



CENTRO UNIVERSITÁRIO LUTERANO DE PALMAS

Recredenciado pela Portaria Ministerial nº 1.162, de 13/10/16, D.O.U. nº 198, de 14/10/2016

AELBRA EDUCAÇÃO SUPERIOR - GRADUAÇÃO E PÓS-GRADUAÇÃO S.A.

DAVID XAVIER BRITO

**UTILIZAÇÃO DE LINKED DATA NO PROCESSO DE ENRIQUECIMENTO DE
CONTEÚDO DE PÁGINAS DO PORTAL ACADÊMICO DO CEULP/ULBRA**

Palmas – TO

2020

DAVID XAVIER BRITO

**UTILIZAÇÃO DE LINKED DATA NO PROCESSO DE ENRIQUECIMENTO DE
CONTEÚDO DE PÁGINAS DO PORTAL ACADÊMICO DO CEULP/ULBRA**

Trabalho de Conclusão de Curso (TCC) II elaborado e apresentado como requisito parcial para obtenção do título de bacharel em Sistemas de Informação pelo Centro Universitário Luterano de Palmas (CEULP/ULBRA).

Orientador: Prof. M.e Jackson Gomes de Souza.

Palmas – TO

2020

DAVID XAVIER BRITO

**UTILIZAÇÃO DE LINKED DATA NO PROCESSO DE ENRIQUECIMENTO DE
CONTEÚDO DE PÁGINAS DO PORTAL ACADÊMICO DO CEULP/ULBRA**

Trabalho de Conclusão de Curso (TCC) II elaborado e apresentado como requisito parcial para obtenção do título de bacharel em Sistemas de Informação pelo Centro Universitário Luterano de Palmas (CEULP/ULBRA).

Orientador: Prof. M.e Jackson Gomes de Souza.

Aprovado em: ____/____/____

BANCA EXAMINADORA

Prof. M.e Jackson Gomes de Souza
Orientador
Centro Universitário Luterano de Palmas – CEULP

Prof. Dra. Parcilene Fernandes de Brito
Centro Universitário Luterano de Palmas – CEULP

Prof. M.e Fabiano Fagundes
Centro Universitário Luterano de Palmas – CEULP

Palmas – TO

2020

RESUMO

BRITO, David Xavier. **Utilização de Linked Data no Processo de Enriquecimento De Conteúdo de Páginas do Portal Acadêmico do CEULP/ULBRA**. 2020. 69 f. Trabalho de Conclusão de Curso (Graduação) – Curso de Sistemas de Informação, Centro Universitário Luterano de Palmas, Palmas/TO, 2020[1].

Resumo

O surgimento das redes de comunicação favoreceu a rápida disseminação de informação no mundo. Dessa forma, foi implementado padrões técnicos que não apenas facilitam a troca de informações, mas também criam uma linguagem mais comum entre equipamentos e sistemas de informação. Para facilitar essa troca de informações, foi desenvolvido o conceito de Web semântica, que é a extensão atual da Web, onde a informação tem significado claro e definido, com intuito de melhorar a interação entre computadores e pessoas. Nessa perspectiva, o presente trabalho apresenta uma proposta de utilização de metadados do Schema.org para enriquecer o conteúdo de páginas do Portal Acadêmico do CEULP/ULBRA, de modo que auxilie motores de busca na extração de conteúdo das páginas de forma estruturada. Para isso foi introduzido microdados no HTML5 agregando informações diretamente em tags, por meio de vocabulários personalizados com algumas propriedades nomeadas no que diz respeito aos Microdados. Os elementos que compõem uma página HTML podem ser vistos como itens que possuem um ou mais pares compostos, que são nome/valor ou até mesmo propriedade/valor. Notou-se que é possível a utilização de schema.org para a representação de dados estruturados de páginas HTML, utilizando para isso documentações gratuitas disponíveis na internet. Entretanto, algumas limitações foram identificadas no âmbito deste trabalho, devido algumas propriedades do schema.org não atenderem a modelagem dos dados estruturados para as páginas do portal.

Palavras chaves: linked data, microdata, dados estruturados.

LISTA DE FIGURAS

INTRODUÇÃO	6
1 REFERENCIAL TEÓRICO	8
1.1 Web Semântica	8
2.2 Linked Data	14
2.2.1 <i>Hypertext Transfer Protocol (HTTP) e Uniform Resource Identifier (URI)</i>	16
2.2.2 <i>Resource Definition Framework (RDF)</i>	18
2.2.3 <i>RDF/XML</i>	20
2.2.4 <i>Microformatos FORTI</i>	22
2.2.5 <i>Resource Definition Framework attributes (RDFa)</i>	25
2.2.6 <i>JavaScript Object Notation (JSON)-Linked Data (LD)</i>	26
2.2.7 <i>Microdata</i>	27
3 TRABALHOS RELACIONADOS	29
4 MATERIAIS E MÉTODOS	30
4.1 Materiais	31
4.2 Métodos	31
5 RESULTADOS E DISCUSSÃO	33
5.1 Página Professor	34
5.1.1 <i>Modelagem da Página Professor</i>	34
5.1.2 <i>Implementação da Página Professor</i>	37
5.1.3 <i>Validação da Página Professor</i>	39
5.2 Página Home	43
5.2.1 <i>Modelagem da Página Home</i>	43
5.2.2 <i>Implementação da Página Home</i>	51
5.2.3 <i>Validação de dados da página Home</i>	57
6 CONSIDERAÇÕES FINAIS	62
REFERÊNCIAS	64

LISTA DE TABELAS

INTRODUÇÃO	6
1 REFERENCIAL TEÓRICO	8
1.1 Web Semântica	8
2.2 Linked Data	14
2.2.1 <i>Hypertext Transfer Protocol (HTTP) e Uniform Resource Identifier (URI)</i>	16
2.2.2 <i>Resource Definition Framework (RDF)</i>	18
2.2.3 <i>RDF/XML</i>	20
2.2.4 <i>Microformatos FORTI</i>	22
2.2.5 <i>Resource Definition Framework attributes (RDFa)</i>	25
2.2.6 <i>JavaScript Object Notation (JSON)-Linked Data (LD)</i>	26
2.2.7 <i>Microdata</i>	27
3 TRABALHOS RELACIONADOS	29
4 MATERIAIS E MÉTODOS	30
4.1 Materiais	31
4.2 Métodos	31
5 RESULTADOS E DISCUSSÃO	33
5.1 Página Professor	34
5.1.1 <i>Modelagem da Página Professor</i>	34
5.1.2 <i>Implementação da Página Professor</i>	37
5.1.3 <i>Validação da Página Professor</i>	39
5.2 Página Home	43
5.2.1 <i>Modelagem da Página Home</i>	43
5.2.2 <i>Implementação da Página Home</i>	51
5.2.3 <i>Validação de dados da página Home</i>	57
6 CONSIDERAÇÕES FINAIS.....	62
REFERÊNCIAS	64

SUMÁRIO

1 INTRODUÇÃO	6
2 REFERENCIAL TEÓRICO	8
2.1 Web Semântica	8
2.2 Linked Data	14
2.2.1 <i>Hypertext Transfer Protocol (HTTP) e Uniform Resource Identifier (URI)</i>	16
2.2.2 <i>Resource Definition Framework (RDF)</i>	18
2.2.3 <i>RDF/XML</i>	20
2.2.4 <i>Microformatos</i>	22
2.2.5 <i>Resource Definition Framework attributes (RDFa)</i>	25
2.2.6 <i>JavaScript Object Notation(JSON)-Linked Data(LD)</i>	26
2.2.7 <i>Microdata</i>	27
3 TRABALHOS RELACIONADOS	29
4 MATERIAIS E MÉTODOS	30
4.1 Materiais	30
4.2 Métodos	31
5 RESULTADOS E DISCUSSÃO	33
5.1 Página Professor	34
5.1.1 <i>Modelagem da Página Professor</i>	34
5.1.2 <i>Implementação da Página Professor</i>	37
5.1.3 <i>Validação da Página Professor</i>	39
5.2 Página Home	43
5.2.1 <i>Modelagem da Página Home</i>	43
5.2.2 <i>Implementação da Página Home</i>	51
5.2.3 <i>Validação de dados da página Home</i>	57
6 CONSIDERAÇÕES FINAIS	62
REFERÊNCIAS	64

INTRODUÇÃO

O surgimento das redes de comunicação favoreceu a rápida disseminação de informação no mundo. Por exemplo, a Internet e a World Wide Web (em inglês: **WWW, A Web**) criaram informações que podem ser usadas por qualquer pessoa a qualquer hora e em qualquer lugar, por meio de celulares e computadores que podem acessar a Internet incluindo a própria Web semântica. Dessa forma, foi implementado padrões técnicos que não apenas facilitam a troca de informações, mas também criam uma linguagem mais comum entre equipamentos e sistemas de informação (BRANDÃO et al., 2002).

Para facilitar essa troca de informações, foi desenvolvido o conceito de Web semântica, que é a extensão atual da Web, onde a informação tem significado claro e definido, com intuito de melhorar a interação entre computadores e pessoas. Esse conceito apresenta uma solução onde a máquina possa auxiliar a humanidade de forma mais efetiva. Contudo, é necessário constituir ferramentas que gerem para máquinas significado lógico e semântico (BERNERS-LEE, 2001).

De acordo com Hendler (2001) a Web Semântica tem foco na estruturação semântica de informações dispostas na Web, de forma que humanos e máquinas possam compreender os dados. Outras possibilidades estão relacionadas à pretensão de fornecer estruturas que possam dar mais sentido aos conteúdos Web. Esta ferramenta é uma extensão da internet atual onde a informação tem significados específicos para facilitar a compreensão de humanos e máquinas (BERNERS-LEE, 2001).

A Web Semântica consiste em implementar padrões tecnológicos para permitir não apenas uma maior facilidade na troca de informações, mas criar uma língua mais comum entre dispositivos e sistemas de informação (SOUZA; ALVARENGA, 2004). Para facultar essa troca de informações, foram desenvolvidos princípios que permitem a criação e a ligação de dados, no intuito de viabilizar o acesso, o uso e o recurso a humanos e não humanos (ARAKAKI, 2016). Esses princípios, denominados *Linked Data*, facilitam a busca e direcionam em diferentes bases a partir dos dados ligados.

Segundo Paletta et al (2014) os princípios de *Linked Data* oferecem às instituições formas mais semânticas e capazes de se interoperar juntamente com as melhores ferramentas de gerenciamento de dados. Assim, ampliam a quantidade e a

qualidade da informação e tornam possível aperfeiçoar a recuperação e navegação dos dados de acordo com as necessidades informacionais de seus usuários.

Nessa perspectiva, considerando uma necessidade de um conjunto de indicadores do Marketing Digital que podem ser extraídos do banco de dados acessível no portal acadêmico, este trabalho, tem como objetivo aplicar os conceitos de Web Semântica no portal do CEULP/ULBRA, possibilitando o aumento na quantidade de acessos ao site para cada vez que ele é exibido como resultado no Google e outros motores de busca. Para isso é necessário a inclusão de metadados do schema.org (uma atividade colaborativa da comunidade com a missão de criar, manter e promover esquemas para dados estruturados na Internet, em páginas da Web, em mensagens de e-mail e muito mais), (ex.: página do professor – com informações pessoais, de contato e cursos onde está alocado; seção de calendário de eventos da página inicial).

Assim, o trabalho conta com a seguinte estrutura: a seção 2 apresenta um breve referencial teórico sobre *Linked Data*, indicando conceitos, técnicas e exemplificações; na seção 3 é apresentado os materiais e métodos necessários para utilização de metadados na página do Portal Acadêmico; a seção 4 exhibe os resultados obtidos a partir dos métodos aplicados; a seção 5 apresenta as conclusões a partir dos estudos e resultados obtidos neste trabalho; e por fim são apresentadas as referências utilizadas para construção do trabalho.

1 REFERENCIAL TEÓRICO

1.1 Web Semântica

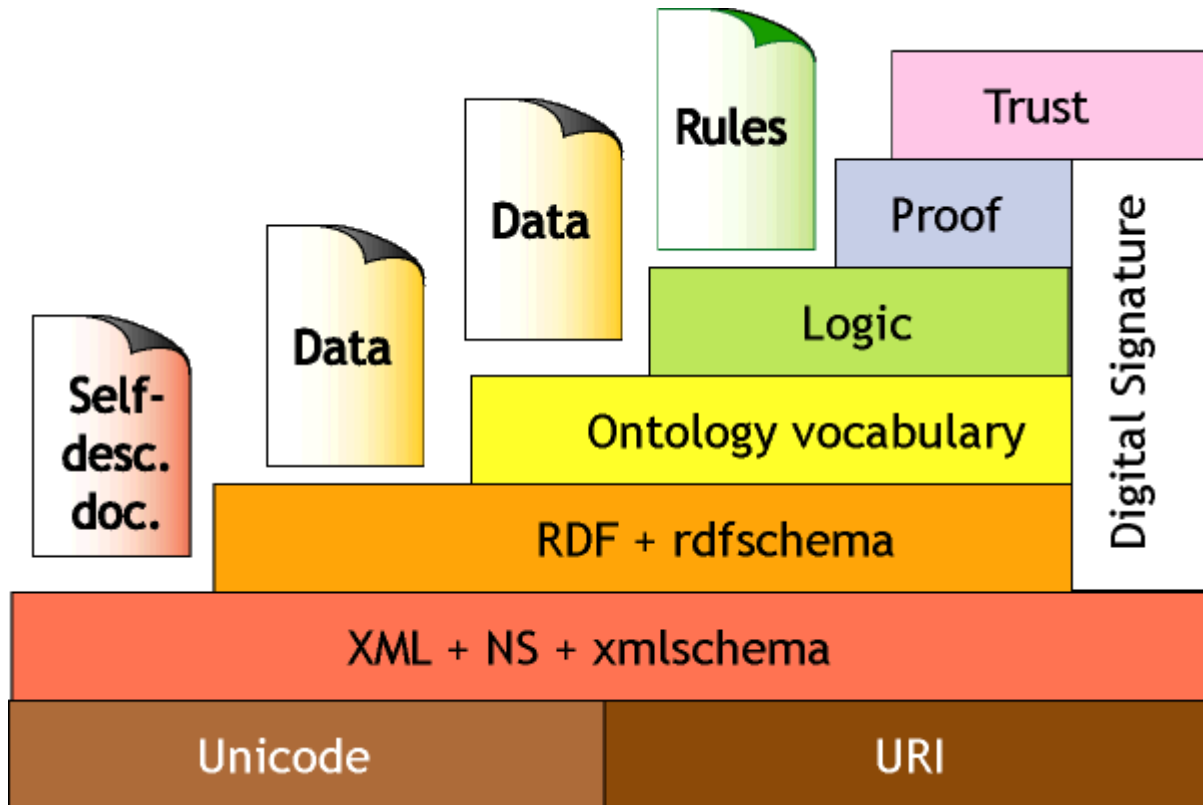
No final do século 20, o surgimento das redes de comunicação e a difusão de informações como a Internet e a Internet global (*World Wide Web*) criaram um banco de dados. Qualquer pessoa pode usar o telefone, o celular e as ferramentas de pesquisa (altavista, yahoo, google, etc.) a qualquer hora e em qualquer lugar. A filtragem de conhecimento com base em meta-informações (como palavras-chave) e a filtragem heurística de informações fornecem a primeira solução para o problema de informações redundantes (BRANDÃO et al., 2002).

O modelo citado anteriormente gera para usuários da Web uma sobrecarga de informação, desorientação e os problemas subsequentes na busca de obtenção de conhecimento. Lins (2013) explica que a internet, criada no início da década de 1960 por cientistas do MIT como um conceito de rede, prometia ser algo inovador para o Departamento de Defesa dos EUA (DARPA). Esta rede funcionava com uma certa quantidade de computadores que mantinham comunicação entre si, ao invés de um sistema de controle centralizado, comunicando-se através de “pacotes”.

Em meados de 1998, começaram a ser formalizadas pesquisas relacionadas ao desenvolvimento da nova geração formalizada da Web (Web Semântica), com o objetivo de combinar *links* semânticos com recursos de informação e, gradativamente, compreender o potencial de utilização desses recursos na informação (RAMALHO, 2010). Bernes-Lee (1998) destaca que o primeiro passo para o desenvolvimento da Web Semântica é a inclusão de dados em um formato de tal modo que os sistemas computacionais possam naturalmente “compreender”. Essa necessidade, onde informações possam ser compreensíveis por máquina, inspirou a primeira pesquisa sobre Web Semântica, publicada no *Journal of Information Science*.

O avanço da Web Semântica está diretamente relacionado com o *World Wide Web Consortium* (W3C), que propõe desenvolvimento e padronização de novas tecnologias baseadas em ambiente Web (RAMALHO, 2006). Assim, em 2000, Bernes-Lee apresenta a primeira proposta de arquitetura semântica da Web conforme ilustra a Figura 1.

Figura 1- Arquitetura da Web Semântica



Fonte: Tim Berners-Lee (2000)¹

¹ Disponível em: <https://www.w3.org/2000/Talks/1206-xml2k-tbl/all.htm>.

A Figura 1 é baseada em uma série de camadas sobrepostas, onde cada camada ou tecnologia deve ser complementar e compatível com a camada inferior, ao mesmo tempo que não deve depender da camada superior, para que possa atingir uma estrutura escalável ideal. Desta forma, a arquitetura visa indicar as etapas e tecnologias necessárias para implementar um projeto de Web Semântica, e apresentar as tecnologias básicas e suas correlações.

Conforme a Figura 1, a primeira camada denominada **característica internacional** é a camada de base da arquitetura da Web Semântica que consiste em *Uniform Resource Identifier* (unificado Identificador de recurso) e UNICÓDICO, que são padrões para descrição e estabelecimento de identificador de recurso universal e código de dados internacionais. Esses dois elementos são responsáveis por estabelecer a identificação mínima de recursos na rede, como a localização do recurso Via URL (*Uniform Resource Locator*) (SANTARÉM SEGUNDO, 2004).

A segunda camada, **camada sintética**, consiste no uso de linguagem XML, *namespace* e XML Schema, A camada é responsável por estabelecer corretamente a

gramática de descrição dos dados. Além disso, a linguagem XML não só fornece recursos informativos melhor estruturados como também inclui dados e metadados que representam estes recursos. Portanto, a linguagem XML é considerada uma linguagem básica porque se concentra mais no conteúdo dos recursos não apenas na forma de apresentação, permitindo que os agentes de software tenham uma melhor visualização de dados (ROSA, 2002).

A terceira camada, **camada de dados**, está diretamente relacionada à apresentação e processamento de metadados. Para este fim, Metadados RDF e *schema* RDF, são as ferramentas responsáveis pela expressão, significado e promoção de interoperabilidade entre metadados e padrões ou formatos Metadados (SANTARÉM SEGUNDO, 2004; MARTINS JÚNIOR, 2003).

A quarta camada, **camada de ontologia**, é responsável por estabelecer o significado dos dados, ou seja, estabelecer a semântica dos dados descritos e representados pelos metadados. Essa camada não é estabelecida apenas com a existência da ontologia, mas também uma comunidade específica, que inclui definição de conceitos usados para representar recursos (SANTARÉM SEGUNDO, 2004).

A camada cinco, **camada lógica**, representa o potencial da Web Semântica porque, com base na hierarquia responsável, constrói, representa e estabelece a semântica dos dados. A camada lógica é responsável por fornecer pesquisa e devido ao uso de agentes de raciocínio, regras e mecanismos, a eficiência de recuperação é maior sobre dados e metadados (ROSA, 2002). Posteriormente, a **camada de prova** é responsável pela troca entre os agentes, está relacionado a várias definições lógicas que serão estabelecidas na camada lógica Lide com o agente para obter evidências (ROSA, 2002).

A última camada da Web Semântica, **camada de validação**, é responsável por estabelecer a verdade, em outras palavras, estabelece autenticidade, confiabilidade e validade dos dados na Web Semântica. Esta camada irá fornecer um *proxy* e garantir informações ou recursos de informação sobre a causa dos dados e se a recuperação é real. (SANTARÉM SEGUNDO, 2004)

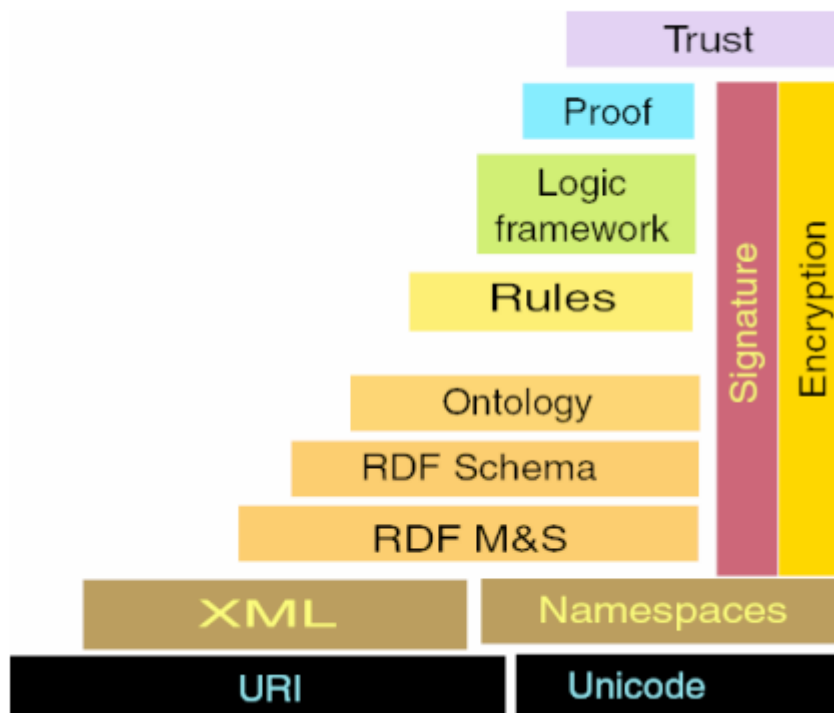
Ainda, de acordo com Hendler (2001), a Web Semântica tem foco na semântica das informações de forma que humanos e máquinas possam compreender os dados. Outras alternativas estão relacionadas à pretensão de fornecer estruturas que possam dar mais sentido semântico aos conteúdos Web. A Web Semântica é uma extensão

da internet atual onde a informação tem significados específicos para facilitar a compreensão de humanos e máquinas (BERNERS-LEE, 2001).

A Web Semântica não é outra Web. Oliveira (2004) apontou que Tim Berners Lee a descreveu como uma extensão da Web atual, que fornecerá uma estrutura que permite o entendimento e gerenciamento do armazenamento. O conteúdo na Web, independentemente do método utilizado, é baseado na avaliação semântica do conteúdo e é apresentado (texto, som, imagem) por meio de agentes de coleta de conteúdo que podem processar informações de diferentes fontes e trocar resultados com outros programas.

Devido ao grande número de pesquisadores participantes do projeto Web Semântica, logo após a divulgação inicial, foi feita uma atualização na estrutura proposta, seguindo as recomendações emitidas pelo W3C em 2002, conforme a Figura 2,

Figura 2: Arquitetura da Semantic Web (2002)



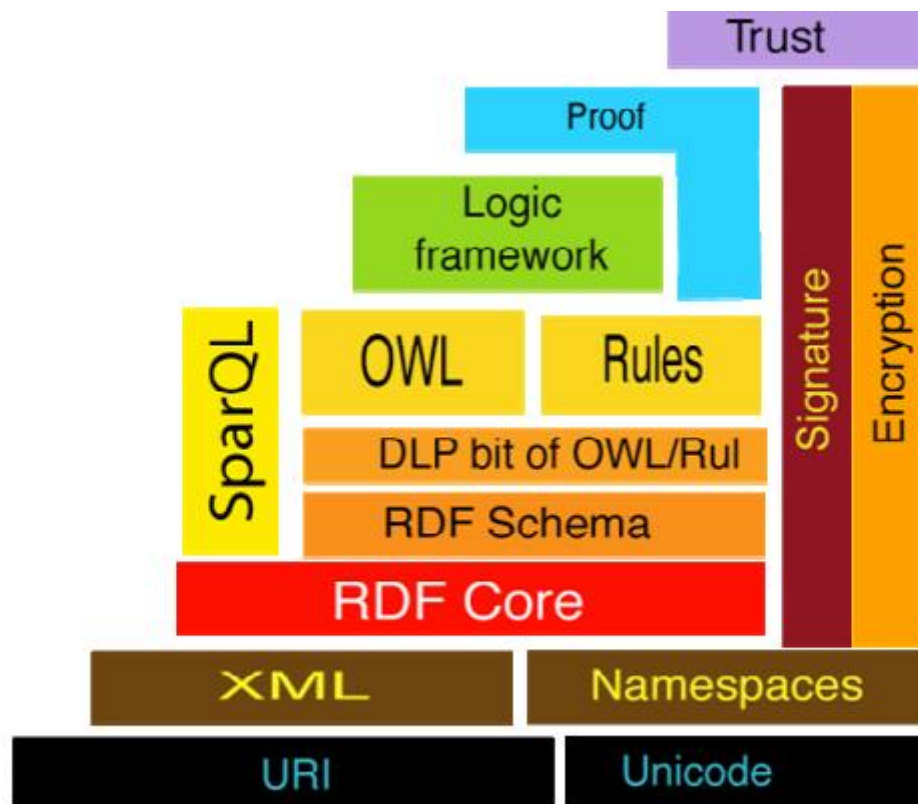
Fonte: <https://www.w3.org/2002/Talks/11-dd-ist2002/all.htm>

As recomendações arquitetônicas emitidas em 2002 não mudaram significativamente em comparação com a arquitetura anterior, apenas alguns detalhes de tecnologia e outros agrupamentos, como XML e XML Schema, porque neste momento, a linguagem XML foi recomendada com complemento do XML Schema.

A camada RDF é subdividida em partes contendo modelos RDF (RDF M&S), o padrão RDF Schema, instruções criptográficas e a assinatura digital, para garantir confidencialidade de informações na Web Semântica.

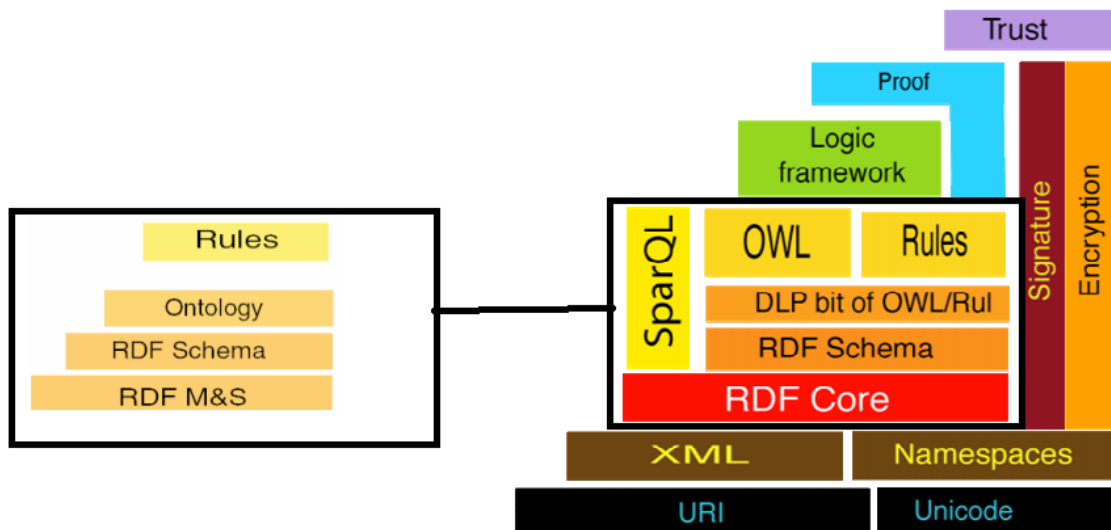
Mesmo utilizando a padronização da linguagem OWL, desenvolver a Ontologia é uma das principais dificuldades que as pessoas encontram para atingir o objetivo nos projetos de Web Semântica, por serem precisos em termos de integração e compatibilidade. O fato é que a camada de ontologia e as outras camadas da arquitetura finalmente atingiram seu pico, então foi lançada, em 2005, uma nova proposta do W3C (RAMALHO, 2006). A Figura 3 a seguir apresenta a arquitetura da Web Semântica.

Figura 3: Arquitetura da Web Semântica(2005)



Portanto, com base nesta proposta de arquitetura, as principais mudanças tecnológicas e camadas inerentes aos projetos da Web Semântica podem ser descritas de forma concisa, como indica a Figura 4.

Figura 4: Mudanças na Arquitetura da Web Semântica entre 2002 e 2005



Fonte: elaborada autor

OWL é uma linguagem de computação recomendada pelo W3C para o desenvolvimento de ontologias. De acordo com McGuinness e Harmelen (2004), a linguagem OWL pode descrever formalmente os aspectos semânticos dos termos usados e seus respectivos relacionamentos de forma mais eficaz, representando assim de forma mais abrangente o RDF e o esquema RDF da linguagem, e suportando uma maior interoperabilidade.

DLP é uma tecnologia candidata recomendada, que constitui a intersecção entre os dois principais paradigmas atualmente utilizados para o desenvolvimento de sistemas em computação baseados em representação de conhecimento, lógica descritiva (OWL DL) e programação lógica (F-Logic), proporcionando uma estrutura muito flexível. Portanto, a DLP atualmente não é considerada uma linguagem de representação do conhecimento, mas sim uma “ponte” que permite a combinação entre os dois principais paradigmas utilizados (GROSOF et al, 2003) .

De acordo com o último trabalho proposto por pesquisadores do W3C, **SparQL** é uma linguagem de computação usada para executar consultas de estruturas RDF,

ajudando assim a recuperar informações de forma mais eficiente. Essa linguagem ainda não foi totalmente padronizada, razão pela qual o W3C não recomendou oficialmente a linguagem, mas foi denominada como uma tecnologia candidata à recomendação pelo W3C (PRUD'HOMMEAUX e SEABORNE, 2005; CLARK, 2005).

O **RDF Core** compreende a especificação do modelo e a sintaxe da estrutura de descrição de recursos, podendo descrever por meio de seus atributos e valores. O RDF, segundo Daum e Merten (2002), pode ser considerado uma técnica de treinamento para modelagem semântica, na qual uma linguagem de computação específica pode ser criada.

Neste novo contexto, a Web semântica representa o desenvolvimento atual da Web. Embora o desenvolvimento da Web tradicional seja apenas para o usuário entender, o objetivo da Web Semântica é fazer com que os computadores nos entendam na forma de agentes computacionais, que podem efetivamente manipular as informações e entender seu significado. Dessa forma, eles ajudarão os usuários a operar na Web.

Esta tecnologia visa, justamente, melhorar a satisfação do usuário no momento da busca, retornando-lhe as informações adequadas às suas necessidades. Contrapondo a Web Sintática, surge a Web Semântica, através da qual se buscam mecanismos que capturem o significado das páginas, criando um ambiente no qual os computadores possam processar e relacionar conteúdos provenientes de várias fontes. Para que isso se torne possível, é necessário introduzir a semântica na estrutura dos documentos disponíveis na Web (BREITMAN, 2005).

2.2 Linked Data

A Web Semântica é o termo aplicado à nova geração de tecnologias que possuem o intuito de caracterizar a informação de maneira que computadores ou agentes de *software* possam compreendê-las (ISOTANI et al., 2009). Isto é, “um conjunto de boas práticas para a implementação da Web de dados, que se refere a publicação de informações semanticamente estruturadas do *World Wide Web*” (REIS, 2013, p. 12).

De acordo com Berners-Lee (2006) a proposta da Web Semântica não é somente sobre colocar dados na Web, mas também formar *links* entre dados relacionados para que pessoas e máquinas possam explorá-los. Esses dados

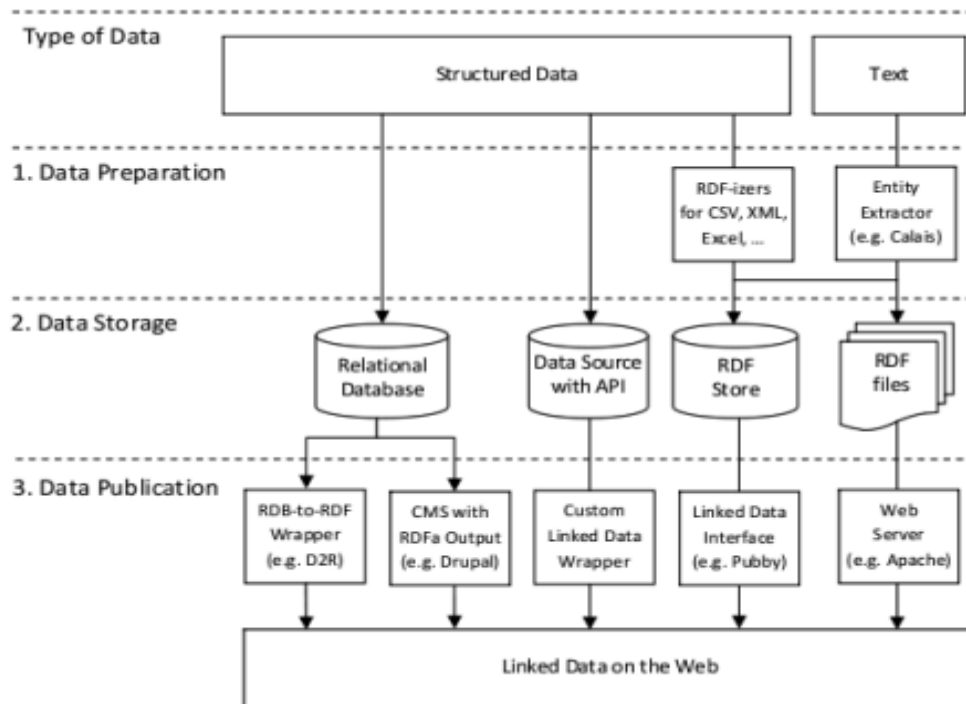
relacionados na Web também são conhecidos pelo termo *Linked Data* (dados vinculados) que, para Bizer (2009):

É uma forma de utilizar a web para criar ligações tipadas entre dados de diferentes fontes. Estas fontes podem ser, por exemplo, de diversas bases de dados mantidas por duas organizações em diferentes regiões geográficas, ou simplesmente sistemas heterogêneos em uma única organização que, historicamente, não provê interoperabilidade no nível de dados. Tecnicamente, *Linked Data* refere-se aos dados publicados na web, de tal forma que pode ser compreendido por máquinas e que seu significado seja explicitamente definido.

Segundo Reis (2013) “a tipagem de ligações entre dados é definida utilizando-se vocabulários que são baseados em ontologia”. Uma ontologia é considerada um elemento fundamental para a Web Semântica, pois permite acesso, interoperação e comunicação baseados em conteúdo (CUNHA, 2002).

Embora vários tipos de fontes de dados possam ser ligados com a ajuda de redes de dados, alguns padrões de publicação precisam ser atendidos. O uso de redes de dados vinculadas pode ser determinado basicamente de duas maneiras: liberação e uso. A publicação de informações na rede vinculada refere-se à estrutura e descrição do conjunto de dados em RDF para disponibilizar no meio (HAUSENBLAS, 2009b). Segundo Bizer e Heath (2011) há um diagrama dos padrões de publicações mais comuns em forma de workflow, como ilustra a Figura 5.

Figura 5: Linked Data Publishing Options and Workflows.



Fonte: <<http://linkeddatatoolkit.com/editions/1.0/>>

Conforme a Figura 5, o diagrama de padrões de publicações possui quatro camadas: Tipo de dados (*Type of Data*), Preparação de dados (*Data Preparation*), Armazenamento de Dados (*Data Storage*) e Publicação de Dados (*Data Publicação*).

A preparação de dados inclui a conversão de informações estruturadas ou não estruturadas (tipo de texto) em linguagem RDF. Após a extração, o armazenamento pode ser realizado por meio do repositório RDF ou arquivo de texto RDF. É importante observar que nem todos os dados precisam ser convertidos para RDF e podem permanecer no banco de dados ou ter acesso limitado à API, para assim serem disponibilizados dados públicos finais (BIZER; HEALTH, 2011).

Para que informações possam ser lidas por máquinas, o *Linked Data* utiliza tecnologias Web como: o protocolo HTTP; *framework* RDF; e URIs (BIZER, 2009). Estas tecnologias são importantes para a aplicação do conceito de *Linked Data*, e por isso as subseções a seguir apresentam essas tecnologias com mais detalhes.

2.2.1 Hypertext Transfer Protocol (HTTP) e Uniform Resource Identifier (URI)

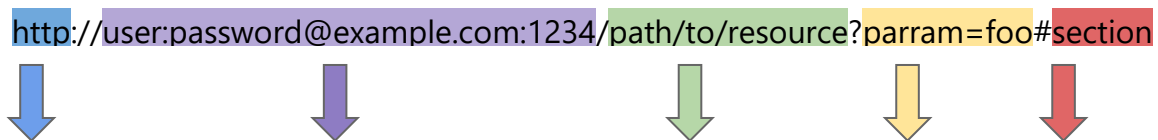
O *Hypertext Transfer Protocol* (HTTP) é um protocolo da camada de aplicação, utilizado como meio de acesso universal à Web, oferecendo uma recuperação do

recurso especificado (MIRANDA; AZEVEDO; MAGALHÃES, 2012). Desta forma, funciona em um modelo requisição/resposta baseado no modelo cliente/servidor. Conforme Reis (2013), o cliente utiliza uma URI para fazer uma requisição para o servidor, solicitando um recurso em um formato especificado, e assim o servidor retorna uma resposta conforme as informações passadas na URI no cabeçalho da requisição.

O *Uniform Resource Identifier* (URI) é um modelo de identificação global e único para recursos na Web como: documentos, arquivos, imagens, caixas de mensagens de email eletrônico entre outros (MIRANDA; AZEVEDO; MAGALHÃES, 2012). Segundo estes autores, no contexto de *Linked Data*, URIs reconhecem objetos do mundo real e conceitos, possibilitando que sejam caracterizáveis para recuperação de informações sobre os mesmos.

Para Reis (2013) a caracterização dos objetos e conceitos de uma URI é dada por sua sintaxe global única, composta pelos elementos posicionais apresentados na Figura 6:

Figura 6: Sintaxe Global URI.



Os elementos apresentados pela Figura 6 são descritos como:

- **Scheme:** É o identificador do esquema de URI. No contexto da internet, o *scheme* identifica o protocolo, como exemplo: **http://**, **ftp://**, **ssh://**.
- **Authority:** É a propriedade que contém as credenciais de usuário, a máquina que hospeda o serviço e a porta onde transitam as informações, como apresentado no exemplo: **ftp://david:password123@host:1234/**.
- **Path:** É o caminho absoluto no servidor. Essa propriedade contém informações de consulta, geralmente em ordenadas, que identificam resource dentro do escopo do *scheme://authority*. Por exemplo: **https://ulbra-to.br/bibliotecadigital/**
- **Query:** Essa propriedade contém dados não hierarquizados que identificam recursos de um escopo *scheme://authority/path/*. Esses dados são responsáveis pela filtragem de informações de determinado recurso, por

exemplo:

<https://ulbrato.br/bibliotecadigital/pesquisar?q=sistemas+de+informacao>.

- **Fragment:** É a propriedade que identifica um resource secundário por meio de um primário, de forma indireta, com base em um escopo *scheme://authority/path?query*. Geralmente em Web sites, o fragment identifica uma parte de uma Web page, por exemplo: <http://example.com/page?query#!state>

2.2.2 Resource Definition Framework (RDF)

O *Resource Definition Framework* (RDF) é um conjunto de modelagem de dados que fornece um vocabulário composto por conceitos básicos e a sintaxe utilizada para estruturação de informações semânticas no intuito e atribuir significado aos dados (DZIEKANIAK; KIRINUS, 2004). Conforme W3 (2019, online) o RDF é uma base para o processamento de metadados que fornece interoperabilidade entre aplicativos que trocam informações compreensíveis por máquina na Web.

Allemang e Hendler (2011, p.49) definem RDF como um “sistema de modelagem de dados”. Para Méndez Rodríguez (2002, p. 176) o “RDF, é mais que um mero formato de metadados”. Segundo o Turcatel(2014), o RDF tem a função de descrever um recurso atribuindo a essas propriedades, esta atribuição é realizada por uma afirmação em forma tripla, contendo recurso, propriedade e valor.

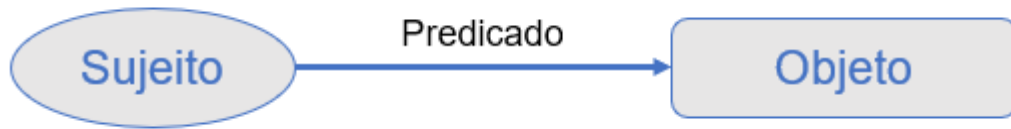
Segundo Faustino e Junio ,(2016) o RDF permite criar declarações de valor por meio de propriedades que representam uma ligação entre recursos. Uma declaração é realizada por meio de triplas do tipo “recurso-propriedade-valor”.

a) Recurso (sujeito). pode ser representado por um Website , ou ainda um objeto não acessível via Web (livro, CD, entre outros). Um recurso pode ser acessado e reconhecido de forma única por meio de um URI.

b) Propriedade (predicado). Trata-se de um atributo usado para descrever um recurso. Um recurso pode conter diversas propriedades: nome do autor, título do artigo, data de publicação, entre outros.

c) Valor (objeto). Representa o conteúdo das propriedades. Sendo os conteúdos referentes ao nome do autor, título do artigo, data de publicação, entre outros. Um exemplo dessa estrutura é apresentado na Figura 7.

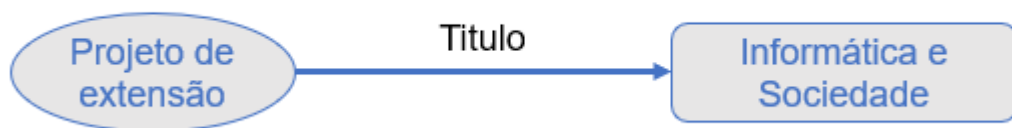
Figura 7: A estrutura do RDF.



Conforme apresentado na Figura 7 as triplas são formadas por: **Sujeito** (recurso) pode ser representado por um documento HTML, uma parte de uma página ou uma imagem, ou seja qualquer objeto que pode ser representado por um URI; **Predicado** (Propriedade) é um atributo do sujeito, aspectos, características, ou seja relacionamentos específicos utilizados para representação do sujeito; e **Objeto** (valor de propriedade) que são valores atribuídos às propriedades dos recursos (BAX; REZENDE, 2001).

De acordo com a estrutura RDF, para representar a informação “o projeto de extensão com o título Informática e Sociedade” em triplas RDF: o sujeito é a expressão “projeto de extensão”, predicado é a propriedade “título” e objeto é a informação “Informática e Sociedade”. A Figura 8 ilustra essa informação na forma tripla.

Figura 8: Informação na forma de tripla do RDF



Essa estrutura de triplas é semelhante ao diagrama de classes da orientação a objetos. Este diagrama possibilita a compreensão das entidades de um contexto e dos relacionamentos entre elas, ressaltando seus dados e atributos (GONÇALVES et al, 2005). Fazendo analogia do modelo RDF (triplas) e o modelo de diagrama de classes, a informação da Figura 8 pode ser interpretada dessa forma: “Projeto de extensão”

como objeto, “Título” como um atributo ao objeto, e “Informática e Sociedade” como valor do atributo.

A sentença “O orientador da página <https://ulbra-to.br/bibliotecadigital/publico/home/orientador/jackson-gomes-de-souza> é Jackson Gomes de Souza” . RDF é o padrão e a tripla é uma abstração. A forma de apresentar a tripla depende do formato concreto (aqui, novamente, usa RDF *Schema*, mas poderia ser OWL ou n3) O código-fonte 1 representação em RDF *Schema*.

O código-fonte 1: Representação em RDF Schema

```

1. <?xml version="1.0">
2. <rdf:RDF
3.   xmlns= http://www.w3.org/1999/02/22-rdf-syntax-ns
4.   xmlns:dc=http://purl.org/dc/elements/1.1>
5.   <rdf:Description rdf:about
6.     "https://ulbra-to-br/bibliotecadigital/publico/home/orientador/Jackson-gomes-de-souza">
7.     <dc:creator>Jackson Gomes de Souza</dc:creator>
8.     <dc:date:2010</dc:date>
9.   </rdf:Description>
10. </rdf:RDF>

```

A nomenclatura da descrição de “recurso”, “propriedade” e “valor” podem variar conforme o autor, mas o conceito do RDF é igual para a descrição de qualquer tripla.

No exemplo do trecho de código acima, a assertiva em RDF utiliza também o que é chamado de namespaces, que é um vocabulário destinado para o predicado e o objeto da assertiva. Os predicados autor e data de publicação são expressões utilizadas no vocabulário de metadados descritivos Dublin Core, “*dc:creator*” e “*dc:date*”; este vocabulário é identificado apenas dentro do documento RDF através da declaração do espaço de nomes Dublin Core, “xmlns:dc=“http://purl.org/dc/elements/1.1”. O recurso espaço de nomes permite que elementos de mesmo nome possam ser utilizados com outro significado dentro do mesmo documento RDF, desde que pertençam a outro espaço de nomes diferentes, como, por exemplo, o autor de um livro junto ao autor de uma ação judicial.

2.2.3 RDF/XML

No ano de 1986, foi lançada a linguagem XML, considerada uma versão simplificada da *Standard Generalized Markup Language* (SGML), para ser utilizada em aplicações Web. Ao contrário da HTML, XML não é uma linguagem escrita em SGML, mas uma linguagem derivada. Desta forma, ela possui recursos manipulados por sua comunidade, possibilitando que seja definida a forma como os segmentos marcados serão interpretados. Isto é viável, pois SGML é uma metalinguagem, ou seja, uma linguagem para a definição de linguagens (SANTANCHÈ, 2003).

Para Furgeri (2006, p. 237) o “RDF provê uma estrutura mais flexível que XML, aproximando-se da forma como os seres humanos relacionam informações, isto é, por associações”.

Atualmente, a principal sintaxe utilizada para serializar um grafo RDF é a sintaxe XML (RDF/XML). Outras sintaxes possíveis são: Notation3 (N3), N-Triples, Turtle e JSON. O modelo de dados no RDF provê uma estrutura abstrata e conceitual para definição e utilização de metadados. Uma sintaxe concreta é também necessária para criação e troca de metadados. RDF utiliza XML para atribuição de valores a propriedades, de propriedades a recursos e com XML permite a diferentes comunidades definir semânticas. É importante, entretanto, evitar a ambiguidade destas semânticas entre as comunidades.

O elemento/propriedade “*creator*”, por exemplo, pode ter diferentes significados de acordo com as diferentes necessidades das comunidades. Para evitar este problema, a RDF identifica unicamente os elementos/propriedades utilizados, através da utilização do mecanismo de *namespaces* da metalinguagem XML. Por exemplo, para se indicar o “*autor*” de um determinado recurso, pode-se utilizar o elemento “*creator*” definido pela iniciativa *Dublin Core*. Para isso, identifica-se unicamente o esquema da *Dublin Core*, acrescentando-se um namespace XML.

O código-fonte 2: Representação em RDF/XML

```
1. <?xml version="1.0">
2. <rdf:RDF>
3.   xmlns= http://www.w3.org/1999/02/22-rdf-syntax-ns
4.   xmlns:dc=http://purl.org/dc/elements/1.1>
5.   <rdf:Description rdf:about "c:\PastaCompartilhada\DM.txt">
6.     <dc:title> As coisas que você só vê quando desacelera</dc:title>
7.     <dc:creator> Haemin Sunim</dc:creator>
8.   </rdf:Description>
9. </rdf:RDF>
```

No exemplo do código-fonte 2, RDF e o esquema da *Dublin Core* são abreviados como RDF e DC respectivamente. Declara-se sempre o esquema antes de utilizá-lo e, tal declaração é feita através da associação de um URI, indicando a localização do esquema. O atributo “*about*” do elemento “*Description*” indica o recurso sendo descrito pelo documento RDF. O valor de uma propriedade pode ainda ser preenchido com um outro recurso:

O código-fonte 3: Representação em RDF/XML utilizando outro recurso

```

1.      <?xml:namespace ns = "http://www.w3.org/RDF/RDF/" prefix = "RDF"?>
2. <?xml:namespace ns = "http://purl.oclc.org/DC/" prefix = "DC"?>
3. <?xml:namespace ns = "http://person.org/BusinessCard" prefix = "CARD"?>
4. <rdf:RDF>
5.   <rdf:Description rdf:about "c:\PastaCompartilhada\DM.txt">
6.     <dc:title> As coisas que você só vê quando desacelera </dc:title>
7.     <dc:creator RDF:HREF = "autor 001"/>
8.   </rdf:Description>
9.   <rdf:Description ID = "autor 001"
10.    <card:name> Haemin Sunim </card:name>
11.    <card:email>exemplo@exemplo.com</card:email>
12.  </rdf:Description>
13. </rdf:RDF>

```

Conforme apresentado no código-fonte 3, na linha (9) do elemento “*Description*” indica o recurso sendo descrito pelo documento RDF, que está referindo as propriedades do autor apresentado na linha a qual herda características do *prefix card*(linha 3).

Contudo isso foi notado que o RDF é um aplicativo XML que impõe as restrições estruturais necessárias para fornecer um método claro para a representação consistente e troca de metadados entre pessoas, máquinas ou organizações

2.2.4 Microformatos FORTI

Segundo Forti (2013), microformatos se caracterizam como modelos de marcações para facilitar a extração de dados e itens. É um conjunto de formatos de dados abertos simples construídos com base em padrões existentes e amplamente usados (MICROFORMATS, 2014). Fernández (2008) apontou que os microformatos fornecem pequenos pedaços de código reutilizável que fornecem aos usuários

informações importantes, como contatos, eventos e marcações específicas. Segundo o autor, isso é introduzido por meio de atributos da linguagem HTML.

Micro formatos precisam usar recursos semiestruturados (como HTML e XML), e inserir novos descritores com recursos amplamente utilizados (como o atributo `class`). Esses descritores inseridos são microformatos e são responsáveis pelo valor semântico do documento (SANTOS, 2014).

A Web também é usada com microformatos para representar e descrever outras tecnologias para recursos na tentativa de alcançar um ambiente mais semântico. De acordo com Kim et al (2011) os microformatos comparado com a função RDF é menos eficiente, devido os microformatos dificilmente requerem qualquer conhecimento de marcação simples em HTML ou XML.

Além disso, hoje em dia, algumas plataformas utilizam microformatos e suas funções para simplificar tarefas. Aplicam semântica ao conteúdo sem o trabalho complicado do produtor (SANTOS, 2014). Há classes generalizadas de Microformatos, as principais são: ***hCalendar*** – eventos e datas; ***hCard*** – cartões de visita e informações de contato; ***hReview*** – comentários e opiniões; ***rel-tag*** – rótulo para elementos; ***hProduct*** – rótulo para produtos.

hCalendar é um formato padronizado aberto usado para descrever eventos na rede com base no padrão iCal, usado em vários aplicativos de calendário. Esse padrão permite que os mecanismos de pesquisa na Web reconheçam eventos (passados, presentes ou futuros) e os convertem para que possam ser incorporados automaticamente às ferramentas de calendário, como o Google Agenda (MICROFORMATS.org, 2012). apresentado a seguir no código-fonte 4:

O código-fonte 4: Representação do formato padronizado ***hCalendar***

```

1. <span class="vevent">
2. <span class="summary"> O site microformats.org foi lançado</span>
3. em <span class="dtstart"> 2005-06-20</span>
4. na Conferência Supernova
5. em <span class="location"> San Francisco, CA, EUA </span>
6. </span>

```

hCard é um microformato usado para publicar informações de contato de pessoas, empresas, organizações e locais (MICROFORMATS.org, 2012). apresentado a seguir no código-fonte 5:

O código-fonte 5: Representação do formato padronizado hCard

```

1. <span class="vcard">
2.   <a class = "fn url" href="http://adactio.com/">Jeremy keith</a>,
3.   <span class = "org">Clearleft</span>
4. </span>

```

hReview é um formato simples, aberto e distribuído adequado para mesclar (produtos, serviços, negócios, eventos, etc.) (MICROFORMATS.org, 2012). apresentado a seguir no código-fonte 6:

O código-fonte 6: Representação do formato padronizado *hReview*

```

1. <div class="hreview">
2.   <span class="reviewer">
3.     <span class="fn">anonymous</span>
4.   <abbr class="dtreviewed" title="20050418">April 18th, 2005</abbr>
5.   </span>
6.   <div class="item">
7.     <a lang="zh" class="url fn" href="http://www.imdb.com/title/tt0299977/">
8.     Ying Xiong (<span lang="en">HERO</span>)
9.     </a>
10.  </div>
11. <div>Rating: <span class="rating">4</span> out of 5</div>
12. <blockquote class="description"><p>
13. This movie has great music and visuals.
14. </p></blockquote>

```

Rel-tag é um *hyperlink*, a página indica o destino do *hyperlink* que se destina a ser a "tag" (ou palavra-chave / tópico) da página atual. Observe que as *tags* podem se referir apenas à maior parte da página atual (como uma postagem de blog)(MICROFORMATS.org, 2012). apresentado a seguir no código-fonte 7:

O código-fonte 7: Representação do formato padronizado *Rel-tag*

```

1. <a href ="http://technorati.com/tag/tech" rel= "tag">tech</a>

```

hProduct é um formato simples e aberto para publicar dados de produtos na web. H *product* é um dos vários rascunhos de padrões de microformatos abertos adequados para incorporar dados HTML) (MICROFORMATS.org, 2012).apresentado a seguir no código-fonte 8:

O código-fonte 8: Representação do formato padronizado - **hProduct**

```

1. <div class="h-product">
2.   <h1 class = "p-name">Microformatos para leigos</h1>
3.   <img class = "u-photo" src = "http://example.org/mfd.png" alt=" "/>
4.   <div class="e-description">
5.     <p>quer começar a usar microformatos, mas se sente intimidado
6.       por hifens e mediawiki? Este livro contém tudo o que você precisa saber!</p>
7.   </div>
8.   <p>Seu hoje por apenas <data class="p-price" value = "20,00"> $20,00</data>
9.     de <a class="p-brand h-
       card" href = "http://exemplo.com/acme">ACME Publishing inc.</a>
10.  </p>
11. </div>

```

Para tornar os dados da Web mais legíveis e orientados para programas e usuários, o XHTML e o HTML 4.01 evoluíram para o HTML 5, trazendo algumas mudanças. Nesse caso, foi desenvolvido o conceito de "microdados", que consiste em um padrão para a criação de atributos de metadados personalizados e sua incorporação em páginas HTML (MAVRODY, 2012).

2.2.5 Resource Definition Framework attributes (RDFa)

RDFa é um formato de serialização que incorpora triplas RDF em documentos HTML (SANTOS, 2016). Com as falhas ao tentar incorporar HTML e RDF em suas primeiras tentativas, agora os dados RDF não estão incorporados em comentários de um documento HTML, eles foram entrelaçados no DOM (*Document Object Model*). Isso significa que o conteúdo existente em uma página pode ser marcado com RDFa, modificando o código HTML (Heath & Bizer, 2011).

A vantagem da utilização do RDFa é que máquinas de buscas podem melhorar seus resultados aumentando a precisão sobre o real significado de um documento. Ou seja, as máquinas de buscas podem agregar os dados de um documento com

dados de outro documento, enriquecendo os resultados de buscas (PRIMER, 2014; CEWEB, 2016).

O código-fonte 9: Representação em RDFa

```

1. <div xmlns:dc="http://purl.org/dc/elements/1.1/"
2.   about="http://www.example.com/books/wikinomics">
3.   <span property="dc:title">Wikinomics</span>
4.   <span property="dc:author">Don Tapscott</span>
5.   <span property="dc:date">2006-10-01</span>
6. </div>

```

A Figura O código-fonte 9, possui algumas *tag*, para implementação do RDFa será levado em consideração as com as seguintes propriedades, *dc:title* que representa o título do livro, *dc:autor* apresenta o nome do autor que livro e por fim, *dc:date* onde apresenta a data de publicação do livro.

2.2.6 JavaScript Object Notation (JSON)-Linked Data (LD)

JSON-LD é uma maneira leve de representar dados de *Linked Data* no formato JSON (SPORNY; KELLOGG; LANTHALER, 2013). É possível adicionar informações semânticas a documentos JSON existentes sem muito esforço. Tem como objetivo proporcionar simplicidade, compatibilidade e expressividade. JSON-LD é compatível com JSON, ou seja, todo documento JSON-LD é um documento JSON válido. Esta compatibilidade permite que bibliotecas e analisadores atuais sejam reutilizados. Seu principal objetivo é usar recursos semânticos para construir APIs da Web, além de armazenar dados de link em um mecanismo baseado em JSON (LANTHALER; GÜTL, 2012). O código-fonte 10 a seguir, apresenta um trecho JSON-LD com informações de um produto.

O código-fonte 10: Representação em JSON-LD

```

1.  {
2.    "@context": "https://schema.org",
3.    "@type": "BreadcrumbList",
4.    "itemListElement": [{
5.      "@type": "ListItem",
6.      "position": 1,
7.      "name": "Books",
8.      "item": "https://example.com/books"
9.    }]

```

```
10.   "@type": "ListItem",
11.   "position": 2,
12.   "name": "The Lord of the Rings",
13.   "item": "https://example.com/books/the-lord-of-the-rings"
14.   }}
15.  }
16. </script>
```

No código-fonte 10, o objeto descrito contém o atributo *@context* com o valor 'http://schema.org', isso indica que qualquer leitor do documento deve tratar qualquer *@type* como um prefixo para aquele contexto. Por exemplo, no caso em que o *@context* acima é 'http://schema.org', o valor *@type* 'Book' acima será analisado como http://schema.org/Book para obter atributos de texto (como name).

Observando o vocabulário do schema.org, podemos saber que os itens com o URI http://schema.org/Book têm os nomes de atributo "nome" e "inLanguage" desses atributos específicos, respectivamente. Um exemplo de tal atributo com um tipo de entrada de http://schema.org/Offer, um novo atributo complexo é incluído e um novo nível de hierarquia JSON é criado sob o atributo "ofertas". Assim como nosso escopo de nível superior usa a palavra-chave reservada *@type* para o tipo de item de 'Livro', novamente usamos a mesma palavra-chave reservada para declarar que nosso novo nível de hierarquia define um item do tipo 'Oferta'.

2.2.7 Microdata

As possibilidades semânticas são realçadas no HTML5 onde é possível agregar informações diretamente em *tags*, por meio de vocabulários personalizados com algumas propriedades nomeadas (FORTI, 2013).

No que diz respeito aos Microdados, os elementos que compõem uma página HTML podem ser vistos como itens que possuem um ou mais pares compostos, que são nome/valor ou até mesmo propriedade/valor, resultando em uma sintaxe composta por *tags* HTML (RONALLO, 2012).

Schema.org também apresenta esquemas para classificação de centenas de novas categorias, incluindo filmes, música, organizações, programas de TV, produtos, lugares e muito mais. Como desenvolvedores Web adicionam essa marcação semântica aos seus sites, os motores de busca podem desenvolver experiências de

pesquisa mais ricas. O código-fonte 11 a seguir, apresenta um trecho HTML com informações de um local.

O código-fonte 11: Exemplo de estrutura de endereço sem marcação em HTML

```

1. <h1>Beachwalk Beachwear & Giftware</h1>
2. A superb collection of fine gifts and clothing to accent your stay in Mexico Beach.
3. 3102 Highway 98
4. Mexico Beach, FL
5. Phone: 850-648-4200

```

As pessoas podem ler essas informações e entender o significado de suas partes individuais, mas um rastreador de mecanismo de pesquisa pode não entender o significado de forma satisfatória. Se forem atribuídos esquemas aos dados acima, o resultado da pesquisa para o rastreador do motor de busca será muito mais legível, conforme apresentado no código-fonte 12.

O código-fonte 12: Exemplo de estrutura de endereço com marcação em HTML

```

1.         <div itemscope itemtype="https://schema.org/LocalBusiness">
2. <h1><span itemprop="name">Beachwalk Beachwear & Giftware</span></h1>
3. <span itemprop="description"> A superb collection of fine gifts and clothing
4. to accent your stay in Mexico Beach.</span>
5. <div itemprop="address" itemscope itemtype="https://schema.org/PostalAddress">
6.   <span itemprop="streetAddress">3102 Highway 98</span>
7.   <span itemprop="addressLocality">Mexico Beach</span>,
8.   <span itemprop="addressRegion">FL</span>
9. </div>
10. Phone: <span itemprop="telephone">850-648-4200</span>
11. </div>

```

As informações descritas usando microdados são estruturadas semanticamente e satisfatoriamente mais legíveis. Em evidência podemos ver que o texto pertence ao gênero de localidade tem seu próprio nome, descrição, endereço, localidade, região é telefone. Por exemplo, o nome da loja está apresentado em um elemento span que contém o atributo itemprop com valor name. É isso que, junto com os atributos itemscope e itemtype com valor 'https://schema.org/LocalBusiness' indicam que o texto trata de uma loja com este nome.

3 TRABALHOS RELACIONADOS

Esta seção apresenta trabalhos relacionados com a aplicação dos conceitos de Web Semântica.

[FORTI, Hericson Ramos, 2013] Aborda em sua pesquisa exploratória, um estudo de caso realizado por meio do desenvolvimento de um site de concessionária de veículos, com intuito de buscar uma melhoria no posicionamento das páginas html. O objetivo de sua pesquisa não foi apenas fornecer uma análise de métodos alternativos de atribuição de significado e informação contida na rede organizacional, mas também apontar a relevância da ciência da informação e seus métodos e ferramentas, as informações, métodos e ferramentas devem ser constantemente buscados por profissionais da área a fim de melhorar as mídias digitais.

[Wilma E. R.; LEOBINO, N. S.; J. A. S. M. 2016] aborda em seu texto que bases de dados abastecidas por conteúdo externo fornecido pela Web tem um potencial tecnológico promissor que pode ser utilizado com muitas finalidades por sites e aplicativos Web. O trabalho proposto tem por objetivo, mostrar como a adoção de estruturas baseadas em metadados podem agregar ainda mais valor para as informações jornalísticas. O estudo envolveu cinco organizações de mídia para identificar níveis de estruturação que proponham um instrumento adequado para a aplicação.

[LIMA, R. R. de., 2018] apresentou a utilização da Web Semântica na criação de portais de conteúdo com conhecimento para a tuberculose, objetivando uma experiência de uso enriquecedora aos usuários por meio de um portal inteligente capaz de entregar informações de maneira mais aderente às necessidades dos usuários e também a computadores que por meio de softwares inteligentes possam interpretar e entender seu conteúdo conforme preconizado pela Web Semântica. Isso foi possível devido ao grande volume de pesquisas e trabalhos realizados sobre a tuberculose gerando um conhecimento consolidado que poderia ser extraído e aproveitado com o uso da tecnologia da informação para o provimento de informações auxiliando no trabalho de profissionais da saúde, bem como apoiando a criação de políticas e estratégias para o controle da tuberculose.

JÚNIOR, V. F., *et. al* (2016) aborda em sua pesquisa uma breve explicação sobre as principais aplicações dos microformatos de forma proeminente, indicando que a tecnologia está se expandindo rapidamente e tem atraído a atenção das

principais empresas de serviços da Web da atualidade. Segundo o autor, ainda é preciso enfrentar o desafio da transformação dos dados publicados na Web, ou seja, dar-lhes significado e semântica, mas a tecnologia do microformato autor, ainda é preciso enfrentar o desafio da transformação dos dados publicados na Web, ou seja, dar-lhes significado e semântica, mas a tecnologia do microformatos torna-se muito importante porque fornece uma solução relativamente simples para a identificação desses dados, para que as máquinas podem manuseá-los de forma transparente, fornecendo-lhes aplicações específicas.

Os pesquisadores estudaram os modelos de programação atuais que usam diferentes paradigmas baseados em componentes de software para construir aplicativos, os quais são projetados para obter uma maior dissociação entre o código funcional da aplicação e a infraestrutura. Em alguns trabalhos houve a necessidade do controle de tráfego em dois grupos: Cliente e Servidor. Do ponto de vista do cliente, os trabalhos apresentaram avaliações de tráfego em determinados horários de uso, por outro lado, na perspectiva do servidor, os trabalhos analisam a quantidade de tráfego gerado pelos usuários, com o intuito de verificar se os provedores estão preparados para uma grande quantidade de tráfego simultaneamente.

Independentemente de ser aplicado técnicas de SEO ou investindo em links patrocinados, certos indicadores são a base para melhorar sua eficácia e o sucesso final de sua campanha de marketing digital. Um deles é a taxa de cliques, que é um indicador muito importante de qualquer estratégia. Quanto mais alta for a taxa de cliques de seus anúncios, mais benefícios sua empresa terá: de custos mais baixos no Google Ads a melhores classificações na pesquisa orgânica.

Visando essa melhoria na taxa de cliques ou acessos na página obtida pelos pesquisadores, replicando os mesmo conceitos, possivelmente acarretará em uma melhora nas buscas realizadas através dos navegadores gerando informações mais detalhadas associada a um professor ou a dados de eventos relacionados ao CEULP/ULBRA.

4 MATERIAIS E MÉTODOS

As subseções a seguir apresentam com mais detalhes os materiais e procedimentos necessários para desenvolvimento do trabalho.

4.1 Materiais

Os materiais utilizados para construção do referencial teórico foram: artigos, dissertações, teses e livros e conteúdos técnicos sobre ferramentas e técnicas, disponibilizados de forma online pela W3C, que é o órgão que regulamenta tecnologias no contexto da WEB. E ainda, foram utilizadas as seguintes ferramentas para o desenvolvimento e validação da proposta:

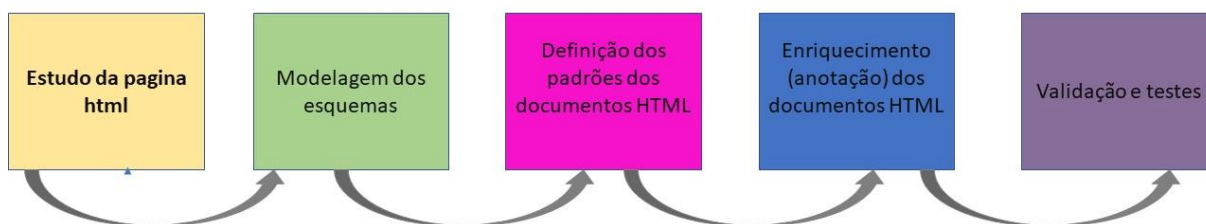
- *Schema.org*: O padrão *Schema.org* é uma colaboração entre Microsoft, Google e Yahoo no sentido de desenvolver um padrão único de marcação de páginas da Web com o intuito de fazê-las mais facilmente reconhecíveis pelos mecanismos de buscas, que poderão analisar semanticamente o conteúdo e fornecer respostas mais precisas aos usuários (GOOGLE, 2012b; SCHEMA.ORG, 2012)
- *Google Structured Data Testing Tool*: Ferramenta da Google para testes de dados estruturados.
- *Microdata.reveal*: Ferramenta para detectar e apresentar os microdados incorporados em uma determinada página HTML.
- HTML5: É a abreviação de *Hypertext Markup Language*, ou seja, Linguagem de Marcação de *Hipertexto*. O HTML é uma linguagem utilizada para a publicação de conteúdo (texto, imagens, vídeos, áudio etc.) na Web.

4.2 Métodos

O presente trabalho foi realizado por meio da execução de etapas de um processo, ilustrado pela Figura 9, que iniciou com o entendimento do domínio e concluiu com a execução de validação e testes:

1. Estudo da página html.
2. Modelagem dos esquemas.
3. Definição dos padrões dos documentos HTML.
4. Enriquecimento (anotação) dos documentos HTML.
5. Validação e testes.

Figura 9: Metodologia do trabalho



Na etapa de **estudo da página html** foram realizadas reuniões com o especialista de domínio do negócio no intuito de coletar informações sobre o contexto do Portal bem como os problemas pertinentes, e conseqüente foi fornecido o conteúdo produzido pela fábrica de software (o *design* e a estrutura das páginas), para inserção de microdados.

Na etapa de **Modelagem dos esquemas**, foi realizado um estudo junto ao especialista de domínio para identificação de atributos das entidades do Portal. Com isso, foi elaborada uma modelagem utilizando o schema (esquema) que consiste em um sistema de hierarquia de dois níveis. No topo, você tem tipos e cada tipo tem um conjunto de propriedades definidas, com isso foi possível identificar as entidades do contexto do Portal, bem como os relacionamento entre eles. Assim, destacando seus dados e atributos.

Em seguida, a etapa **Definição dos padrões dos documentos HTML** consistiu no estudo para definição das *tags* ideais para cada situação.

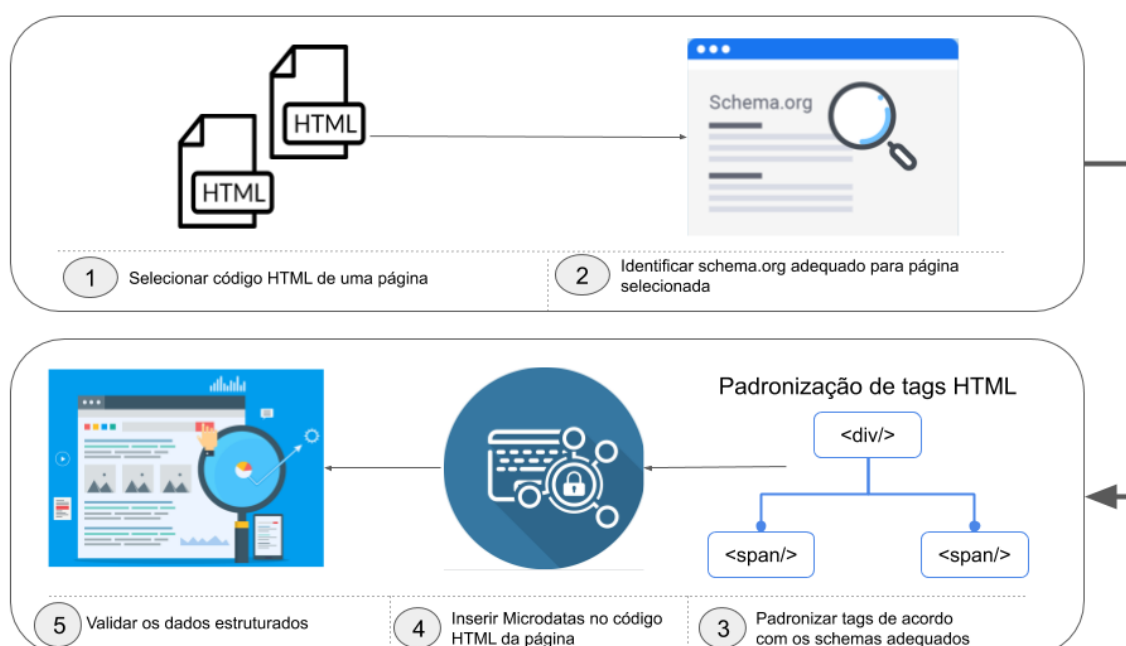
Posteriormente, na etapa **Enriquecimento dos documentos HTML** foi realizada a implementação das *tags* de microdata para estruturação dos dados da página web de Professor e home do Portal Acadêmico do CEULP/ULBRA.

Por fim, a etapa de **Validação e testes** consistiu na utilização das ferramentas *Google Structured Data Testing Tool* e *Microdata.reveal*, para validar os dados estruturados. Essa validação, consistiu na verificação da exibição aprimorada na Pesquisa do Google por meio dos dados estruturados.

5 RESULTADOS E DISCUSSÃO

Conforme os objetivos deste trabalho, foram aplicados Microdados do *schema.org* nas páginas Home e Professor do portal acadêmico do CEULP/ULBRA. Para compreender melhor o processo de aplicação de Microdatas em páginas HTML neste trabalho, foi gerado um diagrama de visão geral do processo apresentado na Figura 10.

Figura 10: Visão geral do processo



fonte: autor

Conforme exibido na Figura 10, inicialmente (**passo 1**) foi necessário obter uma página html com seu código. Em seguida (**passo 2**), foi identificado no *schema.org* quais os esquemas e suas propriedades adequadas para expressar os dados presentes na página. Posteriormente (**passo 3**), foi essencial a organização do código para fazer a padronização das *tags* de acordo com as propriedades selecionadas na etapa anterior. Em seguida (**passo 4**), foram inseridas as *tags* dos microdados no código html da página E por fim (**passo 5**), foi realizada a validação da página estruturada por meio da ferramenta do google anteriormente neste trabalho.

Nas subseções a seguir são apresentados com detalhes os resultados deste trabalho.

5.1 Página Professor

A seguir são apresentados os resultados para a página de Professor do portal acadêmico.

5.1.1 Modelagem da Página Professor

A Figura 11 ilustra a seção da página de professor do Portal Acadêmico, na qual apresenta informações do perfil de um professor da instituição.

Figura 11: Página de professor



fonte: autor

A página ilustrada pela Figura 11 apresenta as informações do professor da instituição com as seguintes estruturas:

- Dados Pessoais (esquema Person): nome (1), especialidade (2), link para currículo lattes (3). link para e-mail (5), foto de perfil (9).
- Dados de cursos ministrados (esquema Course): nome (8), e-mail (4), telefone (7), descrição (6).

Foram selecionadas as propriedades apresentadas acima que constituem o esquema *Person* e *Course* do *schame.org* para a descrição/anotação desta seção da página do professor. Algumas informações são relevantes e não são encontradas na página, porém esses dados são essenciais para que o motor de busca possa

compreendê-las, assim foram inseridas informações adicionais na página: *description* e *name* do Objeto Course, e ocultadas com a *tag html hidden*.

Para melhor entendimento da representação das propriedades dos esquemas, serão apresentadas a seguir tabelas com as propriedades utilizadas na página do professor, juntamente com suas descrições. A Tabela 1 a seguir, apresenta as propriedades do esquema *Person* que foram utilizadas para representação de dados na página do professor.

Tabela 1: Propriedades do esquema Person

Propriedade	Descrição
Name	O nome do objeto
<i>jobTitle</i>	O cargo da pessoa (por exemplo, Gerente Financeiro)
<i>sameAs</i>	URL de uma página da Web de referência que indica inequivocamente a identidade do item. Por exemplo, o URL da página da Wikipedia do item, entrada do Wikidata ou site oficial.
<i>image</i>	Uma imagem do item. Pode ser um <u>URL</u> ou um <i>ImageObject</i> totalmente descrito .
<i>email</i>	Endereço de e-mail
<i>description</i>	Uma descrição do item
<i>hasCredential</i>	Uma credencial concedida à pessoa ou organização.
<i>affiliation</i>	Uma organização à qual essa pessoa é afiliada. Por exemplo, uma escola / universidade, um clube ou uma equipe.

Na propriedade *hasCredential* é esperado receber propriedades do esquema *EducationalOccupationalCredential*, apresentado na Tabela 2 a seguir:

Tabela 2: Propriedades do esquema *EducationalOccupationalCredential*

Propriedade	Descrição
<i>credentialCategory</i>	A categoria ou tipo de credencial que está sendo descrita, por exemplo "diploma", "certificado", "crachá" ou termo mais específico.

A propriedade *affiliation* pertence ao esquema *Person* e de acordo com schema.org ela herda as características do esquema *Organization* descrito na Tabela 3, que apresenta os atributos que foram usados para representação da *affiliation*.

Tabela 3: Propriedades do esquema *Organization*

Propriedade	Descrição
Name	O nome do objeto.

As propriedades do esquema *Course* que foram utilizadas para representação dos cursos que o professor faz parte, são apresentadas na Tabela 4 a seguir:

Tabela 4: Propriedades do esquema *Course*

Propriedade	Descrição
<i>Name</i>	O nome do objeto
<i>Description</i>	Uma descrição do item.
<i>Provider</i>	O provedor de serviços, operadora ou provedor de serviços; o produtor de bens. Outra parte (um vendedor) pode oferecer esses serviços ou mercadorias em nome do fornecedor. Um provedor também pode servir como vendedor. Substitui o operador.

Na propriedade *provider* é esperado herdar propriedades do esquema *Organization*. As propriedades do esquema *Organization* utilizados no esquema *Course*, são apresentadas a seguir na Tabela 5.

Tabela 5: Propriedades do esquema Organization

Propriedade	Descrição
Email	Endereço de e-mail.
Telefone	O número de telefone.
Name	O nome do objeto.

Os dados referentes a página de professor, apresentados nas tabelas acima, foram implementados utilizando *tags* de microdados do *schema.org*. A próxima seção demonstra como foi a implementação dos microdados no HTML da página do professor.

5.1.2 Implementação da Página Professor

O código-fonte 13 representa a implementação de microdados para descrever dados pessoais do professor.

O código-fonte 13: Dados *estrutura Person*

```

1. <section itemscope itemtype="http://schema.org/Person" id="professor-detahes">
2.   <div itemprop="affiliation" itemscope itemtype="https://schema.org/Organization">
3.     <span itemprop="name">CEULP ULBRA</span>
4.   </div>
5.   <div class="container">
6.     <div class="row resumo">
7.       <div class="col-lg-3 col-md-4">
8.         <div class="card mb-3 mr-0 card-professor">
9.           <div class="card-img">
10.            <a href="foofighters-3.jpg" itemprop="image" >
11.              
12.            </a>
13.          </div>
14.        </div>
15.      </div>
16.      <div class="card-body m-3 p-0">
17.        <h1 itemprop="name" class="mb-0">Abizair Antonio Paniago</h1>
18.        <h2 style="font-size: 1rem; margin-
19.        bottom: 0px;"><span hidden></span> Direito</h2>
20.        <span itemprop="jobTitle" hidden> Professor</span>
21.        <div itemprop="hasCredential" itemscope itemtype="http://schema.org/Educ
22.        ationalOccupationalCredential">
23.          <span itemprop="credentialCategory">Especialista</span>
24.        </div>

```

```

23.
24.         <a itemprop="sameAs" target="_blank" href="http://lattes.cnpq.br/424772015
25.         5465925" title="Currículo Lattes"></a>
27.         <a itemprop="email" class="text-
28.         dark" href="mailto:abizair@ceulp.edu.br" title="E-mail"><i
29.             class="fa fa-envelope fa-fw fa-xl" aria-hidden="true"></i></a>
30.         </div>
31.     </div>
32. </div>
33. </div>
34. <div class="col-lg-9 col-md-8">
35.     <div class="row">
36.         <div class="col text-justify">
37.             <p class="detalhes-professor" itemprop = "description">
38.                 Atualmente é professor do Centro Universitário Luterano de Palmas. Tem
39.                 experiência na
40.                 área de Direito, com ênfase em Direito Penal e Processual Penal. </p>
41.             </div>
42.         </div>
43.     </div>
44. </div>

```

Conforme apresentado no código acima, as propriedades inseridas para estruturação da seção de dados pessoais da página de professor foram: *itemprop* que é usado para adicionar propriedades para um item; *itemtype*), usado para definir o escopo de onde na estrutura de dados que o vocabulário definido por *itemtype* estará ativo; e *itemscope* que define o escopo dos metadados associados. As informações inseridas no código acima são propriedades do objeto *Person*: *name*(linha 3); nome do professor, *jobTitle*(linha 19); título de trabalho , *sameAs*(linha 24) referente ao currículo lattes, *emai*(linha 26); email para contato do professor, *description*(linha 34); um breve roteiro sobre o professor , *affiliation*(linha 2) que tem suas característica baseada no esquema *Organization*(linha 2) composto por e *name*(linha 3) da instituição a qual o professor faz parte, *hasCredential*(linha 20); que tem suas característica baseada no esquema *EducationalOccupationalCredential*(linha 20) composto por *credentialCategory*(linha 21). Essas propriedades também foram inseridas para estruturação da seção de cursos ministrados, conforme o código-fonte 14 a seguir.

O código-fonte 14: Dados estrutura Course

```

1.     <div class="col-9 justify-content-center pr-0" itemscope
2.     itemtype="https://schema.org/Course">
3.     <h2 itemprop="name" style="font-size: 1rem; margin-bottom: 0px;" >
4.     Direito</h2>
5.     <span itemprop="description" hidden>De acordo com a tradição do
6.     CEULP/ULBRA, o Curso de Direito prepara pensadores do Direito para o sucesso
7.     no mercado de trabalho. Para tanto,

```



```

4.     é colocada à disposição excelentes professores – especialistas, mestres e
      doutores,
5.     de acordo com as exigências do MEC – acervo bibliográfico completo,
      laboratórios de prática jurídica
6.     (reais e simuladas), excelente estrutura física, frequência de ofertas em
      pesquisa e extensão.
7.     Fatos que comprovam o índice de aprovação na OAB! Pense bem! Pense
      Direito! Venha para o CEULP/ULBRA!</span>
8.     <div     itemprop="provider"     itemscope
      itemType="https://schema.org/Organization">
9.
10.    <small itemprop="email" style="font-size: 75%; white-space: nowrap;"
      class="text-center">direito@ceulp.edu.br</small><br>
11.    <small itemprop="telephone" class="text-uppercase text-center"><i
      class="fas fa-phone"
12.    aria-hidden="true"></i> 32198093</small> <small class="text-uppercase
      text-center"
13.    style="padding-left: 5px;"><i class="fa fa-building fa-fw" aria-
      hidden="true"></i>632</small>
14.    <span itemprop="name" hidden>CEULP ULBRA</span>
15.  </div>
16.</div>

```

As informações inseridas no código acima são propriedades do objeto *Course* (linha 1 1): *name* (linha 2) da referência do curso a qual o professor faz parte; *description* (linha 3), que apresenta uma breve descrição do curso de direito; a propriedade *provider* (linha 8) que tem suas característica baseada no esquema *Organization* (linha 8) composto por email do curso de direito, *telephone* (linha 11) da referência ao número telefônico do curso, *name* (linha 14) da instituição a qual o curso faz parte.

Com a inserção de microdados no código HTML da página de professor, foi realizada a validação dos dados estruturados, conforme apresentado na subseção abaixo.

5.1.3 Validação da Página Professor

Com a finalidade de validar se os dados estão estruturados de forma padronizada, foi inserido o código HTML apresentado na seção anterior na ferramenta *Google Structured Data Testing Tool*. A Figura 12 ilustra o resultado da ferramenta após o código inserido.

Figura 12: Esquemas detectados na página de professor

Detectado	0 ERROS	0 AVISOS	2 ITENS
Course	0 ERROS	0 AVISOS	1 ITEM
Person	0 ERROS	0 AVISOS	1 ITEM

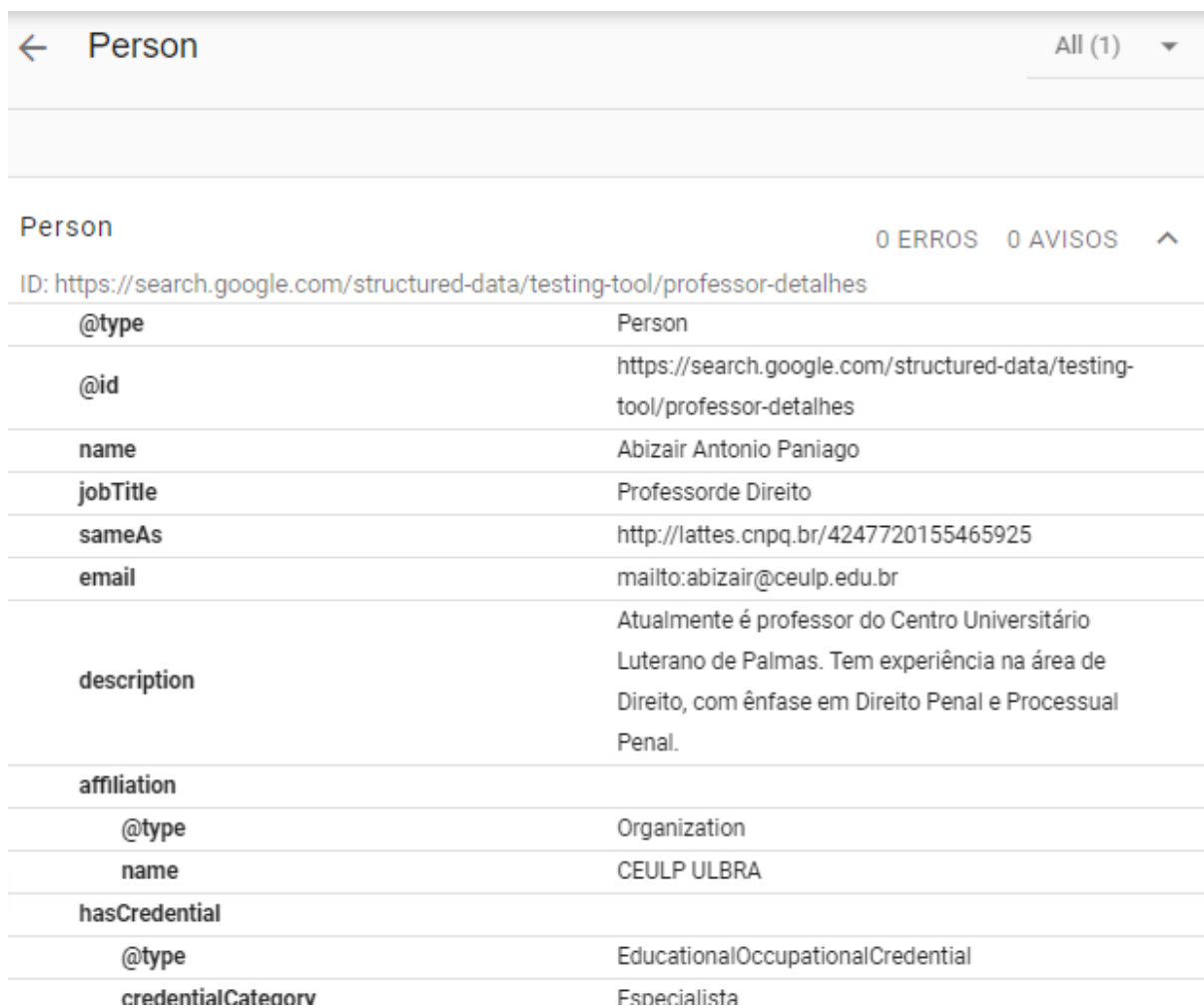
A ferramenta detecta os atributos globais de documentos HTML estruturados, conforme apresentado na Figura 12, bem como suas propriedades e apresentação de erros e avisos. Na página do professor, os atributos globais detectados foram *Course* e *Person*, referentes a pessoa e seus cursos respectivamente. As propriedades do atributo global *Course*, são observadas na Figura 13 abaixo.

Figura 13: Propriedades detectadas em Course

Course	
@type	Course
name	Direito
description	De acordo com a tradição do CEULP/ULBRA, o Curso de Direito prepara pensadores do Direito para o sucesso no mercado de trabalho. Para tanto, é colocada à disposição excelentes professores – especialistas, mestres e doutores, de acordo com as exigências do MEC – acervo bibliográfico completo, laboratórios de prática jurídica (reais e simuladas), excelente estrutura física, frequência de ofertas em pesquisa e extensão. Fatos que comprovam o índice de aprovação na OAB! Pense bem! Pense Direito! Venha para o CEULP/ULBRA!
provider	
@type	Organization
email	direito@ceulp.edu.br
telephone	32198093
name	CEULP/ULBRA

Conforme apresentado na figura 13, ao selecionar o atributo *global Course* na Figura 12, a ferramenta exibe suas propriedades e valores, bem como erros e avisos caso houver. Na Figura 14, a seguir, são apresentadas as propriedades de *Person*.

Figura 14: Propriedades detectadas em Person



Person	
ID:	https://search.google.com/structured-data/testing-tool/professor-detalhes
@type	Person
@id	https://search.google.com/structured-data/testing-tool/professor-detalhes
name	Abizair Antonio Paniago
jobTitle	Professor de Direito
sameAs	http://lattes.cnpq.br/4247720155465925
email	mailto:abizair@ceulp.edu.br
description	Atualmente é professor do Centro Universitário Luterano de Palmas. Tem experiência na área de Direito, com ênfase em Direito Penal e Processual Penal.
affiliation	
@type	Organization
name	CEULP ULBRA
hasCredential	
@type	EducationalOccupationalCredential
credentialCategory	Especialista

Para entidade global *Person*, a ferramenta identificou suas propriedades e valores, sendo elas: *name*, *jobTitle*, *sameAs*, *email*, *description*, e *HasCredential* que tem como propriedade *credentialCategory* que representa a credencial do professor sendo: especialista, mestre ou doutor.

Dessa forma, com os atributos globais e suas propriedades identificadas pela ferramenta, sem detecção de erros e avisos, foi possível validar os dados estruturados na página de professor do Portal Acadêmico do CEULP/ULBRA, conforme desenvolvido neste trabalho.

Além da página de professor, foram também adicionados microdados na página *Home* do portal. O desenvolvimento e resultados para a página de *Home* é apresentado a seguir.

5.2 Página Home

A seguir são apresentados os resultados para a página *Home* do portal acadêmico.

5.2.1 Modelagem da Página Home

A Figura 15 ilustra a primeira seção da página de *Home* do Portal Acadêmico que apresenta informações sobre a instituição.

Figura 15: Seção 1 da página Home.



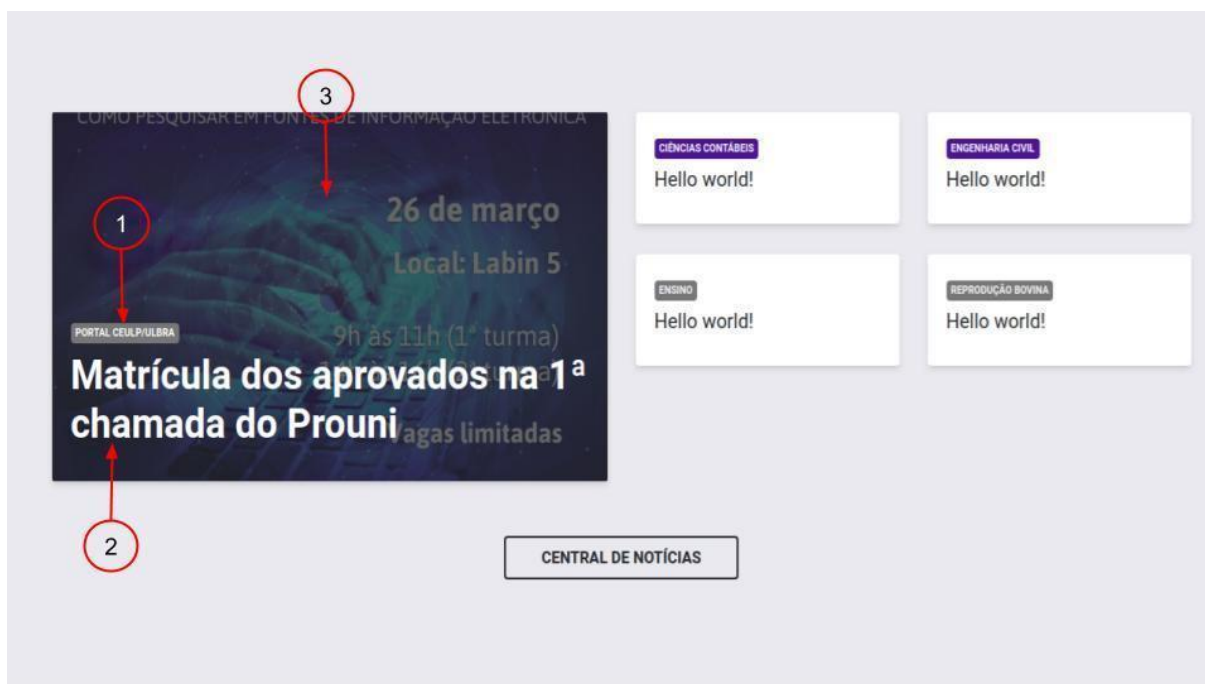
A seção ilustrada pela Figura 5 apresenta as informações relacionada a instituição com as seguintes estruturas:

- Dados da instituição: Telefone da instituição (1)

Inicialmente foi selecionada a propriedade *telephone* do esquema *CollegeOrUniversity* para a informar o telefone da instituição.

A segunda seção da página *Home* ilustrada pela Figura 16, apresenta as informações relacionadas a eventos da instituição:

Figura 16: Seção 2 da página Home.



A seção ilustrada pela Figura 16 apresenta as informações relacionadas a instituição com as seguintes estruturas:

- Dados dos eventos: nome (1), descrição (2), link para acesso, imagem (3).

Foram selecionadas as seguintes propriedades do esquema *CreativeWork* para a descrição/anotação desta seção da página home da instituição: *name*, *image*, *description*.

A terceira seção da página *Home* ilustrada pela Figura 17 apresenta as informações relacionadas a Calendário acadêmico e calendário de eventos presentes na instituição.

Figura 17: Seção 3 da página Home.



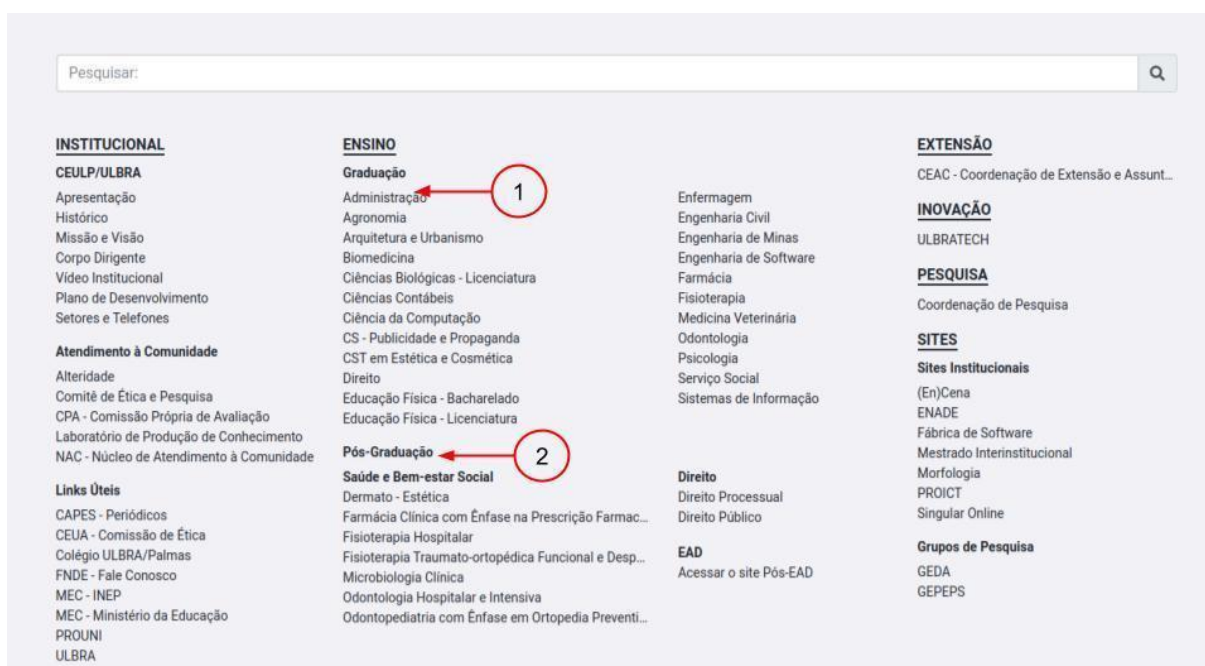
A Seção ilustrada pela Figura 17 apresenta as informações relacionadas a instituição com as seguintes estruturas:

- Dados Calendário Acadêmico: dia (1), mês (1), nome (2).
- Dados Calendário de Eventos: dia (3), mês (3), nome (1), descrição.
- Dados Destaque: nome, breve descrição, link do destaque, imagem (5).

Foram selecionadas as seguintes propriedades do esquema *ScheduleAction* para a descrição/anotação do calendário acadêmico: *byDay*, *ByMonth* e *Name*. Para notação do calendário de eventos, foi utilizado propriedades do esquema *Schedule*: *byDay*, *ByMonth*, *name* e *description*. Para destaque foram aplicadas propriedades do esquema *CreativeWork*: *name*, *abstract*, *url* e *image*. Algumas informações são relevantes, mas não são encontradas na página. Para melhor compreensão dos motores de busca foram inseridas informações adicionais: *description* e *image* e ocultadas com a tag html *hidden*.

A quarta seção da página *Home*, ilustrada pela Figura 18, apresenta as informações sobre cursos fornecidos pela instituição.

Figura 18: Seção 4 da página Home



A seção ilustrada pela Figura 18 apresenta as informações de cursos, sendo eles de graduação ou pós-graduação. Para expor essas informações foi utilizado as seguintes informações:

- Dados Cursos: name (1)(2).

Foi selecionada a propriedade *name* do esquema *Course* para a descrição/anotação desta seção de cursos da página *Home*. Algumas informações são relevantes, mas não são encontradas na página. Para melhor compreensão dos motores de busca foram inseridas informações adicionais na página: *description* do Objeto *Course*, e ocultadas com a *tag html hidden*.

A quinta seção ilustrada pela Figura 19 apresenta as informações relacionadas ao endereço da instituição.

Figura 19: Seção 5 da página Home.



A seção ilustrada pela Figura 19 apresenta as informações relacionadas a instituição com a seguinte estrutura:

- Dados Endereço: nome (1), caixa postal (3), endereço de localização (2), região (4).

Foram selecionadas as seguintes propriedades do esquema *PostalAddress* para a descrição/anotação desta seção de endereço da instituição: *postalCode*, *email*, *name*, *addressLocality*, *addressRegion*. Algumas informações são relevantes, mas não são encontradas na página. Para melhor compreensão dos motores de busca foram inseridas informações adicionais na página, sendo o atributo: *name* do Objeto *Country*, e ocultada com a tag html *hidden*.

Para melhor entendimento da representação das propriedades dos esquemas descritos anteriormente, a seguir são apresentadas tabelas com as propriedades e descrições para os esquemas da página *Home*.

Na Tabela 6 a seguir são mostradas as propriedades do esquema *CollegeOrUniversity* que foram utilizadas para representação de dados na página home.

Tabela 6: Propriedades do esquema *CollegeOrUniversity* para seção 1 da página Home.

Propriedade	Descrição
Telefone	O número de telefone

Na Tabela 7 a seguir, são apresentadas propriedades do esquema *CreativeWork* que foram utilizadas para representação de dados na página home relacionados às notícias da instituição e destaque.

Tabela 7: Propriedades do esquema *CreativeWork* para seção 2 e 3 da página Home.

Propriedade	Descrição
<i>name</i>	O nome do objeto
<i>image</i>	Uma imagem do item. Pode ser um URL ou um <i>ImageObject</i> totalmente descrito .
<i>description</i>	Uma descrição do item.
<i>abstract</i>	Um resumo é uma breve descrição que resume um <i>CreativeWork</i> .
<i>sameAs</i>	URL de uma página da Web de referência que indica inequivocamente a identidade do item. Por exemplo, o URL da página da <i>Wikipedia</i> do item, entrada do <i>Wikidata</i> ou site oficial
<i>url</i>	URL do item.

Na tabela 8 são apresentadas propriedades do esquema *ScheduleAction* que foram utilizadas para representação de dados na página home relacionados aos calendários acadêmicos da instituição.

Tabela 8: Propriedades do esquema *ScheduleAction* para seção 3 da página Home.

Propriedade	Descrição
<i>agent</i>	O executor direto ou condutor da ação (animado ou inanimado). por exemplo, <i>John</i> escreveu um livro

A propriedade *agent* apresentado na tabela acima, espera herdar propriedades do esquema *Schedule*, apresentado na Tabela 9 a seguir:

Tabela 9: Propriedades do esquema *Schedule* de calendário acadêmico para seção 3 da página Home.

Propriedade	Descrição
-------------	-----------

<i>ByMonth</i>	Define o (s) mês (es) do ano em que ocorre um Evento recorrente. Especificado como um número inteiro entre 1-12. janeiro é 1º.
<i>ByDay</i>	Define o (s) dia (s) da semana em que ocorre um Evento recorrente. Pode ser especificado usando <i>DayOfWeek</i> ou, alternativamente, Texto em conformidade com a sintaxe do <i>iCal</i> para regras de recorrência <i>byDay</i>
<i>Name</i>	O nome do objeto.

Na Tabela 10 é mostrado propriedades do esquema *Schedule* que foram utilizadas para representação de dados na página home relacionados aos calendários de eventos da instituição.

Tabela 10: Propriedades do esquema *Schedule* de calendário de eventos para seção 3 da página Home.

Propriedade	Descrição
<i>Image</i>	Uma imagem do item. Pode ser um URL ou um <i>ImageObject</i> totalmente descrito
<i>SameAs</i>	URL de uma página da Web de referência que indica inequivocamente a identidade do item. Por exemplo, o URL da página da <i>Wikipedia</i> do item, entrada do <i>Wikidata</i> ou site oficial
<i>ByDay</i>	Define o (s) dia (s) da semana em que ocorre um Evento recorrente. Pode ser especificado usando <i>DayOfWeek</i> ou, alternativamente, Texto em conformidade com a sintaxe do <i>iCal</i> para regras de recorrência <i>byDay</i>
<i>ByMonth</i>	Define o (s) mês (es) do ano em que ocorre um Evento recorrente. Especificado como um número inteiro entre 1-12. janeiro é 1º.
<i>Name</i>	O nome do objeto.
<i>Description</i>	Uma descrição do item.

Na Tabela 11 a seguir, são apresentadas propriedades do esquema *Course* que foram utilizadas para representação de dados na página home relacionados aos cursos fornecidos pela instituição.

Tabela 11: Propriedades do esquema *Course* para seção 4 da página Home.

Propriedade	Descrição
<i>name</i>	O nome do objeto.
<i>description</i>	Uma descrição do item.
<i>provider</i>	O provedor de serviços, operadora de serviços ou executor de serviços; o produtor de bens. Outra parte (um vendedor) pode oferecer esses serviços ou bens em nome do fornecedor. Um provedor também pode servir como vendedor. Substitui a operadora.

A propriedade *provider* apresentada na tabela acima, espera herdar propriedades de *Organization*, apresentada na Tabela 12 a seguir:

Tabela 12: Propriedades do esquema *Organization* para seção 4 da página Home.

Propriedade	Descrição
<i>name</i>	O nome do objeto.
<i>url</i>	URL do item.

A Tabela 13 apresenta as propriedades do esquema *PostalAddress* e suas descrições, que foram utilizadas para representação dos dados relacionados ao endereço da instituição:

Tabela 13: Propriedades do esquema *PostalAddress*.

Propriedade	Descrição
<i>addressLocality</i>	A localidade em que se encontra o endereço da rua e qual é a região. Por exemplo, <i>Mountain View</i>

<i>PostalCode</i>	O código postal. Por exemplo, 94043.
<i>Email</i>	Endereço de e-mail.
<i>Name</i>	O nome do objeto.
<i>addressRegion</i>	A região em que se encontra a localidade e qual se encontra no país. Por exemplo, Califórnia ou outra divisão administrativa de primeiro nível apropriada.
<i>addressCountry</i>	O país. Por exemplo, EUA. Você também pode fornecer o código de país ISO 3166-1 alfa-2 de duas letras.

A propriedade *addressCountry* da tabela acima, herda propriedades do esquema *Country* apresentado na Tabela 14 seguir:

Tabela 14: Propriedades do esquema *Country*.

Propriedade	Descrição
<i>Name</i>	O nome do objeto.

Todas as propriedades apresentadas nas tabelas acima foram implementadas utilizando *tags* de microdata do *schema.org* para marcação dos dados. A seção a seguir demonstra como foi realizada essa implementação na página Home do portal.

5.2.2 Implementação da Página Home

Com a modelagem dos esquemas apresentados na seção anterior, foram inseridos microdados nas *tags* HTML para estruturação destes esquemas na página Home. O Código-fonte 15 a seguir apresenta os microdatas na seção de dados da página home.

O código-fonte 15: Dados estrutura CollegeOrUniversity.

```

1. <div itemscope itemtype="https://schema.org/CollegeOrUniversity">
2. <aside id="barra-links">
3.   <div class="container">
4.     <div class="clearfix">
5.
6.   <!-- Tela grande -->
```

```

7.     <div class="row align-items-center d-lg-flex d-none tela-
grande">
8.     <div class="col-lg-4 float-left"><li id="menu-item-48"
class="menu-item menu-item-type-custom menu-item-object-custom
menu-item-48"><a href="https://conecta.ulbra-
to.br/logon">Conecta</a></li>
9. <li id="menu-item-49" class="menu-item menu-item-type-custom
menu-item-object-custom menu-item-49"><a
href="https://servicos.ulbra.br/pls/ulbra24/AAMAIN.Login">AutoAtendi
mento</a></li>
10.<li id="menu-item-50" class="menu-item menu-item-type-custom
menu-item-object-custom menu-item-50"><a
href="https://www.ulbra.br/ead">EAD</a></li>
11.<li id="menu-item-51" class="menu-item menu-item-type-custom
menu-item-object-custom menu-item-51"><a href="http://ulbra-
to.br/espaco-academico/2011/02/07/Ouvidoria">Ouvidoria</a></li>
12.</div>
13.     <div class="col-lg-4 text-center link-home">
14.         
15.     </div>
16.
17.     <div class="col-lg-4 text-right float-right d-none d-sm-block
d-sm-none d-md-block"><li id="menu-item-58" class="menu-item
menu-item-type-custom menu-item-object-custom menu-item-58"><a
href="https://www.facebook.com/CEULPULBRA/"><i class="fab fa-
facebook" aria-hidden="true"></i></a></li>
18.<li id="menu-item-59" class="menu-item menu-item-type-custom
menu-item-object-custom menu-item-59"><a
href="https://www.instagram.com/ceulpulbra/"><i class="fab fa-
instagram" aria-hidden="true"></i></a></li>
19.<li id="menu-item-60" class="menu-item menu-item-type-custom
menu-item-object-custom menu-item-60"><a><i class="fas fa-phone"
aria-hidden="true"></i> <span itemprop="telephone">3219
8000</span></a></li>
20.<li id="menu-item-63" class="traco menu-item menu-item-type-
post_type menu-item-object-page menu-item-63"><a
href="http://142.93.207.35/pesquisar/"><i class="fas fa-search" aria-
hidden="true"></i></a></li>
21.</div>     </div>

```

Conforme apresentado no código acima as propriedades inseridas para estruturação da seção de dados da instituição foram: *itemprop*, *itemtype* e *itemscope*. Desta forma, foi inserido a propriedade *telephone* (linha 19) do esquema

CollegeOrUniversity (linha 1), que representa o telefone da instituição, conforme apresentado no Código-fonte 16 acima.

O Código-fonte 16 a seguir, apresenta a estrutura HTML com microdados para seção de notícias.

O código-fonte 16: Dados estrutura CreativeWork Notícia

```

1. <div itemscope itemtype="https://schema.org/CreativeWork" class="card-
img-overlay d-flex align-items-end" >
2.   <div>
3.     <span itemprop="name" class="badge badge-dark mb-2" style="color:
#fff !important;">Portal CEULP/ULBRA</span><br>
4.     <h1 class="card-title">
5.       <a          itemprop="image"
href="http://142.93.207.35/blog/2020/03/02/titulo-da-noticia-sem-
capa/"><span          itemprop="description">Título da notícia sem
capa</span></a>
6.     </h1>
7.   </div>
8. </div>
9.

```

As informações representadas pelo Código-fonte acima, são itens do esquema *CreativeWork* (linha 1), que possui as propriedades: *name* (linha 3), que representa nome a qual está referente a notícia; *image* (linha 5) que apresenta a logo da notícia; e *description* (linha 5) aponta uma breve descrição da notícia. Para descrever o esquema para calendário acadêmico foi utilizado o Código-fonte 17 a seguir.

O código-fonte 17: Dados estrutura ScheduleAction Calendário acadêmico.

```

1. <div class="row row-calendarios">
2.   <div class="col academico" itemscope
itemtype="https://schema.org/ScheduleAction">
3.     <h2 class="mb-3">Calendário Acadêmico</h2>
4.     <ul class="list-unstyled">
5.       <div itemprop="agent" itemscope
itemtype="https://schema.org/Schedule" >
6.         <li class="media">
7.           <div class="text-center data" >
8.             <h6 class="mb-0" ><span
itemprop="byDay">07</span><br><span
itemprop="byDay">02</span></h6>

```

```

9.         </div>
10.        <div class="media-body">
11.            <h6 class="titulo mt-0 mb-1"><span itemprop="name" >Dia
Letivo </span>                </h6>
12.        </div>
13.    </li>
14. </div>
15. <div itemprop="agent" itemscope
itemtype="https://schema.org/Schedule">
16.     <li class="media">
17.         <div class="text-center data" >
18.             <h6 class="mb-0"><span
itemprop="byDay">06</span><br><span
itemprop="byMonth">02</span></h6>
19.         </div>
20.         <div class="media-body">
21.             <h6 class="titulo mt-0 mb-1"><span itemprop="name">Dia
Letivo </span>                </h6>
22.         </div>
23.     </li>
24. </div>
25. </ul>
26. <a href="http://ulbra-to.br/download/2018/12/19/Calendario-2019-
Web/download" type="button" class="button">
27.     <i class="fas fa-download fa-fw" aria-hidden="true"></i>
Calendário Acadêmico
28. </a>
29. </div>

```

As informações inseridas no código são propriedades do objeto do *ScheduleAction* (linha 2), onde é apresentado dados relacionados ao calendário acadêmico da instituição. As propriedades utilizadas foram, *agent* (linha 5) do tipo *Schedule* (linha 5), contendo *byDay* (linha 18) que representa um dia do calendario, *byMonth* (linha 18) representando o mês, e o *name* (linha 21) que indica se o dia é letivo ou não.

O Código-fonte 18 a seguir apresenta informações relacionadas ao calendário de eventos da universidade.

O código-fonte 18: Dados estrutura Schedule Calendário Eventos.

```

1. <div itemscope itemtype="http://schema.org/Schedule">
2.
3. <a href="foofighters-3.jpg" itemprop="image" hidden>

```



```

4.     
5.   </a>
6.   <a href="foofighters-3.jpg" itemprop="sameAs" hidden>
7.     
8.   </a>
9.   <a href="https://ulbra-to.br:8051/sig/extensao/acoes/452/detalhes/">
10.    <li class="media">
11.      <div class="text-center data">
12.        <h6 class="mb-0"><span itemprop="byDay">25</span><br><span
        itemprop="byMonth">08</span></h6>
13.      </div>
14.      <div class="media-body">
15.        <h6 class="titulo mt-0 mb-1"><span itemprop="name">III Jornada
        Acadêmica do Curso de Agronomia do Ceulp/Ulbra</span></h6>
16.      </div>
17.      <div itemprop="description" hidden>Um tipo adicional para o item,
        normalmente usado para adicionar tipos mais específicos de vocabulários
        externos na sintaxe de microdados. Este é um relacionamento entre algo e
        uma classe em que a coisa está. Na sintaxe RDFa, é melhor usar a sintaxe
        RDFa nativa - o atributo 'typeof' - para vários tipos. As ferramentas do
        Schema.org podem ter apenas um entendimento mais fraco de tipos extras,
        em particular aqueles definidos externamente</div>
18.    </li>
19.  </a>
20.</div>

```

As informações representadas pelo Código-fonte 18, são itens da propriedade de *Schedule* (linha 1), com as propriedades: *image* (linha 3), que representa uma imagem do evento; *sameAs* (linha 6) que se refere a página oficial do evento; *byDay* (linha 12) e *byMonth* (linha 12), representando o dia e mês de acontecimentos do evento; *name* (linha 15) que apresenta nome do evento com sua edição; e por fim *description* (linha 15), que apresenta uma breve descrição do evento. Para descrever a seção “Em destaque” foi utilizado o Código-fonte 19 a seguir:

O código-fonte 19: Dados estrutura Schedule Calendário Eventos.

```

1. <div itemprop="actionableFeedbackPolicy" itemscope
  itemType="https://schema.org/CreativeWork">
2.   <span itemprop="abstract" hidden>Noticia</span>
3.   <span itemprop="name" hidden>Noticia</span>
4.   <span itemprop="sameAs" hidden>Noticia</span>
5.   <span itemprop="url" hidden>http://142.93.207.35/</span>
6.   <a itemprop="image" href="http://142.93.207.35/">
7.     <div class="card bg-dark text-white border-0 ml-lg-auto"
      style="background: url(http://142.93.207.35/wp-

```

```

content/uploads/2020/02/9617b62b3c5d7ca88c8b455be2654638-
768x576-1-200x100.jpg);">
8.     </div>
9.   </a>
10.</div>

```

Para representar os Destaques da página home foi utilizado o código acima. Foi necessário a inserção das propriedades do esquema de *CreativeWork* (linha 1), sendo elas: *abstract* (linha 2) apresentando uma breve descrição do destaque; *name* (linha 3), apresentando o nome referente ao destaque; *sameAs* (linha 4) apontando diretamente para o site oficial; *url* (linha 5) apontando para página do CEULP/ULBRA; e por fim uma logo representada por *image* (linha 6).

A implementação da seção de cursos é apresentada no Código-fonte 20 a seguir.

O código-fonte 20: Dados estrutura Course

```

1. <div itemscope itemtype="https://schema.org/Course">
2.   <li>
3.     <a href="http://ulbra-to.br/cursos/Administracao">Administração</a>
4.   </li>
5.   <span itemprop="name" hidden>Administração</span>
6.   <span itemprop="description" hidden>Descrição</span>
7.   <div itemprop="provider" itemscope
   itemtype="https://schema.org/Organization" hidden>
8.     <span itemprop="name">asdsadsadas</span>
9.     <span itemprop="URL">http://ulbra-
to.br/cursos/Administracao</span>
10.  </div>
11.</div>

```

O Código-fonte 20, apresenta informações sobre os cursos da instituição, representadas pelo esquema *Course* (linha 1) que possui as propriedades: *name* (linha 5) representando o nome o curso (linha 5); *description* (linha 6) apresenta uma breve descrição; *provider* (linha 7) que se refere ao esquema de *Organization* (linha 7) com as propriedades *name* (linha 8) e *url* (linha 9).

O Código-fonte 21 apresentado a seguir, representa a implementação de microdados para o rodapé da página Home com os dados de endereço da instituição.

O código-fonte 21: Dados estrutura PostalAddress.

```

1. <div class="col-lg-6 text-center order-2 order-lg-1 order-md-1 footer-portal"
   itemprop="address" itemscope
   itemtype="https://schema.org/PostalAddress">
2.   <strong><span>CENTRO UNIVERSITÁRIO LUTERANO DE
   PALMAS</span></strong><br>
3.   <span itemprop="addressLocality">Avenida Teotônio Segurado
   </span>, 1501 Sul,</span><span itemprop="addressRegion">:Palmas -
   TO</span><br>
4.   CEP<span itemprop="postalCode ">77.019-900</span> <span>Caixa
   Postal nº 85</span>
5.   <span itemprop="addressCountry"hidden>BR</span>
6.   <span itemprop="email"hidden>contato@ceulp.edu.br</span>
7.   <span itemprop="name" hidden>CEULP/ULBRA</span>
8. </div>

```

No Código-fonte acima, as informações inseridas são propriedades do objeto *PostalAddress* (linha 1), sendo elas: *addressLocality* (linha 3) que apresenta avenida que está localizada a instituição; *postalCode* (linha 4) que é referente ao cep; *addressCountry* (linha 5) que representa o país que a instituição faz parte; *email* (linha 6) referente para contato; e *name* (linha 7) que retrata nome do objeto.

Com a inserção de microdados no código HTML da página de Home, foi realizada a validação dos dados estruturados seguindo a metodologia do projeto, conforme apresentado na subseção abaixo.

5.2.3 Validação de dados da página Home

A Figura 20 a seguir ilustra o resultado da ferramenta do Google após o código inserido.

Figura 20: Esquemas detectados na página Home

Detectado	0 ERROS	0 AVISOS	39 ITENS
PostalAddress	0 ERROS	0 AVISOS	1 ITEM
ScheduleAction	0 ERROS	0 AVISOS	1 ITEM
CollegeOrUniversity	0 ERROS	0 AVISOS	1 ITEM
Schedule	0 ERROS	0 AVISOS	4 ITENS
CreativeWork	0 ERROS	0 AVISOS	3 ITENS
Course	0 ERROS	0 AVISOS	29 ITENS

Conforme apresentado na figura acima, os atributos globais detectados pela ferramenta do Google para página *home* do portal foram: *PostalAddress*, *ScheduleAction*, *CollegeOrUniversity*, *Schedule*, *CreativeWork* e *Course* referentes aos dados da instituição. As propriedades do atributo global *ScheduleAction*, são observadas na Figura 21 abaixo.

Figura 21: Propriedades detectadas em ScheduleAction

ScheduleAction		0 ERROS	0 AVISOS	^
@type	ScheduleAction			
agent				
@type	Schedule			
byDay	07			
byMonth	2			
name	Dia Letivo			
agent				
@type	Schedule			
byDay	06			
byMonth	2			
name	Dia Letivo			

Conforme a figura acima, para entidade global *ScheduleAction*, a ferramenta identificou suas propriedades e valores, sendo elas: *agent* que tem como propriedade *Schedule* que possui *byDay*, *byMonth* referentes ao dia e mês do evento respectivamente; e *name*. As propriedades do atributo global *CollegeOrUniversity*, são observadas na Figura 22 a seguir.

Figura 22: Propriedades detectadas em CollegeOrUniversity

CollegeOrUniversity		0 ERROS	0 AVISOS	^
@type	CollegeOrUniversity			
telephone	3219 8000			

Na figura acima, foi identificado pela ferramenta do Google as propriedades e valores do esquema *CollegeOrUniversity*, sendo elas: *telephone* representando telefone da instituição. As propriedades do atributo global *Schedule*, são observadas na Figura 23 a seguir.

Figura 23: Propriedades detectadas em Schedule

Schedule		0 ERROS	0 AVISOS	^
@type	Schedule			
image	https://search.google.com/structured-data/testing-tool/foofighters-3.jpg			
sameAs	https://search.google.com/structured-data/testing-tool/foofighters-3.jpg			
byDay	25			
byMonth	8			
name	III Jornada Acadêmica do Curso de Agronomia do Ceulp/Ulbra			
description	Um tipo adicional para o item, normalmente usado para adicionar tipos mais específicos de vocabulários externos na sintaxe de microdados. Este é um relacionamento entre algo e uma classe em que a coisa está. Na sintaxe RDFa, é melhor usar a sintaxe RDFa nativa - o atributo 'typeof' - para vários tipos. As ferramentas do Schema.org podem ter apenas um entendimento mais fraco de tipos extras, em particular aqueles definidos externamente			

Na figura acima, foi identificado para o esquema *Schedule* as propriedades e valores: *image*, *sameAs*, *byDay*, *byMonth*, *name* e *description*.

As propriedades do atributo global *CreativeWork*, são observadas na Figura 24 abaixo.

Figura 24: Propriedades detectadas em CreativeWork

CreativeWork		0 ERROS	0 AVISOS	^
@type	CreativeWork			
name	Portal CEULP/ULBRA			
image	http://142.93.207.35/blog/2020/03/02/titulo-da-noticia-sem-capa/			
description	Título da notícia sem capa			

CreativeWork		0 ERROS	0 AVISOS	^
@type	CreativeWork			
abstract	Noticia			
name	Noticia			
sameAs	Noticia			
url	http://142.93.207.35/			

Conforme a figura acima, foi identificado as propriedades e valores do esquema *CreativeWork*, sendo elas: *abstract*, *name*, *sameAs*, *url* e *image*, *description* sendo elas utilizadas para destaque é para notícias.

As propriedades do atributo global *Course*, são observadas na Figura 25 abaixo:

Figura 25: Propriedades detectadas em Course.

Course		0 ERROS	0 AVISOS	^
@type	Course			
name	Administração			
description	Descrição			
provider				
@type	Organization			
name	asdsadsadas			
url	http://ulbra-to.br/cursos/Administracao			

Para entidade global *Course*, conforme a figura acima, foram identificadas suas propriedades e valores, sendo elas: *name*, *description*, *provider* que tem como propriedade *Organization* que possui *name* e *url*.

As propriedades do atributo global *PostalAddress*, são observadas na Figura 26 abaixo.

Figura 26: Propriedades detectadas em PostalAddress.

PostalAddress		0 ERROS	0 AVISOS	^
@type	PostalAddress			
addressLocality	Avenida Teotônio Segurado			
postalCode	77.019-900			
email	contato@ceulp.edu.br			
name	CEULP/ULBRA			
addressCountry				
@type	Country			
name	BR			

Para a entidade global *PostalAddress*, a ferramenta identificou suas propriedades e valores, sendo elas: *addressLocality*, *postalCode*, *email*, *name* e *addressCountry* que tem como propriedade *Country* com que possui nome que descreve o país da instituição.

Dessa forma, com os atributos globais e suas propriedades identificadas pela ferramenta, sem detecção de erros e avisos, é possível validar os dados estruturados também na página de Home do Portal Acadêmico do CEULP/ULBRA, conforme desenvolvido neste trabalho.

6 CONSIDERAÇÕES FINAIS

Neste trabalho foi proposta a implementação de microdata nas páginas de professor e home do Portal Acadêmico do CEULP/ULBRA. Para tanto, foram utilizados conceitos e técnicas de *linked data*, a fim de enriquecer semanticamente o conteúdo das páginas Web apresentadas neste trabalho. Para avaliar essa proposta, foram realizados testes com a ferramenta Structured Data Testing Tool da Google, que analisa e valida dados estruturados em páginas HTML.

Conforme a avaliação apresentada, foi identificado que a implementação de microdata nas páginas de professor e home do Portal Acadêmico foi satisfatória para

o reconhecimento e exibição avançada de mecanismos de busca. Dessa forma, promoveu um enriquecimento no conjunto de dados dessas páginas, possibilitando melhor compreensão por parte de usuários humanos e aplicações de consumo de dados.

Em vista disso, este trabalho apresentou que é possível a utilização de schema.org para possibilitar representação de dados estruturados de páginas HTML, utilizando para isso documentações gratuitas disponíveis na internet. Entretanto, algumas limitações foram identificadas no âmbito deste trabalho, devido algumas propriedades de esquemas do schema.org não atenderem a modelagem dos dados estruturados para as páginas do portal. A exemplo disso, o esquema *Person* apresentado para a página de professor, neste trabalho, não possibilita relacionamento com *Course*, pois este não é uma propriedade de *Person*, segundo os padrões do schema.org. Por não atender as necessidades do portal, se faz necessário então a criação de uma ontologia própria para atender todas as necessidades das páginas HTML do portal. Dessa forma, é possível estruturar todos os dados sem a necessidade de seguir somente os esquemas e propriedades restritas do schema.org.

REFERÊNCIAS

ALLEMANG, D.; HENDLER, J. **Semantic Web for the working ontologist**: effective modeling in RDFS and OWL. 2nd.ed. Amsterdam: Morgan Kaufmann, 2011. Disponível em: <http://www.kevenlw.name/downloads/Ontologist.pdf>. Acesso em: 29 nov. 2020.

ARAKAKI, F. A. **Linked data**: ligação de dados bibliográficos.2016. Dissertação (Mestrado em Ciência da Informação - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista - Marília, 2016.

ARAYA, Elizabeth Roxana Mass; VIDOTTI, Silvana Aparecida Borsetti Gregorio. Direito autoral e tecnologias de informação e comunicação no contexto da produção, uso e disseminação de informação: um olhar para as Licenças Creative Commons. **Informação & Sociedade**: estudos. João Pessoa: Universidade Federal da Paraíba (UFPB), v. 19, n. 3, p. 39-51, 2009. Disponível em: <https://repositorio.unesp.br/handle/11449/10564>. Acesso em: 9 dez. 2020.

BAX, M. P.; REZENDE, B.V. **Projeto Indexa**: ferramenta de auxílio à divulgação de informações na Web. 2001. Disponível em: http://www.paradigma.com.br/artigos/artigos_04.pdf. Acesso em: 20 dez. 2020.

BERNERS-LEE, T. **Semantic Web road map**. 1998. Disponível em: <http://www.w3.org/DesignIssues/Semantic.html>. Acesso em: 18 set. 2020.

BIZER, Christian; HEATH, Tom; BERNERS-LEE, Tim. Linked data - The story so far. **International Journal on Semantic Web and Information Systems**, [S. l.], v. 5, n. 3, p. 1–22, 2009. Disponível em: https://www.researchgate.net/publication/225070216_Linked_Data_The_Story_so_Far. Acesso em: 3 jan. 2021.

BRANDÃO, A. A. F.; LUCENA, C. J. P. Uma Introdução à Engenharia de Ontologias no contexto da Web Semântica. **PUC-RioInf**, Rio de Janeiro, 2002. Disponível em: http://www.dbd.puc-rio.br/depto_informatica/02_29_brandao.pdf. Acesso em: 3 jan. 2021.

BRANSKI, Regina Meyer. Recuperação de informações na Web. **Perspect. Ciênc. Inf.**, Belo Horizonte, v. 9, n. 1, p. 70-87, jan./jun. 2004. Disponível em: <http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/351>. Acesso em: 10 dez. 2021.

BREITMAN, Karin Koogan. **Web Semântica: a Internet do futuro**. 1. ed. Rio de Janeiro: LTC, 2005.

CLARK, K. G. SPARQL Protocol for RDF. Criação, proteção e uso legal de informação em ambientes da World Wide Web. **W3C Working Draft**. 2005. São Paulo. Disponível em: <http://www.w3.org/TR/2005/WD-rdf-sparql-protocol-20050914>. Acesso em: 29 nov. 2020.

DAUM, Berthold; MERTEN, Udo. **Arquitetura de sistemas com XML**. Rio de Janeiro: Campus, 2002.

DZIEKANIAK, Gisele Vasconcelos; KIRINUS, Josiane Boeira. Web Semântica. **Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação**, Rio Grande, v. 9, n. 18, p. 20–39, 2004. Disponível em: <https://periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2004v9n18p20>. Acesso em: 11 jan. 2021.

FERREIRA, J. A.; SANTOS, P. L. V. A. DA C. O modelo de dados Resource Description Framework (RDF) e o seu papel na descrição de recursos. **Informação & Sociedade: Estudos**, João Pessoa, v. 23, n. 2, p. 13-23, maio/ ago. 2013. Disponível em: <https://periodicos.ufpb.br/ojs/index.php/ies/article/view/15436/9681>. Acesso em: 29 nov. 2020.

FILHO, S. N. V., Moura, A. M. C., and Cavalcanti, M. C. R. (2010). Armazenamento e manipulação de ontologias utilizando sistemas gerenciadores de banco de dados. Technical report, Instituto Militar de Engenharia (IME), Rio de Janeiro. 2012.

FORTI, Hericson Ramos. Um estudo de caso de utilização de microdados e microformatos para tornar conteúdo na Web compreensível por máquina. Trabalho de Conclusão de Curso (Especialização) - Universidade Tecnológica Federal do Paraná, Londrina, 2013. Disponível em: <http://repositorio.utfpr.edu.br/jspui/handle/1/20155>. Acesso em: 12 jan. 2021.

(Rene Faustino Gabriel Junior)FREE, Graphdb. XVII Encontro Nacional de Pesquisa em Ciência da Informação (XVII ENANCIB) GT 7 – Produção e Comunicação da Informação em Ciência , Tecnologia & Inovação UTILIZAÇÃO DA WEB SEMANTICA E RDF EM ESTUDOS MÉTRICOS DA INFORMAÇÃO .

GONÇALVES, V.; CARRAPATOSO, E. Web semântica e cérebro global juntos por uma boa causa. **EduSer: Revista de Educação**, Bragança, v.1, n.1, p. 70-87, 2009. Disponível em: <https://bibliotecadigital.ipb.pt/handle/10198/1219> . Acesso em: 28 out. 2020.

GROSOFF, B. N. *et al.* Description logic programs: Combining logic programs with description logics. **Social Science Research Network**, Budapest, mai. 2003, p. 48-57. Disponível em: <http://www2003.org/cdrom/papers/refereed/p117/p117-grosof.html>. Acesso em: 16 set. 2020.

HAUSENBLAS, Michael. Exploiting Linked Data to Build Web Applications. National University of Ireland. **IEEE Internet Computing**, v. 13, Issue: 4, July-Aug. 2009. Disponível em: <https://ieeexplore.ieee.org/document/5167270/authors#authors>. Acesso em: 13 jan. 2021.

HUBNER, J. F; BRANDÃO, A. A. F., SILVEIRA, R.; MARCHI, J. Uma Introdução à Engenharia de Ontologias no contexto da Web Semântica. **World Wide Web Internet And Web Information Systems**. 2nd. ed. Florianópolis, 2012. Disponível em:

https://www.academia.edu/31882188/Proceedings_of_Wesaac_2012. Acesso em: 3 dez. 2021.

ISOTANI, S. *et al.* Estado da Arte em Web Semântica e Web 2.0: Potencialidades e Tendências da Nova Geração de Ambientes de Ensino na Internet. **Revista Brasileira de Informática na Educação**, [S. l.], v. 17, n. 1, p. 30–42, 2009. Disponível em: <https://www.br-ie.org/pub/index.php/rbie/article/view/4/4>. Acesso em: 4 jan. 2021.

ISOTANI, S. *et al.* Web 3.0-Os Rumos da Web Semântica e da Web 2.0 nos Ambientes Educacionais. In: Simpósio Brasileiro de Informática na Educação- SBIE. **Anais do Simpósio** [S. l.], n. Sbie 2008, p. 785–795, 2008. Disponível em: <http://br-ie.org/pub/index.php/sbie/article/view/767>. Acesso em: 12 jan. 2021.

FERNANDEZ, Jorge González. Um olhar para o mundo de microformas. **Localización: Revista Cubana de Información en Ciencias de la Salud**, v.17, n.3, mar, 2008. ISSN 1024-9435, ISSN-e 2307-2113.

LANTHALER, M.; GÜTL, C. On using json-ld to create evolvable restful services. **WS-REST**. [S.l.]: ACM, 2012. p. 25–32. ISBN 978-14503-1190-8. Disponível em: <https://www.bibsonomy.org/bibtex/469b8fe7e6cb81f486c6fa85cf6d5168>. Acesso em: 11 jan. 2021.

MARSHALL, Catherine C.; SHIPMAN, Frank M. Which Semantic Web? HIPERTEXTO '03: Procedimentos da décima quarta conferência ACM sobre hipertexto e hiperídia. Nova York, NY, EUA, p. 57-66, 2003. **Microformats.org**. Disponível em: <https://dl.acm.org/doi/10.1145/900051.900063>. Acesso em: 3 jul. 2020.

MARTINS JÚNIOR, J. **Classificação das páginas na Internet**. 2003. Dissertação (Mestrado em Ciência da Computação e Matemática Computacional) – Instituto de Ciências e Matemáticas e de Computação, Universidade de São Paulo, São Paulo, 2003. Disponível em: https://teses.usp.br/teses/disponiveis/55/55134/tde-12092003-101358/publico/Martins_Dissertacao.pdf. Acesso em: 21 set. 2020.

MAVRODY, Sergey. **Sergey's html5 e CSS3 Quick Reference**. 2. ed. [S. l.]: Belisso, 2012. E-book.

MILLER, E. An Introduction to the Resource Description Framework. **Boletim da Sociedade Americana de Ciência e Tecnologia da Informação**. [S.l.], 1. ed., v. 25, out./nov. 1998. Disponível: <https://asistdl.onlinelibrary.wiley.com/doi/epdf/10.1002/bult.105>. Acesso em: 18 set. 2020.

Miranda, D. S.; Azevedo, L.L.S.; Magalhães, R.P. Consumindo Linked Data na Web. In: Encontro UNificado de computação em Parnaíba, 2011, Parnaíba. Consumindo Linked Data na Web, 2011. Disponível em: <http://docplayer.com.br/1861923-Consumindo-linked-data-na-web.html>. Acesso em: 12 jan. 2021.

MOREIRA, Alexandra; ALVARENGA, Lídia; OLIVEIRA, Alcione de Paiva. O nível do conhecimento e os instrumentos de representação: tesauros e ontologias.

DataGramaZero, v. 5, n. 6, 2004, v. 5, n. 6, dez. 2004. Disponível em: <https://brapci.inf.br/index.php/res/v/6563>. Acesso em: 10 ago. 2020.

ORA, Lassila; SWICK, Ralph R. Resource Description Framework (RDF) Model and Syntax Specification. **World Wide Web Consortium**, [S.l.], fev. 1999. Disponível em: <https://www.w3.org/TR/1999/REC-rdf-syntax-19990222/>. Acesso em: 13 jan. 2021.

MUCHERONI, Marcos Luiz; PALETTA, Francisco Carlos. O desenvolvimento da Web 3.0: Linked Data e DBpedia. **Anais..** São Paulo: [s.n.], 2014. Disponível em: <https://repositorio.usp.br/item/002659100>. Acesso em: 13 jan. 2021.

PAULHEIM, Heiko. What the adoption of schema.org tells about linked open data. **CEUR Workshop Proceedings**, [S. l.], v. 1362, p. 85–90, 2015. Disponível em: <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.1057.1285&rep=rep1&type=pdf>. Acesso em: 13 jan. 2021.

PRUD'HOMMEAUX, E.; SEABORNE, A. SPARQL Working Draft: Sparql Query Language For RDF. **W3C Working Draft**, 2005. Disponível em: <https://www.w3.org/TR/2005/WD-rdf-sparql-query-20050217/>. Acesso em: 16 nov. 2020.

RAMALHO, R. A. S. VIDOTTI, S. A. B. G.; FUJITA, M. S. L. Semantic Web: an inquiry under the view of the Information Science. **DataGramaZero**, [S. l.], v. 8, n. 6, p. A04. 2007. Disponível em: https://www.marilia.unesp.br/Home/Pos-Graduacao/CienciadaInformacao/Dissertacoes/ramalho_ras_do_mar.pdf. Acesso em: 23 nov. 2020.

REIS, T. K. B. **Linked data utilizando o vocabulário schema.org: uma aplicação no portal (EN)Cena**. 2013. Trabalho de Conclusão de Curso (Bacharelado em Sistemas de Informação). Centro Universitário Luterano de Palmas, Palmas, Tocantins, 2013. Disponível em: <http://ulbrato.br/bibliotecadigital/publico/home/documento/56>. Acesso em: 29 jan. 2021

RONALLO, Jason. HTML5 Microdata and Schema.org. **Code{4}lib journal**, [S. l.], n. 16, 3. 2012. Disponível em: <https://journal.code4lib.org/articles/6400>. Acesso em: 29 jan. 2021.

ROSA, P. A. Web Semântica. [S. l.: S. n.], 2002. Disponível em: http://www.ime.usp.br/~yw/ano2002/mac5701/sem2/rosa_final . Acesso em: 26 jun. 2020.

SANTAREM SEGUNDO, J. E. **Recursos tecno-metodológicos para a descrição e recuperação de informações na web**. 2004, 157 f. Dissertação (Mestrado em Ciência da Informação) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2004. Disponível em: <https://repositorio.unesp.br/handle/11449/93618>. Acesso em: 29 jan. 2021.

SANTOS, Fabio Rodrigues Dos.. **Um modelo semântico para integração automática de conteúdo com um agente conversacional**. [S. l.], 2016. Disponível em: <http://www.repositorio.jesuita.org.br/handle/UNISINOS/5243> .

SANTOS, N. H. Dos. **MICROGISBORNE: Uma ferramenta para recuperação da informação em blogs baseada em microformatos**. Trabalho de conclusão de curso (Gestão da Informação) - Universidade Federal de Pernambuco Recife, 2014.

SHADBOLT, N., BERNERS-LEE, T. The semantic Web revisited. **IEEE Intelligent Systems**, v. 21(3), 96–101. Disponível em: <https://ieeexplore.ieee.org/document/1637364/authors#authors>. Acesso em: 16 jan. 2021.

SOUZA, Renato Rocha; ALVARENGA, Lídia. **A Web Semântica e suas contribuições para a ciência da informação**. *Ci. Inf.* [online]. 2004, vol.33, n.1, pp.132-141. ISSN 1518-8353. Disponível em: <https://doi.org/10.1590/S0100-19652004000100016>. Acesso em: 16 jan. 2021.

SPORNY, Manu; KELLOGG, Gregg; LANTHALER, Markus. Json-Ld 1.0. A JSON based Serialization for Linked Data. **W3C [S. l.]**, n. January, p. 1–33, 2013. Disponível em: <https://json-ld.org/spec/FCGS/json-ld-syntax/20130222/diff-20130202.html>. Acesso em: 18 jan. 2021.

BERNERS-LEE, T.; HENDLER, J.; LASSILA, O. The Semantic Web: A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities, **Scientific American**, [S.l.], (2001). Disponível em: https://www.sop.inria.fr/acacia/cours/essi2006/Scientific%20American_%20Feature%20Article_%20The%20Semantic%20Web_%20May%202001.pdf. Acesso em: 18 jan. 2021.

HEATH, T.; BIZER, C. **Linked Data: Evolving the Web into a Global Data Space**, 1st ed. New York, New York, USA: **Morgan & Claypool Publishers**, 2011, p. 137. Disponível em: <http://linkeddatabook.com/editions/1.0/>. Acesso em: 18 jan. 2021.

TURCATEL, Isadora Oliveira. **Ontologias para descoberta de recursos na ciência**. 2014. Trabalho de Conclusão de Curso (Bacharelado em Biblioteconomia) - Universidade Federal do Rio Grande do Sul, Porto Alegre, 2014. Disponível em: <http://www.lume.ufrgs.br/handle/10183/112165>. Acesso em: 18 jan. 2021.

WILMA, E. R.; LEOBINO, N. S.; J. A. S. M. 2016. Recomendações de Características Ergonômicas para Interfaces de Sistemas de Monitoramento de Redes Baseadas em Critérios de Usabilidade. In Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos (SBRC). **Anais de Simpósio**, [S.l.], 2016.