



MINISTÉRIO DA CIÊNCIA E TECNOLOGIA  
INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS



## **IMPLANTAÇÃO E USO DO OPENSEARCH NA DISSEMINAÇÃO DE DADOS GEOGRÁFICOS PRODUZIDOS NO INPE**

**RELATÓRIO FINAL DE PROJETO DE INICIAÇÃO TECNOLÓGICA  
(PIBITI/CNPq/INPE)**

**Matheus Cavassan Zaglia (FATEC, Bolsista PIBITI/CNPq)  
E-mail: mzaglia@dpi.inpe.br**

**Lúbia Vinhas (DPI/OBT/INPE, Orientadora)  
E-mail: lubia@dpi.inpe.br**

**Julho de 2015**



## **AGRADECIMENTOS**

Gostaria de agradecer a minha família por me acompanhar durante a minha jornada.



## RESUMO

A utilização de dados de observação da Terra, produzidos e gerenciados pela agências espaciais de diferentes países é de extrema importância em diversas áreas de aplicação, porém o acesso a eles nem sempre é fácil e simples. Em geral, as agências espaciais adotam diferentes formas para a disponibilização de seus dados, com isso para acessar dados de diferentes agências é necessário que o usuário tenha conhecimento de diferentes tecnologias de acessos e catálogos especializados. Mais recentemente, as organizações internacionais que congregam diferentes agências espaciais têm procurado definir padrões, especificações e melhores práticas para a construção de catálogos mais interoperáveis e de mais simples acesso. Este trabalho tem como principal objetivo aplicar as especificações de um desses padrões, chamado OpenSearch na disseminação dos dados de observação da Terra produzidos pelo INPE, de modo a tornar a sua disseminação mais simples e mais padronizada.



# **IMPLEMENTATION AND USE OF THE OPENSEARCH IN THE DISSEMINATION OF THE GEOGRAPHIC DATA PRODUCED IN INPE**

## **ABSTRACT**

The usage of Earth observation data, produced and managed by space agencies around the world, is extremely important in many areas of application, however the access is not always easy or simple. Space agencies adopt different forms for the provision of their data, so that users have to understand and be able to interact with different technologies and different catalogs. More recently, international organizations that bring together different space agencies have sought to set standards, specifications and best practices for building catalogs that are more simple and interoperable. This work aims to apply the standard specifications of the OpenSearch in the Earth observation data produced by INPE in order to make its dissemination simpler and more standardized.



## LISTA DE FIGURAS

	<b><u>Pág.</u></b>
Figura 2.1.1 – Cliente/Servidor.....	19
Figura 2.1.2 – Catálogo de Imagens do INPE.....	20
Figura 2.3.1 – Stateless Service .....	22
Figura 2.4.1 – Leitor de Feed RSS .....	22
Figura 2.5.1 – ATOM Feed .....	23
Figura 4.1 – ATOM Feed com OpenSearch e os dados do INPE.....	26
Figura 4.2 – OpenSearch com dados do INPE na API JavaScript do Google Maps.....	27
Figura 4.3 – OpenSearch Feed com dados do INPE na API JavaScript do ArcGis.....	28
Figura 4.4 – OpenSearch Feed com dados do INPE no Mozilla Firefox.....	28



## LISTA DE TABELAS

	<b><u>Pág.</u></b>
Tabela 3.2.1 – OpenSearch Best Practices .....	24



## LISTA DE SIGLAS E ABREVIATURAS

INPE	Instituto Nacional de Pesquisas Espaciais
IBGE	Instituto Brasileiro de Geografia e Estatística
OGC	Serviço de Informação e Documentação
ISO	International Organization for Standardization
HTTP	Hypertext Transfer Protocol
CEOS	Committee on Earth Observation Satellites
URL	Uniform Resource Locator
RSS	Really Simple Syndication
XML	eXtensible Markup Language
SOAP	Simple Object Access Protocol
REST	Representational State Transfer
SQL	Structured Query Language



## SUMÁRIO

	<b><u>Pág.</u></b>
<b>1</b>	<b>INTRODUÇÃO ..... 17</b>
<b>2</b>	<b>REVISÃO DE TECNOLOGIAS RELEVANTES ..... 19</b>
2.1.	ARQUITETURA CLIENTE/SERVIDOR ..... 19
2.2.	SERVIÇOS WEB ..... 21
2.3.	REPRESENTATIONAL STATE TRANSFER ..... 21
2.4.	OPENSEARCH..... 22
2.5.	ATOM..... 23
<b>3</b>	<b>MATERIAIS E MÉTODOS ..... 23</b>
3.1.	AMBIENTE DE DESENVOLVIMENTO ..... 23
3.2.	OPENSEARCH CEOS BEST PRACTICES ..... 24
<b>4</b>	<b>RESULTADOS ..... 25</b>
<b>5</b>	<b>CONCLUSÃO E TRABALHOS FUTUROS..... 29</b>



## 1 INTRODUÇÃO

Nos últimos anos, a utilização de informação geoespacial tem sido cada vez mais frequente na elaboração de planos governamentais, estudos socioeconômicos e ambientais por parte dos órgãos e agências do governo.

No Brasil, diversas instituições produzem este tipo de informação, entre elas podemos citar o INPE, a Embrapa, o IBGE, o Exército Brasileiro e o Ministério do Meio Ambiente. Um dos grandes desafios dessas instituições é disponibilizar seus dados via internet, para possibilitar seu compartilhamento e uso efetivo.

A internet está cheia de informações geoespaciais, dispostas em diferentes formatos e locais. Para isso foram criados serviços de catálogos que permitem a publicação de informação descritiva sobre dados (metadados), serviços e informação relacionada a um objeto. Os metadados quando utilizados em dados geoespaciais, disponibilizam informações como imagens de satélite, localização espacial, autor, data de aquisição, instrumentos utilizados para aquisição (NAKAMURA, 2010). Aplicações que expõe metadados sobre as informações contidas em um dado repositório podem ser chamadas genericamente de catálogos.

Um Serviço de Catálogo ajuda os usuários e aplicações a encontrar os metadados contidos em um ambiente distribuído de computação através de um conjunto de interfaces de serviços que permitem a organização, descoberta e acesso a dados geoespaciais (OGC, 1999).

Para isso oferecem uma camada de acesso onde um dado e seus metadados possam ser descobertos, ou seja, um mecanismo de busca interno, por uma aplicação cliente-servidor, e posteriormente enviado para o usuário. Para que esse fluxo de uso seja interoperável, ou seja, independente de tecnologia usada no lado do cliente ou do servidor, ou das instituições envolvidas utiliza-se interfaces que seguem padrões abertos.

Na área de dados geoespaciais, os padrões adotados pelas organizações e pela comunidade são aqueles desenvolvidos pelo Open Geospatial Consortium (OGC)<sup>1</sup>. O OGC é uma organização formada por membros da academia e indústria nos moldes da Internacional Organization for Standardization (ISO)<sup>2</sup>. Muitos dados espaciais, incluindo imagens de satélites, estão cada vez mais sendo descobertos e acessados através de aplicações clientes que se valem de interfaces que implementam padrões OGC. Interfaces de catálogos especializadas como as definidas pelo OGC tem grande valor para permitir, porém são complexas de serem implementadas e são usadas em geral em aplicações específicas como os Sistemas de Informação Geográfica (SIG).

Por isso, muitos desses dados, talvez a sua maioria, ainda não são descobertos através de aplicações mais genéricas como os modernos Navegadores Web (ex. Chrome, Firefox, Thunderbird, etc.). Uma maneira de incluir as capacidades de busca especializadas em Navegadores Web usa uma coleção de formatos simples chamada genericamente de OpenSearch.

Recentemente, em 2014, o Committee on Earth Observation Satellites (CEOS)<sup>3</sup> reconhecendo o valor da especificação OpenSearch, criou um projeto buscando adaptar e estabelecer um conjunto de melhores práticas de implementação do OpenSearch no acesso a metadados e dados de Observação da Terra a serem usados pelas agências participantes do CEOS, visando aumentar a interoperabilidade e simplificar o acesso aos dados espaciais que elas produzem.

O INPE faz parte do CEOS e tem colaborado com seu Working Group on Information Systems and Services (WGISS)<sup>4</sup>. O INPE é reconhecido nacional e internacionalmente pela sua política de acesso livre a dados de observação da Terra e também de dados temáticos resultantes do processamento sobre eles.

---

<sup>1</sup> <http://www.opengeospatial.org/>

<sup>2</sup> <http://www.iso.org/>

<sup>3</sup> <http://www.ceos.org/>

<sup>4</sup> <http://ceos.org/ourwork/workinggroups/wgiss/>

Esse trabalho visa estudar e implementar a especificação CEOS OpenSearch no acesso e disseminação dos dados de observação da Terra produzidos e gerenciados pelo INPE.

## 2 REVISÃO DE TECNOLOGIAS RELEVANTES

Esse trabalho visa o desenvolvimento de uma aplicação de disseminação de dados e interoperabilidade no ambiente da Internet. Essa seção faz uma revisão das tecnologias relevantes para esse desenvolvimento.

### 2.1. ARQUITETURA CLIENTE/SERVIDOR

A Internet é uma rede mundial de computadores que funciona a partir de um conjunto de protocolos abertos que permite que essas máquinas heterogêneas se comuniquem formando um ambiente compartilhado. A maioria das aplicações que rodam sobre a internet são baseadas em uma arquitetura chamada de cliente/servidor. Essa arquitetura permite o compartilhamento de recursos através de requisições e resposta (Figura 2.1).

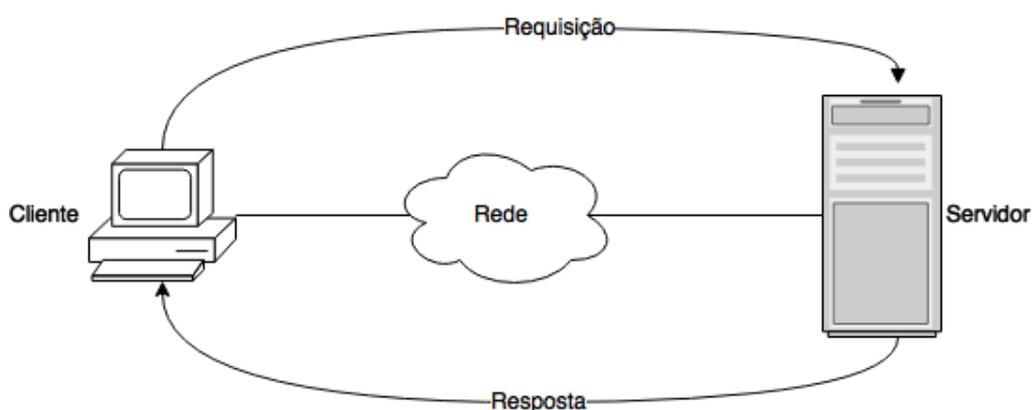


Figura 2.1.1 - Cliente/Servidor

Existem muitos clientes que acessam os recursos de um servidor concomitantemente, independentemente de sua localização, podendo este

estar no mesmo edifício ou até mesmo do outro lado do mundo (TANENBAUM, 2003).

Baseadas na arquitetura cliente-servidor inúmeras aplicações têm sido desenvolvidas na internet. Entre as mais importantes podem ser citadas o e-mail para troca de mensagens e World Wide Web para a publicação de documentos como hipertextos (documentos que embutem ligações para outros documentos, figuras, vídeos, etc.). Essas aplicações são desenvolvidas para um usuário humano.

Um exemplo de aplicação web é o atual catálogo para acesso as imagens distribuídas pelo INPE (Figura 2.1.2).

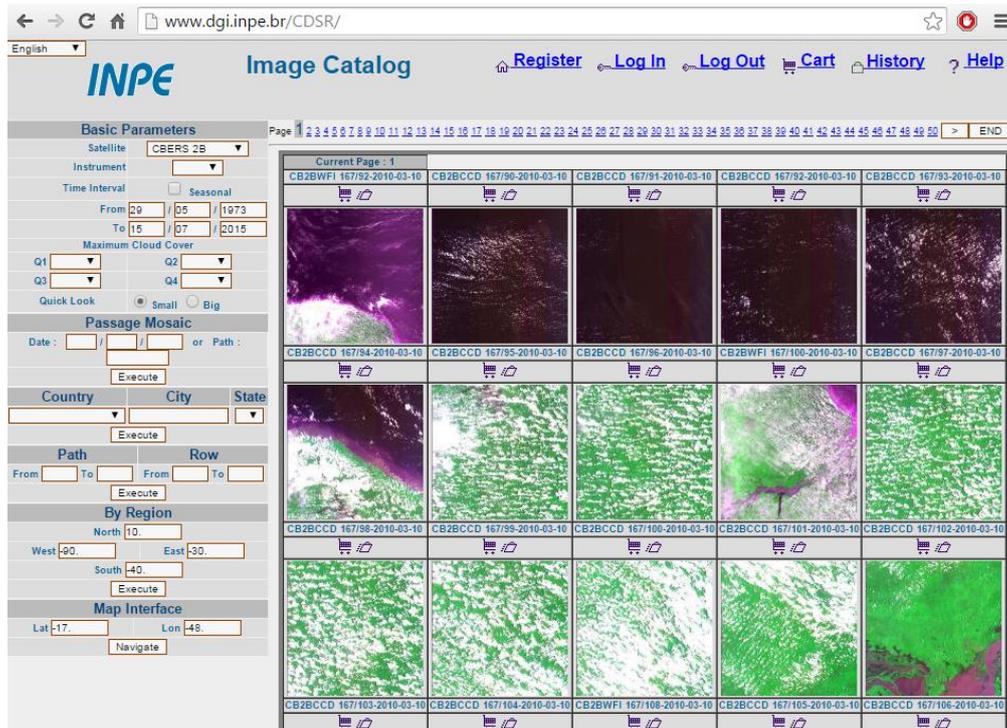


Figura 2.1.2 – Catálogo de Imagens do INPE

O sucesso e a evolução da internet trouxeram a demanda por uma maneira de comunicação entre aplicações. Isso pode ser feito através da tecnologia de Web Services.

## **2.2. SERVIÇOS WEB**

A necessidade de comunicação e interoperabilidade entre aplicações tem crescido de uma maneira muito rápida e com isto cresce a criação e utilização de serviços web. Os Serviços web permitem que aplicações se comuniquem de forma dinâmica entre si através de protocolos padrões presentes da internet (HTTP, XML, SOAP, etc.)

Um Serviço web pode ser definido como um sistema de software com o objetivo de suportar interoperabilidade entre máquinas. Este possui uma interface descrita um formato processável por uma máquina (especificamente WSDL). Outros sistemas interagem com o Serviço web de maneira prescrita em sua descrição utilizando mensagens SOAP, tipicamente são transportados usando o protocolo HTTP com XML em conjunto com outros padrões (W3C, 2004)

## **2.3. REPRESENTATIONAL STATE TRANSFER**

O Representational State Transfer (REST) define uma arquitetura de princípios que possibilitam o desenvolvimento de Serviços web orientado a recursos. Cada recurso deve ser identificado utilizando um URI que utiliza explicitamente os métodos HTTP: GET, POST, PUT e DELETE. Um serviço web REST não deve manter estado, ou seja, o servidor não deve manter informações sobre a requisição e para que isto aconteça, toda as informações necessárias são enviadas junto com a resposta. Por exemplo, um resultado com múltiplas páginas, o cliente informa a página atual para que o servidor gere uma resposta informando a próxima página. Isto permite que o próprio cliente mantenha seu estado (Figura 2.3.1).

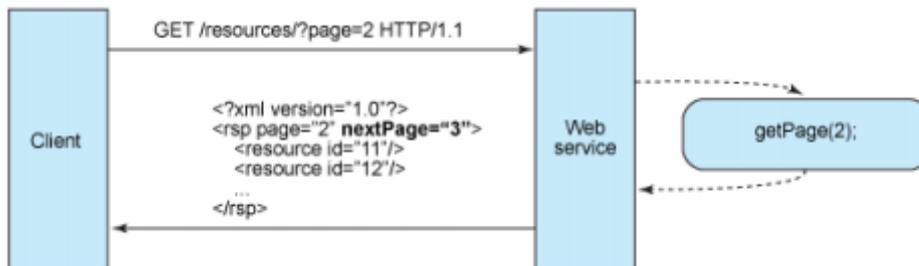


Figura 2.3.1 – Stateless Service  
 Fonte: RODRIGUEZ (2008)

## 2.4. OPENSEARCH

O OpenSearch é uma coleção de formatos para o compartilhamento de resultados de pesquisa que ajuda na comunicação entre os mecanismos de buscas no servidor e as aplicações cliente. O OpenSearch é utilizado para descrever uma interface de pesquisa de uma maneira auto descritiva tanto para o usuário quanto para a aplicação. Ele é constituído por dois elementos, um documento descritor OpenSearch (OSDD) e por um uma resposta padronizada que é obtida através de uma requisição HTTP/GET. Sua utilização permite que as respostas em ATOM obtidas através da requisição possam ser utilizadas em clientes genéricos como os browsers e leitores de feeds RSS (Figura 2.4.1).

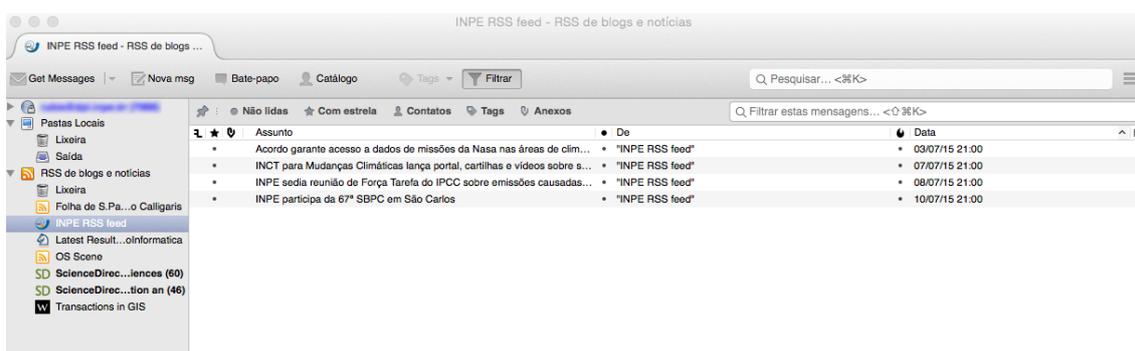


Figura 2.4.1 – Leitor de Feed RSS

O OpenSearch possui uma extensão geográfica que propõe a padronização dos resultados de uma pesquisa em um geo webservice, com isto é possível

descobrir dados geoespaciais em APIs de mapas disponíveis em browsers, como o GoogleMaps e ArcGIS.

Para que o OpenSearch possa ser melhor utilizado ele se baseia nos protocolos acima, que permitem a interoperabilidade se tornar possível. Isso torna a disseminação de dados geográficos na internet muito mais simples.

## 2.5. ATOM

Atom é um documento baseado no formato XML que apresenta uma lista de informações relacionadas conhecidas como “feeds”. Os feeds são compostos por diferentes entradas. Cada entrada possui uma série de metadados como: título, id, data de aquisição, links relacionados, etc. (NOTTINGHAM, 2005)

```
1. <?xml version="1.0"?>
2. <feed xmlns="http://www.w3.org/2005/Atom">
3. <title>O nome do feed de dados</title>
4. <link href="http://www.example.com" rel="alternate" type="text/html" />
5. <updated>2006-06-11T18:30:02Z</updated>
6. <author>
7. <name>Example</name>
8. </author>
9. <id>tag:example.com,2006-06-11:/support/products</id>
10. <entry>
11. <title>Título do item</title>
12. <id>1a</id>
13. <link href="http://www.example.com/item1-info-page.html" />
14. <summary>Informações do item 1</summary>
15. <updated>2005-10-13T18:30:02Z</updated>
16. </entry>
17. </feed>
```

Figura 2.5.1 - ATOM Feed

## 3 MATERIAIS E MÉTODOS

### 3.1. AMBIENTE DE DESENVOLVIMENTO

Para o desenvolvimento dese trabalho foi escolhida a linguagem de programação Python 2.7 em conjunto com o microframework Flask que possibilitaram a elaboração do serviço web. Para acesso a base de dados do INPE foi utilizado o pacote Python SQL Alchemy que permite recuperar dados

em uma base de dados MySQL de uma maneira simples, facilitando o desenvolvimento do código.

### 3.2. OPENSEARCH CEOS BEST PRACTICES

Para melhor aplicar o OpenSearch em um ambiente com dados geoespaciais o CEOS criou um documento onde são definidas boas práticas. Neste documento são ditados quais atributos são requeridos, recomendados ou opcionais em sua aplicação OpenSearch. Na Tabela 3.2.1 estão resumidas as boas práticas a serem utilizadas na aplicação.

Tabela 3.2.1 – OpenSearch Best Practices

<b>Categoria</b>	<b>Título</b>	<b>Classificação</b>
Busca em dois passos	Suporte a busca em dois passos	Recomendado
	dc:identifier na resposta da coleção	Requerido (caso possua suporte a busca em dois passo)
OSDD	Suporte a extensão OpenSearch Parameter (Draft 2)	Recomendado
	Atributo rel da URL no OSDD	Recomendado
	Identificador do CEOS OpenSearch Best Practice	Recomendado
Requisição de Busca	Suporte a parâmetros de pesquisa	Requerido

	Múltiplas palavras para o searchTerms	Recomendado
	Uso de startPage ao invés de startIndex	Recomendado
	Busca com geo:name	Recomendado
	Formato de saída na URL	Opcional
Resposta da Busca	Suporte ao formato ATOM	Requerido
	Representação dos metadados na resposta	Recomendado
	atom:summary	Recomendado
	GeoRSS	Recomendado
	Pesquisar imagem	Recomendado
	Acesso a dados	Recomendado
Exceções	Códigos de exceção	Recomendado

Fonte: Adaptado de CEOS (2014)

#### 4 RESULTADOS

Foi desenvolvida uma biblioteca em Python com funções capazes de ler o banco de dados do INPE e a partir dos dados recuperados estruturar a resposta no formato ATOM para que o cliente consiga utilizar o mesmo em uma aplicação genérica.

Por ser um protótipo, o serviço foi hospedado no site Python Anywhere<sup>5</sup>, que permite realizar a distribuição de aplicações desenvolvidas utilizando Flask e Python gratuitamente. O acesso ao serviço está disponível em <http://osinpe.pythonanywhere.com/>.

O serviço atual é um espelho do banco de dados de metadados já criado no contexto do CSDR.

Dos critérios definidos na Tabela 3.2.1, a implementação atual do serviço permite: Utilização de parâmetros de pesquisa, suporte ao formato ATOM, pesquisar imagens, aplicação do GeoRSS, representação dos metadados, seleção da página inicial e itens por página, acesso a dados.

Com serviço online, ao acessar a URL onde está hospedado e passando o parâmetro bbox em uma requisição GET é possível obter. Ao utilizar os valores -70,-50,-20,0 no bbox é obtida a seguinte resposta (Figura 4.1):

```
<?xml version="1.0" encoding="UTF-8" ?>
<feed xmlns="http://www.w3.org/2005/Atom" xmlns:content="http://purl.org/rss/1.0/modules/content/"
xmlns:dc="http://purl.org/dc/elements/1.1/" xmlns:geo="http://a9.com/-/opensearch/extensions/geo/1.0/"
xmlns:georss="http://www.georss.org/georss" xmlns:os="http://a9.com/-/spec/opensearch/1.1/"
xmlns:params="http://a9.com/-/spec/opensearch/extensions/parameters/1.0/" xmlns:time="http://a9.com/-
/opensearch/extensions/time/1.0/" xml:lang="en">
  <title>OS Scene</title>
  <id>
    http://osinpe.pythonanywhere.com/atom?bbox=-70,-50,-20,0
  </id>
  <link href="http://osinpe.pythonanywhere.com/osdd" rel="search" title="OS Scene"
type="application/opensearchdescription+xml"/>
  <author>
    <name>INPE - National Institute for Space Research</name>
    <email>atus@dgi.inpe.br</email>
  </author>
  <os:totalResults>3046</os:totalResults>
  <os:startIndex>0</os:startIndex>
  <os:itemsPerPage>10</os:itemsPerPage>
  <entry>
    <title>
      Entity ID: L5TM21806420070622, Satellite: L5, Sensor: TM, Acquisition date: 2007/06/22
    </title>
    <id>
      http://osinpe.pythonanywhere.com/atom?sceneid=L5TM21806420070622
    </id>
    <dc:date>2007/06/22</dc:date>
    <updated>2014/01/31 15:33:44</updated>
    <author>INPE - National Institute for Space Research</author>
    <link href="http://www.dgi.inpe.br/cwic/sceneid.php?sceneid=L5TM21806420070622" rel="via" title="Original source
metadata" type="text/xml"/>
    <link href="http://www.dgi.inpe.br/cwic/sceneid.php?sceneid=L5TM21806420070622" rel="alternate" title="Alternate
metadata URL" type="text/xml"/>
    <link href="http://www.dgi.inpe.br/CDSR/display.php?TABELA=Browse&PREFIXO=Landsat&INDICE=L5TM21806420070622"
rel="icon" title="Browse image URL" type="image/jpeg"/>
    <link href="http://www.dgi.inpe.br/CDSR/cart-cwic.php?SCENEID=L5TM21806420070622" rel="enclosure" title="Product
download URL" type="text/html"/>
    <georss:polygon>
      -4.88724 -41.7839 -5.13358 -40.0901 -6.68324 -40.4185 -6.43614 -42.1175 -4.88724 -41.7839
    </georss:polygon>
    <content:encoded>
      <div><a href="http://osinpe.pythonanywhere.com/atom?sceneid=L5TM21806420070622" target="_blank"></a></div>
    </content:encoded>
  </entry>
</feed>
```

Figura 4.1 – ATOM Feed com OpenSearch e os dados do INPE

<sup>5</sup> <https://www.pythonanywhere.com/>

Este feed em ATOM pode ser utilizado em APIs com suporte a leitura de GeoRSS como o GoogleMaps JavaScript API, ArcGis JavaScript API, Mozilla Firefox (Figuras 4.2 a 4.4).

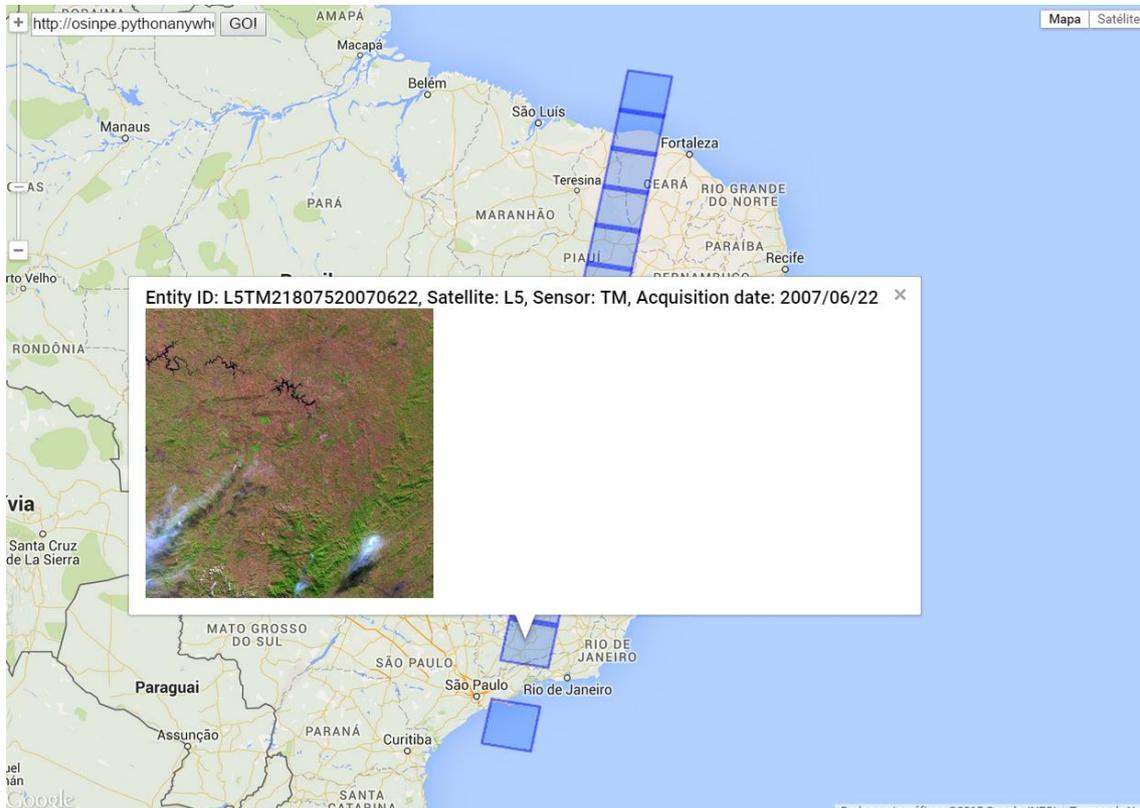


Figura 4.2 - OpenSearch com dados do INPE na API JavaScript do Google Maps

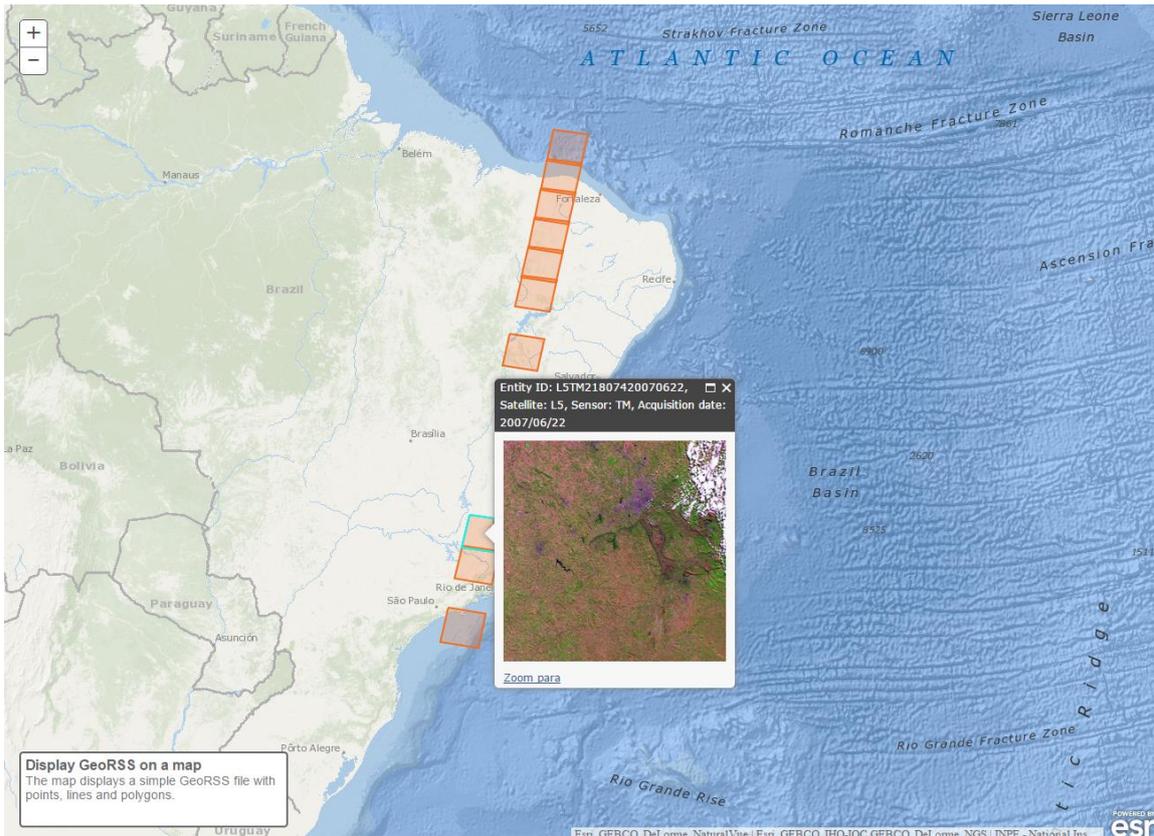


Figura 4.3 - OpenSearch Feed com dados do INPE na API JavaScript do ArcGis

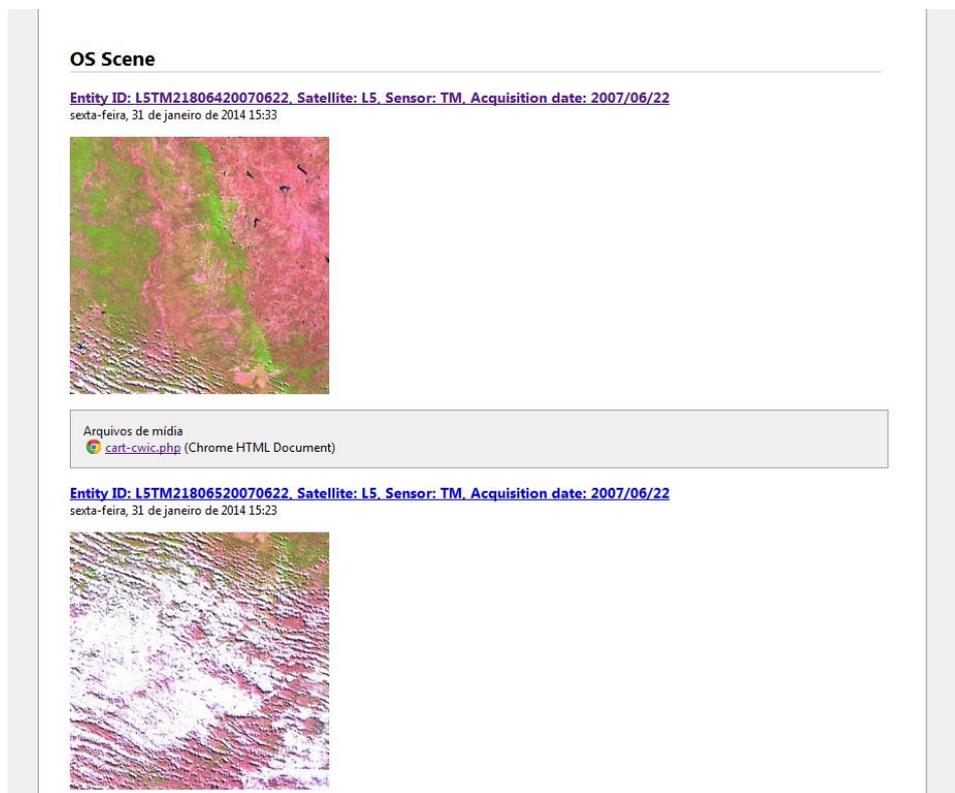


Figura 4.4 - OpenSearch Feed com dados do INPE no Mozilla Firefox

## **5 CONCLUSÃO E TRABALHOS FUTUROS**

Durante o período de julho de 2014 a julho de 2015, foi possível adquirir conhecimentos sobre dados geográficos, sensoriamento remoto, serviços web e programação.

Nesse trabalho foi implementado a especificação CEOS OpenSearch no acesso e disseminação nas produzidos e gerenciados pelo INPE, como uma complementação ao sistema mostrado na Figura 2.1.2.

Isto permitiu com que a implementação do OpenSearch se tornasse uma realidade. Com a utilização do OpenSearch no serviço web utilizando a base de dados do INPE, foi possível que a descoberta dos dados fosse visualizada em diferentes ferramentas disponíveis na internet, sem necessidade de realizar downloads ou instalações.

A versão protótipo ainda não explora as capacidades do OpenSearch por não implementar todos critérios de buscas, mesmo assim já é possível perceber seu potencial ao ser utilizado.

Para trabalhos futuros o serviço web pode ser melhorado com implementação de novos formatos de resposta comuns na internet como o JSON e o KML. Também pode ser adicionada uma interface com caixas de texto e botões onde o usuário irá inserir os parâmetros de busca, que no atual momento é feito diretamente na URL do serviço web.



## REFERÊNCIAS BIBLIOGRÁFICAS

OGC - **The OpenGIS™ Abstract Specification Topic 13: Catalog Services**  
Disponível em: <[http://portal.opengeospatial.org/files/?artifact\\_id=901](http://portal.opengeospatial.org/files/?artifact_id=901)> Acesso em: 28/05/2015

NAKAMURA, Eduardo Tomio. **Infraestrutura de Dados Espaciais em Unidades de Conservação: uma proposta para disseminação da informação geográfica do Parque Estadual de Intervales-SP**. 2010. Tese de Doutorado. Universidade de São Paulo.

TANENBAUM, Andrew S. **Redes de Computadores**. 4ª Edição. Rio de Janeiro: Editora Campus, 2003.

W3C – WORLD WIDE WEB CONSORTIUM et al. **Web Services Architecture, W3C Working Group**. 2004. Disponível em: <<http://www.w3.org/TR/ws-arch/wsa.pdf>> Acesso em: 30/06/2015

NOTTINGHAM, Mark; SAYRE, Robert. **The atom syndication format**. 2005. Disponível em: <<http://www.ietf.org/rfc/rfc4287.txt>> Acesso em: 30/06/2015

RODRIGUEZ, Alex. **RESTful Web services: The basics**. 2008. Disponível em: <<http://cs.kennesaw.edu/~bsetzer/4320su15/extra/examples/ws-restful-pdf.pdf>> Acesso em: 02/07/2015

CEOS. **CEOS OpenSearch Best Practices Document**. 2014. Disponível em: <<http://ceos.org/wp-content/uploads/2014/12/CEOSOpenSearchBestPracticeDocument-PublicComment.pdf>> Acesso em: 03/05/2015