configura uma economia baseada na informação, comunicação, telecomunicações e tecnologias da informação. Esta informação serve como matéria prima do processo, as telecomunicações como meio de distribuição e as tecnologias da informação como infraestrutura de

armazenamento, processamento e acesso.

A sociedade moderna vem de uma realidade em que o sigilo das informações era considerado como regra. Atualmente, vive-se ou caminha-se para uma realidade em que a regra é a transparência e o sigilo tornou-se a exceção. De forma geral, as tecnologias vigentes não permitiam que os cidadãos tivessem informações sobre o que os governos municipais, estaduais e federais estavam realizando em seus mandatos.

Em 2011, o Brasil deu um grande salto na busca da transparência pública com a sanção da Lei nº 12.527, de 18 de novembro de 2011, conhecida como Lei de Acesso à Informação (LAI). A LAI regulamenta o direito constitucional de acesso a informações públicas pelos cidadãos brasileiros, sendo aplicada aos 3 poderes do Governo Federal, Governo Estadual e Governo Municipal.

Tendo como base as mudanças propostas pela Lei de Acesso à Informação, Garcia e Targino (2012) apontam que é preciso vencer o desafio de caráter administrativo relativo à carência de recursos financeiros e humanos capacitados com a meta de derrubar a “cultura do silêncio” ou a “cultura do segredo”.

A LAI propõe uma situação ideal na qual governo e cidadão podem dialogar por meios eletrônicos, mudando assim a forma de prestação de serviços à população, estimulando a interação e discussão sobre os dados governamentais. Porém, a LAI é omissa quanto à especificação de padrões e formatos como as informações deverão estar disponibilizadas.

Idealmente, as informações de governo disponibilizadas à sociedade devem (ou deveriam) seguir normas e padrões internacionais, permitindo, assim, uma ampla e irrestrita utilização de tais informações. Nesse sentido, em dezembro de 2007, na Califórnia (Estados Unidos), um grupo de trabalho conhecido pelo nome de *Open Government Working Group* (<http://www.whitehouse.gov/open/about/working-group>), reuniu-se para desenvolver um conjunto de princípios que poderiam conceituar um dado aberto governamental (*open government data*).

Porém, mesmo que as informações sejam efetivamente distribuídas, atendendo a tais princípios, é dever do Estado fornecer aos cidadãos informações de forma que possam ser facilmente interpretadas e analisadas. A LAI só terá sentido se o cidadão fizer uso das informações disponibilizadas de forma a gerar

conhecimento.

Desta forma, os governos devem criar mecanismos que possibilitem a prestação de contas dos resultados de suas ações garantindo, assim, maior transparência e a exposição pública de seus feitos. A esse contexto dá-se o nome de *Accountability*, o qual norteia os trabalhos em busca de maior interação entre governos e cidadãos (JARDIM, 2000).

Por meio da Web vislumbra-se uma situação sem precedentes, na qual é possível, por um lado, os gestores captarem e entenderem os anseios da população. Por outro lado, é possível uma efetiva participação dos cidadãos na gestão pública.

A LAI afirma que é dever dos órgãos e entidades públicas promover, independentemente de requerimentos, a divulgação em local de fácil acesso, no âmbito de suas competências, informações de interesse coletivo ou geral por eles produzidas ou custodiadas. Porém, mesmo tornando-se lei, só a divulgação das informações não basta para o efetivo uso destas. É preciso criar mecanismos que permitam descrever, armazenar e recuperar de forma eficiente tais recursos informacionais¹.

Atualmente, a principal fonte de informação do governo federal é o “Portal da Transparência”. O Portal da Transparência do Governo Federal é uma iniciativa da Controladoria-Geral da União (CGU) com o “objetivo é aumentar a transparência da gestão pública, permitindo que o cidadão acompanhe como o dinheiro público está sendo utilizado e ajude a fiscalizar”.

O Portal da Transparência, em sua página principal ², apresenta seis grandes grupos, divididas em diferentes páginas representadas por abas de cores diferentes: “Despesas”, “Receitas”, “Convênios”, “CEIS”, “CEPIM” e “Servidores”. As consultas realizadas nesses diferentes grupos são parametrizadas utilizando-se os critérios disponíveis. Após selecionados os parâmetros e realizada uma consulta, uma nova página é apresentada, contendo os dados referentes ao parâmetros escolhidos.

¹ Segundo Giannasi-Kaimen e Carelli (2007), um recurso informacional abrange textos, imagens e sons disponíveis para acesso do usuário. Para este trabalho, considera-se um recurso informacional digital gerado dinamicamente apenas as páginas *Web*.

² <http://www.portaldatransparencia.gov.br/>

Verifica-se que a página contendo os resultados da consulta é gerada (montada) em tempo real, dinamicamente, obedecendo os critérios e parâmetros da consulta. Não se trata de um arquivo físico, material, sendo acessível apenas enquanto estiver sendo visualizada no navegador (browser). Caso o navegador seja fechado, o resultado da consulta se perde.

Esses tipos de recursos, denominado nesse trabalho como “Recursos Informacionais Digitais gerados dinamicamente”, não permitem o acesso automatizado por sistemas externos, não sendo legível por máquina, o que depõe contra a lei de acesso à informação e às características de dados abertos. Diante disso, cria-se a necessidade de persistir (materializar) os Recursos Informacionais Digitais gerados dinamicamente.

O termo “persistência” é usado comumente na área de Ciência da Computação e refere-se ao armazenamento não volátil de dados, por exemplo, o armazenamento em um dispositivo físico como um disco rígido. Pode-se dizer que de maneira geral, o termo persistência é associado a uma ação que consiste em manter em meio físico recuperável, como banco de dados ou arquivo, de modo a garantir a permanência das informações de um determinado estado de um objeto lógico.

Por meio da ferramenta HTTRACK Website Copier, que permite baixar um sítio para um diretório local, construindo recursivamente todos os diretórios, incluindo imagens e outros arquivos do servidor para o seu computador; é possível promover a persistência de um Recurso Informacional Digital gerado dinamicamente.

Para Alves (2005, p. 115), “[...] os metadados são conjuntos de atributos, mais especificamente dados referenciais, que representam o conteúdo informacional de um recurso que pode estar em meio eletrônico ou não”.

Vários padrões de metadados foram pesquisados, entre eles Dublin Core e e-PMG, que serviram como base para a proposta de um novo conjunto de metadados capaz de descrever o acesso aos Recursos Informacionais Digitais gerados dinamicamente, de forma mais precisa os caminhos que levam às informações, já que estas não podem ser utilizadas, pois não atendem às leis e princípios de dados abertos citados acima. Além disso, foi desenvolvido o sistema Web MDDIR Tool para validar o conjunto de metadados proposto e sua efetividade na recuperação de

informação.

Serviu como base para este trabalho, uma pesquisa feita em sítios oficiais dos Ministérios, órgãos e secretarias que possuem *status* de ministério. Desses sítios oficiais, nenhum apresenta as informações efetivamente em formato de dado aberto.

1.1 DELIMITAÇÃO DO TEMA

Este trabalho ficará restrito aos sítios oficiais do governo federal, podendo extrapolar outros contextos, que deveriam disponibilizar informações atendendo ao que rege a Lei de Acesso à Informação. A descrição dos recursos será feita apenas de fontes de dados gerados dinamicamente e o conteúdo / resultado das buscas será descrito pelo conjunto de metadados proposto.

1.2 PROBLEMA DE PESQUISA

A LAI regulamenta o direito constitucional de acesso a informações públicas pelos cidadãos brasileiros, sendo aplicada aos três poderes do Governo Federal, Governo Estadual e Governo Municipal.

É possível observar que os sítios oficiais dos governos não apresentam uma estrutura padrão que facilite o acesso às fontes de dados, além de que tais dados e informações não são disponibilizadas em formato de Dado Aberto o que fere a lei citada acima. É facilmente observado que a desestruturação das informações e a dificuldade de acessá-las é recorrente, o que faz com que o acesso se torne difícil para o cidadão. Ainda que a maioria das informações são geradas em tempo de execução da consulta, ou seja, a partir de parâmetros selecionados e informados, é gerada uma página Web dinâmica que só existe enquanto o navegador estiver aberto; ao fechá-lo estas informações se perdem e para acessá-las novamente é necessário percorrer exatamente o mesmo caminho realizado pela primeira vez. Todavia, nem sempre é possível percorrer o mesmo caminho, seja por falta de conhecimento ou por esquecimento do caminho percorrido anteriormente.

Sendo assim, é possível acessar novamente resultados de pesquisas com dados gerados dinamicamente a partir de parâmetros informados ou selecionados em sítios oficiais do governo federal a ponto de garantir novas formas de

recuperação?

1.3 OBJETIVOS

1.3.1 Objetivo Geral

O objetivo deste trabalho é propor um modelo de representação de Recursos Informativos Digitais gerados dinamicamente, disponibilizados pelo Governos Federal em seus sítios oficiais, usando metadados para descrever o acesso e persistir tais recursos a fim de possibilitar sua reutilização e garantir novas formas de recuperação.

1.3.2 Objetivos Específicos

- Realizar um levantamento da forma de acesso a dados em sítios oficiais do Governo Federal;
- Analisar o grau de aderência dos padrões de metadados Dublin Core e e-PMG para descrição de dados gerados dinamicamente;
- Elaborar um conjunto de metadados para descrição de Recursos Informativos Digitais gerados Dinamicamente;
- Criar histórico dos Recursos Informativos Digitais gerados dinamicamente acessados nos sítios oficiais do Governo Federal;
- Desenvolver um sistema Web para validar o conjunto de metadados proposto e sua efetividade na recuperação de informação.

1.4 JUSTIFICATIVA

A melhoria nos serviços de disponibilização de informações por meio dos órgãos públicos é uma busca que se tornou uma necessidade social. Ajudar a promover esta melhoria é obrigação de todo cidadão principalmente dos pesquisadores ligados à área de Ciência da Informação.

No Brasil existem leis que garantem o acesso aos dados governamentais e que estes devam estar disponíveis em formatos passíveis de interação e utilização.

Porém isso não é uma realidade, o que justifica o desenvolvimento de um trabalho desta natureza, que busca meios para descrever de forma efetiva o caminho que leva às informações disponibilizadas nos sítios oficiais do governo federal.

Com isso, o cidadão poderá ter acesso direto ao resultado de uma consulta sem precisar a cada nova necessidade passar por toda a estrutura hierarquizada dos sítios oficiais, selecionando diretamente o caminho (URL) que leva ao resultado esperado na consulta. Tal caminho (URL) e os resultados estarão descritos pelos metadados informados.

Atualmente, a cada consulta feita em um sítio oficial do Governo Federal e dependendo do caminho percorrido os resultados podem e serão diferentes. Para contornar essa situação, é apresentada uma proposta de metadados para descrição de Recursos Informacionais Digitais gerados dinamicamente; esses metadados descreverão o caminho do recurso informacional digital e permitirão o acesso direto a tal recurso.

Como contribuições deste trabalho pode-se descrever:

- definição de um conjunto de metadados para descrição de Recursos Informacionais Digitais gerados dinamicamente;
- novas possibilidades para o processo de recuperação de dados, em especial dados dinâmicos.

1.5 HIPÓTESE

Quando uma pesquisa é feita em um sítio oficial do governo federal a partir de parâmetros informados, os recursos informacionais digitais gerados dinamicamente não podem ser reutilizados por meio da recuperação de informação pois não possuem acesso direto por meio de um endereço eletrônico, só existindo enquanto o navegador estiver aberto e, além disso, as informações não são disponibilizadas em formato de dados abertos.

Caso a persistência ocorra e os dados estejam disponibilizados em formato aberto, será possível uma maior interação com tais informações possibilitando um maior entendimento das ações executadas pelo governo, e conseqüentemente essas ações governamentais poderão ter um maior acompanhamento por parte do

cidadão.

1.6 TESE

A partir da persistência / materialização dos resultados gerados dinamicamente e com metadados agregados é possível garantir o reuso de informação dos sítios oficiais do Governo Federal.

1.7 PROCEDIMENTOS METODOLÓGICOS

Este trabalho possui uma abordagem qualitativa e consiste em uma pesquisa exploratória dos sítios oficiais do Governo Federal, identificando os Recursos Informacionais Digitais gerados dinamicamente e aplicando os metadados (Dublin Core e e-PMG) para verificar se estes são suficientes para descrever tais fontes de dados.

Foi feita uma pesquisa descritiva dos padrões de metadados citados acima, a fim de cruzar informações coletadas na descrição das fontes de dados e identificar quais campos de cada um dos padrões podem ser úteis e aplicáveis na descrição de Recursos Informacionais Digitais gerados dinamicamente. Com isso, foi proposto um novo conjunto de metadados diferenciado capaz de descrever os recursos informacionais digitais gerados dinamicamente (Página Resultado), bem como as páginas que os geraram (Página Menu), unindo-as em um único conjunto.

Foi apresentado ainda o sistema MDDIR Tool para validar o conjunto de metadados proposto responsável pela descrição do acesso aos Recursos Informacionais Digitais gerados dinamicamente.

1.8 ORGANIZAÇÃO DO TRABALHO

No Capítulo 1, encontra-se a introdução do trabalho apresentando o problema de pesquisa, os objetivos, a justificativa, delimitação do tema, hipótese, procedimentos metodológicos e organização do trabalho.

O capítulo 2 trata da gestão pública, dados abertos, Lei de Acesso à Informação e da estrutura dos sítios oficiais do governo federal.

O capítulo 3 aborda os padrões de metadados Dublin Core e e-PMG, e também apresenta conceitos sobre a representação de recursos informacionais digitais.

O capítulo 4 apresenta os conceitos relacionados à Recuperação e Visualização de Informação a fim de identificar quais informações atendem às necessidades do usuário e o uso de imagens para representação destas de forma significativa.

O capítulo 5 apresenta a proposta de metadados para descrição de Recursos Informacionais Digitais gerados dinamicamente.

No Capítulo 6, destaca-se o sistema Web MDDIR (Metadata for Describing Digital Information Resources - MDDIR - TOOL) para descrição de Recursos Informacionais Digitais gerados dinamicamente, bem como suas funcionalidades e particularidades.

Por fim, apresentam-se as considerações finais e as referências do trabalho.

2 ACESSO A DADOS DA GESTÃO PÚBLICA

O acesso a dados sobre gestão pública está intimamente relacionado à questão da democracia, pois estabelece uma relação com a disponibilização de informações que servem como subsídio para a tomada de decisão por parte do cidadão. É um direito de todo cidadão brasileiro ter acesso à informação pública, direito esse garantido pela Constituição Federal de 1988, em seu artigo 5º, inciso XXXIII, que aponta que todos os brasileiros têm direito a receber dos órgãos públicos informações de seu interesse particular, ou de interesse coletivo ou geral, que serão prestadas no prazo da lei, sob pena de responsabilidade, ressalvadas aquelas cujo sigilo seja imprescindível à segurança da sociedade e do Estado.

A Constituição Federal garante, ainda, no 2º parágrafo do artigo 216, que cabe à administração pública, na forma da lei, a gestão da documentação governamental e as providências para franquear sua consulta a quantos dela necessitem. Sendo assim, todas as informações que estejam de posse do estado são públicas a todos os cidadãos brasileiros, com raras exceções em que a segurança nacional ou o próprio cidadão possam estar em risco.

O acesso aos dados sobre a gestão pública é uma forma de o cidadão se tornar agente na busca de melhores condições sociais com vistas a cobrar dos governantes a prestação de contas de suas ações, podendo ser considerado ainda como uma área de intermediação e comunicação entre os órgãos públicos e os cidadãos. Para tanto, é necessário criar uma forma de comunicação pública que possibilite informar e estabelecer um diálogo que promova a discussão das ações governamentais, lembrando que a Constituição Federal obriga o Estado a estabelecer esta comunicação, apontada acima.

Até o ano de 2012, antes de a Lei de Acesso à Informação (LAI) ser sancionada, a Lei Federal 11.111 de 5 de maio de 2005 que foi revogada pela LAI, garantia que todos os documentos da gestão pública deveriam ser apropriados para consulta com exceção daqueles cujo sigilo é necessário para a segurança da sociedade e do Estado, sendo o sigilo de tais documentos avaliado pela Comissão de Averiguação e Análise de Informações Sigilosas que fazia parte do Ministério da Casa Civil. Esta comissão não tinha poder de análise sobre os poderes Legislativo e Judiciário, o Ministério Público da União e o Tribunal de Contas da União, que

disciplinam internamente sobre o sigilo de seus próprios documentos.

O décimo inciso do 5º parágrafo da Lei 11.111 garantia o sigilo a arquivos públicos que comprometessem a intimidade, vida privada, honra e imagem de pessoas, considerando que, para a Lei Federal 8.159 de 1991, arquivo público são os documentos produzidos e recebidos, no exercício de suas atividades, por órgãos públicos de âmbito federal, estadual, do Distrito Federal e municipal em decorrência de suas funções administrativas, legislativas e judiciárias, garantindo ainda o direito a qualquer cidadão o acesso a esses documentos públicos.

Pela LAI, todas as instituições, esferas, autarquias, empresas públicas e qualquer entidade controlada pelo Estado, União, Distrito Federal, incluindo as Cortes de Contas, ficam subordinadas ao seu regime a partir de 16 de maio de 2012. Esta lei determina ainda a subordinação de entidades sem fins lucrativos com ações de interesse público que recebam qualquer aporte financeiro oriundo de recursos financeiros públicos. A nova lei tem o propósito de regulamentar o direito constitucional de acesso dos cidadãos às informações públicas e seus dispositivos são aplicáveis aos três Poderes da União, Estados, Distrito Federal e Municípios.

Atualmente, o Gestor Público exerce papel de extrema importância na sociedade e no processo de democratização, pois tem a função de maximizar e aperfeiçoar as demandas necessárias para atingir os objetivos estabelecidos por órgãos governamentais respaldando principalmente os anseios e necessidades de todos os cidadãos, e para isso a disponibilização de dados e informações da gestão pública bem como seu acesso é de extrema importância.

2.1 ASPECTOS LEGAIS DO ACESSO A DADOS GOVERNAMENTAIS

Muito se tem falado nos últimos anos sobre a transparência do governo no que diz respeito ao acesso às informações governamentais públicas, como um direito de todo cidadão, garantido pela Lei 12.527/2011. Para formalizar o acesso a dados governamentais, foi criado o Programa de Governo Eletrônico, que tem como princípio a utilização das modernas tecnologias de informação e comunicação (TIC) para democratizar o acesso à informação, ampliar discussões e dinamizar a prestação de serviços públicos com foco na eficiência e efetividade das funções

governamentais.

As políticas do Programa de Governo Eletrônico seguem um conjunto de diretrizes que atuam em três frentes fundamentais:

- junto ao cidadão;
- na melhoria da sua própria gestão interna;
- na integração com parceiros e fornecedores.

O principal objetivo do Programa de Governo Eletrônico brasileiro é a transformação das relações do Governo com os cidadãos, empresas e também entre órgãos do próprio governo de forma a:

- aprimorar a qualidade dos serviços prestados;
- promover a interação com empresas e indústrias;
- fortalecer a participação cidadã por meio do acesso à informação e a uma administração mais eficiente.

As diretrizes gerais de implantação e operação do Governo Eletrônico funcionam no âmbito dos Comitês Técnicos de Governo Eletrônico e servem de referência para estruturar as estratégias de intervenção, sendo adotadas como orientação para todas as ações de governo eletrônico, gestão do conhecimento e gestão da TI em toda a Administração Pública Federal.

As diretrizes do Governo Eletrônico são:

- a prioridade do Governo Eletrônico é a promoção da cidadania;
- a Inclusão Digital é indissociável do Governo Eletrônico;
- o Software Livre é um recurso estratégico para a implementação do Governo Eletrônico;
- a gestão do conhecimento é um instrumento estratégico de articulação e gestão das políticas públicas do Governo Eletrônico;
- o Governo Eletrônico deve racionalizar o uso de recursos;
- o Governo Eletrônico deve contar com um arcabouço integrado de políticas, sistemas, padrões e normas;

- a integração das ações de Governo Eletrônico com outros níveis de governo e outros poderes.

Com essas diretrizes e por meio de portais Web, é possível diminuir as distâncias entre o governo e o cidadão. Com essa proximidade, cria-se uma esfera de "diálogo eletrônico" entre as duas partes. Lemos (2004) defende que o governo eletrônico objetiva regenerar o espaço público, otimizar os serviços prestados à população e estimular a interação e discussão dos problemas locais.

Todavia, ainda não é possível analisar e interpretar os dados fornecidos pelo governo federal devido ao fato de que, na maioria das vezes, estes dados não estão disponíveis em formatos eletrônicos, abertos e não proprietários.

Para solução desse problema, o ideal é que as informações sejam disponibilizadas em formato aberto e completas, permitindo a análise e utilização resultando em uma forma que o cidadão entenda.

Um grande avanço nesse sentido foi a criação da LAI, que garante a todos os cidadãos o acesso à informação pública, estabelece ainda a obrigatoriedade de os órgãos e entidades públicas divulgarem, independente de solicitação por parte dos cidadãos, informações de interesse geral ou coletivo, garantindo a confidencialidade prevista no texto legal. A Lei determina que estejam acessíveis na internet dados relacionados à estrutura, aos gastos, aos processos licitatórios e contratos, entre outros, o que é denominado de Transparência Ativa.

Esta lei ainda garantiu a inversão dos conceitos referentes à disponibilização e acesso às informações governamentais, uma vez que as legislações anteriores entendiam que quase a totalidade das informações era considerada sigilosa e que apenas algumas exceções poderiam ser acessíveis. A partir da sanção da Lei de Acesso à Informação, os documentos do governo seguem o princípio da publicidade e o sigilo passa a ser a exceção e não mais a regra.

Sant'ana (2008) afirma que a garantia do acesso a dados possibilita ampliar a participação do cidadão em um controle vertical na dinâmica entre Estado e Sociedade.

Para o efetivo acesso às informações governamentais, o cidadão não necessita fazer uma solicitação formal mas sim deve ter a possibilidade de fazê-lo por meios eletrônicos. Com isso, o conceito de Governo Eletrônico permite a

automação de processos pré-existentes em papel e escritórios, criando e possibilitando às ferramentas computacionais tecnológicas o acesso e disponibilização das informações necessárias ao cidadão.

Essa disponibilização de informações faz com que os governos passem a prestar conta de suas ações, garantindo maior transparência das decisões governamentais. Para tanto, é necessária a descrição de um conjunto de mecanismos e procedimentos conhecido como "*accountability*".

Considera-se "*accountability*" o conjunto de mecanismos e procedimentos que levam os tomadores de decisões governamentais a prestar contas dos resultados de suas ações, garantindo-se maior transparência e a exposição pública das políticas públicas. Quanto maior a possibilidade de os cidadãos poderem discernir se os governantes estão agindo em função do interesse da coletividade e sancioná-los apropriadamente, mais *accountable* é um governo. Trata-se de um conceito fortemente relacionado ao universo político administrativo anglo-saxão (JARDIM, 2000).

Em linhas gerais, o conceito de *Accountability* pode ser encarado como Prestar Contas, pois engloba ações de responsabilização, prestação de contas e sanções, o que no caso deste trabalho diz respeito à transparência na disponibilização dos dados governamentais, pois pode incluir responsabilidades dos eleitos na transparência e na fiscalização.

Para Schedler (1999), *Accountability* seria, então, um conceito que abarca, genericamente, três formas de prevenir e corrigir os abusos de poder, quais sejam: obrigação de se abrir ao público, obrigação de se explicar e justificar suas ações e subordinação à possibilidade de sanções.

Akutsu (2002) afirma que *Accountability* envolve duas partes:

[...] a primeira que delega responsabilidade para que a segunda proceda à gestão dos recursos e, ao mesmo tempo, gera a obrigação daquele que administra os recursos de prestar contas de sua gestão, demonstrando o bom uso desses recursos. [Esclarece, ainda, que] quando os recursos a serem geridos são públicos, a parte que delega é a Sociedade, representada pelo Poder Legislativo, e a parte delegada é o governo, a quem cabe a responsabilidade final pela gestão dos recursos.

Schedler (1999) afirma que *Accountability* deve contemplar os aspectos de natureza obrigatória, apoiar-se nos três pilares (informação, justificativa e responsabilização), manter uma “modéstia relativa”, ter caráter público, reconhecer-se de natureza multifacetada, atentar à complementaridade dos controles vertical e horizontal e adotar a recursividade intransitiva.

Baseando nos conceitos da *Accountability*, uma pergunta fica no ar: Por que é importante acessar dados governamentais? Muitos pontos devem ser levados em consideração antes de respondê-la.

O primeiro ponto é que muitos dados produzidos pelo governo não são utilizados criando uma quantidade enorme de dados inúteis, tendo em vista que os governos fazem uso apenas daquilo que é necessário em suas gestões, enquanto a maior parte não se transforma em informação, conhecimento ou serviços para os cidadãos.

O segundo ponto é que as informações disponíveis podem e devem ser usadas pelo cidadão da forma que lhe for mais conveniente. O objetivo é superar as restrições existentes para que os cidadãos possam encontrar, acessar, entender, cruzar, combinar, utilizar e consolidar os dados públicos de acordo com suas necessidades. A representação dos dados e informações devem ocorrer de forma que os cidadãos possam reutilizá-las, sendo este o ponto mais importante para a caracterização das informações e dados como governamentais abertos.

O terceiro ponto é que disponibilizar dados na Web sempre foi possível, porém isso não os caracteriza como dados abertos. O conceito de dados abertos vai muito além e se caracteriza pelo uso de tecnologias que garantam que estes possam ser acessados e reutilizados por máquinas; o valor dos dados abertos está no interesse do cidadão em utilizá-los.

Além disso, existe ainda a necessidade de prover mais transparência ao disponibilizar dados que podem ser reutilizados livremente, permitindo uma visão mais ampla das ações de governo. E se os dados governamentais abertos forem “confrontados” com outros dados de outras fontes, será possível obter novos significados sobre o desempenho do governo, o que demanda maior responsabilidade dos agentes públicos (DINIZ, 2010).

O acesso a dados governamentais é importante, pois ao acessá-los podemos

entendê-los, cruzá-los, combiná-los e utilizá-los na intenção de verificar se os dados e informações estão corretas e, dessa forma, analisar os serviços prestados pelo governo, tornando a sociedade mais consciente e crítica.

2.2 LEIS E PRINCÍPIOS DE DADO ABERTO

Em dezembro de 2007, na Califórnia, nos Estados Unidos da América, um grupo de trabalho conhecido pelo nome de *Open Government Working Group* (<http://www.whitehouse.gov/open/about/working-group>), reuniu-se para desenvolver um conjunto de princípios que poderiam conceituar um dado aberto governamental (*open government data*). Este grupo foi composto por membros de vários setores: representantes de entidades governamentais, de empresas de tecnologia da informação, de organizações não-governamentais e pesquisadores, totalizando 30 membros.

Dentre as discussões, chegou-se à conclusão de que a Internet possibilita aos Estados uma oportunidade sem igual de "entender" de forma precisa os anseios dos cidadãos, além de proporcionar mecanismos que possibilitam a participação efetiva na gestão pública. Além disso, o grupo entendeu que a informação tem maior peso quando é compartilhada e que os dados abertos tornam-se mais valiosos quando promovem melhorias no diálogo e na qualidade e bem-estar da sociedade, promovendo o uso consciente e eficiente dos dados e informações das esferas públicas.

Para o *Open Knowledge Foundation*, os dados são abertos quando qualquer pessoa pode usá-los livremente, reutilizá-los e redistribuí-los, estando sujeito, no máximo, à exigência de creditar a sua autoria e compartilhá-los pela mesma licença. Essas exigências são geralmente satisfeitas pela publicação dos dados em formato aberto e sob uma licença aberta.

Dados abertos são fundamentados em três leis e oito princípios que quando seguidos comprovam sua natureza. Tais leis foram propostas por David Eaves, especialista em políticas públicas e ativista dos dados abertos; lembrando que o foco são dados abertos governamentais, mas que podem ser aplicadas a dados abertos em geral. Tais leis são vistas abaixo:

- Se o dado não pode ser encontrado e indexado na Web, ele não existe;

- Se não estiver aberto e disponível em formato compreensível por máquina, ele não pode ser reaproveitado; e
- Se algum dispositivo legal não permitir sua replicação, ele não é útil.

Nesta mesma reunião de 2007, o grupo definiu os princípios dos Dados Abertos Governamentais, a saber:

Completo: Todos os dados públicos devem ser disponibilizados. Dados são informações eletronicamente gravadas, incluindo, mas não se limitando a, documentos, bancos de dados, transcrições e gravações audiovisuais. Dados públicos são dados que não estão sujeitos a limitações válidas de privacidade, segurança ou controle de acesso, reguladas por estatutos.

Primários: Os dados devem ser publicados na forma coletada na fonte, com a mais fina granularidade possível e não de forma agregada ou transformada.

Atuais: Os dados devem ser disponibilizados o quanto rapidamente seja necessário para preservar o seu valor.

Acessíveis: Os dados devem ser disponibilizados para o público mais amplo possível e para os mais variados propósitos.

Processáveis por máquina: Os dados devem ser razoavelmente estruturados para possibilitar o seu processamento automatizado.

Acesso não discriminatório: Os dados devem estar disponíveis a todos, sem que seja necessária identificação ou registro.

Formatos não proprietários: Os dados devem estar disponíveis em um formato sobre o qual nenhuma entidade tenha controle exclusivo.

Livres de licenças: Os dados não devem estar sujeitos a regulações de direitos autorais, marcas, patentes ou segredo industrial. Restrições razoáveis de privacidade, segurança e controle de acesso podem ser permitidas na forma regulada por estatutos.

Apesar de os princípios terem sido pensados para os Dados Abertos Governamentais, podem ser aplicados a Dados Abertos de modo geral.

Por meio da Lei de Acesso à Informação, os sítios oficiais do governo federal devem possibilitar a gravação de relatórios em diversos formatos eletrônicos,

abertos e não proprietários, tais como planilhas e texto, de modo a facilitar a análise das informações e possibilitar o acesso automatizado por sistemas externos em formatos abertos, estruturados e legíveis por máquina.

A seção seguinte aborda de forma detalhada a Lei de Acesso à Informação.

2.3 LEI DE ACESSO À INFORMAÇÃO (LAI)

Importantes organismos internacionais, como a Organização das Nações Unidas (ONU) e a Organização dos Estados Americanos (OEA), reconhecem que o acesso à informação é um direito fundamental a todos os cidadãos, e o Brasil, a partir de novembro de 2011, também passou a ter esse entendimento. Toda informação que esteja sob tutela do estado é sempre pública e seu acesso não pode ser restrito, salvando-se apenas alguns poucos casos específicos. Com isso, o Estado tem o dever de produzir, guardar, organizar, gerenciar e dar acesso à informação já que é considerada um bem público.

Segundo a Controladoria Geral da União - CGU, a Lei de Acesso à Informação no Brasil prevê como exceções o acesso aos dados pessoais e às informações classificadas por autoridades como sigilosas.

Há um entendimento internacional de que a consolidação da democracia passa pelo acesso aos dados e às informações que estão sob tutela do Estado e, desta forma, há um fortalecimento da capacidade de os cidadãos participarem e entenderem as tomadas de decisões que dizem respeito a eles. Isso pode ser explicado pelo fato de que o cidadão bem informado tem mais condições de conhecer e acessar informações sobre seus direitos primários como saúde, educação e benefícios sociais. Atualmente, cerca de 90 países já possuem leis que regulamentam o acesso à informação e entendem este processo como um direito do cidadão.

A seguir encontram-se transcritos trechos de alguns tratados, convenções e declarações que asseguram o direito do cidadão às informações públicas e que serviram como base para a criação da LAI:

Toda pessoa tem direito à liberdade de opinião e expressão; este direito inclui a liberdade de, sem interferência, ter opiniões e de procurar, receber e transmitir informações e idéias por quaisquer meios e independentemente de fronteiras. (DECLARAÇÃO..., 1948,

Art. XIX).

[...] cada Estado Parte [...] adotará medidas que sejam necessárias para aumentar a transparência em sua administração pública [...]. Essas medidas poderão incluir [...] informação sobre a organização, o funcionamento e os processos de adoção de decisões de sua administração pública [...]. (BRASIL, 2006, Art. 10).

[...] Cada Estado Parte adotará medidas adequadas [...] para fomentar a participação ativa de pessoas e grupos que não pertençam ao setor público, como a sociedade civil, as organizações não-governamentais e as organizações com base na comunidade, na prevenção e na luta contra a corrupção, e para sensibilizar a opinião pública a respeito à existência, às causas e à gravidade da corrupção, assim como a ameaça que esta representa. (BRASIL, 2006, Art. 13).

O acesso à informação em poder do Estado é um direito fundamental do indivíduo. Os Estados estão obrigados a garantir o exercício desse direito. (DECLARAÇÃO..., 2000, Art. 4)

Toda pessoa terá direito à liberdade de expressão; esse direito incluirá a liberdade de procurar, receber e difundir informações e idéias de qualquer natureza, independentemente de considerações de fronteiras, verbalmente ou por escrito, em forma impressa ou artística, ou qualquer outro meio de sua escolha. (BRASIL, 1992, Art. 18, item 2).

A LAI regulamenta o direito constitucional de acesso às informações públicas pelos cidadãos brasileiros, sendo aplicada aos 3 poderes do Governo Federal, Governo Estadual e Governo Municipal.

Ao sancionar uma lei dessa natureza, confirmou-se a consolidação do regime democrático do Brasil, segundo o qual o cidadão brasileiro tem o direito de ter acesso às informações públicas de acordo com sua necessidade e expectativa e, dessa forma, fortalece a política de transparência pública que pode acontecer de forma ativa ou passiva.

Segundo o Manual sobre Transparência e Lei de Acesso à Informação Pública do Tribunal de Contas dos municípios do estado do Ceará (2013), a Transparência Pública refere-se à obrigação imposta ao administrador público em promover a prestação de contas para a população. O governo deve regularmente divulgar o que faz, como faz, por que faz, quanto gasta e apresentar o planejamento para o futuro. A transparência ativa ocorre quando os governos divulgam dados por iniciativa própria, sem terem sido solicitados; já a transparência passiva é entendida como o acesso aos dados públicos fornecidos pelos governos, quando solicitados.

Esta lei toma como princípio básico e fundamental que o acesso à informação

pública é a regra e o não fornecimento de tais informações é a exceção, assim como já foi dito acima. Para a garantia do cumprimento desta lei, há mecanismos, prazos e procedimentos para entrega das informações solicitadas à administração pública, seja em esfera municipal, estadual ou federal, sem contar que a lei determina que os organismos públicos devem divulgar um número mínimo de dados e informações de forma natural via internet. O acesso à informação é um direito do cidadão previsto pela Constituição Federal, no Art. 5º, inciso XIV, Art. 37, § 3º, inciso II e Art. 216, § 2º. A lei prevê ainda medidas que responsabilizam os agentes públicos que retardarem ou impossibilitarem o acesso às informações por parte do cidadão.

Como apresentado acima, algumas informações consideradas "sigilosas" podem ter o acesso restrito por um determinado período de tempo e sua restrição justifica-se pela salvaguarda da segurança nacional e / ou do próprio cidadão. O maior desafio para implementação da lei são questões de ordem técnica, tecnológica, cultural, social e administrativa, pois para que isso ocorra é necessária uma mudança de paradigmas e a capacitação de servidores, já que a conscientização da importância destes no processo ocasionará o seu sucesso. O controle efetivo da implementação e cumprimento desta lei fica sob responsabilidade da Controladoria Geral da União.

Informações classificadas como sigilosas são aquelas cuja Lei de Acesso à Informação prevê alguma restrição de acesso, mediante classificação por autoridade competente, visto que são consideradas imprescindíveis à segurança da sociedade (à vida, segurança ou saúde da população) ou do Estado (soberania nacional, relações internacionais, atividades de inteligência).

As informações pessoais não são públicas e terão seu acesso restrito, independentemente de classificação de sigilo, pelo prazo máximo de 100 (cem) anos a contar da sua data de produção. Elas sempre podem ser acessadas pelos próprios indivíduos e por terceiros apenas em casos excepcionais previstos na Lei.

Dados Pessoais são aquelas informações relacionadas à pessoa natural identificada ou identificável. Seu tratamento deve ser feito de forma transparente e com respeito à intimidade, vida privada, honra e imagem das pessoas, bem como às liberdades e garantias individuais.

Conforme a Lei de Acesso à Informação, a informação pública pode ser

classificada como:

- Ultrassecreta: prazo de segredo 25 anos (renovável uma única vez);
- Secreta: prazo de segredo 15 anos;
- Reservada: prazo de segredo 5 anos.

Em se tratando da Lei de Acesso à Informação, a transparência pública é sua alma e para tanto é importante que tais informações e dados possam ser acessados e utilizados a fim de ser útil para o cidadão. Para isso, é importante que os dados estejam abertos e atendam aos seus princípios que foram vistos na seção 2.2 deste trabalho.

É importante frisar que sítios oficiais dos governos, em muitos casos, não seguem os princípios e leis previstos na LAI e nem tão pouco possuem uma estrutura para exibição de informação. A LAI delibera de forma simplificado sobre tal estrutura, podendo ser observado na seção seguinte.

2.4 ESTRUTURA DOS SÍTIOS OFICIAIS DO GOVERNO FEDERAL

De acordo com o Artigo 8º da LAI, é dever dos órgãos e entidades públicas promover, independentemente de requerimentos, a divulgação em local de fácil acesso, no âmbito de suas competências, de informações de interesse coletivo ou geral por eles produzidas ou custodiadas.

O parágrafo 1º do mesmo artigo afirma que na divulgação das informações, deverão constar, no mínimo:

I - registro das competências e estrutura organizacional, endereços e telefones das respectivas unidades e horários de atendimento ao público;

II - registros de quaisquer repasses ou transferências de recursos financeiros;

III - registros das despesas;

IV - informações concernentes a procedimentos licitatórios, inclusive os respectivos editais e resultados, bem como a todos os contratos celebrados;

V - dados gerais para o acompanhamento de programas, ações, projetos e obras de órgãos e entidades; e

VI - respostas a perguntas mais freqüentes da sociedade.

O parágrafo 2º atenta que, para cumprimento efetivo da divulgação das informações, os órgãos e entidades públicas deverão utilizar todos os meios e instrumentos legítimos de que dispuserem, sendo obrigatória a divulgação em sítios oficiais da rede mundial de computadores (internet).

Além disso, a LAI aponta ainda no 3º parágrafo do artigo 8º, que os sítios oficiais deverão, na forma de regulamento, atender, entre outros, aos seguintes requisitos:

I - conter ferramenta de pesquisa de conteúdo que permita o acesso à informação de forma objetiva, transparente, clara e em linguagem de fácil compreensão;

II - possibilitar a gravação de relatórios em diversos formatos eletrônicos, inclusive abertos e não proprietários, tais como planilhas e texto, de modo a facilitar a análise das informações;

III - possibilitar o acesso automatizado por sistemas externos em formatos abertos, estruturados e legíveis por máquina;

IV - divulgar em detalhes os formatos utilizados para estruturação da informação;

V - garantir a autenticidade e a integridade das informações disponíveis para acesso;

VI - manter atualizadas as informações disponíveis para acesso;

VII - indicar local e instruções que permitam ao interessado comunicar-se, por via eletrônica ou telefônica, com o órgão ou entidade detentora do sítio;

VIII - adotar as medidas necessárias para garantir a acessibilidade de conteúdo para pessoas com deficiência, nos termos do art. 17 da Lei no 10.098, de 19 de dezembro de 2000, e do art. 9º da Convenção sobre os Direitos das Pessoas com Deficiência, aprovada pelo Decreto Legislativo nº 186, de 9 de julho de 2008.

Mesmo havendo as garantias e obrigações criadas pela LAI, não há uma definição de um layout de interfaces que os portais públicos devam seguir no que diz respeito a quais seções e informações devem estar disponíveis. Essa falta

de padronização acaba gerando dificuldades no ato das consultas nestes ambientes.

Se levarmos em consideração os Recursos Informativos Digitais gerados dinamicamente, o problema se agrava ainda mais, pois não possuem acesso direto por meio de um endereço eletrônico, ou seja, são páginas de internet criadas a partir da seleção de parâmetros e que só existem enquanto o navegador estiver aberto e não permite o acesso direto sem passar pelas páginas de parametrização. Como não há estrutura definida e descrição de tais recursos, sua recuperação torna-se muito difícil.

2.5 RESUMO E DISCUSSÃO

O capítulo 2 apresenta todo levantamento bibliográfico sobre o acesso a dados na gestão pública, *accountability*, dados abertos, Lei de Acesso à Informação e a arquitetura da informação dos sítios oficiais do governo federal.

A Constituição Federal de 1988, em seu artigo 5º, inciso XXXII, garante a todo brasileiro o direito de acesso à informação pública, sendo possível estabelecer uma relação estreita com a democracia, uma vez que a disponibilização de informações serve como subsídio para a tomada de decisão por parte do cidadão.

É garantido também pelo 2º parágrafo do artigo 216 da Constituição Federal que cabe à administração pública a gestão e a franquia da documentação governamental para consulta. Diante disso, todas as informações que estejam de posse do estado são públicas a todos os cidadãos brasileiros, com raras exceções em que a segurança nacional ou o próprio cidadão possam estar em risco.

Em 2012, com a sanção da Lei Federal 12.527, foi regulamentado o direito constitucional de acesso dos cidadãos às informações públicas e seus dispositivos são aplicáveis aos três Poderes da União, Estados, Distrito Federal e Municípios. Para formalizar o acesso a dados governamentais, foi criado o Programa de Governo Eletrônico, que tem como princípio a utilização das modernas tecnologias de informação e comunicação (TIC) para democratizar o acesso à informação, ampliar discussões e dinamizar a prestação de serviços públicos com foco na eficiência e efetividade das funções governamentais.

Estudiosos da área de dados abertos entendem que a informação tem maior

peso quando é compartilhada e que dados abertos tornam-se mais valiosos quando promovem melhorias no discurso civil e na qualidade e bem-estar da sociedade promovendo o uso consciente e eficiente de dados e informações das esferas públicas. Para que um dado seja considerado aberto, deve ser encontrado e indexado na Web; deve estar disponível em formato compreensível por máquina; e deve ser permitida sua replicação. Além disso, devem atender a 8 princípios: Completos; Primários; Atuais; Acessíveis; Processáveis por máquina; Acesso não discriminatório; Formatos não proprietários; e Livres de licenças.

Para que esses dados considerados públicos estejam à disposição do cidadão para consulta, a Lei de Acesso à Informação, em seu 8º artigo, afirma que é dever dos órgãos e entidades públicas promover, independentemente de requerimentos, a divulgação, em local de fácil acesso, no âmbito de suas competências, de informações de interesse coletivo ou geral por eles produzidas ou custodiadas. Para isso, deve constar em seus sítios oficiais informações como estrutura organizacional, endereços e telefones das respectivas unidades e horários de atendimento ao público, repasses ou transferências de recursos financeiros, despesas, procedimentos licitatórios, entre outros.

Com este trabalho, é possível verificar que a gestão pública vai muito além de gerir as ações exigidas pelos cargos, é necessária também uma postura de transparência para que os resultados de tais ações possam servir como indicador para análise do desempenho do gestor público. Para isso, os dados precisam ser efetivamente abertos e as informações precisam estar de fato disponíveis.

Após analisar cada um dos sítios oficiais dos Ministérios do Governo Federal, foi possível detectar que as informações disponibilizadas, exigidas pela Lei de Acesso à Informação, não estão em formato aberto e nem disponibilizadas de acordo com as obrigações da lei. Com isso, surge o problema de pesquisa que é como reutilizar resultados de pesquisas gerados dinamicamente a partir de parâmetros informados ou selecionados em sítios oficiais do governo federal a ponto de garantir novas formas de recuperação.

Como solução, a proposta é utilizar em mecanismos de recuperação o novo conjunto metadados apresentado neste trabalho, a fim de localizar o ponto exato em que é gerada uma tabela de resultado de uma determinada consulta sem passar pela estrutura (caminho) que hoje é hierarquizado.

Os metadados servirão como um conjunto de atributos ou dados referenciais, que representarão o conteúdo informacional de um recurso informacional digital disponibilizado pelo Governo Federal.

3 METADADOS

Castro e Santos (2005) apontam que o uso cada vez maior dos computadores contribuiu para o avanço da tecnologia e da ciência em geral, o que fez com que a automação de bibliotecas e o desenvolvimento de formas de representação que agilisassem o processamento dos dados bibliográficos e catalográficos por máquinas fossem possíveis.

Aliado ao avanço das tecnologias, podemos citar a criação e evolução da internet na década de 1990, que acabou possibilitando um grande crescimento de informações disponibilizadas em meio digital. Com isso, houve a necessidade de uma união de bibliotecários, profissionais de computação e profissionais de lingüística em busca da criação de novas técnicas e instrumentos tecnológicos nos processos de produção, armazenagem, representação, descrição e recuperação de Recursos Informacionais Digitais que podem ser disponibilizados em diversas mídias / formatos, tais como, som, imagem, texto etc.

Quando se fala em formatos, é necessário definir estruturas que possibilitem a descrição de documentos digitais e, para isso, são utilizados os metadados que se referem a dados sobre dados.

Para Alves (2005, p. 115), “[...] os metadados são conjuntos de atributos, mais especificamente dados referenciais, que representam o conteúdo informacional de um recurso que pode estar em meio eletrônico ou não”.

Segundo Gilliland-Swetland (1999), “[...] para os profissionais da Ciência da Informação, o termo metadados está relacionado com o tratamento da informação, em especial às formas de representação que caracterizam um recurso informacional para fins de identificação, de localização e de recuperação, ou seja, dados sobre catalogação e indexação que servem para organizar e tornar a informação mais acessível”.

Também conhecida como representação descritiva, a catalogação tem como objetivo principal descrever de forma padronizada um recurso informacional, identificando de modo individualizado e possibilitando sua recuperação.

Catalogação é a atividade que consiste, basicamente, em registrar um conjunto de informações sobre um determinado documento ou recurso. Tais

informações são registradas e variam de acordo com o tipo de documento que está sendo catalogado.

A catalogação descreve um conjunto de informações determinadas por meio de regras de descrição ou representação, onde são definidos os elementos que irão descrevê-lo. Os padrões de representação disponíveis facilitam a localização de informações fornecendo o mínimo de uniformidade aos registros bibliográficos.

Um dos elementos do conjunto de metadados proposto neste trabalho é o parâmetro que é utilizado para compor o resultado das consultas. Esses parâmetros garantem a persistência e o reuso de um Recurso Informacional Digital gerado dinamicamente.

3.1 REPRESENTAÇÃO DE RECURSOS INFORMACIONAIS DIGITAIS

Nos dias atuais, a necessidade por padronização da representação de Recursos Informacionais Digitais tem feito com que instituições, órgãos e institutos ligados à Ciência da Informação unam esforços com a intenção de desenvolver e melhorar a utilização dos metadados.

A busca pela definição de um formato de descrição documental de Recursos Informacionais Digitais que facilite a recuperação de informação, passa por uma padronização que garanta a total interoperabilidade entre sistemas de recuperação da informação.

Carvalho (2003) aponta que as informações geradas atualmente estão, cada vez mais, sendo armazenadas no formato digital. As vantagens desta forma de armazenamento de informações são muitas, destacando-se, entre elas, a flexibilidade oferecida para a sua recuperação e a possibilidade de armazenamento e veiculação em diferentes tipos de mídia.

Do ponto de vista específico da representação da informação, Alvarenga (2001) expõe com clareza:

A representação do conhecimento, em nossos dias, não compreende somente a substituição do documento primário por uma informação catalográfica. Considerando-se que o documento não se acha fisicamente em outro espaço, mas no próprio meio que lhe

proporciona materialidade, novas formas de se criar índices de recuperação foram ensejadas. No novo contexto de produção, organização e recuperação de Recursos Informacionais Digitais, as metas de trabalho não se restringem à criação de representações simbólicas dos objetos físicos constantes de um acervo, mas compreendem estabelecimento dos denominados metadados, muitos dos quais podem ser extraídos diretamente dos próprios objetos, constituindo-se esses em chaves de acesso a serviço dos internautas.

[...] Na catalogação bibliográfica tradicional, o documento é representado por um conjunto de informações relativas à sua descrição física e pontos de acesso, representação esta preparada e armazenada em um contexto físico independente do documento primário.

[...] À medida que as tecnologias da informação foram sendo criadas, disponibilizadas e aperfeiçoadas, os sistemas de representação e recuperação de informações documentais assistiram a uma extrapolação dos limites dos tradicionais catálogos referenciais em fichas, alcançando as bases de dados online.

A seção seguinte aponta as principais características dos padrões de metadados abordados acima.

3.2 PADRÕES DE METADADOS

Os padrões de metadados têm como objetivo principal estabelecer um grau de normalização, padronização de representação descritiva automatizada dos dados informacionais, favorecendo a compatibilidade e a transferência de dados legíveis por computador, possibilitando o compartilhamento e a cooperação de recursos e serviços entre diferentes unidades de informação (FUSCO, 2010).

Para Smith e Schirling (2006), os padrões de metadados possuem elementos descritores organizados para descrever objetos do domínio de informação do usuário. Eles possibilitam melhor armazenamento, utilização, recuperação e troca de informações entre aplicações e entre organizações, de forma rápida, automatizada e eficiente. Com isso, possibilita uma padronização que é elemento essencial para que ocorra um melhor entendimento das estruturas, sintaxes e significados dos metadados, permitindo assim uma melhora na comunicação, reutilização e compartilhamento das informações, promovendo uma queda no curso e trabalho para tais atividades.

Com a utilização de padrões de metadados, é possível garantir que informações, sobre os mesmos temas ou áreas sejam comuns, facilitando a

compreensão, integração e compartilhamento de informações entre usuários de formações, propósitos e experiências diferentes. Determinar padrões é permitir a utilização das definições estabelecidas por um conjunto de elementos descritores. Os dados, como nomes, informações ou grupos de dados utilizados para descrever determinado tipo de acervo são conteúdos que geralmente são padronizados (WESTERMANN, 2003).

Nascimento (2008) aponta alguns dos principais objetivos da utilização de padrões de metadados:

- auxiliar na automatização da recuperação das informações;
- estabelecer a comunicação consistente entre bancos de dados;
- garantir o registro de informações que sejam importantes;
- facilitar o uso e o compartilhamento das informações e conhecimentos;
- auxiliar a troca entre informações de diferentes bancos de dados.

Cada padrão de metadados possui uma quantidade de elementos descritores que varia de acordo com suas características e finalidade e sua complexidade pode ser observada levando em consideração a variedade de informações que seus elementos são capazes de descrever. Um conjunto de elementos descritores deve ser apropriado o suficiente para que permita a interpretação de suas informações por humanos e máquinas, uma vez que essa interpretação pode ser feita por aplicativos computacionais que fazem a busca e a recuperação de informações.

Pesquisas apontam para uma variedade significativa de padrões de metadados, cada um com características específicas e intenção de facilitar o entendimento de relacionamentos e utilidade de informações. Entre esses padrões podemos citar: Dublin Core e o e-PMG, que poderão ser vistos neste capítulo.

3.2.1 Dublin Core

O padrão de metadados Dublin Core teve origem em Outubro de 1994, durante a 2ª Conferência Internacional sobre WWW realizada em Chicago. Yuri Rubinsky, Stuart Weibel e Eric Miller, os três da *Online Computer Library Center* (OCLC) e Joe Hardin do *National Center for Supercomputing Applications* (NCSA) conduziram uma discussão sobre semântica e Web. Deste Workshop surgiu a

proposta de um evento direcionado ao assunto.

Em 1995, a OCLC e a NCSA organizaram um evento chamado de *OCLC/MCSA Metadata Workshop*, que contou com a participação de pesquisadores de várias áreas (ciência da computação, bibliotecários, profissionais de serviços de informação online, indexadores, catalogadores, profissionais envolvidos com tratamento de dados geo-espaciais, imagens, museus e arquivos, entre outros) cujo objetivo era discutir como um conjunto semântico para recursos baseados na Web, poderia ser útil para uma pesquisa e recuperação de recursos na Internet. Este evento tinha como principal objetivo chegar à definição de um conjunto mínimo de descritores para recursos da Web.

Em 1997, Desay afirmava que a intenção era tratar o problema da catalogação de recursos Web com a adoção, a extensão ou modificação de padrões existentes e de protocolos para facilitar a recuperação e acesso a tais recursos e atender a várias comunidades que utilizassem metadados. Para chegar a um conjunto mínimo de elementos, foram seguidas algumas suposições durante o *OCLC/MCSA Metadata Workshop*, a saber:

- os elementos deveriam descrever um documento como objeto (DLO);
- chegar a um conjunto de elementos comuns;
- os elementos deveriam suportar a recuperação dos recursos da rede;
- todos os elementos de metadados poderiam ser repetitivos;
- todos os elementos seriam opcionais;
- todos os elementos descreveriam o recurso por si próprio, com exceção do elemento fonte;
- os elementos deveriam descrever características intrínsecas do recurso;
- nenhuma suposição seria feita para dizer se o recurso era acessível pela rede ou especificamente eletrônico;
- elementos não incluídos não seriam necessariamente excluídos.

Entre as principais características do Dublin Core, podemos destacar:

- Simplicidade: pode ser facilmente gerado pelo responsável do

documento sem a necessidade de extensos treinamentos;

- Interoperabilidade semântica: existência de um modelo comum que aumenta a possibilidade de interoperabilidade entre as áreas;
- Consenso internacional: participação de muitos países no DCMI para a busca de escopo internacional na WEB e de uma infraestrutura adequada que contribui para um consenso internacional;
- Extensibilidade: permite que novos elementos possam ser acrescentados para atender a uma necessidade de descrição de um determinado recurso;
- Flexibilidade: seus elementos são opcionais, podem ser repetidos se necessário.

Dentre as principais dificuldades encontradas pelos participantes do *OCLC/MCSA Metadata Workshop*, havia a questão de trabalhar com padrões que possuísem muitos elementos. Diante disso, foi determinada uma relação de elementos não exaustiva para descrever os recursos da rede, tornando-se consenso que a relação de elementos não deveria ser extensa. Este padrão de metadados, receberia o nome de Dublin Core, pois o *OCLC/MCSA Metadata Workshop* aconteceu em Dublin, Ohio.

Ainda durante o OCLC/MCSA Metadata Workshop, surgiu o DCMI - *Dublin Core Metadata Initiative* que é uma organização cuja missão é promover a difusão da adoção de padrões de metadados e o desenvolvimento de vocabulário especializado de metadados para descrever recursos. É importante salientar que o DCMI é aberto a qualquer pessoa, órgão ou instituição cujo interesse seja buscar um consenso mínimo no desenvolvimento de vocabulários de metadados (DCMI, 2013).

Entre as atividades desenvolvidas pelo DCMI, destacam-se (DCMI, 2013):

- Desenvolvimento de padrões de metadados para a recuperação de informações nos domínios da Internet;
- Definição de estruturas para a interoperabilidade de conjuntos de metadados;
- Organização de workshops internacionais e de grupos de trabalhos para o desenvolvimento e a manutenção das recomendações do

DCMI;

- Disponibilizar ferramentas, serviços e infraestrutura em metadados;
- Alcançar outras comunidades de metadados, desenvolvendo a interação entre elas e o DCMI.

O DCMI é composto de (DCMI, 2013):

- Conselho de Administradores: que aconselha a Diretoria nas questões estratégicas e de alocação de recursos financeiros, sendo seus membros de diversos países;
- Diretoria: composta por um Diretor Executivo e outro Administrativo, que supervisionam o gerenciamento e a coordenação das atividades dos Grupos de Trabalho;
- Conselho Consultivo: composto pelos presidentes dos Grupos de Trabalho, dando consultoria à Diretoria;
- Grupos de Trabalho: criados de acordo com as necessidades e composto por voluntários, que contribuem na resolução de problemas específicos.

Como este padrão de metadados tem o objetivo de descrever um recurso eletrônico, a descrição dos elementos de metadados é normalizada. Esta normalização ajuda a melhorar a consistência com outras comunidades de metadados, permitindo ainda uma melhor clarificação, definição do alcance e consistência interna nas definições dos elementos de metadados "Dublin Core".

O Dublin Core é composto de um conjunto de quinze elementos, com o objetivo principal de descrever um recurso eletrônico, sendo cada elemento "Dublin Core" definido usando um conjunto de dez atributos da norma ISO/IEC 11179 [ISO11179], para a descrição de elementos de dados. Estes consistem em (DCMI, 2013):

- Nome: A etiqueta atribuída ao elemento de dado;
- Identificador: O identificador único atribuído ao elemento de dado;
- Versão: A versão do elemento de dado;
- Autoridade de Registro: A entidade autorizada a registrar o elemento

de dado;

- Língua: A linguagem na qual o elemento de dado é definido;
- Definição: Uma afirmação que representa claramente o conceito e a natureza do elemento de dado;
- Obrigação: Indica se o elemento de dados é sempre obrigatório ou não (contém um valor);
- Tipo dos Dados: Indica o tipo de dados que pode ser representado no valor do elemento de dado;
- Máxima Ocorrência: Indica qualquer limite à repetição do elemento de dado;
- Comentário: Uma nota relativa à aplicação do elemento de dado.

Dentre os 10 atributos apresentados acima, 6 são comuns a todos os elementos do padrão de metadados Dublin Core, abaixo são apresentados com seus respectivos valores:

- Versão: 1.1;
- Autoridade de Registro: *Dublin Core Metadata Initiative*;
- Língua: en;
- Obrigação: Opcional;
- Tipo dos Dados: Cadeia de caracteres;
- Máxima Ocorrência: Ilimitada.

As definições são apresentadas de duas formas: conceituais e de representação dos elementos. O atributo Definição captura o conceito semântico, enquanto os atributos Tipo dos Dados e Comentário capturam a representação dos dados. Por sua vez, o atributo Identificador deverá ser entendido não como um nome mas apenas como uma etiqueta ou código único identificador do elemento, o qual será sempre representado pela palavra em língua Inglesa equivalente ao valor do atributo Nome.

Cada elemento "Dublin Core" refere-se ao recurso a ser descrito. Um recurso é definido como tudo o que tiver uma identidade. Sendo assim, um recurso será

tipicamente um serviço ou um recurso de informação, mas o conceito poderá ser aplicado de forma mais abrangente. Abaixo serão apresentados cada elemento de forma individualizada (DCMI, 2013):

- Elemento: Título.

Nome: Título.

Identificador: Title.

Definição: O nome dado ao recurso.

Comentário: Um Título será o nome pelo qual o recurso é formalmente conhecido.

Exemplo: Análise sobre metadados para descrição de Recursos Informativos Digitais gerados dinamicamente (texto).

- Elemento: Criador.

Nome: Criador.

Identificador: Creator.

Definição: A entidade responsável em primeira instância pela existência do recurso.

Comentário: O nome de um Criador deve ser usado para indicar uma entidade. Exemplos de Criador incluem uma pessoa, uma organização ou um serviço.

Exemplo: MODESTO, Lisandro Rogério (autor).

- Elemento: Assunto.

Nome: Assunto e Palavras-Chave.

Identificador: Subject.

Definição: Tópicos do conteúdo do recurso.

Comentário: Um Assunto deverá ser expresso por palavras-chave, frases ou códigos de classificação que descrevem o conteúdo do recurso. Recomenda-se a seleção de termos de vocabulários controlados ou de sistemas de classificação formais.

Exemplo: metadados, padrões de metadados, sítios oficiais do governo federal.

- Elemento: Descrição.

Nome: Descrição.

Identificador: Description.

Definição: Uma descrição do conteúdo do recurso.

Comentário: Descrições podem incluir, sem estarem limitadas a tal: um resumo, um índice, uma referência a uma representação gráfica do conteúdo ou uma descrição textual.

Exemplo: "O objetivo desse trabalho é definir um conjunto e um formato de metadados capaz de representar a estrutura de dados, servindo como um elemento mediador entre eles e sistemas e aplicações que possam utilizar tais dados para as mais variadas finalidades." (Resumo)

- Elemento: Editor.

Nome: Editor.

Identificador: Publisher.

Definição: Uma entidade responsável por tornar o recurso acessível.

Comentário: O nome de um Editor deve ser usado para indicar a entidade. Exemplos de um Editor incluem uma pessoa, uma organização ou um serviço. Exemplo: Saraiva (editora).

- Elemento: Colaborador.

Nome: Colaborador.

Identificador: Contributor.

Definição: Uma entidade responsável por qualquer contribuição para o conteúdo do recurso.

Comentário: O nome de um Outro Contribuinte deve ser usado para indicar a entidade. Exemplos de Outro Contribuinte incluem uma pessoa, organização ou serviço.

Exemplo: FAP - Faculdade de Apucarana (Entidade).

- Elemento: Data

Nome: Data

Identificador: Date

Definição: Uma data associada a um evento do ciclo de vida do recurso.

Comentário: Uma Data deve ser associada à criação ou disponibilidade do recurso. Recomenda-se para codificação de valores de datas um perfil da norma ISO 8601 [W3CDTF], segundo o formato AAAA-MM-DD.

Exemplo: 2013-02-10.

- Elemento: Tipo.

Nome: Tipo do Recurso.

Identificador: Type.

Definição: A natureza ou gênero do conteúdo do recurso.

Comentário: Tipos incluem termos descrevendo categorias genéricas, funções, gêneros ou níveis de agregação para o conteúdo. Recomenda-se a seleção de valores a partir de vocabulários controlados (por exemplo, a lista do documento de trabalho "Dublin Core Types" [DCT1]). Para descrever a manifestação física ou digital do recurso, deve ser usado o elemento Formato.

Exemplo: Interactive Resource.

- Elemento: Formato

Nome: Formato.

Identificador: Format.

Definição: A manifestação física ou digital do recurso.

Comentário: O Formato deve incluir o tipo de meio do recurso, ou as suas dimensões. Este elemento deve ser usado para determinar as aplicações informáticas ou qualquer tipo de equipamento necessário para reproduzir ou operar com o recurso. Exemplos de dimensões incluem tamanho e duração. Recomenda-se a seleção de valores a partir de vocabulários controlados (como por exemplo a lista de "Internet Media Types" [MIME] definindo formatos e meios).

Exemplo: text/pdf.

- Elemento: Identificador

Nome: Identificador do Recurso.

Identificador: Identifier.

Definição: Uma referência não ambígua ao recurso, definida num determinado contexto.

Comentário: Recomenda-se a identificação do recurso por meio de uma cadeia de caracteres ou por um número de acordo com um sistema de identificação formal. Exemplos de sistemas de

identificação formais incluem o "Uniform Resource Identifier" (URI) (incluindo o "Uniform Resource Locator" (URL)); o "Digital Object Identifier" (DOI); e o "International Standard Book Number" (ISBN).

Exemplo: <http://www.dublincore.org/documents/1999/07/02/dces/>.

- Elemento: Fonte.

Nome: Fonte.

Identificador: Source.

Definição: Uma referência a um recurso de onde o presente recurso possa ter derivado.

Comentário: O presente recurso pode ter derivado do recurso Fonte na sua totalidade ou apenas em parte. Recomenda-se a referência ao recurso fonte através de um identificador em conformidade com um sistema de identificação formal.

Exemplo: <http://www.dublincore.org>.

- Elemento: Língua

Nome: Língua

Identificador: Language

Definição: A língua do conteúdo intelectual do recurso.

Comentário: Recomenda-se para valores do elemento Língua a utilização do RFC 1766 [RFC1766], o qual inclui um código de língua de duas letras (retirado da norma ISO 639 [ISO639]), seguido opcionalmente por um código de duas letras para o país (retirado da norma ISO 3166 [ISO3166]). Por exemplo, 'en' para Inglês, 'fr' Francês, ou 'en-uk' para o Inglês do Reino Unido.

Exemplo: pt-br.

- Elemento: Relação

Nome: Relação.

Identificador: Relation.

Definição: Uma referência a um recurso relacionado.

Comentário: Recomenda-se referir o recurso por meio de uma cadeia de caracteres ou número em conformidade com um sistema de identificação formal.

Exemplo: é referente ao XXX.

- Elemento: Cobertura

Nome: Cobertura.

Identificador: Coverage.

Definição: A extensão ou alcance do recurso.

Comentário: Cobertura inclui uma localização espacial (o nome de um lugar ou coordenadas geográficas), um período no tempo (a sua designação, data, ou intervalo de tempo), ou jurisdição (o nome de uma entidade administrativa). Recomenda-se a seleção de valores de vocabulários controlados (como por exemplo o "Thesaurus of Geographic Names" [TGN]), devendo ainda ser usados, quando apropriado, preferencialmente nomes de lugares e designações de períodos no tempo, em vez de identificadores numéricos tais como coordenadas ou intervalos de datas.

Exemplo: FAP, Apucarana, Paraná, Brasil.

- Elemento: Direitos

Nome: Gestão de Direitos.

Identificador: Rights.

Definição: Informação de direitos sobre o recurso ou relativos a este.

Comentário: Este elemento deverá conter uma declaração de gestão de direitos sobre o recurso ou uma referência a um serviço que fornecerá essa informação. Poderá compreender informação sobre direitos de propriedade intelectual, direitos de autor ou outros. A ausência deste elemento não permite formular qualquer hipótese válida sobre quaisquer direitos que possam incidir sobre o recurso.

Exemplo: <http://www.dublincore.org/about/organization>.

Para explicar com mais riqueza de detalhes, abaixo será apresentado um exemplo da descrição de um recurso utilizando o padrão de metadados Dublin Core:

Quadro 1 - Descrição de recurso utilizando o padrão de metadados Dublin Core

Elemento	Conteúdo
DC.Title	Ministério do Trabalho e Emprego
DC.Creator	Ministério do Trabalho e Emprego
DC.Subject	Trabalho

DC.Subject	Emprego
DC.Description	O portal do Ministério do Trabalho e Emprego apresenta informações institucionais, conselhos e comissões, emprego e renda, inspeção do trabalho, economia solidária, relações de trabalho, dados e estatísticas, entre outras.
DC.Publisher	
DC.Contributor	
DC.Date	2013-02-10
DC.Type	Interactive Resource
DC.Format	Text/html
DC.Identifier	http://portal.mte.gov.br/portal-mte/
DC.Source	http://portal.mte.gov.br/portal-mte/
DC.Language	pt-br
DC.Relation	
DC.Coverage	1997-2008
DC.Rights	MTE

Fonte: Elaborado pelo autor.

3.2.2 e-PMG

O Governo Federal utiliza o padrão de metadados chamado e-PMG, cujo objetivo é assegurar às pessoas, que pesquisam as informações do governo brasileiro na Web, o acesso rápido e eficiente a descrições dos recursos. Os elementos e-PMG têm o propósito de facilitar a localização dos recursos que as pessoas necessitam, mesmo que elas não possuam conhecimento detalhado da localização ou da entidade responsável pela informação.

O Padrão de Metadados do Governo Eletrônico (e-PMG) foi desenvolvido pelo Comitê Executivo de Governo Eletrônico e estabelece a semântica dos elementos, qualificadores e a forma de usá-los para descrição dos recursos informacionais (e-PMG, 2010).

O e-PMG foi desenvolvido com base no Padrão Dublin Core (DC) do Dublin Core Metadata Initiative (DCMI), grupo responsável pelo desenvolvimento do Dublin Core. O e-PMG estabelece um conjunto mínimo de elementos que contém dados necessários para a recuperação e o gerenciamento de informações. Cada elemento representa informações de um aspecto específico do recurso, como por exemplo, “Título” ou “Criador”.

No decorrer desta seção, são apresentadas as principais características do e-PMG, todas as informações foram extraídas do guia desenvolvido pelo Governo Federal. As razões para o desenvolvimento do e-PMG são:

- a modernização do governo exige melhor uso da informação oficial, unindo sistemas, políticas e serviços destinados às necessidades do cidadão;
- os metadados facilitam gerenciar e recuperar informação, esteja na forma eletrônica (páginas Web, documentos digitais, imagens, áudio etc.) ou não eletrônica.

O e-PMG difere do padrão DC em outros pontos importantes:

- pode ser utilizado para descrever tanto recursos eletrônicos (por exemplo, páginas Web, arquivos ou outros Recursos Informativos Digitais) quanto recursos não eletrônicos (por exemplo, livros, acervos de museu, pinturas, documentos impressos etc.);
- pretende descrever mais do que recursos informativos, podendo descrever serviços disponíveis na Web;
- a adoção de qualificadores adicionais; e
- diferentes critérios na obrigatoriedade ou não do uso de elementos e qualificadores.

A adoção deste padrão propõe benefícios para os órgãos do governo e para a sociedade. Embora haja necessidade de considerável esforço e investimento inicial, os benefícios são muitos, como por exemplo:

- permitir que as pessoas localizem os recursos disponíveis na Web sem o conhecimento detalhado de como o governo está organizado;
- dispor de métodos eficientes de acesso às informações e serviços;
- facilitar o compartilhamento de informações e serviços entre órgãos governamentais e entre governo e sociedade, o que resulta em maior eficiência do governo;
- apoiar o desenvolvimento de serviços com maior qualidade, com base em semântica comum e padronização de formatos e esquemas de codificação dos dados; e
- eficiência nos processos de gestão da informação das organizações.

Os órgãos do governo precisam avaliar os seus principais objetivos e o

público-alvo, a fim de decidir quais recursos devem ser descritos prioritariamente com o uso do e-PMG. Excetuadas as determinações regulamentares, cada organização deve decidir quais recursos requerem os metadados e-PMG prioritariamente.

Em 2010, o Governo Brasileiro lançou um documento sobre o e-PMG - Padrão de Metadados do Governo Eletrônico, que recomenda a identificação dos recursos com maior frequência de utilização para que estes também sejam prioritariamente descritos. Nem todo recurso precisa ser descrito. Num processo de adoção gradativa, um órgão pode inicialmente utilizar metadados e-PMG, agregando os documentos de uma determinada coleção (um sítio da Web por exemplo), então gradativamente reduzir a granularidade, chegando no nível individual se for o caso. Para os recursos legados (já existentes), caberá à organização a decisão quanto à necessidade de descrever esses recursos com metadados e-PMG. A legislação ou as políticas governamentais poderão orientar a necessidade de descrição desses recursos.

O conjunto de elementos do e-PMG consiste em 20 elementos: 15 elementos do Dublin Core e 5 elementos adicionais identificados como necessários para o contexto do governo eletrônico brasileiro, são eles:

- Nome: Contexto Jurídico-administrativo
Identificador: Mandate
Definição: Legislação e normas que regulam e autorizam o funcionamento e o desempenho das atividades da organização responsável pela produção e/ou disponibilização do recurso.
- Nome: Destinação
Identificador: Disposal
Definição: Decisão quanto ao encaminhamento dos documentos arquivísticos para guarda permanente ou eliminação.
- Nome: Destinatário
Identificador: Addressee
Definição: Entidade (pessoa ou organização) para quem a informação contida no documento foi dirigida.
- Nome: Localização
Identificador: Location

Definição: Local de armazenamento atual do recurso.

- Nome: Preservação

Identificador: Preservation

Definição: Informação para apoiar o acesso, uso e preservação de longo prazo de um recurso.

Para cada elemento são descritos os seguintes dados:

- Nome: Termo atribuído ao elemento em português;
- Identificador: Termo atribuído ao elemento de acordo com o padrão originário;
- Definição: Uma descrição do que trata o elemento;
- Obrigatoriedade: Define o grau de uso do elemento, divididos em: obrigatório (este elemento deve obrigatoriamente ter um valor); obrigatório se aplicável (a este elemento deve ser fornecido um valor se o tipo de recurso assim o requerer); e opcional (a este elemento pode ser fornecido um valor se o dado estiver disponível e apropriado ao recurso). A obrigatoriedade aplica-se ao elemento e a seus qualificadores quando for o caso;
- Objetivo: Fornece a informação essencial sobre a aplicabilidade do elemento;
- Comentários: Informação adicional para a compreensão e utilização do elemento ou dos qualificadores do elemento;
- Não confundir com: Fornece informações comparativas com elementos que eventualmente possam gerar dúvidas ou semelhanças no nome e suscitar uso inadequado;
- Qualificadores: Usado para tornar o significado de um elemento mais específico, podendo ser utilizado para informação adicional de um recurso;
- Exemplos: Indica como os elementos podem ser acrescentados para diferentes situações. Os exemplos são fictícios, pois são destinados somente para demonstrar a forma de utilização do elemento ou

qualificador;

- **Sintaxe de HTML:** Usualmente os metadados aparecem em forma de tag em um arquivo HTML. Este campo visa facilitar a forma de uso dos elementos e-PMG para páginas da Web. Por exemplo: <meta name="dc.element" content="valor">;
- **Esquemas:** Os esquemas indicam a fonte que especifica a forma ou informa os valores possíveis para um elemento. Incluem vocabulários controlados, normas brasileiras e internacionais;
- **Mapeamento:** Correlação com elementos de outros padrões de metadados como: e-GMS (Government Metadata Standard), AGLS (Australian Government Locator Service), GILS (Government Locator Service), FGDC (Federal Geographic Data Committee (FGDC) Metadata), etc.

Para explicar com mais riqueza de detalhes, abaixo será apresentado um exemplo da descrição de um recurso utilizando o padrão de metadados e-PMG:

Quadro 2 - Descrição de recurso utilizando o padrão de metadados e-PMG

Elemento	Conteúdo
Mandate	
Disposal	
Addressee	Lisandro Rogério Modesto
Location	
Preservation	
Title	Ministério do Trabalho e Emprego
Creator	Ministério do Trabalho e Emprego
Subject	Trabalho
Subject	Emprego
Description	O portal do Ministério do Trabalho e Emprego apresenta informações institucionais, conselhos e comissões, emprego e renda, inspeção do trabalho, economia solidária, relações de trabalho, dados e estatísticas, entre outras.
Publisher	
Contributor	
Date	2013-02-10
Type	Interactive Resource
Format	Text/html
Identifier	http://portal.mte.gov.br/portal-mte/
Source	http://portal.mte.gov.br/portal-mte/
Language	pt-br
Relation	
Coverage	1997-2008
Rights	MTE

Fonte: Elaborado pelo autor.

3.3 RESUMO E DISCUSSÃO

No capítulo 3, foi feito um apanhado histórico e conceitual sobre metadados, apresentando alguns padrões que serviram como fonte para a proposta de um novo padrão de metadados para descrição de Recursos Informativos Digitais gerados dinamicamente.

O crescimento da computação contribuiu para o avanço da tecnologia e da ciência em geral, o que fez com que a automação de bibliotecas e o desenvolvimento de formas de representação que agilisassem o processamento dos dados bibliográficos e catalográficos por máquinas fossem possíveis. Além disso, a criação e evolução da internet na década de 1990 acabou possibilitando um grande crescimento de informações disponibilizadas em meio digital.

Esses avanços criaram a necessidade da união de bibliotecários, profissionais de computação e profissionais de lingüística em busca da criação de novas técnicas e instrumentos tecnológicos nos processos de produção, armazenagem, representação, descrição e recuperação de Recursos Informativos Digitais que podem ser disponibilizados em diversas mídias / formatos, tais como, som, imagem, texto etc.

Quando se fala em padrões de metadados, é necessário definir estruturas que possibilitem a descrição de documentos digitais e, para isso, são utilizados os metadados que referem-se a dados sobre dados.

Segundo Alves (2005), os metadados são conjuntos de atributos, mais especificamente dados referenciais, que representam o conteúdo informativo de um recurso que pode estar em meio eletrônico ou não, permitindo assim sua recuperação.

Padrões de metadados objetivam estabelecer um grau de normalização, padronização de representação descritiva automatizada dos dados informativos, favorecendo a compatibilidade e a transferência de dados legíveis por computador, possibilitando o compartilhamento e a cooperação de recursos e serviços entre diferentes unidades de informação e, neste caso, tais informações estão presentes nos sítios oficiais dos Ministérios do Governo Federal.

A descrição de metadados permite que a recuperação de informações seja

mais efetiva e que realmente se consiga recuperá-las e manipulá-las no intuito de analisá-las para que novos conhecimentos possam ser gerados.

4 RECUPERAÇÃO E VISUALIZAÇÃO DE INFORMAÇÃO

A Recuperação de Informação (RI) refere-se ao armazenamento e recuperação de documentos, tema de grande importância nos dias de hoje devido a enorme quantidade de dados e informações que as organizações precisam armazenar.

O surgimento da web no início dos anos 90 impulsionou um crescimento exponencial no volume de informação eletrônica, trazendo muitas vantagens no que tange à possibilidade de troca, difusão e transferência de dados. Entretanto, esse crescimento ocasionou problemas relacionados com o acesso, busca, localização e recuperação de informações relevantes em grandes volumes de dados.

Mazza (2009) contextualiza esse grande volume de informação a partir de atividades rotineiras, como exemplo, os e-mails que chegam em nossos computadores, extratos de cartões de créditos enviados todos os meses pelas operadoras, ofertas de última hora, variações de índices de mercados, panfletos publicitários.

Nas situações apontadas acima, o tema recorrente é a enorme quantidade de informações que temos que lidar diariamente. Cada uma das situações descritas anteriormente quase sempre envolve tomar uma decisão.

O detalhe é que não percebemos o aumento na quantidade de informações. A globalização da economia e o avanço da tecnologia fez eclodir uma massa enorme de informação.

O cientista da computação Mark Weiser vislumbrava, ainda no final da década de 80, uma presença cada vez maior dos computadores em nosso dia a dia, em que a tecnologia estaria integrada ao nosso ambiente sem chamar nossa atenção, a qual denominou de "computação ubíqua". (WEISER, 1991)

Weiser (1991) usou o termo ubíquo pela primeira vez para designar a terceira onda na computação. Primeiro foram os computadores de grande porte, compartilhado por inúmeras pessoas; logo após tem-se a era da computação pessoal, pessoa e máquina frente a frente sobre uma mesa; e por último, a era da computação ubíqua, ou a era da tecnologia calma, quando a tecnologia desaparece, recua para segundo plano em nossas vidas.

Nesse contexto, Frakes (1992) explica que, com o crescimento do volume de publicações ao longo dos anos, foram desenvolvidas técnicas de recuperação de informação para responder às necessidades dos usuários de bibliotecas, tradicionais ou digitais.

Para Cardoso (2000), não é possível falar em um simples crescimento do volume de publicações; mas o que tem ocorrido é uma verdadeira explosão. As bibliotecas digitais aparecem como um paradigma para melhorar a busca e a apresentação de informações desejadas, sendo que o índice é a ferramenta mais importante para auxiliar o processo de recuperação, que nada mais é que uma coleção de termos que indicam o local onde a informação desejada pode ser localizada.

Desta forma, faz-se necessário o aprofundamento no estudo de técnicas de digitalização de objetos originados de fontes heterogêneas, técnicas de armazenamento, processos de busca, recuperação e apresentação de forma amigável das informações.

Os estudos sobre recuperação de informação dividem-se em duas estratégias que abordaremos em seguida, que são: o querying e o browsing. O Querying e o Browsing, paradigmas básicos da recuperação da informação, representam dois tipos de comportamento na busca de informação. Cada um apresentando vantagens e desvantagens e complementam-se na RI (XAVIER, 2009, p. 15).

4.1 QUERYING

Pesquisa analítica baseada no modelo clássico, no querying, o usuário faz a busca expressando sua necessidade informacional através de palavras chaves, de tal forma que o sistema, após ter realizado a busca nos documentos, retorna ao usuário uma lista de resultados que, em seguida, inspeciona e seleciona o que interessa nos documentos recuperados. (XAVIER, 2009).

Nos sistemas baseados em querying, é importante fazer a distinção entre sistemas de recuperação de dados e recuperação de informação. Esta diferenciação é motivada, principalmente, pela técnica utilizada para as correspondências entre uma consulta e documento, denominadas respectivamente de correspondência exata e parcial (HERRERO-SOLANA; HASSAN, 2006).

De acordo com Herrero-Solana e Hassan (2006), a correspondência exata é a comparação do método determinista entre uma consulta e um conjunto de documentos, em que somente os resultados válidos são resolvidos naqueles documentos que atendam plenamente as necessidades expressas na consulta. Em outras palavras, somente é devolvido como resultado aqueles documentos que aparecem todos e cada um dos termos expressos na consulta de forma binária.

Herrero-Solana e Hassan (2006) explicam que a correspondência parcial é um método que possibilita a obtenção tanto de resultados parcialmente válidos, como a ordenação dos dados em função do seu grau de relevância para a consulta introduzida. Para tanto, o sistema deve fazer uso de algum modelo que proporcione:

- a) um método para transformar os documentos textuais em representações numéricas e computáveis;
- b) um método que, utilizando as representações numéricas, realize uma ponderação automática em cada termo do documento;
- c) um método de correspondência entre consulta e documento que determine em que grau o documento é relevante para a consulta com base no peso de cada um dos termos do documento.

Entre os métodos propostos que cumprem com tais requisitos, podemos ressaltar o Modelo de Espaço Vetorial, o Modelo Probabilístico e o Modelo Booleano.

No modelo espaço vetorial é possível recuperar resultados parciais, os dados são representados a partir da vetorização do conjunto documental. É considerado no modelo que cada documento é descrito por um conjunto de palavras chaves, chamadas termos de indexação. Associa-se a cada termo de indexação t_i em um documento d_j um peso $w_{ij} \geq 0$, que quantifica a correlação entre os termos e o documento (CARDOSO,2000)

Xavier (2009) aponta alguns pontos fortes na estrutura baseada em vetores, ressaltando que o modelo é adequado para a representação de objetos com vários atributos. A idéia de pesos complementa a indexação após a extração de termos relevantes que caracterizam o documento quantificando os termos de maior representatividade para o objeto, tornando a consulta aprimorada e flexível. Permite também uma gama de cálculos de similaridade resultando em escolhas mais

refinadas. O modelo ainda permite a implementação de métodos e técnicas sofisticados de processamento tais como self-organization maps (SOM), redes associativas Pathfinder, modelos de escalonamento multidimensional (MDS).

Xavier (2009) explica que uma das dificuldades encontradas no espaço vetorial é a alta dimensionalidade inerente, que faz com que sua eficiência dependa de coleções relativamente menores. Problemas de ambigüidade também podem ocorrer tendo em vista que os termos selecionados são usados na construção de uma matriz para uma futura recuperação; neste período, os contextos semânticos dos atributos do documento podem ser perdidos. O modelo de espaço vetorial pressupõe que os descritores dos documentos são independentes.

No modelo booleano, uma consulta é uma expressão booleana formada com a utilização dos operadores lógicos AND, OR e NOT. Nesse modelo, os documentos recuperados são somente aqueles nos quais os termos encontrados satisfazem a expressão lógica da consulta.

Cardoso (2000) aponta como principais problemas do modelo booleano a ausência de ordem na resposta, e as respostas podem ser nulas ou muito grandes. As vantagens desse modelo estão na facilidade de implementação e na expressividade completa das expressões.

Para Cardoso (2000), no modelo probabilístico, documentos são descritos considerando pesos binários que indicam a presença ou não de termos; é gerado um vetor resultado utilizando o teorema de Bayes como principal ferramenta matemática. Baseado no princípio probabilístico de ordenação, garante que o modelo pode ser utilizado de forma ótima. O conjunto resultado para cada requisição são documentos classificados na ordem decrescente com base na probabilidade de relevância para o usuário. O modelo considera a interação de estimativas da probabilidade de relevância.

Cardoso (2000) aponta como vantagem para o modelo probabilístico o bom desempenho e o princípio probabilístico de ordenação. Todavia, a desvantagem é que este comportamento depende da precisão das estimativas de probabilidade. Além disso, o método não explora a frequência do termo no documento e ignora o problema de filtragem de informação.

4.2 BROWSING

O Browsing é uma atividade não programada, não orientada, informal, assistemática, casual, o que pode justificar a dificuldade de uma definição precisa. Alguns autores consideram como um tipo de busca sem critérios ou objetivos definidos previamente.

Em contraposição ao querying, o browsing é uma resposta natural aos sentidos humanos, como: visão, tato, cheiro, audição; o que leva a pessoa a examinar títulos, páginas, índices, cor, entre outros aspectos cognitivos e sensoriais.

De acordo com Herrero-Solana e Hassan (2006), a estratégia de busca por browsing é utilizada em muitas situações do cotidiano, tal como quando são exploradas as estantes de uma biblioteca na busca de um livro. No meio digital, o melhor exemplo de recuperação por browsing se dá na navegação de hipertextos, em que o usuário explora visual e espacialmente o conjunto de hiperdocumentos com o objetivo de localizar a informação de seu interesse.

Na busca de um conceito para browsing, Naves (1998, p. 2) explica que:

[...] browsing é o processo de se expor a um espaço de recursos pelo esquadramento (scanning) de seu conteúdo de seu conteúdo (objetos ou representações) e/ou estrutura, possivelmente resultando em um conhecimento inesperado, novo conteúdo ou caminho nesse espaço de recursos.

Para Naves (1998, p.2), o usuário tem três tipos diferentes de atenção ao fazer browsing: no primeiro caso, o usuário pesquisa ao acaso em uma coleção desconhecida; no segundo caso, o usuário pesquisa no quase ao acaso através de um coleção já explorada; no terceiro caso, o usuário já tem determinado parcialmente o domínio da pesquisa.

Uma tipologia multidimensional é sugerida por Naves (1998, p.2), que tem como objetivo diferenciar essa atividade de outros tipos de busca de informação caracterizando, desta forma, os diversos tipos. A primeira dimensão é contextual, pois parte do princípio que as pessoas constroem significados a partir de contextos. A segunda dimensão é comportamental envolvendo esquadramento e movimento, o que leva à descoberta de novas informações e assim à aprendizagem e à descoberta. A terceira dimensão é motivacional e envolve a motivação e propósito do empenho em algumas atividades. A quarta dimensão diz respeito aos

aspectos cognitivos que incluem planos, conhecimento e experiência. A quinta dimensão envolve aspectos de recursos que inclui a forma e o foco da pesquisa.

Xavier (2009, p. 18) salienta que na pesquisa por browsing o usuário deve se posicionar em uma base com uma estrutura definida, na qual pode mover-se no interior dessa estrutura com base nas informações recebidas do sistema. O usuário tem a possibilidade de se deslocar livremente para a área de interesse. Destaca ainda dois níveis diferentes de informações a serem recuperadas, em que o primeiro é o nível micro refere-se a objetos individuais, como exemplo: palavras-chave e resumo; ao passo que no nível macro as informações são geradas a partir da junção de objetos individuais, que formam o nível micro. Desta forma, a pesquisa em nível micro é óbvia e direta e em nível macro, indireta e sofisticada. Xavier traça um contraste entre as estratégias analíticas (querying) e as estratégias baseadas no browsing, salientando que, na primeira estratégia, a definição tem o maior peso para o sucesso da busca; já a segunda estratégia pode ser aplicada a objetivos mais gerais ou informais e dependem muito mais da informação encontrada durante a busca, além disso também aponta algumas razões para realizar o browsing, sendo elas:

- obter uma visão Geral;
- monitorar um processo;
- dividir/compartilhar uma carga cognitiva;
- tornar mais clara uma necessidade informacional;
- desenvolver uma estratégia formal;
- descobrir/aprender.

Em suma, browsing e querying são estratégias de busca diferentes, porém se complementam, uma vez que o browsing pode apresentar problemas no que diz respeito à perda de foco de busca pelo usuário, não se colocando como substituto do querying, mas complementar.

4.3 VISUALIZAÇÃO DE INFORMAÇÃO

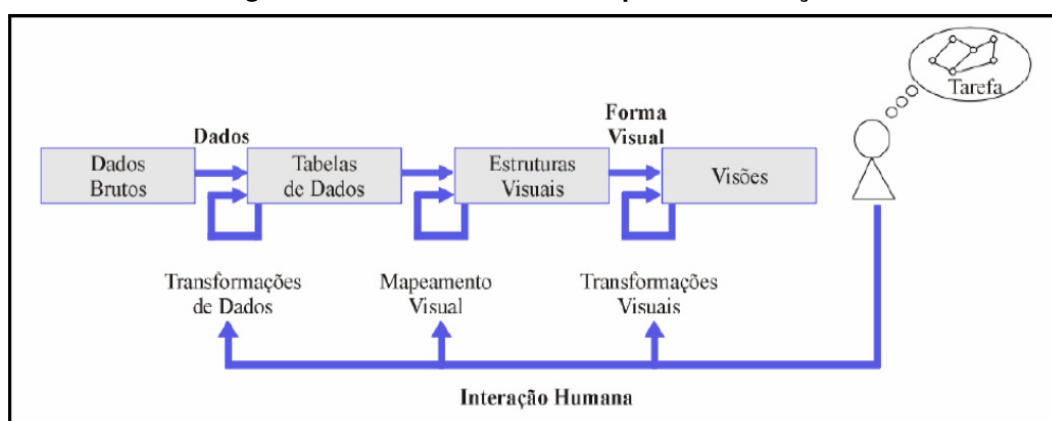
Quando se tem dados nos quais é necessário ilustrar conceitos, idéias e

propriedades intrínsecas aos dados, a representação visual fornece uma excelente ferramenta de visualização, por oferecer uma interface amigável que facilita a compreensão do usuário acerca de determinada informação.

A visualização de informação tem se apresentado como um excelente campo de pesquisa e cresce rapidamente devido a sua característica interdisciplinar; trabalhos nessa área são encontrados em bibliotecas e na interação homem-computador.

Vieira e Corrêa (2010) explicam que a Visualização de Informação (VI) é uma ciência que combina aspectos de computação gráfica, interação humano-computador, cartografia e mineração de dados. O processo de Visualização da Informação está relacionado com a transformação de dados abstratos em imagens que possam ser visualizadas.

Figura 1 - Modelo de referência para visualização.



Fonte: Vieira; Corrêa(2010)

Vieira e Corrêa (2010), observam que sistemas que utilizam linguagem gráfica devem apresentar os conceitos de expressividade e efetividade. A visualização é expressiva quando permite expressar os dados que interessam ao usuário dentro de um contexto. Já uma visualização efetiva facilita a compreensão dos dados apresentados, de tal forma a levar o usuário a localizar a informação de interesse sem erros semânticos.

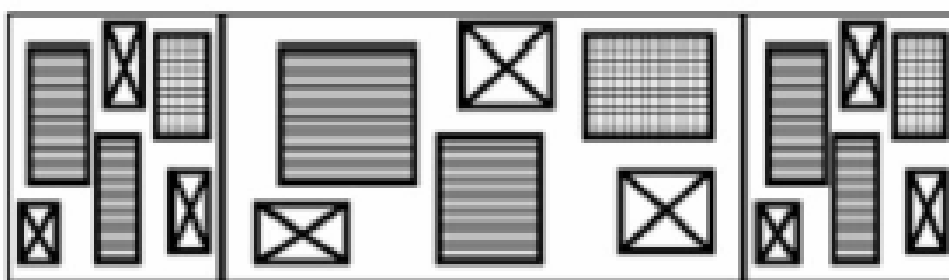
Em seguida, serão apresentadas algumas técnicas de VI dentro das muitas técnicas de visualização existentes.

As técnicas hierárquicas subdividem um espaço de dimensão k e apresentam os sub-espacos obtidos segundo um critério hierárquico que descendem de um nó

raiz. Na representação hierárquica, os elementos são apresentados em níveis, ramificações ou agrupamentos descendentes de um nodo raiz.

De acordo com Furnas (1986, apud XAVIER, 2009, p.38), a técnica Bifocal Display apresenta três áreas diferentes para apresentação dos elementos informacionais, em que a área central apresenta a informação de maior relevância, ao passo que informações de contexto geral são apresentadas nas laterais da região focal.

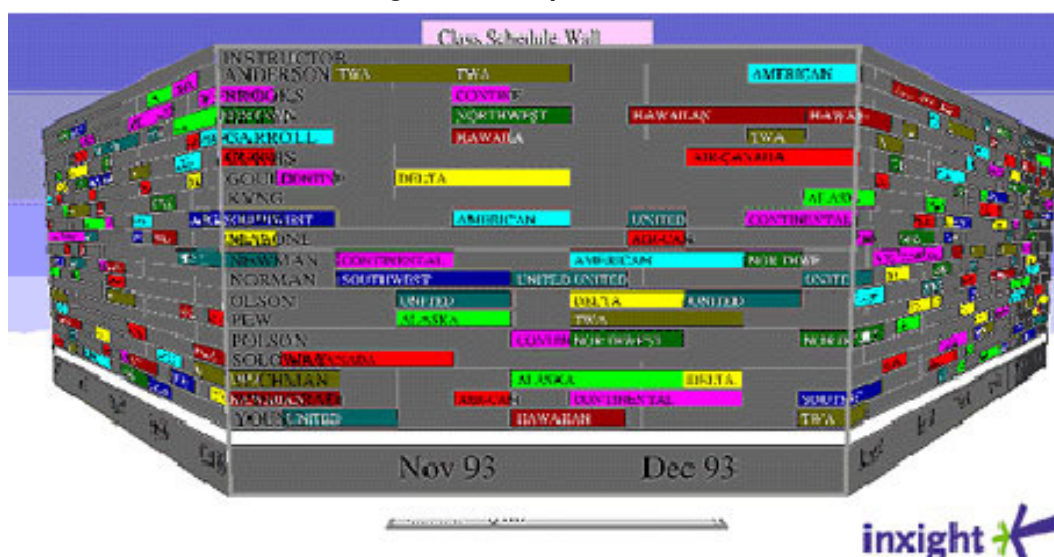
Figura 2 - Representação na técnica Bifocal Display.



Fonte: Freitas et. al.(2001, apud Xavier 2009, p.38)

A XEROX PARC se utilizou dos mesmos conceitos do Bifocal Display no Perspective Wall que apresenta o espaço informacional em uma parede com os muros laterais contendo informação de menor relevância, permitindo a visualização de muitas informações lineares em uma planilha na forma de retângulo horizontal, o qual é dobrado para trás nos limites à direita e à esquerda da região de interesse, fornecendo, assim, uma perspectiva 3D do resultado (VIEIRA, CORRÊA, 2010).

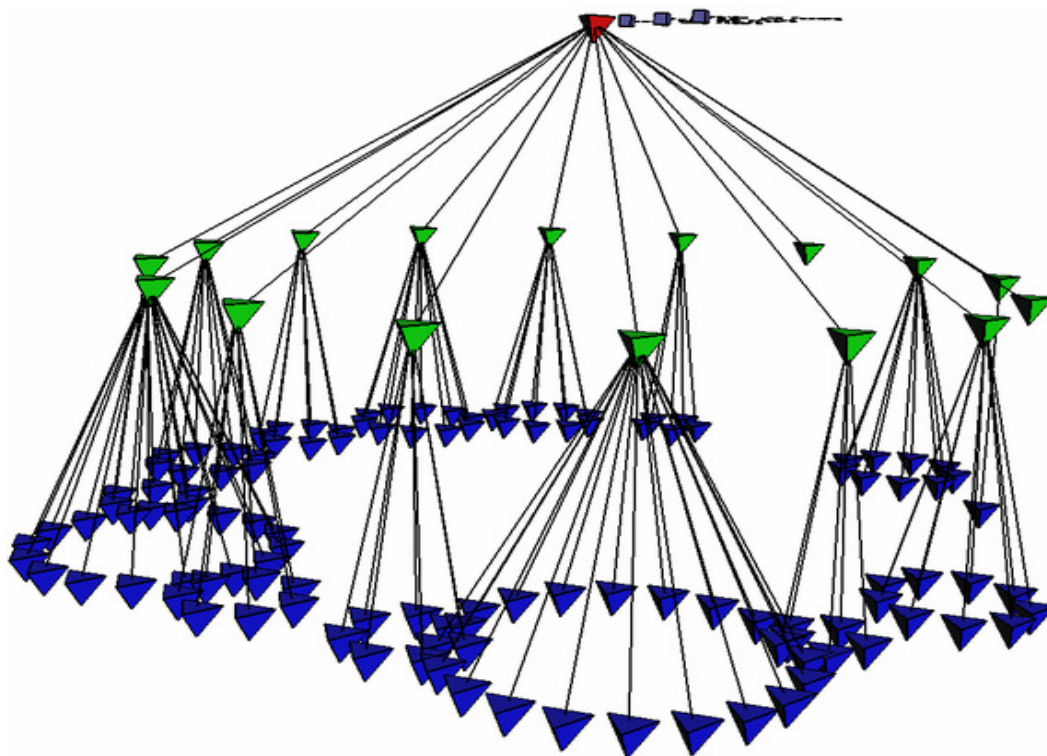
Figura 3 - Perspective Wall.



Fonte: SAP <http://www.sapdesignguild.org/goodies/controls/PerspectiveWall.htm>

A técnica Cone Tree representa em 3D uma informação hierárquica e permite a visualização da estrutura na íntegra sem necessidade de scrolling, permitindo ocultar e exibir os nodos durante a navegação, apresentando uma disposição espacial cônica. O problema dessa técnica é que as árvores podem se tornar confusas em curto espaço de tempo (XAVIER, 2009).

Figura 4 - Cone Tree.

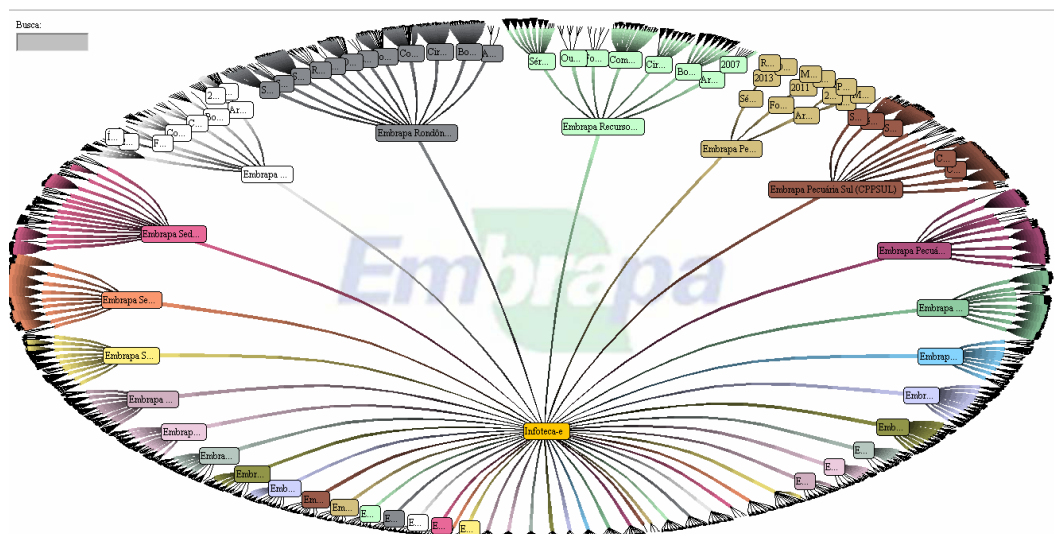


Fonte: <http://www-graphics.stanford.edu/papers/webviz/htmlnosplit>

A representação Hyperbolic Tree através de visualização radial em plano hiperbólico apresenta hierarquias em um aspecto de construção chamado de espinha de peixe, pela indicação de um nodo de interesse exibido no centro da interface, ao passo que o contexto é mantido no restante do diagrama. Ou seja, os nodos mais próximos do foco central são mais detalhados que os distantes. Os nodos são representados por retângulos dispostos de forma radial com foco no centro (XAVIER, 2009).

A representação de Hyperbolic Tree é utilizada pelo EMBRAPA, no serviço de informação tecnológica e agricultura, com o objetivo de reunir e permitir acesso a respeito das tecnologias produzidas pelo EMBRAPA. A representação permite ver a organização do acervo com relação ao local, temática e ano, tornando a navegação amigável (VIEIRA, CORRÊA, 2010).

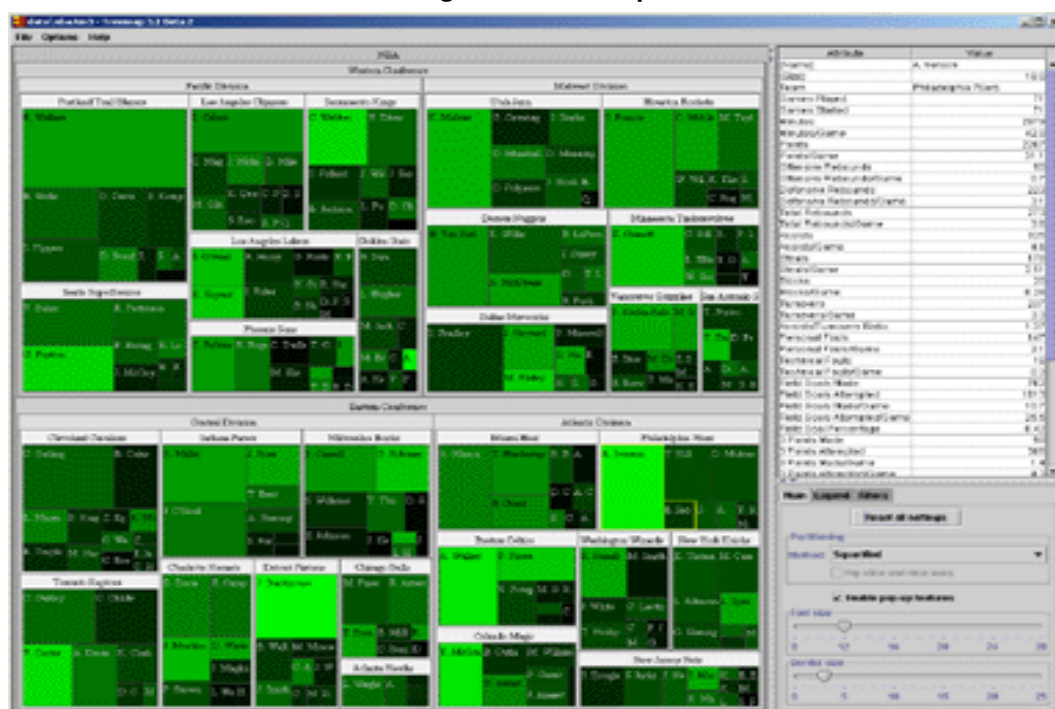
Figura 5 - Hyperbolic Tree.



Fonte: <http://www.infoteca.cnptia.embrapa.br/hypertree/hypertree.html>

Treemap é uma visualização com restrições de espaço de estruturas hierárquicas. É muito eficaz em mostrar atributos de nós folha utilizando um tamanho e um código de cores. Treemap permite aos usuários comparar os nós e sub-árvores, mesmo em profundidade variável na árvore, e ajudá-los a padrões e exceções pontuais. A estrutura hierárquica é apresentada utilizando-se de subdivisões sucessivas na tela, em que cada subdivisão representa um diretório que também possui outras subdivisões.

Figura 6 - TreeMap.



Fonte: <http://www.cs.umd.edu/hcil/treemap/>

O browsing gráfico é uma área em constante avanço, principalmente nos estudos com enfoque em representações de interfaces visuais para a recuperação de informação. Com foco na criação de interfaces ricas para auxiliar o usuário na compreensão e navegação de espaços de informações complexos (XAVIER, 2009).

4.4 RESUMO E DISCUSSÃO

Os sistemas de recuperação de informação são de grande complexidade, pois envolvem a busca de informações em base de dados ou coleções organizados previamente.

Ao se tratar de sistemas de recuperação de informação, pode-se destacar os modelos clássicos de recuperação (querying), tais como booleano, vetorial e probabilísticos, que apresentam estratégias que são relevantes para uma consulta. Os modelos clássicos consideram que cada documento é descrito por um conjunto de palavras chaves, denominado índices, ou ainda, termos de indexação.

A recuperação de informação não tem muito significado para o usuário sem a apresentação de uma interface amigável; desta forma, a visualização da informação tem como objetivo transformar o dado em algo com significado para o usuário, pois o fato de representar um dado abstrato de forma visual diminui a sobrecarga cognitiva. Entretanto, para uma efetividade na representação visual, o mapeamento deve preservar o dado, sendo aplicadas várias transformações no dado bruto transformando-o em uma visualização.

No capítulo 5 é apresentado o sistema Web MDDIR Tool que, dentre outras funções, possibilita a recuperação de informação que é apresentada de forma visual seguindo o conceito de mapeamento em árvore.

5 PROPOSTA DE METADADOS PARA DESCRIÇÃO DO ACESSO AOS RECURSOS INFORMACIONAIS DIGITAIS GERADOS DINAMICAMENTE

Castro e Santos (2010) afirmam que a nossa sociedade é marcada atualmente por constantes mudanças nos mais variados segmentos; uma delas de cunho tecnológico, no contexto que conhecemos hoje por sociedade da informação. Nos últimos tempos, houve um aumento desordenado e caótico na quantidade de informações produzidas e disponibilizadas em meio digital, requerendo uma mudança e um repensar nas formas de armazenamento, de representação, de descrição e de preservação dos recursos informacionais digitais. Aponta ainda que a Web atual, caracterizada como espaço aberto e de fácil socialização do conhecimento, vem enfrentando mudanças significativas em sua estrutura e na forma de representar e apresentar os recursos informacionais digitais. Isto exige e pressupõe um novo olhar de diversos profissionais, especialmente, os da área de Ciência da Informação (CI) e de Biblioteconomia, que possuem papel importante como agentes transformadores, no século XXI.

Nesse contexto e de acordo com a LAI, todas as informações constantes nos sítios oficiais do governo devem ser acessadas de forma fácil possibilitando a gravação de relatórios em diversos formatos eletrônicos, inclusive abertos e não proprietários, tais como planilhas e texto, de modo a facilitar a análise das informações; tornando possível o acesso automatizado por sistemas externos em formatos abertos, estruturados e legíveis por máquina e a divulgação em detalhes dos formatos utilizados para estruturação da informação.

Diante dessa necessidade e tentando supri-la, foram explorados os 39 sítios oficiais do Governo Federal que possuem status de ministérios, estando divididos em 5 órgãos, 24 ministérios e 10 secretarias. Com esta análise, foi possível verificar que nenhum dos órgãos citados acima apresentam as informações em formato de dado aberto e que tais informações são apresentadas como Recursos Informacionais Digitais gerados dinamicamente o que dificulta usos futuros.

Sendo assim, foi proposto um conjunto de metadados capaz de descrever tais recursos, e além disso foi desenvolvido um sistema capaz de persisti-los e recuperá-los.

Na seqüência será apresentada a proposta de metadados e no capítulo 6 o Sistema Web para descrição de Recursos Informacionais Digitais gerados dinamicamente.

5.1 PROPOSTA DE METADADOS

As propostas acima definidas são extremamente relevantes, porém, quando um cidadão entra em um dos sítios oficiais dos governos a realidade é outra, uma vez que o acesso às informações é complicado e o caminho é muito extenso para se chegar ao objetivo, conforme pode ser observado no quadro abaixo:

Quadro 3 - Procedimentos necessários para a recuperação de um determinado conjunto de dados

Passo	Atividade	Decisão	Observações
1	Acessar a URL do ministério		Pode ser feita de forma direta ou por meio de um sítio de busca
2	Identificar e clicar no ícone de Acesso à Informação (figura 7)		Todos os sítios oficiais do governo devem possuir este ícone
3	Acessar a página de Despesas	Decidir quais informações serão analisadas	Em cada ministério as informações são apresentadas de forma diferente.
4	Optar por informação sobre diárias e passagens		É requisitado o portal da Transparência Pública
5	Abertura do portal da Transparência Pública	Optar se será feita uma busca por Data, Servidor, Unidade Gestora ou Busca Avançada	
6	Acessar busca por data		É solicitada a informação de ano e mês
7	Informar ano e mês	Decidir entre Pesquisar ou refinar Busca	Caso a opção seja pesquisa, uma lista em HTML ³ será apresentada. Caso a opção seja refinar busca, será apresentada uma nova tela solicitando novamente Ano e Mês, além de Unidade Gestora e Servidor

³ HTML (abreviação para a expressão inglesa HyperText Markup Language, que significa Linguagem de Marcação de Hipertexto) é uma linguagem de marcação utilizada para produzir páginas na Web. Documentos HTML podem ser interpretados por navegadores. A tecnologia é fruto do "casamento" dos padrões HyTime e SGML.

8	Apresentação da tela de resultados	Decidir clicar em Período, Servidor, Unidade Gestora e Valor Total para ordenação em ordem alfabética	Na tela em HTML, são apresentados os resultados em forma de tabelas com as colunas: Período, Servidor, Unidade Gestora e Valor Total
9	Clicar no nome do Servidor desejado		São apresentados os dados do servidor selecionado
10	Abertura da tela com os resultados esperados		A URL desta página permite acesso direto a tais dados

Fonte: Elaborado pelo autor.

O quadro 3 apresenta o acesso feito ao sítio do Ministério da Educação, apontando quais os passos necessários para que a recuperação de um determinado conjunto de dados seja possível e, além disso, é identificado o número de decisões necessárias para a conclusão de tais atividades, a página a ser analisada é Despesas.

Figura 7 - Ícone de Acesso à Informação.



Fonte: http://portalsaude.saude.gov.br/portalsaude/arquivos/jpg/2012/Mai/02/icone_ai

A figura 7 apresenta o ícone que representa o acesso à informação, ou seja, ao encontrar este ícone em um sítio é capaz de obter informações a respeito do órgão responsável por ele.

Figura 8 - Página de despesas do Ministério da Educação.



Fonte:

http://portal.mec.gov.br/index.php?option=com_content&view=article&id=17656&Itemid=1168

A figura 8 apresenta a página de acesso às despesas do Ministério da Educação e após executar a análise do quadro 3, é possível observar que o acesso às informações é altamente complicado e não permite o acesso automatizado por sistemas externos devido ao formato apresentado, não sendo legível por máquina, o que depõe contra a LAI e as características de dados abertos.

Cada consulta executada apresenta os resultados de acordo com as decisões e parâmetros selecionados, o que gera um recurso informacional digital que é gerado em tempo de execução, ou seja, só existe este recurso enquanto o navegador está aberto. Caso o navegador seja fechado, o resultado da consulta se perde sendo necessário executar novamente todos os passos apresentados acima para acessar o mesmo recurso.

A cada nova consulta e dependendo dos parâmetros selecionados os resultados serão diferentes. Para contornar essa situação, é apresentada uma proposta de metadados para descrição do acesso aos Recursos Informacionais Digitais gerados dinamicamente; esses metadados descreverão o caminho do recurso informacional digital e permitirão o acesso direto a tal recurso.

Segundo Giannasi-Kaimen e Carelli (2007), um recurso informacional abrange textos, imagens e sons disponíveis para acesso do usuário. Para este trabalho, considera-se um recurso informacional digital apenas páginas *Web*.

A coleta de dados deste trabalho foi realizada nos 39 sítios oficiais do Governo Federal que possuem status de ministérios, estando divididos em 5 órgãos, 24 ministérios e 10 secretarias. O quadro 4 apresenta todos os sítios bem como seus endereços eletrônicos.

Quadro 4 - Sítios oficiais analisados

Entidade	Endereço Eletrônico
Advocacia-Geral da União	http://www.agu.gov.br
Banco Central do Brasil	http://www.bcb.gov.br
Casa Civil da Presidência da República	http://www.casacivil.gov.br
Controladoria Geral da União	http://www.cgu.gov.br
Gabinete de Segurança Institucional da Presidência da República	http://www.gsi.gov.br
Ministério da Agricultura, Pecuária e Abastecimento	http://www.agricultura.gov.br
Ministério da Ciência, Tecnologia e Inovação	http://mct.gov.br
Ministério da Cultura	http://www.cultura.gov.br
Ministério da Defesa	http://www.defesa.gov.br
Ministério da Educação	http://www.mec.gov.br
Ministério da Fazenda	http://www.fazenda.gov.br

Ministério da Integração Nacional	http://www.integracao.gov.br
Ministério da Justiça	http://portal.mj.gov.br
Ministério da Pesca e Aquicultura	http://www.mpa.gov.br
Ministério da Previdência Social	http://www.previdencia.gov.br
Ministério da Saúde	http://www.saude.gov.br
Ministério das Cidades	http://www.cidades.gov.br
Ministério das Comunicações	http://www.mc.gov.br
Ministério das Relações Exteriores	http://www.itamaraty.gov.br
Ministério de Minas e Energia	http://www.mme.gov.br
Ministério do Desenvolvimento Agrário	http://www.mda.gov.br
Ministério do Desenvolvimento Social e Combate à Fome	http://www.mds.gov.br
Ministério do Desenvolvimento, Indústria e Comércio Exterior	http://www.mdic.gov.br
Ministério do Esporte	http://www.esporte.gov.br
Ministério do Meio Ambiente	http://www.mma.gov.br
Ministério do Planejamento Orçamento e Gestão	http://www.planejamento.gov.br
Ministério do Trabalho e Emprego	http://www.mte.gov.br
Ministério do Turismo	http://www.turismo.gov.br
Ministério dos Transportes	http://www.transportes.gov.br
Secretaria da Micro e Pequena Empresa	
Secretaria de Assuntos Estratégicos da Presidência da República	http://www.sae.gov.br
Secretaria de Aviação Civil da Presidência da República	http://www.aviacaocivil.gov.br/
Secretaria de Comunicação Social da Presidência da República	http://www.secom.gov.br
Secretaria de Direitos Humanos da Presidência da República	http://www.direitoshumanos.gov.br
Secretaria de Políticas de Promoção da Igualdade Racial da Presidência da República	http://www.portaldainigualdade.gov.br
Secretaria de Políticas para as Mulheres da Presidência da República	http://www.spm.gov.br/
Secretaria de Portos da Presidência da República	http://www.portosdobrasil.gov.br
Secretaria de Relações Institucionais da Presidência da República	http://www.relacoesinstitucionais.gov.br
Secretaria-Geral da Presidência da República	http://www.secretariageral.gov.br

Fonte: Elaborado pelo autor.

Analizando cada um dos sítios acima é facilmente observado que a desestruturação das informações e a dificuldade de acessá-las é recorrente, o que faz com que o acesso se torne difícil para o cidadão. Ainda que a maioria das informações sejam geradas em tempo de execução da consulta, ou seja, a partir de parâmetros selecionados e informados, é gerada uma página Web dinâmica que só existe enquanto o navegador estiver aberto; ao fechá-lo estas informações se perdem e para acessá-las novamente é necessário percorrer exatamente o mesmo caminho realizado pela primeira vez. Todavia, nem sempre é possível percorrer o mesmo caminho, seja por falta de conhecimento ou por esquecimento do caminho percorrido anteriormente.

Diante disso, torna-se necessária a criação de formas que façam com que o usuário final possa localizar o ponto exato em que é gerado o resultado de sua

consulta sem passar pela estrutura (caminho) que é hierarquizado, propondo metadados capazes de descrever os Recursos Informativos Digitais e amenizar tal situação.

Para o desenvolvimento desta proposta de metadados para descrição de Recursos Informativos Digitais gerados dinamicamente, foram estudados os padrões de metadados Dublin Core e e-PMG. Para tanto, foi feita uma análise detalhada de cada um desses padrões de metadados estudados, a fim de identificar quais elementos de cada um poderiam ser utilizados na descrição de Recursos Informativos Digitais gerados dinamicamente. Essa decisão de quais elementos seriam úteis foi baseada nas características dos Recursos Informativos Digitais gerados dinamicamente e quais metadados seriam importantes para descrevê-los. Para fazer essa análise, foi criada uma legenda para informar a indicação deste elemento dentro da formulação da proposta do novo padrão de metadados:

- AT - Atende totalmente: se um elemento de metadados possui esta indicação, significa que pertencerá à proposta do novo padrão de metadados;
- AP - Atende parcialmente: se um elemento de metadados possui esta indicação, significa que pertencerá à proposta do novo padrão de metadados com adaptações à nova realidade;
- NA - Não atende: se um elemento de metadados possui esta indicação, significa que não possui nenhuma relação com a proposta do novo padrão de metadados.

Para que a análise de cada um dos padrões de metadados estudados fosse efetiva, foi feita a descrição do sítio oficial do Ministério do Trabalho e Emprego nos padrões de metadados Dublin Core e e-PMG, apontando a indicação de cada um de seus elementos na composição do conjunto de metadados proposto neste trabalho.

Quadro 5 - Descrição do acesso ao sítio do Ministério do Trabalho e Emprego no Padrão de Metadados Dublin Core

Elemento	Conteúdo	Indicação
Title	Ministério do Trabalho e Emprego	AT
Creator		AT
Subject	Trabalho e Emprego	AT
Description	O portal do Ministério do Trabalho e Emprego apresenta informações institucionais, conselhos e comissões,	AT

	emprego e renda, inspeção do trabalho, economia solidária, relações de trabalho, dados e estatísticas, entre outras.	
Publisher		NA
Contributor		AT
Date	2013-02-10	AT
Type	Interactive Resource	NA
Format	Text/html	AT
Identifier	http://portal.mte.gov.br/portal-mte/	AT
Source	http://portal.mte.gov.br/portal-mte/	NA
Language	pt-br	NA
Relation		NA
Coverage	1997-2008	NA
Rights	MTE	NA

Fonte: Elaborado pelo autor.

O quadro 5 apresenta a descrição do sítio do Ministério do Trabalho e Emprego no Padrão de Metadados Dublin Core, podendo observar o grau de aderência desse padrão com o conjunto de metadados proposto neste trabalho.

Quadro 6 - Descrição do acesso ao sítio do Ministério do Trabalho e Emprego no Padrão de Metadados e-PMG

Entidade	Conteúdo	Indicação
Mandate		NA
Disposal		NA
Addressee		NA
Location		AP
Preservation		NA
Title	Ministério do Trabalho e Emprego	AT
Creator		AT
Subject	Trabalho e Emprego	AT
Description	O portal do Ministério do Trabalho e Emprego apresenta informações institucionais, conselhos e comissões, emprego e renda, inspeção do trabalho, economia solidária, relações de trabalho, dados e estatísticas, entre outras.	AT
Publisher		NA
Contributor		AT
Date	2013-02-10	AT
Type	Interactive Resource	NA
Format	Text/html	AT
Identifier	http://portal.mte.gov.br/portal-mte/	AT
Source	http://portal.mte.gov.br/portal-mte/	NA
Language	pt-br	NA
Relation		NA
Coverage	1997-2008	NA
Rights	MTE	NA

Fonte: Elaborado pelo autor.

O quadro 6 apresenta a descrição do sítio do Ministério do Trabalho e Emprego no Padrão de Metadados e-PMG, podendo observar o grau de aderência desse padrão com o conjunto de metadados proposto neste trabalho.

Após realizar um levantamento de exemplos de acesso a dados em sítios oficiais do governo federal e analisar o grau de aderência de modelos de metadados na descrição de Recursos Informativos Digitais gerados dinamicamente, foi possível chegar à proposta do novo padrão de metadados cujo foco é descrever o acesso aos Recursos Informativos Digitais gerados dinamicamente. Para criação desse novo padrão de metadados foram tomados como base os padrões Dublin Core e e-PMG.

Como o principal objetivo desta nova proposta de metadados é descrever o acesso aos Recursos Informativos Digitais gerados dinamicamente, é necessário definir dois tipos de recursos informativos digitais:

- **Página Menu:** define-se como a página Web onde constam os parâmetros que serão selecionados para formar o recurso informativo digital, é uma página Web que pode ser acessada diretamente via endereço eletrônico;
- **Página Resultado:** define-se como a página Web formada pela seleção dos parâmetros na Página Menu, é uma página Web que geralmente não pode ser acessada via endereço eletrônico direto, seja pela complexidade do endereço seja pela característica desta. Esta página só existe enquanto o navegador estiver ativo, ao fechá-lo esta referência é perdida.

De acordo com a explicação feita acima, tornou-se necessário definir elementos de metadados específicos para as Páginas Menu e Páginas Resultado, abaixo tais elementos podem ser analisados:

Página Menu:

- **Elemento:** URL.

Identificador: URL.

Definição: endereço eletrônico de acesso à página menu.

Comentário: Recomenda-se a identificação do recurso por meio da cadeia de sistemas de identificação "Uniform Resource Identifier" (URI) (incluindo o "Uniform Resource Locator" (URL)).

Exemplo: <http://www.portaldatransparencia.gov.br/>

- Elemento: Parâmetro.
Identificador: Parameter.
Definição: Parâmetros utilizados para compor o resultado das consultas.
Comentário: O parâmetro define a formação da página resultado.
Exemplo: Exercício e Programa.
- Elemento: Título.
Identificador: Title.
Definição: Nome dado à página menu.
Comentário: Um Título será o nome pelo qual a página menu é formalmente conhecida.
Exemplo: Transferências de Recursos.
- Elemento: Assunto.
Identificador: Subject.
Definição: Tópicos do conteúdo do recurso.
Comentário: Um Assunto deverá ser expresso por palavras-chave ou frases que descrevam o conteúdo do recurso.
Exemplo: metadados, padrões de metadados, sítios oficiais.
- Elemento: Descrição.
Identificador: Description.
Definição: Uma descrição do conteúdo do recurso.
Comentário: Descrições podem incluir, sem estarem limitadas a tal: um resumo, um índice, uma referência a uma representação gráfica do conteúdo ou uma descrição textual.
Exemplo: Esta pesquisa permite acompanhar os recursos públicos transferidos pela União ao exterior, a estados e municípios brasileiros, ao Distrito Federal, a instituições privadas e aos cidadãos, bem como conferir os gastos diretos do Poder Executivo Federal. A atualização dos dados é mensal.
- Elemento: Autor.
Identificador: Author.
Definição: Responsável pela descrição do recurso.
Comentário: Um autor pode ser uma pessoa, departamento,

ministério, órgão etc., responsável pela descrição do recurso.

Exemplo: Controladoria Geral da União.

- Elemento: Data.

Identificador: Date.

Definição: Data de criação do recurso.

Comentário: Uma Data deve ser associada à criação ou disponibilidade do recurso. Recomenda-se para codificação de valores de datas um perfil da norma ISO 8601 [W3CDTF], segundo o formato AAAA-MM-DD.

Exemplo: 2013-07-27.

- Elemento: Colaborador

Identificador: Contributor.

Definição: Uma pessoa responsável pela descrição de um recurso.

Comentário: O colaborador é uma pessoa que contribui para a descrição de recursos utilizando este conjunto de metadados.

Exemplo: Lisandro Rogério Modesto.

Página Resultado:

- Elemento: URL

Identificador: URL.

Definição: endereço eletrônico da página resultado.

Comentário: Recomenda-se a identificação do recurso por meio da cadeia de sistemas de identificação "Uniform Resource Identifier" (URI) (incluindo o "Uniform Resource Locator" (URL)).

Exemplo: <http://www.portaldatransparencia.gov.br/PortalTransparenciaPesquisaAcao.asp?Exercicio=2013>

- Elemento: Campo.

Identificador: Field.

Definição: Campos que irão compor uma tabela de resultados na Página Resultado.

Comentário: Este campos podem ser do mais variados tipos, não há uma definição, pois para cada recurso novos campos de resultados pode ser apresentados.

Exemplo: Linguagem Cidadã, Ação Governamental, Função e Total no Ano (R\$).

- Elemento: Formato

Identificador: Format.

Definição: Formas de representação dos recursos informacionais gerados dinamicamente.

Comentário: Este elemento depende diretamente do tipo de saída apresentada nas páginas resultados.

Exemplo: HMTL, XML, PDF etc.

- Elemento: Título

Identificador: Title.

Definição: Nome dado à página resultado.

Comentário: Um Título será o nome pelo qual a página resultado é formalmente conhecida.

Exemplo: Transferências de Recursos por Ação de Governo.

- Elemento: Assunto

Identificador: Subject.

Definição: Tópicos do conteúdo do recurso.

Comentário: Um Assunto deverá ser expresso por palavras-chave ou frases que descrevam o conteúdo do recurso.

Exemplo: Recursos, Poder Executivo, Federal.

- Elemento: Descrição

Identificador: Description.

Definição: Uma descrição do conteúdo do recurso.

Comentário: Descrições podem incluir, sem estarem limitadas a tal: um resumo, um índice, uma referência a uma representação gráfica do conteúdo ou uma descrição textual.

Exemplo: Total destinado pelo Governo Federal em âmbito nacional em 2013.

- Elemento: Autor

Identificador: Author.

Definição: Responsável pela descrição do recurso.

Comentário: Um autor pode ser uma pessoa, departamento, ministério, órgão etc., responsável pela descrição do recurso.

Exemplo: Controladoria Geral da União.

- Elemento: Data

Identificador: Date.

Definição: Data de criação do recurso.

Comentário: Uma Data deve ser associada à criação ou disponibilidade do recurso. Recomenda-se para codificação de valores de datas um perfil da norma ISO 8601 [W3CDTF], segundo o formato AAAA-MM-DD.

Exemplo: 2013-07-27.

- Elemento: Colaborador

Identificador: Contributor.

Definição: Uma pessoa responsável pela descrição de um recurso.

Comentário: O colaborador é um pessoa que contribui para a descrição de recursos utilizando este conjunto de metadados.

Exemplo: Lisandro Rogério Modesto.

Para exemplificar o uso da proposta de metadados para descrição de recursos informacionais digitais, pode-se analisar o quadro abaixo:

Quadro 7 - Representação de um recurso informacional digital gerado dinamicamente

Elemento	Descrição
PM.URL	http://www.portaldatransparencia.gov.br/
PM.Autor	Controladoria Geral da União
PM.Título	Transferências de Recursos
PM.Assunto	Transferências de recursos do Poder Executivo Federal
PM.Descrição	Esta pesquisa permite acompanhar os recursos públicos transferidos pela União ao exterior, a estados e municípios brasileiros, ao Distrito Federal, a instituições privadas e aos cidadãos, bem como conferir os gastos diretos do Poder Executivo Federal. A atualização dos dados é mensal.
PM.Data	2013-07-24
PM.Colaborador	Lisandro Rogério Modesto
PM.Parâmetro	Exercício; 2013
PM.Parâmetro	Programa
PR.URL	http://www.portaldatransparencia.gov.br/PortalTransparenciaTRProgramaPesquisaPrograma.asp?Exercicio=2013
PR.Autor	Controladoria Geral da União
PR.Título	Transferências de Recursos por Programa
PR.Assunto	Recursos do Poder Executivo Federal
PR.Descrição	Total destinado pelo Governo Federal em âmbito nacional em 2013

PR.Data	2013-07-24
PR.Colaborador	Lisandro Rogério Modesto
PR.Campo	Total no Ano (R\$)
PR.Campo	Programa
PR.Formato	HTML

Fonte: Elaborado pelo autor.

O quadro 7 apresenta os elementos de um recurso informacional digital gerado dinamicamente a partir do sítio da Controladoria Geral da União. É importante salientar que os metadados referentes à Página Menu serão identificados por "PM" e os metadados referentes às Páginas Resultados serão identificados por "PR".

5.2 RESUMO E DISCUSSÃO

A LAI determina que todas as informações constantes no sítios oficiais do governo devem possuir acesso fácil, possibilitando a gravação de relatórios em diversos formatos eletrônicos, inclusive abertos e não proprietários, tais como planilhas e texto, de modo a facilitar a análise das informações; viabilizando o acesso automatizado por sistemas externos em formatos abertos, estruturados e legíveis por máquina e a divulgação em detalhes dos formatos utilizados para estruturação da informação.

Atualmente, o acesso às informações é altamente complicado e não permite o acesso automatizado por sistemas externos devido ao formato apresentado, não sendo legível por máquina, o que depõe contra a lei de acesso à informação e às características de dados abertos.

Cada consulta executada em um sítio oficial apresenta os resultados de acordo com as decisões e filtros selecionados, o que gera um recurso informacional digital que é gerado em tempo de execução, ou seja, só existe este recurso enquanto o navegador está aberto, sendo perdido caso o navegador seja fechado, tendo que executar novamente todos os procedimentos apresentados no quadro 3 para acessar o mesmo recurso.

É importante salientar que a cada nova consulta e dependendo do caminho percorrido os resultados podem e serão diferentes, e para contornar essa situação, está sendo proposto um conjunto de metadados capaz de descrever os Recursos Informacionais Digitais gerados dinamicamente.

Este capítulo apresenta e exemplifica a proposta de metadados proposto por este trabalho, que possibilitarão localizar o ponto exato em que é gerado um resultado de uma determinada consulta sem passar pela estrutura (caminho) que hoje é hierarquizado.

No capítulo seguinte, será apresentada a ferramenta desenvolvida com a finalidade de automatizar o processo de descrição e recuperação dos Recursos Informacionais Digitais gerados dinamicamente, e também apresentará o sistema web que persisti tais recursos digitais.

6 SISTEMA WEB PARA DESCRIÇÃO DE RECURSOS INFORMACIONAIS DIGITAIS GERADOS DINAMICAMENTE

Neste capítulo, serão apresentadas as funcionalidades e particularidades do sistema MDDIR (Metadata for Describing Digital Information Resources - MDDIR - TOOL) para descrição do acesso aos de Recursos Informativos Digitais gerados dinamicamente. Este sistema foi desenvolvido utilizando as seguintes tecnologias:

- Linguagem de programação: PHP.
- Banco de Dados: MySQL.
- Sistema Operacional do servidor: Linux.
- Servidor Web: Apache.
- Software para gerar objeto material das páginas Web: HTTRACK Website Copier.

Por ser um sistema web, pode ser aberto em qualquer navegador e em qualquer sistema operacional, sem nenhum prejuízo de informações para o usuário.

6.1 MODELO DE DADOS LÓGICO DO SISTEMA WEB MDDIR TOOL

Segundo Pressman (2002, p. 296), a modelagem de dados corresponde a um conjunto de questões específicas que são relevantes para qualquer aplicação de processamento de dados.

Um modelo de dados lógico representa as regras de negócio do contexto em questão e não é um banco de dados especificamente, ou seja, independe do modelo físico. Este modelo deve ser independente da tecnologia aplicada para o desenvolvimento devido à dinamicidade da área tecnológica.

Os atributos de cada objeto de dados, que figuram no modelo de dados lógico, possui tipos de dados específicos que podem ser verificados no próprio modelo. Isso significa que para cada atributo existe um tipo de dado permitindo o seu preenchimento.

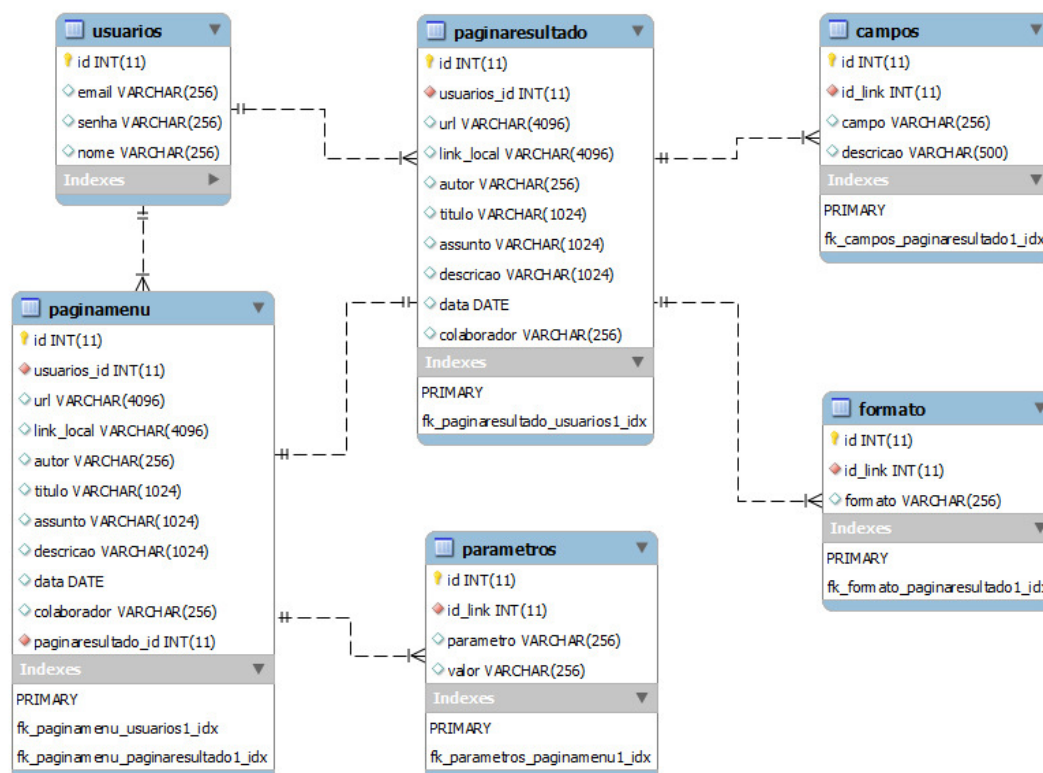
Mayer (2007) afirma que o modelo de dados lógico é o retrato de todas as informações necessárias para a realização dos negócios. Ele é representado de

diversas maneiras diferentes, utilizando metodologias e ferramentas diferentes para implementações diferenciadas. O modelo entidade-relacionamento (MER) é uma técnica de modelagem amplamente utilizada pelos administradores de dados. Os componentes do MER são representados pelas entidades, os relacionamentos e os atributos. Cada entidade representa um conjunto de pessoas, coisas ou conceitos sobre os quais o domínio precisa de informações. Cada relacionamento representa a associação entre duas entidades. Cada atributo é a característica ou parte da informação de uma entidade. Um nome é uma definição textual que descreve cada um desses componentes. Estes nomes e definições provêm da documentação das regras de negócio e informações as quais são armazenadas e mantidas num repositório de dados garantindo assim a padronização conceitual dos dados.

Na figura 9, apresenta-se o modelo entidade-relacionamento (MER) do sistema Web MDDIR Tool. Este modelo é composto pelas seguintes entidades: Usuários, Parâmetros, Campos, Formatos, Paginamenu e Paginaresultado.

A partir deste modelo serão construídas as regras para a modelagem e o banco de dados utilizados pelo sistema MDDIR Tool.

Figura 9 - Modelo entidade-relacionamento (MER) do Sistemas Web MDDIR Tool.



Fonte: Elaborado pelo autor.

Por meio do MER, é possível verificar as relações existentes entre cada um dos elementos que compõe o sistema WEB MDDIR - Tool, tornando mais fácil seu entendimento e a análise do fluxo de informação.

A próxima seção apresentará as funcionalidades do sistemas bem como suas particularidades.

6.2 FUNCIONALIDADES DO SISTEMA WEB MDDIR TOOL

Sistema WEB pode ser definido, de forma geral, como sistemas de informática projetados para utilização através de um navegador, na internet ou em redes privadas (Intranet). Trata-se de um conjunto de programas que é executado em um servidor HTTP (Web Host). O desenvolvimento da tecnologia web está relacionado, entre outros fatores, à necessidade de simplificar a atualização e manutenção mantendo o código-fonte em um mesmo local, de onde ele é acessado pelos diferentes usuários. Pode-se ainda definir um sistema web como uma aplicação de software que utiliza a web, através de um browser, como ambiente de execução.

O sistema MDDIR Tool é caracterizado por funcionalidades que possibilitam o processo de descrição de Recursos Informativos Digitais gerados dinamicamente; para tanto, é necessário definir dois tipos destes recursos:

- **Página Menu:** define-se como a página Web onde constam os parâmetros que serão selecionados para formar o recurso informativo digital, é uma página Web que pode ser acessada diretamente via endereço eletrônico;
- **Página Resultado:** define-se como a página Web formada pela seleção dos parâmetros na Página Menu, é uma página Web que geralmente não pode ser acessada via endereço eletrônico direto, seja pela complexidade do endereço, seja pela característica desta. Esta página só existe enquanto o navegador estiver ativo, ao fechá-lo esta referência é perdida.

Ao iniciar o sistema, é solicitado o preenchimento dos dados do usuário, caso ainda não esteja cadastrado, é possível efetuar o cadastro, conforme pode ser verificado na figura 10:

Figura 10 - Tela inicial do sistema MDDIR Tool

Fonte: Elaborado pelo autor.

Na seqüência, o usuário poderá optar por uma das funcionalidades do sistema, conforme pode ser visto na figura 11, entre elas:

- Cadastro: Esta funcionalidade permite que o usuário efetue o cadastro do recurso informacional digital em questão. É importante salientar que uma Página Resultado só poderá ser descrita depois da descrição de uma Página Menu. Para cada Página Menu é possível existirem várias Páginas Resultado;
- Consulta por parâmetro: Esta funcionalidade permite que o usuário selecione os valores de Busca (Descrição, Campos e Parâmetros), Metadados (de acordo com a seleção feita em Busca), Parâmetros (de acordo com a seleção feita em Busca e Metadados) e, por fim, Valor (onde o usuário poderá definir um valor específico de acordo com a seleção feita em Busca, Metadados e Parâmetros);
- Consulta por Palavras-Chave: Esta funcionalidade permite que o usuário digite dados presentes em um dos recursos cadastrados no banco de dados do sistema, a fim de localizar um recurso específico.

Figura 11 - Tela de menu do sistema MDDIR Tool



Fonte: Elaborado pelo autor.

Ao clicar em Cadastro, o usuário poderá efetuar o lançamento dos metadados que descreverão os Recursos Informativos Digitais da Página Menu, conforme pode ser visto na figura 12.

Ao clicar no campo Parâmetros, é permitida a criação de quantos parâmetros forem necessários, não havendo limite para este elemento. Após serem preenchidos os elementos da descrição da Página Menu, o usuário poderá optar por FINALIZAR a descrição não possibilitando a descrição de uma página resultado, ou então poderá optar por prosseguir com a descrição de uma página resultado. Para isso, basta optar por PRÓXIMO, e só será permitido dar sequência caso o usuário preencha os elementos URL e Título.

Figura 12 - Tela de cadastro de Página Menu do sistema MDDIR Tool

Captura de URL - http://lo x

localhost/mddirtool4/paginamenu.php?id=3

TOOL
mddir

[Sair](#)

[Início](#)

[Página Menu](#)

URL:

+ Parâmetros

Título:

Assunto:

Descrição:

Autor:

Data: 19/11/2013

Colaborador:

Fonte: Elaborado pelo autor.

Caso o usuário tenha prosseguido para a descrição de uma Página Resultado, será apresentada a tela de descrição deste recurso, conforme pode ser visto na Figura 13.

Ao clicar nos campos Formato e Campos, é permitida a criação de quantos campos e formatos forem necessários, não havendo limite para estes elementos. Ao clicar em Campos é possível informar os nome dos campos bem como a descrição dos campos que descrevem qual informação compõem a coluna da tabela. Após serem preenchidos os elementos da descrição da Página Resultado, o usuário poderá optar por FINALIZAR a descrição não possibilitando a descrição de uma nova página resultado, ou então poderá optar por prosseguir com a descrição de uma nova página resultado. Para isso, bastar optar por ADICIONAR, e só será

permitido dar seqüência caso o usuário preencha os elementos URL e Título.

Figura 13 - Tela de cadastro de Página Resultado do sistema MDDIR Tool

Captura de URL - http://lo x

localhost/mddirtool4/paginaresultado.php?id_paginamenu=234

TOOL
mddir

[Sair](#)

[Inicio](#)

[Página Resultado](#)

URL:

[+ Campos](#)

[+ Formatos](#)

Titulo:

Assunto:

Descrição:

Autor:

Data: 19/11/2013

Colaborador:

MDDIR Tool (c) Todos os Direitos Reservados - 2013

Fonte: Elaborado pelo autor.

Ao clicar em Consulta por parâmetro, a tela de seleção de parâmetros será apresentada, conforme pode ser visto na figura 14. Esta consulta acontece de forma estruturada uma vez que os valores em cada caixa de seleção vão sendo montados de acordo com a seleção. A primeira caixa de seleção oferece a opção de consultar por metadados (descrição), Campos ou Parâmetros. Caso seja selecionada descrição na primeira caixa de seleção, na segunda caixa aparecerão todos os elementos que compõem o conjunto de metadados proposto neste trabalho, com exceção de Campos e Parâmetros. Caso seja selecionado Campos, na segunda caixa de seleção aparecerão todos os campos cadastrados pelos usuários na descrição de uma Página Resultado. Ao selecionar Parâmetros, na segunda caixa

de seleção aparecerão todos os Parâmetros cadastrados pelos usuários na descrição de uma Página Menu. A terceira caixa de seleção apresenta com parâmetros de consulta podendo ser Antes de, Contém, Depois de, Diferente, Igual e Não contém, sendo apresentado de acordo com a segunda caixa de seleção (Metadados). E, por fim, o campo de preenchimento Valor, onde pode ser colocado um valor específico que se conheça ou se queira consultar. Ao parametrizar toda a consulta basta clicar em Buscar para se obter os resultados.

Figura 14 - Tela de consulta por parâmetro do sistema MDDIR Tool

Fonte: Elaborado pelo autor.

Ao clicar em Consulta por Palavras-chave, é apresentada a tela onde se pode digitar o(s) parâmetro(s) de consulta desejado(s), conforme pode ser visto na figura 15. No caso de se efetuar a consulta por mais de um parâmetro, é possível utilizar os conectores lógicos "e" e "ou".

Para a lógica, um conector lógico é um símbolo ou palavra usada para conectar duas ou mais palavras ou sentenças de uma maneira gramaticalmente válida. Os conectores lógicos mais comuns são os binários, que juntam duas sentenças, que podem ser consideradas os operandos da função. Várias palavras e pares de palavras expressam conectores lógicos e algumas delas são sinônimos, no

caso da ferramenta, pode-se selecionar um dos conectores abaixo:

- e - conjunção: com este conector uma consulta é feita buscando a ocorrência de dois ou mais termos. Exemplo: Programa de Governo "e" Programa Saúde da Família;
- "ou" - disjunção: com este conector uma consulta é feita buscando a ocorrência de um termos informados. Exemplo: Programa de Governo "ou" Programa Saúde da Família.

Figura 15 - Tela de consulta por Palavras-chave do sistema MDDIR Tool

http://localhost/mddirtc x

localhost/mddirttool4/loogle.php

TOOL
mddir

[Sair](#)

[Inicio](#)

[Executar busca utilizando: ☐ ou ☐ e]

Digite um ou mais termos, separados por ";" (ponto-e-virgula)
Ex: idade;sexo;estado

MDDIR Tool (c) Todos os Direitos Reservados - 2013

Fonte: Elaborado pelo autor.

Tanto na consulta por parâmetros ou palavras-chaves, a página de resultado da consulta pode ser observada na Figura 18. Após efetuar uma consulta, pode-se observar os resultados de forma gráfica por meio dos ícones apresentados abaixo.

Figura 16 - Ícone para representar uma Página Menu



Fonte: Elaborado pelo autor.

A figura 16 informa, graficamente, se o resultado da consulta é uma Página Menu.

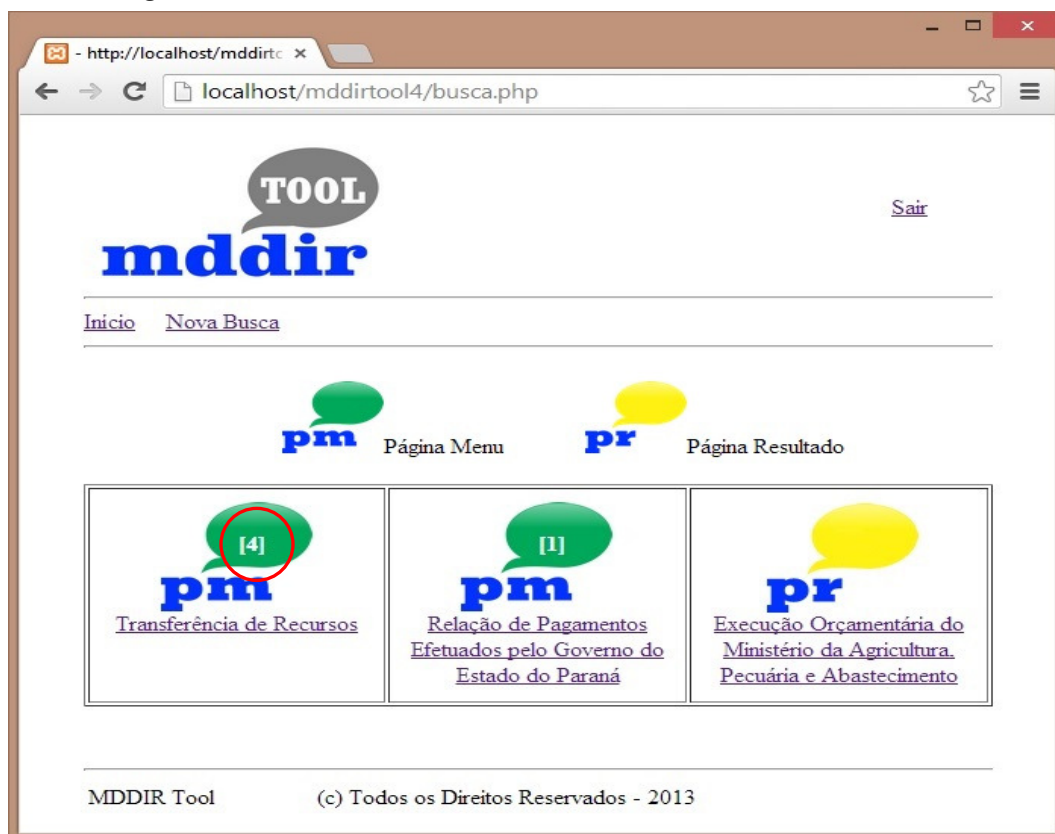
Figura 17 - Ícone para representar uma Página Resultado



Fonte: Elaborado pelo autor.

A figura 17 informa, graficamente, se o resultado da consulta é uma Página Resultado.

Figura 18 - Tela de resultado de consultas do sistema MDDIR Tool



Fonte: Elaborado pelo autor.

A figura 18 apresenta o resultado de uma consulta feita no sistema; neste resultado, é possível observar 3 resultados, sendo duas Páginas Menu e uma Página Resultado. É possível observar que sobre o ícone das Páginas Menu aparece um número (em destaque) que representa o número de Páginas Resultado relacionadas. No caso da Página Resultado, este número não aparece, pois este tipo de página não possui nenhuma outra relacionada a ela.

Ainda na tela de resultado de consulta do sistema, tela de Detalhes, há a possibilidade de ver graficamente a Página Menu com suas respectivas Páginas Resultados, isso pode ser observado na figura 19.

Figura 19 - Tela de apresentação da Página Menu e Páginas Resultado



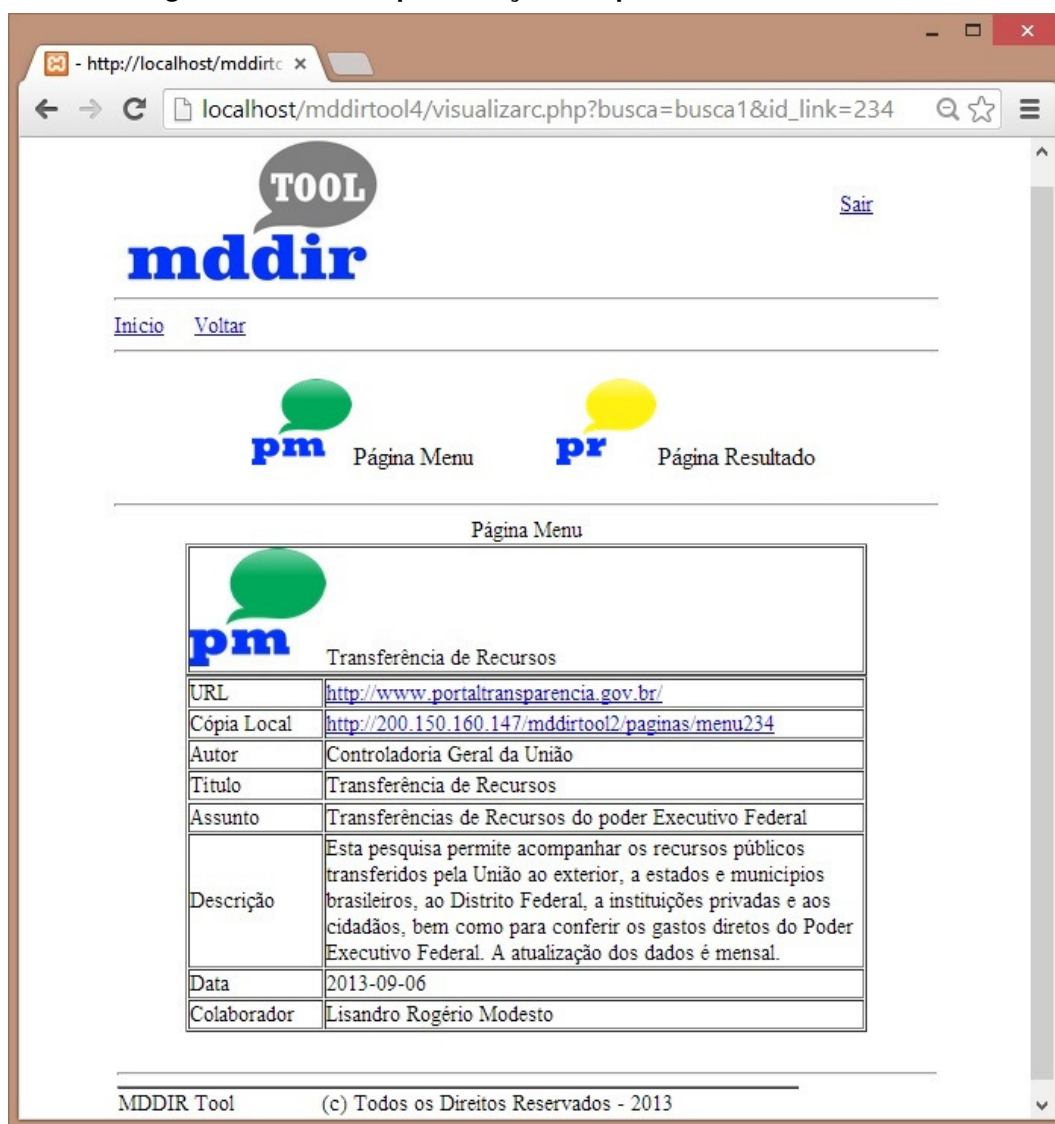
Fonte: Elaborado pelo autor.

Por meio da página de detalhes, é possível adicionar novas Páginas Resultado conforme pode ser observado na figura 19, lembrando que os dados da

Página Menu não sofrerão alterações.

Para se conhecer as informações de cada uma dessas páginas, Menu e Resultado, apresentadas na figura 19, basta clicar sobre o título da página. Os detalhes de cada recurso podem ser observados na figura 20.

Figura 20 - Tela de apresentação completa do recurso descrito

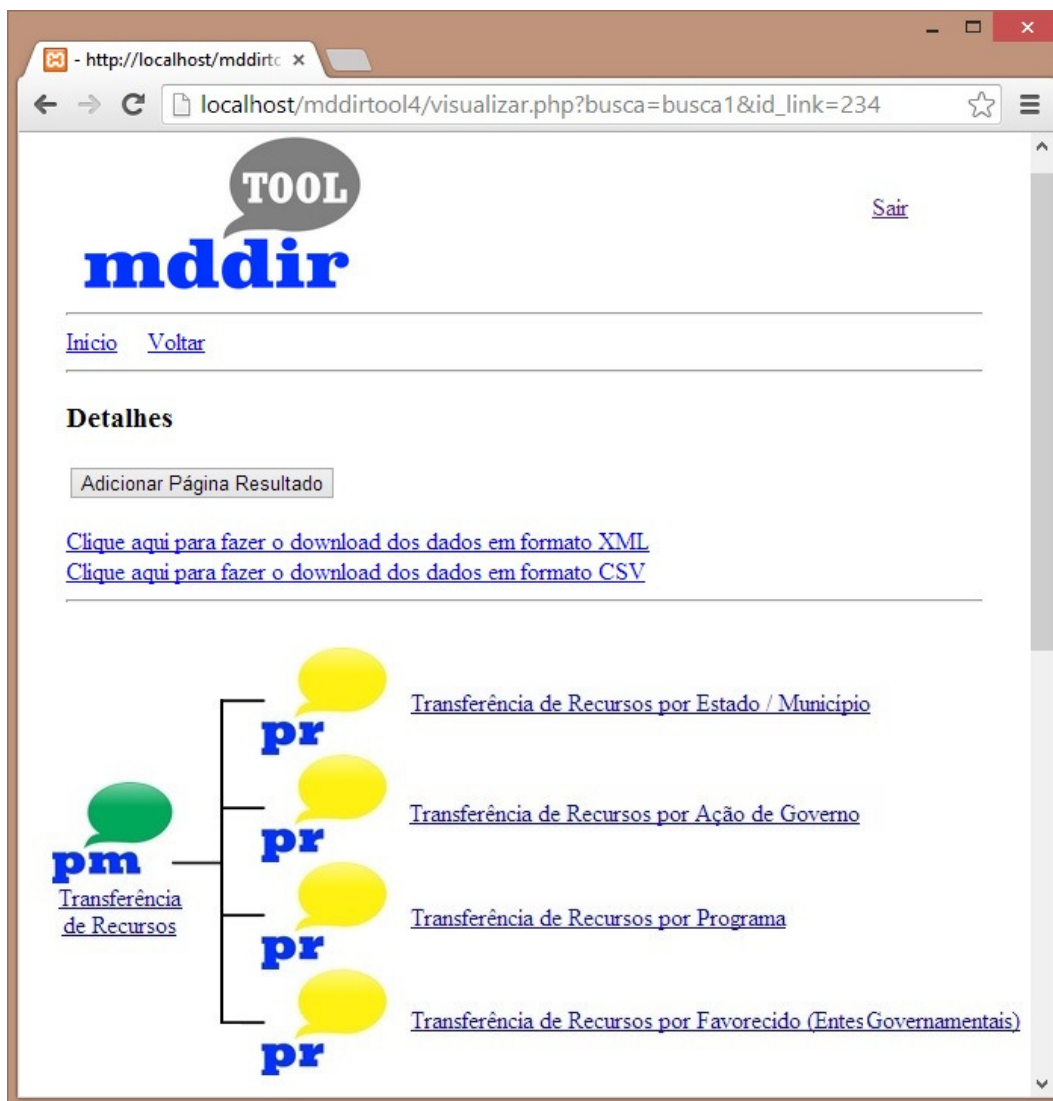


Fonte: Elaborado pelo autor.

Como principal particularidade do sistema, é importante salientar que todos os Recursos Informativos Digitais descritos são persistidos e ficam armazenados em disco, o que antes dessa proposta não acontecia devido à dinamicidade dos recursos informativos que só existiam em tempo de execução, ou seja, enquanto o navegador estivesse aberto. Assim, caso o Recurso Informativo Digital deixe de existir ou não possa mais ser localizado, basta efetuar uma consulta e clicar no link

da Cópia Local para que a mesma seja apresentada, conforme pode ser verificado na tela de detalhes da Figura 21.

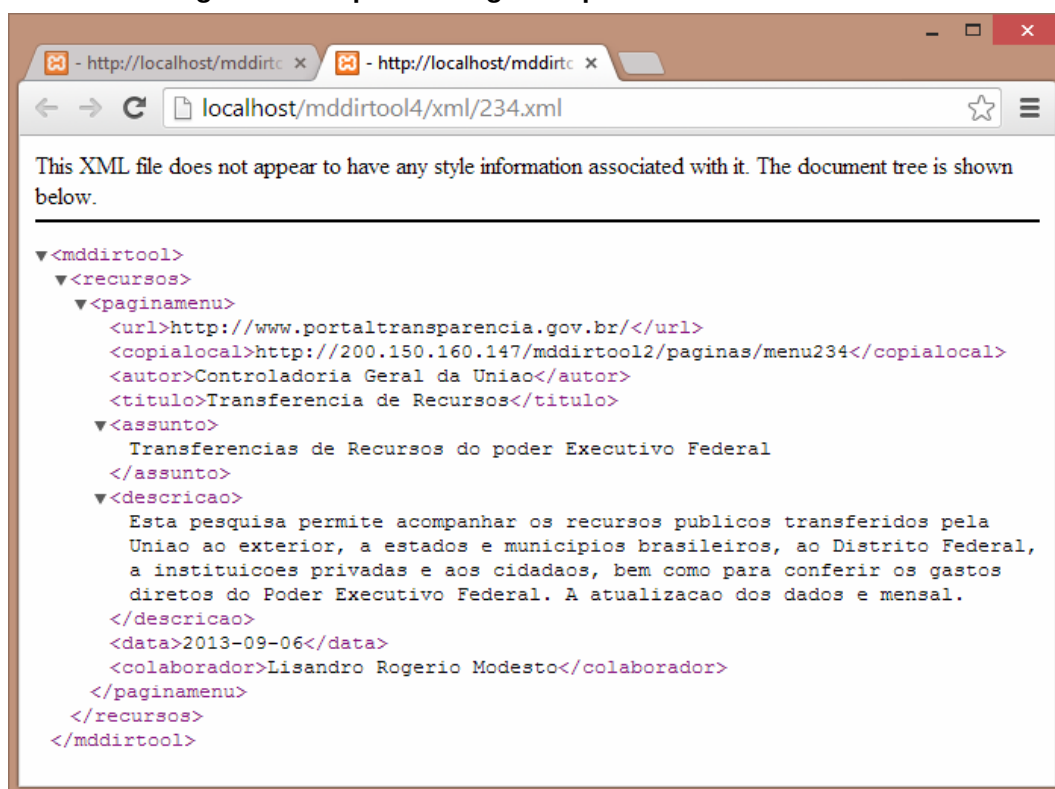
Figura 21 - Tela de detalhes do sistema MDDIR Tool



Fonte: Elaborado pelo autor.

Na figura 21, aparecem dois links denominados "Clique aqui para fazer o download dos dados em formato XML" e "Clique aqui para fazer o download dos dados em formato CSV", que permitem a geração dos recursos descritos e armazenados em formato XML e CSV, respectivamente. Os resultados gerados em formato XML e CSV podem ser observados nas figuras 22 e 23, respectivamente.

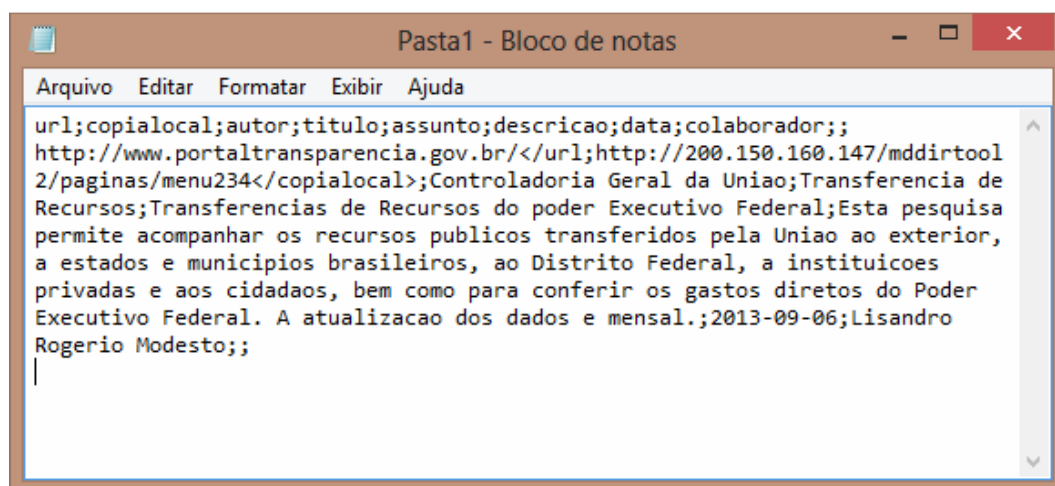
Figura 22 - Arquivo XML gerado pelo sistema MDDIR Tool



Fonte: Elaborado pelo autor.

A figura 22 apresenta uma Página Menu descrita e armazenada no sistema em formato XML.

Figura 23 - Arquivo CSV gerado pelo sistema MDDIR Tool



Fonte: Elaborado pelo autor.

A figura 23 apresenta uma Página Menu descrita e armazenada no sistema em formato CSV.

6.3 RESUMO E DISCUSSÃO

O resultado obtido com o sistema MDDIR Tool é a automação, de parte do processo da descrição de Recursos Informacionais Digitais gerados dinamicamente. Além disso, cada objeto armazenado passa a ter uma cópia material no disco o que permite a visualização desse objeto mesmo que o recurso deixe de existir ou o caminho não tenha sido localizado novamente, o que permite criar histórico dos Recursos Informacionais Digitais gerados dinamicamente.

O sistema MDDIR Tool é caracterizado por funcionalidades que possibilitam o processo de descrição de Recursos Informacionais Digitais gerados dinamicamente, tais como: Cadastro, Consulta por Parâmetro e Consulta por Palavras-Chave.

Dentre as funcionalidades do sistema, podemos destacar a possibilidade de persistir um Recurso Informacional Digital gerado dinamicamente. Para esta funcionalidade, foi usado o HTTRACK Website Copier, que permite baixar um sítio para um diretório local, construindo recursivamente todos os diretórios, incluindo imagens e outros arquivos do servidor para o seu computador.

7 CONSIDERAÇÕES FINAIS

Atualmente, a sociedade vive uma realidade em que as pessoas estão cada vez mais críticas e cada vez buscando mais informações pela Internet, ainda que muitas vezes não saibam muito bem o que fazer com elas. Mesmo assim, as pessoas querem e precisam se sentir como guardiões e fiscais do que acontece no âmbito da gestão pública.

Por meio de leis, como a Lei de Acesso à Informação, os cidadãos estão amparados legalmente e sabem que podem exigir que os órgãos públicos disponibilizem as informações mesmo que ninguém as tenha solicitado e, com isso, têm uma falsa impressão de que poderão "controlar" o que acontece dentro dos gabinetes e as ações de cada gestor público.

Falsa impressão porque ainda que os dados e informações estejam disponíveis, geralmente estão em formatos impossíveis de serem utilizados e conseqüentemente entendidos. Para solucionar esses problemas, é necessário que sejam criados mecanismos que facilitem a busca e a recuperação da informação e que estas possam ter sentido e não sejam apenas um amontoado de números, nomes e valores que não têm nenhum significado.

Para que leis, como a Lei de Acesso à Informação, surtam efeito, é necessário que o cidadão cobre pela disponibilização efetiva das informações e que sejam capazes de utilizá-las a seu próprio tempo e de acordo com sua necessidade.

Na Web, é necessário que os recursos informacionais estejam cada vez mais bem descritos para que se possa recuperá-los de forma eficiente e que atendam os critérios de dados abertos; a esses elementos de descrição dá-se o nome de metadados.

Como os sítios oficiais do Governo Federal não fazem essa descrição dos dados disponibilizados, este trabalho criou um conjunto de metadados capaz de descrever o acesso aos Recursos Informacionais Digitais gerados dinamicamente e persistir tais recursos a fim de possibilitar sua reutilização e garantir novas formas de recuperação.

Para desenvolver o conjunto de metadados foi realizado um levantamento da forma de acesso a dados em sítios oficiais do Governo Federal e analisado o grau

de aderência dos padrões de metadados Dublin Core e e-PMG para descrição de dados gerados dinamicamente.

Para criar histórico dos Recursos Informativos Digitais gerados dinamicamente acessados nos sítios oficiais do Governo Federal foi desenvolvido um sistema Web que faz a recuperação e visualização dos Recursos Informativos Digitais armazenados e descritos na proposta de metadados deste trabalho.

O sistema MDDIR Tool tem ainda a possibilidade de persistir / materializar um recurso digital descrito, com isso caso tal recurso esteja impossibilitado de ser acessado ou o cidadão não lembre mais as páginas e decisões que levaram a formação do recurso ou ainda não lembrem o caminho completo de acesso ao recurso, ele estará arquivado no próprio sistema o que permite a criação de um histórico completo e atende questões sociais no que diz respeito à promoção da democratização da informação por meio da análise de tais recursos digitais armazenados.

Como trabalho futuro, pretende-se criar mecanismos que automatizem o processo de descrição de um Recurso Informativo Digital gerado dinamicamente para que este ocorra independente da vontade humana e sem sua interferência, ou seja, possibilitar que o sistema web seja executado junto ao navegador e, ao acessar um sítio oficial, este sistema permita a descrição dos recursos de forma automatizada.

REFERÊNCIAS

- ALVARENGA, Lidia. A teoria do conceito revisitada em conexão com ontologias e metadados no contexto das bibliotecas tradicionais e digitais. **DataGramaZero**: revista de Ciência da Informação, v. 2, n.6, dez. 2001. Disponível em: <http://www.dgz.org.br/dez01/Art_05.htm>. Acesso em: 01 fev. 2013.
- ALVES, Rachel Cristina Vesu. **Web semântica**: uma análise focada no uso de metadados. 2005. 180 f. Dissertação (Mestrado em Ciência da Informação) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2005.
- BAUMAN, Zygmunt. **Modernidade líquida**. Rio de Janeiro: Zahar Ed., 2001.
- BAX, Marcelo Peixoto. Introdução às linguagens de marcas. *Ciência da Informação*, v.30, n.1, p.32-38, jan./abr. 2001.
- BELL, Daniel. **O advento da sociedade pós-industrial**: uma tentativa de previsão social. São Paulo: Cultrix, 1977.
- BRASIL. **Decreto n. 5.687, de 31 de janeiro de 2006**. 2006. Disponível em: http://www.planalto.gov.br/ccivil_03/_Ato2004-2006/2006/Decreto/D5687.htm. Acesso em: 16 nov. 2013. Promulga a Convenção das Nações Unidas contra a Corrupção, adotada pela Assembléia-Geral das Nações Unidas em 31 de outubro de 2003 e assinada pelo Brasil em 9 de dezembro de 2003.
- BRASIL. **Decreto n. 592, de 6 de julho de 1992**. Disponível em: http://portal.mj.gov.br/sedh/ct/legis_intern/pacto_dir_politicos.htm. Acesso em: 16 nov. 2013.
- BRASIL. **Lei número 12.527, de 18 de novembro de 2011**. Disponível em: <http://www.planalto.gov.br/ccivil_03/_Ato2011-2014/2011/Lei/L12527.htm>. Acesso em: 01 dez. 2012.
- CARDOSO, Olinda Nogueira Paes. **Recuperação de Informação**. INFOCOMP Journal of Computer Science, vol. 2, no. 1, pp.27-32, 2000.
- CARVALHO, João Oscar Fontanini de. **O papel da interação humano computador na inclusão digital**. Transinformação, Campinas, v. 15, p. 75-89, set./dez. 2003.
- CASTELLS, Manuel. **A sociedade em rede**. São Paulo: Paz e Terra, 1999.
- CASTRO, Fabiano Ferreira de. **Conversão retrospectiva de registros bibliográficos**. 2002. 87 f. Trabalho de Conclusão de Curso (Graduação em Biblioteconomia) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2002.

CASTRO, Fabiano Ferreira de; SANTOS, Plácida Leopoldina Ventura Amorim da Costa. **Conversão retrospectiva de registros bibliográficos**. In: FUJITA, M. S. L. (Org.). A dimensão social da biblioteca digital na organização e acesso ao conhecimento: projetos, produtos e relatos de experiência. São Paulo: Sistema Integrado de Bibliotecas da USP, Dep. Técnico: IBICT, 2005. v. 2, p. 63-87.

CASTRO, Fabiano Ferreira de; SANTOS, Plácida Leopoldina Ventura Amorim da Costa. Representação e descrição de recursos informacionais: aspectos estruturantes no delineamento de ambientes informacionais digitais. *Pesq. bras. ci. inf.*, Brasília, v.3, n.1, p.155-168, jan./dez. 2010

CEARÁ. Tribunal de Contas dos Municípios do Estado. **Manual sobre transparência e lei de acesso à informação pública**. / Tribunal de Contas dos Municípios do Estado do Ceará. - Fortaleza: TCM-CE, 2013

COUTINHO, Clara; LISBÔA, Eliana. Sociedade da Informação, do Conhecimento e da Aprendizagem: desafios para educação no século XXI. **Revista de Educação**, Vol. XVIII, nº 1, 2011

DCMI. **Dublin Core Metadata Initiative**. 2013. Disponível em: <http://dublincore.org/>. Acesso em: 08 nov. 2013.

DECLARAÇÃO DE PRINCÍPIOS SOBRE LIBERDADE DE EXPRESSÃO. 2000. Disponível em: <http://www.cidh.oas.org/basicos/portugues/s.Convencao.Libertade.de.Expressao.htm>. Acesso em: 16 nov. 2013

DECLARAÇÃO UNIVERSAL DOS DIREITOS HUMANOS. 1948. Disponível em: http://portal.mj.gov.br/sedh/ct/legis_intern/ddh_bib_inter_universal.htm. Acesso em: 16 nov. 2013.

DESAI, Bipin C. **Supporting discovery in virtual libraries**. *Journal of the American Society for Information Science*, v. 48, n. 3, p. 190-204, 1997.

DINIZ, Vagner. Como conseguir Dados Governamentais Abertos. Paineis 13/049. III **Congresso Consad de Gestão Pública**. 2010.

DRUCKER, Peter. **Sociedade pós-capitalista**. São Paulo: Pioneira, 1994.

e-PMG. **Padrão de Metadados do Governo Eletrônico**. Disponível em: <http://www.governoeletronico.gov.br/biblioteca/arquivos/e-pmg>. Acessado em: 16 nov. 2013.

FRAKES, William B.; BAEZA-YALES, Ricardo, R. **Information Retrieval Data Structures & Algorithms**. Prentice Hall, 1992.

FUSCO, Elvis. Modelos conceituais de dados como parte do processo da catalogação: perspectiva de uso dos FRBR no desenvolvimento de catálogos bibliográficos digitais. 2010.

251 f. Tese (Doutorado em Ciência da Informação) - Universidade Estadual Paulista, Marília. 2010.

GARCIA, Joana Coeli Ribeiro; TARGINO, Maria das Graças. Conceitos de inédito e original: uso e implicações na comunicação científica. **DataGramaZero**: Revista de Ciência da Informação, v. 13, n.6, dez12. Disponível em: <http://www.dgz.org.br/dez12/Art_02.htm#R1>. Acesso em: 31 mai. 2013.

GIANNASI-KAIMEN, Maria Júlia; CARELLI, Ana Esmeralda (Orgs). **Recursos informacionais para compartilhamento da informação: redesenhando acesso, disponibilidade e uso**. Rio de Janeiro: Epapers, 2007. 226 p. ISBN 9788576501404.

GILLILAND-SWETLAND, A. J. **La definición de los metadatos**. In: INTRODUCCIÓN a los metadatos: vías a la información digital. [S. l.]: GETTY, 1999. p. 1-9.

GRACIO, José Carlos Abbud. **Metadados para a descrição de recursos da Internet: o padrão Dublin Core, aplicações e a questão da interoperabilidade**. 2002. 127f. Dissertação (Mestrado em Ciência da Informação) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília. 2002.

Herrero-Solana, Víctor; Hassan, Yusef. **Metodologías para el desarrollo de interfaces visuales de recuperación de información: análisis y comparación**. Grupo SCImago, Universidad de Granada Facultad de Biblioteconomía y Documentación Colegio Máximo de Cartuja s.n., 18071 Granada, España. 2006.

JARDIM, José Maria. Capacidade governativa, informação, e governo eletrônico. **DataGramaZero** - Revista de Ciência da Informação - v.1 n.5 out/00.

LEMOS, André (coord.). Cidade, tecnologias e interfaces. Análise de interfaces de portais governamentais brasileiros. Uma proposta metodológica. Revista Fronteiras. Estudos Midiáticos. Vol. VI no.2 p.117-136. São Leopoldo, 2004.

LYOTARD, Jean-François. **O pós-moderno**. Rio de Janeiro: J.Olympio, 1990.

MATTELART, Armand. **História da sociedade da informação**. São Paulo: Loyola, 2002.

MAYER, Marceliz. Por quê construir um Modelo de Dados Lógico?. Devmedia, 2007. Disponível em <http://www.devmedia.com.br/por-que-construir-um-modelo-de-dados-logico-parte-i/368>. Acessado em 15 de novembro de 2013.

MEY, Eliane Serrão Alves; SILVEIRA, Naira Christofoletti. **Catalogação no plural**. Brasília, DF: Briquet de Lemos, 2009.

MOURA, Ana Maria de Carvalho e FERNANDES, Cezar S. A Metadata Approach to Represent and Visualize Sites on the Web. **International Workshop on Information**

Integration Technologies on the Web. Itaipava, RJ, 9-11 abril 2001. Disponível em: <<http://www.ipanema.ime.eb.br/~anamoura/METADATA.html>> Acesso em: 11 jan. 2013.

NASCIMENTO, Luciana Umburanas. Um padrão de metadados para indexação e recuperação de objetos multimídia. 2008. 85 f. Dissertação (Mestrado em Informática) – Universidade Federal do Paraná, Curitiba. 2008.

Naves, Madalena Martins Lopes. **Aspectos conceituais do browsing na recuperação da informação.** Ci. Inf. vol.27 n.3 Brasília Sept. 1998, disponível em: http://www.scielo.br/scielo.php?pid=S0100-19651998000300005&script=sci_arttext

O'DONNELL, Guillermo. Accountability horizontal e novas poliarquias. Tradução de Clarice Cohn e Alvaro Augusto Comin. **Lua Nova: Revista de Cultura e Política**, n. 44, 1998. Disponível em: <http://www.scielo.br/scielo.php?pid=S0102-64451998000200003&script=sci_arttext>. Acesso em: 11 fev 2013.

PRESSMAN, Roger S. **Engenharia de Software.** 5 ed. Rio de Janeiro: McGraw-Hill, 2002.

SANT'ANA, Ricardo César Gonçalves. Mensuração da disponibilização de informações e do nível de interação dos ambientes informacionais digitais da administração municipal com a sociedade. Tese de Doutorado em Ciência da Informação – Universidade Estadual Paulista “Júlio de Mesquita Filho”, Marília, São Paulo, 2008. 153 p.

SILVA, Edna Lúcia da; LOPES, Marili Isensee Lopes. A internet, a mediação e a desintermediação da informação. **DataGramZero: Revista de Ciência da Informação**, v.12 n.2 abr/11. Disponível em: <http://www.dgz.org.br/abr11/Art_04.htm>. Acesso em: 31 mai. 2013.

SMITH, John R.; SCHIRLING, Peter. **Metadata Standards Roundup.** <http://ieeexplore.ieee.org>. 2006. Acesso em: 02 fev. 2013.

SOUZA, Terezinha Batista de; CATARINO, Maria Elizabete; SANTOS, Paulo Cesar. **Metadados: Catalogando dados na Internet.** Transinformação, v.9, nº 2, p. 93-105, maio/agosto, 1997.

VIEIRA, Jessica Monique de Lira; CORRÊA, Renato Fernandes. **Recuperação de Informação através de recursos visuais.** Universidade Federal da Paraíba, 2010.

WEISER, Mark. **The Computer for the 21st Century.** Scientific American, v. 265, n. 3, p.94-104, setembro, 1991. Disponível em <<https://www.cs.cmu.edu/afs/cs/Web/People/jasonh/courses/ubicomp-sp2007/papers/02-weiser-computer-21st-century.pdf>>. Acesso em: 16 de Novembro de 2013.

WESTERMANN, Utz; KLAS, Wolfgang.. **An Analysis of XML Database Solutions For The Management Of MPEG-7 Media Descriptions.** <http://portal.acm.org>. 2003. Acesso em: 02 fev. 2013.

XAVIER, Raphael Figueiredo. Análise de Métodos de Produção de Interfaces Visuais para Recuperação da Informação. Marília, 2009.