

## **Técnicas de Marcação Estruturada na Otimização de Sites para Buscas Semânticas: Um Estudo de Caso com o Site do Curso de Ciência da Computação do IFSULDEMINAS – Câmpus Muzambinho**

**Everson F. Leite, Ramon G. T. M. Silva, Aracele G. O. Fassbinder**

Laboratório de Tecnologias de Software e Computação Aplicada à Educação - Instituto Federal de Educação, Ciência e Tecnologia do Sul de Minas Gerais – Câmpus Muzambinho (IFSULDEMINAS)

Estrada de Muzambinho, km 35 - Bairro Morro Preto - Cx. Postal 02 - CEP: 37890-000 – Muzambinho – MG – Brasil

eversonfelix\_mb@hotmail.com, ramon.silva@muz.ifsuldeminas.edu.br, aracele.garcia@muz.ifsuldeminas.edu.br

**Abstract.** *The main way of searching for Web sites is through search engines, requiring a good position to achieve good results. Optimization techniques for search engines are used for this purpose. However, given the high competitiveness of Web sites and of technological advancement, these techniques may not be enough. The current Web implies the need to create semantic content so that the pages can be better interpreted by search robots and they bring rich results in information for users. In this context, this article discusses a study on the relationship of the semantic web with search engines, considering mainly the influence that using SEO techniques in conjunction with structured markup of information has on these mechanisms. Thus, enforcement activities were conducted, analysis and evaluation techniques of SEO and marking structured in a case study on the website of the Computer Science IFSULDEMINAS - Campus Muzambinho. After application of the techniques noted on the results of consultations the occurrence of rich semantic information and an improvement in the site's positioning.*

**Resumo.** *A principal forma de procura de sites na Web ocorre por meio de mecanismos de buscas, sendo necessário um bom posicionamento para o alcance de bons resultados. Técnicas de Otimização para mecanismos de buscas são utilizadas para esta finalidade. Contudo, diante da alta competitividade dos sites e do avanço tecnológico da Web, essas técnicas podem não ser suficientes. A Web atual implica na necessidade de criação de conteúdos semânticos para que as páginas possam ser melhor interpretados pelos robôs de busca e estes trazerem resultados ricos em informações para os usuários. Neste contexto, este artigo aborda um estudo sobre a relação da web semântica com os motores de busca, considerando, principalmente, a influência que o uso de técnicas de SEO em conjunto com marcação estruturada das informações exerce sobre estes mecanismos. Assim, foram conduzidas atividades de aplicação, análise e avaliação de técnicas de SEO e marcação estruturada em um estudo de caso com o website do curso de Ciência da Computação do IFSULDEMINAS – Câmpus Muzambinho. Após a aplicação das técnicas observou-se nos resultados de consultas a ocorrência de informações semânticas ricas e uma melhora no posicionamento do site.*

## 1. Introdução

O uso da *Web* como meio de comunicação e interação impulsionou a produção de conteúdo, o que levou a necessidade de mecanismos capazes de encontrar uma informação precisa dentre todas as outras, conforme destaca (Pickler, 2007). Com a finalidade de atender essa demanda, surgiram os motores de busca, tais como *Google*<sup>1</sup>, *Yahoo*<sup>2</sup> e *Bing!*<sup>3</sup>, que baseiam-se em robôs lançados na internet pelo servidor da ferramenta com o objetivo de registrar os dados do maior número de documentos possível, para indexá-los a sua base de dados (Cendón, 2001).

Para (Ledford, 2008), a maioria das pessoas utiliza um motor de busca para começar sua navegação e encontrar as informações que procuram. Assim, é evidente que, para uma página conseguir o maior número de acessos possíveis, ela precisa aparecer entre as primeiras posições nos resultados dos motores de busca. Para isso, faz-se necessário o uso de técnicas de *Search Engine Optimization (SEO)*, que buscam otimizar as aplicações *Web* para o alcance de melhores resultados em buscas orgânicas (sem pagamento). Porém, novas tendências e desafios surgem periodicamente e o uso de técnicas *SEO* tende a não ser suficiente para a obtenção de bons resultados nas buscas. (Cendón, 2001) relatou que um dos problemas mais graves enfrentados pelos motores de busca consiste na falta de um tratamento adequado ao conteúdo, como a ausência de linguagens de marcação que possam expressar de forma eficiente o significado dos conteúdos. Contudo, novas tecnologias de marcação estruturada surgiram nos últimos anos e viabilizaram novos tipos de buscas, mais semânticas e ricas em conteúdos.

A *Web Semântica*, segundo (Dziekaniak e Kirinus, 2004), é a evolução da *Web* atual, projetada para ser compreendida não só por humanos, mas também pelas máquinas. Isto é possível por meio de marcações estruturais que podem operar eficientemente sobre os conteúdos, com a capacidade de transmitir os seus significados, auxiliando os usuários nas diversas atividades que eles desempenham na *Web*, assim como na cooperação entre eles.

Com o objetivo de melhorar a visibilidade do curso de Ciência da Computação do IFSULDEMINAS – Câmpus Muzambinho, assim como acompanhar as tendências *Web*, o *website* do curso, criado em 2012, passou a ser monitorado em Fevereiro de 2014 pela ferramenta *Google Analytics*. Em análises de monitoramento realizadas no ano de 2014, observou-se que a maioria dos acessos vinha de Muzambinho e região, o que implica em ações de otimização de forma a ampliar o alcance do site do curso.

Dentre as medidas a serem tomadas, algumas foram consideradas primordiais, como a reformulação do Site, tanto *front-end*, com design responsivo, quanto *back-end*, com framework sob a arquitetura *MVC (Model-View-Controller)*. A Engenharia reversa do banco de dados, para abranger novos processos da aplicação, também foi conduzida. E, como meta principal, a adoção de técnicas *SEO* e marcação semântica estruturada de páginas *Web*, de forma a qualificar (ou não) o site para os motores de busca.

Considerando o contexto descrito anteriormente, este trabalho teve como objetivo avaliar a otimização em buscas semânticas do *website* do curso de Ciência da

---

<sup>1</sup> <https://www.google.com.br/>

<sup>2</sup> <https://br.yahoo.com/>

<sup>3</sup> <http://www.bing.com/?setlang=pt-BR>

Computação – IFSULDEMINAS – Câmpus Muzambinho, aplicando técnicas de SEO e marcação estruturada semântica do dicionário *schema*<sup>4</sup> em sua reformulação.

## 2. Revisão de Literatura

### 2.1 Motores de Busca e a Web Semântica

Segundo (Quintana, 2012), em geral os motores de busca realizam, rotineiramente, uma verificação incremental da web, em busca de índices relevantes que podem ser acessados posteriormente, para auxiliar em consultas de usuários. A página de resultados de um mecanismo de busca, *Search Engines Results Pages (SERP)*, é dividida em duas categorias: resultados patrocinados (pagos) e orgânicos (obtidos por meio de técnicas de SEO). Para (Soares, 2006), os motores de busca de 3ª geração têm necessidades especiais, como ter conhecimento semântico do conteúdo de um documento. Com isso, conseguem devolver resultados ricos em conteúdo, conhecidos como *Rich Snippets*. De acordo com o *Website* de suporte para desenvolvedores do Google<sup>5</sup>, os *Rich Snippets* tornam a tarefa de compreender o conteúdo de uma página algo mais intuitivo, pois fornecem uma noção do conteúdo de uma página específica.

Em uma pesquisa feita em junho de 2010, por (Goel e Gupta, 2010), para estimar a quantidade de marcação semântica e estruturada que existia na *Web*, uma amostra aleatória de um milhão de páginas foi coletada. Posteriormente, eles examinaram quantas páginas apresentavam *rich snippets*. Poucas páginas utilizavam marcação semântica e, além disso, uma fração ainda menor estava sendo usada para a exibição dos *rich snippets*. A única exceção foram os tipos de marcação destinados a análises, onde a presença de *rich snippets* apresentou altos índices. Os *rich snippets* são resultados da marcação de conteúdo semântico e asseguram a existência de motores de busca mais eficientes no atendimento aos usuários. Os motores de busca tendem a utilizar dados estruturados como fator de relevância de um *Website* (Ronallo, 2012).

### 2.2 Técnicas de Marcação Estruturada

Dentre as tecnologias utilizadas para marcação estruturada semântica de conteúdo pode-se citar: *XML*, *RDF*, *Microformats* e *Microdata*.

A linguagem de marcação XML permite a criação de tags personalizadas para a marcação dos conteúdos, de forma similar ao HTML (Figura 1). Contudo, segundo (Filho e Lóscio, 2009), possui uma grande desvantagem, uma vez que a mesma informação pode ser estruturada de várias formas diferentes, levando ao problema de falta de padrão na extração das informações.

<p>a) &lt;!--HTML--&gt; &lt;TABLE BORDER=1&gt; &lt;TR&gt; &lt;TH&gt;Professor ID&lt;/TH&gt; &lt;TH&gt;Nome&lt;/TH&gt; &lt;TH&gt;Área&lt;/TH&gt; &lt;TR&gt; &lt;TD&gt;12345&lt;/TD&gt; &lt;TD&gt;Fernando Wagner&lt;/TD&gt; &lt;TD&gt;Banco de Dados&lt;/TD&gt; &lt;TR&gt; &lt;/TABLE&gt;</p>	<p>&lt;!--XML--&gt;  &lt;Professor&gt; &lt;id&gt;12345&lt;/id&gt; &lt;Nome&gt;Fernando Wagner&lt;/Nome&gt; &lt;Área&gt;Banco de Dados&lt;/Área&gt; &lt;/Professor&gt;</p>	<p>b) &lt;Professor id='12345'&gt; &lt;Descricao&gt; &lt;Nome&gt;Fernando Wagner&lt;/Nome&gt; &lt;Area&gt;Banco de Dados&lt;/Area&gt; &lt;/Descricao&gt; &lt;/Professor&gt;</p>
--	---	---

<sup>4</sup> <http://schema.org/docs/documents.html>

<sup>5</sup> <https://support.google.com/Webmasters/answer/99170?hl=pt-BR>

**Figura 1. Exemplos de marcação utilizando XML (Filho e Lóscio, 2009). (a) Exemplo para representação dos dados de um professor. (b) Forma alternativa de marcação XML para o mesmo conteúdo da figura 1a.**

Na figura 1b é possível observar o problema da falta de padrão, onde a mesma informação mostrada na figura 1a é apresentada de forma diferente. Tais problemas levaram a iniciativas que buscavam estabelecer um dicionário padrão para a marcação semântica na *Web*, como o *Resource Description Framework (RDF)*.

O *RDF* é escrito em *XML*, representa metadados em forma de sentenças sobre propriedades e seus relacionamentos. Para (Morais e Soares, 2006), o *RDF* utiliza *XML* para representar dados em forma de triplas no formato <objeto, atributo, valor>. Apresentava diversas qualidades, porém não obteve o suporte necessário por parte de grandes empresas do setor, principalmente os motores de busca. Surgiram os Microformatos, uma coletânea de especificações que busca estabelecer as relações entre informações e agentes humanos/computacionais. Eles têm como objetivo resolver os problemas de menor complexidade, pois descrevem a informação de modo mais simples. Apesar dos esforços da comunidade *Web*, ainda não havia uma tecnologia padrão para a marcação de conteúdo semântico na *Web*. Foi quando surgiram os *Microdatas*.

Os *Microdatas* utilizam diversos vocabulários para descrição semântica com pares chave-valor, para atribuir valores as suas respectivas propriedades. O uso de uma técnica de marcação estruturada baseada em *microdata* é, segundo (Ronallo, 2012), a forma mais óbvia para a obtenção de *rich snippets* em buscas realizadas na *Web*. Para (Mika e Potter, 2012) em *microdata*, apenas dois vocabulários (*schema.org* e *datavocabulary.org*) ganharam força significativa até o presente momento, e a tendência é que o último seja substituído pelo primeiro.

De acordo com (Pastore, 2012), o *Schema* fornece um vocabulário que adiciona informações ao conteúdo *HTML*, padronizando um conjunto de esquemas para marcação estruturada em páginas *Web*. De acordo com (Ronallo, 2012), os motores de busca (*Bing, Google, Yahoo!*) conceberam o *Schema.org* e concordaram em apoiá-lo e entendê-lo, pois seria inviável oferecer suporte a todo o vocabulário em uso na *Web*, tornando-se, assim, numa tentativa de definir um amplo (na escala da *Web*) vocabulário compartilhado e com foco em conceitos populares. Consideremos o seguinte exemplo de conteúdo marcado utilizando o dicionário *schema*: imagine que um usuário possui uma página sobre o filme *Avatar*, com um *link* para o *trailer*, informação sobre o diretor, dentre outros dados.

Em *schema*, para identificar qual parte da página fala sobre o filme *Avatar*, deve-se adicionar um elemento “*itemscope*” dentro da *tag HTML* que contém as informações do item. Porém o tipo de item deve ser especificado, e para isso basta adicionar “*itemtype*” logo após “*itemscope*”. Tipos de itens são fornecidos em forma de *URL's*, neste caso <http://schema.org/Movie>. Informações adicionais podem ser dadas por meio da classificação de itens de propriedade “*itemprop*”.

De acordo com (Ronallo, 2012), *microdata* com *Schema.org* é atualmente o método preferido para o uso de dados estruturados em páginas *Web*, enquanto outros podem surgir no futuro, atualmente esta técnica apresenta os resultados mais convincentes, especialmente para obtenção de *rich snippets*.

No entanto é importante destacar que, segundo a página de perguntas e respostas para desenvolvedores da empresa, o *Google*<sup>6</sup> não oferece garantias de que os *rich snippets* sejam mostrados nos resultados de pesquisa de um site específico, mesmo que possam ser corretamente extraídos por sua ferramenta de teste<sup>7</sup>.

### 3. Materiais e Métodos

#### 3.1 Monitoramento

O *Website* do curso de Ciência da Computação do IFSULDEMINAS – Câmpus Muzambinho, selecionado para o estudo de caso, foi cadastrado no serviço *Google Analytics*<sup>8</sup>, com o objetivo de gerar relatórios de monitoramento do tráfego e guiar o planejamento visando a otimização. Além disso, através das Ferramentas do Google para *Webmasters*<sup>9</sup> foi possível acompanhar o andamento do processo de indexação pelo Google, fazendo correções quando necessário.

#### 3.2 Seleção de Palavras-chave

Foram elaborados alguns questionários com o intuito de levantar informações relevantes ao desenvolvimento da nova versão do *Website*, estes buscavam conhecer as necessidades dos usuários e possíveis usuários, assim como as melhorias que poderiam ser feitas em uma possível atualização.

Os questionários foram aplicados após o término da prova no vestibular interno da instituição para vestibulandos do curso de Ciência da Computação, e em outro momento também para os discentes do curso, totalizando 104 pessoas. Dentre as questões com foco nas palavras-chave que os usuários utilizam ou utilizariam para uma eventual busca, as palavras de maior frequência para encontrar o curso de computação no sul de Minas Gerais foram: Computação (80), sul de minas (38), ciência (17), muz (12), ifsuldeminas (10), bacharelado (5), muzambinho (5), ccomp(5). Quanto à presença de *rich snippets* nos resultados de buscas relacionadas ao *Website* do curso, 92% das pessoas informaram que os resultados não apresentavam *rich snippets*.

Através da Teoria da Competitividade de Felipini (2010) foi possível estipular quais termos teriam a menor concorrência nos motores de busca. Adotou-se o critério de menor concorrência em conjunto com o bom senso no sentido de escolher as palavras que melhor definissem o Website utilizado como estudo de caso. Desse modo, as seguintes palavras foram escolhidas: ifsuldeminas, campus muzambinho, ccomp, bacharelado, tecnologia da informação, curso de computação, ciência da computação. O conteúdo das páginas foi orientado para o uso de tais palavras. Este alinhamento teve o objetivo de contribuir com a otimização do Website.

#### 3.3 Aplicação de Técnicas de SEO

As técnicas de SEO foram aplicadas na reformulação do site, sendo elas: otimização de imagens e CSS; uso de *metatags* dinâmicas para descrição, autor do conteúdo e palavras-chave; menu rápido para acesso direto a certas áreas da página com apenas um clique; uso de *breadcrumbs* para orientação durante a navegação; arquivo

<sup>6</sup> <http://www.google.com>

<sup>7</sup> <https://developers.google.com/structured-data/testing-tool/>

<sup>8</sup> <http://www.google.com/analytics/>

<sup>9</sup> <https://www.google.com/Webmasters/tools/home?hl=pt-BR>

*sitemaps.xml* que auxilia os motores de busca a encontrarem as *URL*'s pertencentes ao website; criação do arquivo *robots.txt* indicando os diretórios que devem ou não ser indexados pelos buscadores; uso de conteúdo relevante e original; design responsivo, com isso, o website passou a adaptar-se ao dispositivo em que está sendo exibido, inclusive *smartphones*.

### 3.4 Aplicação de Marcação Estruturada Semântica

Nessa etapa, o objetivo foi aplicar a marcação semântica utilizando os vocabulários *schema.org*. No caso de eventos acadêmicos, existem algumas peculiaridades, como a necessidade de fornecer data de início do evento, término e local.

O *schema* possui um vocabulário considerável para marcação estruturada de informações, porém quando certo tipo não possui suporte específico é necessário recorrer a um vocabulário genérico, por exemplo, na seção de documentos, foi necessário marcar o conteúdo como “*Creative Work*” na falta de algo que descreva de forma mais específica esse tipo de conteúdo. Certas páginas podem conter maior nível de detalhes na marcação estruturada, onde um vocabulário específico engloba outro a fim de enriquecer o conteúdo semântico. Esse comportamento pode ser visto na Figura 2, onde a propriedade “*address*” poderia conter apenas um texto normal. Ao invés disso, é enriquecida com o tipo “*PostalAddress*”, adquirindo um nível maior de especialização.

```
<p>
  <span itemprop="address" itemscope itemtype="http://schema.org/PostalAddress">
    <strong>Local de Funcionamento: </strong> <br>
    <span itemprop="streetAddress">Prédio Pedagógico de Informática Estrada de Muzambinho, Km 35,
    <span itemprop="telephone">(35) 3571-5130</span><br>
    Caixa Postal: <span itemprop="postOfficeBoxNumber">02</span><br>
    <span itemprop="addressLocality">Muzambinho</span>,
    <span itemprop="addressRegion">Minas Gerais</span> -
    <span itemprop="addressCountry">Brasil</span><br>
    CEP: <span itemprop="postalCode">37890-000</span>
  </span>
</p>
```

Figura 2. Marcação semântica utilizando schema com informações do curso

### 3.5 Análise e Avaliação da Otimização para Mecanismos de Buscas

Diversas ferramentas foram utilizadas para a análise e avaliação de técnicas *SEO*. Uma delas é o *app* para o navegador *Chrome* conhecido como *SEOquake*<sup>10</sup>, com esse aplicativo é possível extrair diversos dados sobre a indexação do site a ser monitorado, como *Page Rank*<sup>11</sup>, número de páginas indexadas até o momento, popularidade nas redes sociais e etc. No entanto, a capacidade de listar palavras-chave de maior densidade e o diagnóstico de otimização para motores de busca foram utilizados com maior frequência. Como complemento ao *SEOquake*, outra aplicação para o navegador *Chrome* conhecido como *SEO e Análise Web*<sup>12</sup> da empresa *WooRank* foi utilizada, com ela foi possível realizar uma análise *SEO* mais abrangente com relação ao *SEOquake*.

### 3.6 Análise de Marcação Estruturada Schema

A marcação semântica em *schema.org* presente no código foi validada por meio da ferramenta *Structured Data Testing Tool*<sup>13</sup> oferecida pelo *Google*. Ao inserir o código

<sup>10</sup> <http://www.seoquake.com/>

<sup>11</sup> Índice utilizado pelo Google para medição de relevância, numa escala de 0 a 10.

<sup>12</sup> <https://chrome.google.com/Webstore/detail/seo-Website-analysis/hlmgmdolgbdnmbmlfhhndibidipaf?hl=pt-BR>

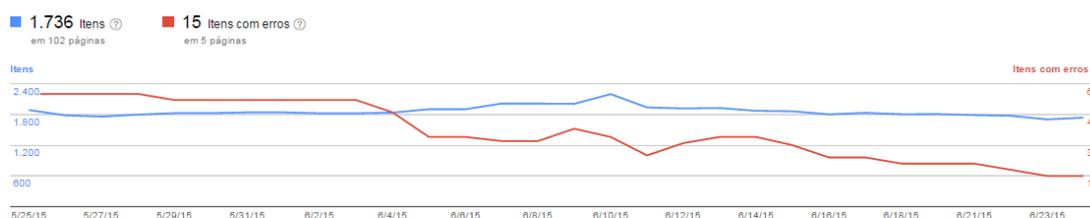
<sup>13</sup> <https://developers.google.com/structured-data/testing-tool/>

na ferramenta ela passa a varrer a página em questão a procura de conteúdo semântico, retornando os itens que foram identificados corretamente, ou seja, atendem aos requisitos. Além disso, também indica possíveis erros na aplicação da marcação estruturada, como a falta de uma propriedade obrigatória ou presença de atributos embutidos que não pertencem a um vocabulário em especial. Os itens identificados são ordenados e exibidos ao usuário para que ele possa inspecionar como a marcação semântica de suas páginas está sendo compreendida pelos buscadores.

Como complemento, utilizou-se o módulo das Ferramentas do *Google* para *Webmasters* responsável por mostrar estatísticas da marcação estruturada. Este módulo auxiliou no processo de gerenciamento dos dados estruturados contidos nas páginas do *Website*, exibindo relatórios dos itens identificados, inclusive aqueles que contêm erros.

#### 4. Resultados e Discussão

O website reformulado com o conteúdo marcado semanticamente foi publicado. De acordo com análises feitas em ferramentas do *Google para Webmasters*, foram identificados 1736 itens semanticamente corretos. Na figura 3 é possível observar a queda brusca de itens semanticamente incorretos e uma manutenção dos itens corretos.



**Figura 3. Representação do número de itens semanticamente corretos e incorretos do website do curso, no período de 25 de maio a 23 de junho de 2015.**

Para avaliar a influência das técnicas de SEO e marcação estruturada sobre a otimização em motores de busca, foram comparados os períodos de monitoramento via ferramenta *Google Analytics*, antes e após a reformulação do site. O primeiro período (antes da reformulação) vai de 15 de março a 3 de maio de 2015 (Figura 4) e o segundo período (após a reformulação) vai de 4 de maio a 23 de junho de 2015 (Figura 5). Ambos compreendem aproximadamente 50 dias. A Figura 6 apresenta um gráfico comparativo entre os resultados das figuras 4 e 5. Alguns itens apresentaram um aumento considerável, como o número de sessões que passou de 395 para 953 (aumento de 141%). O número de usuários saltou de 263 para 496 (aumento de 88%), e o tempo médio de duração da sessão passou de 3 minutos e 17 segundos para 5 minutos e 30 segundos (aumento de 67%).



**Figura 4. Relatório de monitoramento do período entre 15 de março e 3 de maio de 2015.**

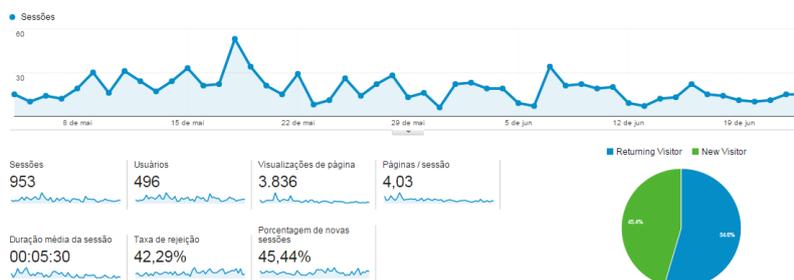


Figura 5. Relatório de monitoramento do período entre 4 de maio e 23 de junho de 2015.

Ainda na figura 6 é possível observar que a taxa de rejeição teve um leve aumento, levando-se em consideração o incremento expressivo do número de usuários, sessões e visualizações de páginas. Sobre os itens que apresentaram diminuição (páginas/sessão e novas sessões) é necessário fazer algumas observações. O *Website* passou por um período sem atualizações ao final do ciclo de monitoramento (15 dias) que influenciou negativamente o número de páginas por sessão, pois não havia conteúdo novo a ser consumido por seus usuários nesse período. Quanto à queda no índice de porcentagem de novas sessões, considerou-se que o período antes da atualização coincidiu com a entrada de novos alunos ao curso, aumentando consideravelmente a porcentagem de novas sessões em comparação ao período após a atualização.

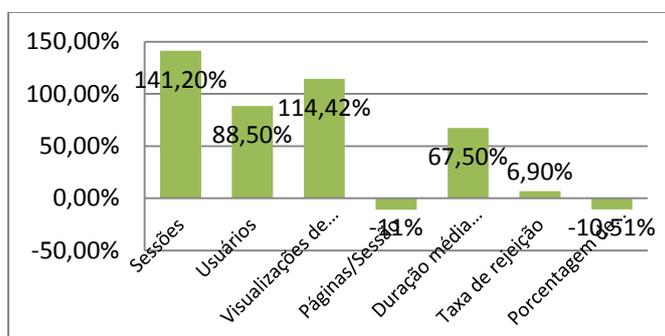


Figura 6. Gráfico comparativo entre os resultados dos monitoramentos antes e após a reformulação do site.

Com relação a acessos provenientes de buscas no período, foram totalizados 249 cliques ao *Website* do curso. O número de impressões, ou seja a quantidade de links do *Website* que foram visualizadas pelos usuários na *SERP* do Google, mesmo que não tenham sido exibidos na rolagem, totalizaram 3.869 impressões. O *Click Through Rate (CTR)* ou taxa de cliques consiste na contagem de cliques dividida pelo número de impressões e obteve o percentual de 6.44%. A oscilação diária destes índices pode ser vista na Figura 7.

Total de cliques	Total de impressões	CTR méd.
249	3.869	6,44%



Figura 7. Tráfego de pesquisa durante o período de monitoramento.

No que se refere às alterações no *ranking* da página de resultados do *Google*, observou-se que ao pesquisar as palavras “ciência da computação ifsuldeminas” o site aparecia na terceira ou quarta posição da primeira página, com as melhorias aplicadas ele passou para a segunda posição, e no final do ciclo de monitoramento passou a ocupar a primeira posição da primeira página na *SERP* do *Google*. O resultado não apresentou um impacto tão relevante, visto que o site já aparecia bem ranqueado pelo *Google*, mesmo sem o uso das técnicas de SEO. Isso é explicado, certamente, pelo fato de que o *website* do curso (*PageRank* 3) é subdomínio do site da instituição (IFSULDEMINAS), que por sua vez possui grande relevância (*PageRank* 6) perante os motores de busca. Um domínio de grande relevância apontando para um subdomínio com relevância menor acaba contribuindo para o posicionamento do segundo. Caso contrário a escalada de posições seria possivelmente maior.

Sobre o suporte a *rich snippets*, apenas cerca de um mês depois esse recurso tornou-se disponível na página de agendas do curso (Figura 8b). Depois de 40 dias, outro *rich snippet* de um evento (Figura 8a) passou a ser exibido nos resultados de busca.



**Figura 8. Rich snippets de Eventos para a página de agendas do site. (a) Rich snippet para agenda. (b) Rich snippet para evento. (c) Rich snippet para localização e fone no buscador Bing! (d) Rich snippet com informações a página na rede social Google Plus e a localização no Google Maps.**

Embora o vocabulário *schema* seja um acordo entre os maiores motores de busca do mercado, aqueles que possuem uma fatia menor não tendem a se concentrar atualmente no suporte aos *rich snippets*. O *Bing* apresentou este recurso para o *Website* utilizado no estudo de caso somente para informações de localização e contato presentes na marcação estruturada da aplicação (Figura 8c). O buscador *Yahoo* não havia atualizado suas informações referentes ao *website* até o término do monitoramento.

Adicionalmente, o site obteve, por meio do cadastro e verificação física no *Google Meu Negócio*, um *rich snippet* para a exibição customizada com informações e fotos do perfil do *Website* na rede social *Google Plus*, assim como a criação de um local físico no *Google Maps* (Figura 8d).

## 5. Considerações Finais

Este trabalho teve o objetivo de investigar, analisar e avaliar a aplicação de técnicas de SEO e marcação estruturada de conteúdo semântico em um estudo de caso com o site do curso de Ciência da Computação do IFSULDEMINAS – Câmpus Muzambinho, de forma a ampliar o alcance do mesmo por meio da obtenção de um melhor posicionamento e informações ricas em resultados de consultas em mecanismos de

buscas. De acordo com as análises e avaliações feitas foi possível perceber o impacto positivo da marcação semântica de conteúdo, por meio do enriquecimento das informações vindas nas buscas, além de uma ligeira melhora no ranqueamento do site nos resultados de consultas.

Espera-se, ainda, que este trabalho sirva de inspiração para outros estudantes da área, de forma que possam realizar estudos semelhantes no sentido da avaliação do uso de marcação semântica e técnicas de SEO na construção de páginas para a Web. É importante também um monitoramento constante do *Website*, de forma a avaliar sob outros prazos os impactos dessas técnicas.

## Referências

- Cendón, B. V. (2001) “Ferramentas de busca na Web”, *Ciência da Informação*, 30 (1), 39-49.
- Dziekaniak, G. V. & Kirinus, J. B. (2004). “Web semântica”, *Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação*, Florianópolis, 18.
- Filho, F. W. & Lóscio, B. F. (2009). *Web Semântica: Conceitos e Tecnologias*.
- Goel, K.; Gupta, P. (2010). Google rich snippets. In: *Semantic Technology Conference*, 21-25.
- Felipini, D. (2010). *Google top 10: Como colocar seu site ou blog na primeira página do Google*. Rio de Janeiro: Brasport, 112p.
- Ledford, J. L. (2008). *SEO Search Engine Optimization Bible*, Indianapolis: Wiley Publishing Inc., 381 p.
- Mika, P., & Potter, T. (2012). Metadata Statistics for a Large Web Corpus, *LDOW*, 937.
- Morais, E. F., & Soares, M. B. (2006). *Web Semântica para Máquinas de Busca*.
- Pastore, S. (2012). Website development and Web standards in the ubiquitous world: where are we going. *WSEAS Transactions on Computers*, 11, 309-318.
- Pickler, M. E. V. (2007). Web Semântica: ontologias como ferramentas de representação do conhecimento. *Perspectivas em Ciência da Informação*, 12 (1), 65-83.
- Quintana, A. D. C. (2012). Avaliação das técnicas de otimização para motores de busca.
- Ronallo, J. (2012). HTML5 Microdata and Schema. org. *Code4Lib Journal*, 16.