



**RODRIGO CASÇÃO ARAÚJO**

**ALINHAMENTO DE METADADOS DA INDÚSTRIA DE *BROADCAST* MULTIMÍDIA  
NO CONTEXTO DA TV DIGITAL COM A WEB SEMÂNTICA**

**CAMPINAS  
2013**





**UNIVERSIDADE ESTADUAL DE CAMPINAS  
FACULDADE DE ENGENHARIA ELÉTRICA E DE COMPUTAÇÃO**

**RODRIGO CASÇÃO ARAÚJO**

**ALINHAMENTO DE METADADOS DA INDÚSTRIA DE *BROADCAST* MULTIMÍDIA  
NO CONTEXTO DA TV DIGITAL COM A WEB SEMÂNTICA**

**Orientador: Prof. Dr. Ivan Luiz Marques Ricarte**

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da Faculdade de Engenharia Elétrica e de Computação da Universidade Estadual de Campinas para obtenção do título de Doutor em Engenharia Elétrica, na área de Engenharia de Computação.

ESTE EXEMPLAR CORRESPONDE À VERSÃO FINAL DA TESE  
DEFENDIDA PELO ALUNO RODRIGO CASÇÃO ARAÚJO  
E ORIENTADA PELO PROF. DR. IVAN LUIZ MARQUES RICARTE

---

**CAMPINAS  
2013**

Ficha catalográfica  
Universidade Estadual de Campinas  
Biblioteca da Área de Engenharia e Arquitetura  
Rose Meire da Silva - CRB 8/5974

Ar15a Araújo, Rodrigo Cascão, 1975-  
Alinhamento de metadados da indústria de broadcast multimídia no contexto da TV digital com a web semântica / Rodrigo Cascão Araújo. – Campinas, SP : [s.n.], 2013.

Orientador: Ivan Luiz Marques Ricarte.  
Tese (doutorado) – Universidade Estadual de Campinas, Faculdade de Engenharia Elétrica e de Computação.

1. Televisão digital. 2. Web semântica. 3. Multimídia interativa. 4. sistemas multimídia. 5. Ontologias (Recuperação da informação). I. Ricarte, Ivan Luiz Marques, 1962-. II. Universidade Estadual de Campinas. Faculdade de Engenharia Elétrica e de Computação. III. Título.

Informações para Biblioteca Digital

**Título em inglês:** Alignment of broadcast multimedia industry metadata in the context of digital tv with the semantic web

**Palavras-chave em inglês:**

Digital television

Semantic web

Interactive multimedia

Multimedia systems

Ontologies (Information retrieval)

**Área de concentração:** Engenharia de Computação

**Titulação:** Doutor em Engenharia Elétrica

**Banca examinadora:**

Ivan Luiz Marques Ricarte [Orientador]

Daniel Moutinho Pataca

José Eduardo Santarém Segundo

Léo Pini Magalhães

Maurício Ferreira Magalhães

**Data de defesa:** 08-03-2013

**Programa de Pós-Graduação:** Engenharia Elétrica

## COMISSÃO JULGADORA - TESE DE DOUTORADO

**Candidato:** Rodrigo Cascão Araújo

**Data da Defesa:** 8 de março de 2013

**Título da Tese:** "Alinhamento de Metadados da Indústria de Broadcast Multímídia no Contexto da TV Digital com a Web Semântica"

Prof. Dr. Ivan Luiz Marques Ricarte (Presidente): \_\_\_\_\_

Dr. Daniel Moutinho Pataca: \_\_\_\_\_

Dr. José Eduardo Santarém Segundo: \_\_\_\_\_

Prof. Dr. Léo Pini Magalhães: \_\_\_\_\_

Prof. Dr. Maurício Ferreira Magalhães: \_\_\_\_\_



Dedico este trabalho  
à minha mãe, Glória;  
ao meu pai, José Henrique;  
à minha esposa Simone.





# Agradecimentos

Ao meu orientador, Prof. Dr. Ivan Luiz Marques Ricarte, pela sua disponibilidade, pelas discussões teóricas e técnicas e pelas valiosas sugestões no encaminhamento de meu trabalho.

À Faculdade de Engenharia Elétrica e de Computação, que me acolheu como aluno, e a seus professores pela oportunidade de conviver com pessoas sábias que me ajudaram a abrir a mente para novas possibilidades e desafios de aprendizado.

Aos meus pais, José Henrique e Glória, pelo seu amor e lições de vida. A eles eu agradeço os valores ensinados e o incentivo e apoio em todas as minhas iniciativas para aprimorar a minha educação.

A minha esposa, Simone, e a minhas filhas, Maria Fernanda e Maria Clara, que foram as âncoras que me deram paz e equilíbrio e que me ajudaram a manter o meu curso de vida e o meu foco no trabalho nos inúmeros imprevistos ocorridos no período de meu doutorado.

Agradeço acima de tudo a Deus que é base de tudo em nossas vidas que nos dá a força interior que necessitamos para alcançar nossos objetivos e cumprir nosso destino.



# Resumo

A integração da Internet e das tecnologias de comunicação móveis com as plataformas de televisão têm provido aos telespectadores novos serviços interativos de conteúdo digital. Devido a estes fatores, os equipamentos para o consumidor têm se tornado cada vez mais sofisticados, suportando uma variedade de conteúdos e conectividade com outras redes e dispositivos.

A TV digital é uma plataforma híbrida que combina elementos da televisão tradicional com a Internet, provendo ao usuário o acesso a uma diversidade de conteúdos de mídia interativa. Com o crescimento do volume e da diversidade de serviços e conteúdos multimídia, a televisão está enfrentando os mesmos desafios de complexidade e excesso de informações que já vinham sendo encarados por outras mídias digitais relacionadas com a Internet.

A tecnologia de metadados pode ser uma alternativa para lidar com esta complexidade de serviços e conteúdos digitais de forma prática e eficiente. Metadados são dados que complementam as informações digitais dos conteúdos multimídia com o objetivo de descrevê-los de forma sintática e semântica, facilitando a estruturação e o gerenciamento de grandes volumes de informação. O uso de metadados em TV digital não se restringe a construção de um ferramental de busca e indexação de conteúdos multimídia, e abre oportunidade para o desenvolvimento de uma gama de serviços inovadores.

Atualmente existem diversas especificações de metadados utilizadas pela indústria de broadcast multimídia em redes de TV digital. Além disso, existem na Internet diversos repositórios de informação baseados em metadados que complementam as informações de metadados da TV digital. Contudo, como os padrões de metadados da TV digital e da Internet são baseados em diferentes especificações não relacionadas, surge o problema de como integrar estas informações, visando criar novos serviços para telespectadores que utilizem tanto informações de metadados da TV digital como informações de metadados da Internet.

Esta tese de doutorado propõe um processo para alinhamento das especificações de metadados existentes em redes abertas de transmissão e recepção de TV digital terrestre com ontologias orientadas para a descrição de domínios de conhecimento específicos existentes em repositórios da Internet, utilizando tecnologias propostas pelo W3C para a Web Semântica.

O processo proposto permitirá que o usuário da TV digital possa facilmente pesquisar conteúdos de interesse a partir da grade de programação dos canais existentes e dos conteúdos já gravados em seu receptor; receber sugestões de conteúdos para exibição ou gravação conforme o seu perfil e interesse; enriquecer sua experiência de assistir televisão acessando informações complementares sobre os programas transmitidos como sinopses, críticas especializadas, histórico do elenco e direção, premiações recebidas, fotos, vídeos e conteúdos relacionados disponíveis para livre acesso via Internet; entre outras funções.

A presente proposta foi validada através de uma prova de conceito implementada em um receptor híbrido de TV digital, que demonstrou a viabilidade de sua operacionalização sem a necessidade de impactar os padrões utilizados no Brasil para transmissão de sinal de TV digital terrestre (ISDB-T).

**Palavras chave:** televisão digital, web semântica, multimídia interativa, sistemas multimídia, ontologias



# *Abstract*

*The integration of the Internet and communication technologies with mobile TV platforms has provided viewers with new interactive services for digital content. Due to these factors, the equipment for the consumer have become increasingly sophisticated, supporting a variety of content and connectivity with other networks and devices.*

*Digital TV is a hybrid platform that combines elements of traditional television with the Internet, providing the user access to a variety of interactive media content. With the growth in the volume and diversity of services and multimedia content, television is experiencing the same challenges of complexity and information overload that were already being seen by other digital media related to the Internet.*

*The metadata technology can be an alternative to deal with this complexity of digital content and services in a practical and efficient way. Metadata is data that supplement the information of digital multimedia contents in order to describe them in a syntactic and semantic form, facilitating the structuring and management of large volumes of information. The use of metadata in digital TV is not restricted to building a tool for search and indexing of multimedia content, and opens opportunities to develop a range of innovative services.*

*Currently there are several metadata specifications used by the broadcast industry in multimedia digital TV networks. Moreover, there are many metadata-based information repositories in the Internet that complement the metadata information of digital TV. However, as the metadata standards of digital TV and Internet are based on different unrelated specifications, a problem arises of how to integrate this information in order to create new services for viewers using both metadata information of digital TV as metadata information of Internet.*

*This thesis proposes a method for alignment of existing metadata specifications in open networks for transmission and reception of digital terrestrial TV with ontologies oriented for describing specific domains of knowledge existing in Internet repositories, using technologies proposed by the W3C for the Web semantics.*

*The proposed process allows the digital TV user to easily search for content of interest from the program schedule of the existing channels and content already recorded in the receiver; to receive suggestions of content for viewing or recording according to his interest and profile; to enrich the experience of watching television by accessing additional information about the transmitted programs as synopses, specialized reviews, history of cast and direction, awards received, photos, videos and related content available for free access via Internet; among other functions.*

*This proposal has been validated through a proof of concept implemented in a hybrid digital TV receiver, which demonstrated the feasibility of its implementation without impacting the standards used in Brazil for signal transmission of digital terrestrial TV (ISDB-T).*

**Keywords:** *digital television, semantic web, interactive multimedia, multimedia systems, ontologies*



# Lista de Figuras

Figura 2.1 – Fragmento de arquivo XMLTV publicado pelo site Revista Eletrônica.....	8
Figura 2.2 – Fragmento de arquivo no padrão MPEG-7.....	9
Figura 2.3 – Fragmento de arquivo no padrão MPEG-21.....	11
Figura 2.4 – Fragmento de arquivo no padrão TVA.....	13
Figura 2.5 – Página de visualização da entidade “Pulp Fiction” do repositório Freebase.....	21
Figura 2.6 – Página de visualização da entidade “Star Wars” do repositório DBpedia.....	22
Figura 2.7 – Entidade relacionada ao filme “King Kong” da ontologia LinkedMDB.....	23
Figura 3.1 – Componentes do processo proposto.....	29
Figura 3.2 – Interdependência entre os componentes do processo desenvolvido.....	30
Figura 3.3 – Representação gráfica parcial da ontologia OWLTV obtida a partir do padrão XMLTV.....	34
Figura 3.4 – Diagrama de classes da ontologia OWLTV obtida a partir do padrão XMLTV.....	35
Figura 3.5 – Instâncias da classe Freebase.type.object.name para a entidade “Pulp Fiction”.....	36
Figura 3.6 – Instância da classe Freebase.type.object.name para a entidade “Jornal Nacional”.....	37
Figura 3.7 – Classe sameAsQuery que alinha entidades OWLTV com entidades de ontologias distintas.....	38
Figura 3.8 – Exemplo de instância da classe sameAsQuery.....	38
Figura 3.9 – Resultado da consulta MQL da Figura 3.7.....	39
Figura 3.10 – Página de visualização da entidade “/en/jornal_nacional” do repositório Freebase.....	39
Figura 3.11 – Exemplos de instâncias da classe sameAsQuery.....	40
Figura 3.12 – Guia eletrônico de programação (EPG) da TV digital.....	41
Figura 3.13 – Informações adicionais dos repositórios de conteúdo da Web Semântica.....	42
Figura 4.1 – Visão geral do processo.....	43
Figura 4.2 – Fragmento de arquivo XMLTV publicado pelo site Revista Eletrônica.....	48
Figura 4.3 – Modelo de arquivo XMLTV a ser armazenado localmente na memória do receptor de TV digital.....	49
Figura 4.4 – Arquivo XMLTV gerado a partir dos metadados da tabela EIT.....	50
Figura 4.5 – Ontologia OWLTV obtida a partir do arquivo XMLTV da Figura 4.2.....	53
Figura 4.6 – Ontologia OWLTV obtida a partir do arquivo XMLTV da Figura 4.3.....	55
Figura 4.7 – Inclusão das instâncias da classe sameAsQuery na ontologia da Figura 4.5.....	58
Figura 4.8 – Obtenção das propriedades owl:sameAs a partir da execução das consultas sameAsQuery.....	60
Figura 5.1 – Cenário do programa “Chocolate com Pimenta”.....	65
Figura 5.2 – Cenário do programa “Táxi”.....	66
Figura 5.3 – Telas de exemplos de aplicações EPG.....	67
Figura 5.4 – Estrutura do sistema desenvolvido em Gíngua-J.....	68
Figura 5.5 – Tela da inicial da aplicação para o usuário final.....	69
Figura 5.6 – Tela inicial da pesquisa de informações na Web Semântica.....	69
Figura 5.7 – Tela de resultados da pesquisa ao repositório Freebase.....	70
Figura 5.8 – Página do conteúdo no repositório Freebase.....	71
Figura 5.9 – Tela inicial da pesquisa de informações na Web Semântica.....	71
Figura 5.10 – Tela de resultados da pesquisa ao repositório DBpedia.....	72
Figura 5.11 – Página do conteúdo no repositório DBpedia.....	72
Figura 5.12 – Tela inicial da aplicação no cenário do programa “Táxi”.....	73
Figura 5.13 – Tela de resultados da pesquisa ao repositório Freebase.....	74
Figura 5.14 – Página do conteúdo no repositório Freebase.....	74
Figura 5.15 – Tela de resultados da pesquisa ao repositório Freebase.....	75
Figura 5.16 – Página do conteúdo no repositório Freebase.....	75
Figura 5.17 – Página da entidade relacionada ao diretor “Quentin Tarantino” no repositório Freebase.....	76
Figura 5.18 – Instância da classe sameAsQuery.....	77
Figura A.1 – Ontologia OWLTV utilizada na implementação da prova de conceito.....	93





# Lista de Tabelas

Tabela 2.1 – Comparativo dos trabalhos apresentados .....	20
Tabela 3.1 – Exemplificação da aplicação da metodologia proposta por Thuy et al. (2009).....	33
Tabela 4.1 – Mapeamento entre metadados EIT e XMLTV utilizando o pacote javax.tv.service.guide .....	49
Tabela 4.2 – Comparativo dos arquivos OWLTV apresentados nas Figuras 4.4 e 4.5 .....	55



# Glossário

**Alta definição:** (High definition) Termo usado para caracterizar tecnologia que possibilita alta resolução de imagem para vídeo. Para TV, temos hoje duas resoluções disponíveis 720 linhas (HD Ready) e 1080 linhas (Full HD).

**Antena UHF:** Modelo de antena que capta o sinal de TV transmitido em UHF - frequência usada para transmissão da TV digital aberta no Brasil.

**ATSC:** (*Advanced Television System Committee*) Padrão Americano de TV Digital. Foi o primeiro sistema a ser desenvolvido e tem como principal aplicação a TV de alta definição (HDTV). É o único a usar a modulação 8-VSB. Não é compatível com o sistema brasileiro de TV Digital.

**Broadcast:** (Do Inglês, "transmitir") ou Radiodifusão é o processo pelo qual se transmite ou difunde determinada informação, tendo como principal característica que a mesma informação está sendo enviada para muitos receptores ao mesmo tempo. As transmissões de TV aberta são via broadcast, onde uma ou mais antenas de transmissão enviam o sinal televisivo (ou, radiodifusor) e qualquer aparelho de TV (ou, rádio) que conseguir captar poderá sintonizar o sinal.

**Canal de Retorno:** Possibilita o tráfego de informações entre o telespectador e a emissora. Pode ser via rede de telefonia comum (utilizando um modem parecido com o de um computador), rede de telefonia celular, sistemas de internet por cabo ou por DSL (Speedy, Velox, etc.), internet sem fio (Wi-Fi, Wimax) ou qualquer outra tecnologia disponível na casa do telespectador.

**Canal Digital:** Canal de 6MHz destinado à transmissão do sinal digital. Será emprestado pelo governo as emissoras que já possuem licença de TV aberta que atenderem as exigências estabelecidas.

**COFDM:** Tipo de modulação usado pelos sistemas europeu (DVB), japonês (ISDB) e brasileiro.

**Conversor ISDB-T:** (ou Conversor Digital ISDB-T) Componente que converte o sinal da TV Digital para exibição das imagens no televisor. Em inglês, é conhecido como set-top-box. Pode ser vendido separadamente, no formato de conversor "set-top box", ou estar incorporado (integrado) ao televisor. Pode oferecer diversos tipos de saídas, dentre elas: HDMI, Vídeo Componente, S-Vídeo ou Vídeo Composto, além de saídas de áudio analógicas e digitais.

**Datacasting:** É a transmissão de dados para interatividade, relacionados ou não à programação, para as TVs Digitais. As transmissões de datacasting são do tipo one-to-many, ou seja, um mesmo conteúdo é transmitido para diversos espectadores, que podem escolher que informações preferem acessar.

**Definição:** É o nível de detalhamento que a imagem pode possuir. É medida em número de linhas horizontais x número de pontos em cada linha (pixels), padronizada no sistema ISDB para

receptores fixos em: 480 linhas x 720 pixels, 720 linhas x 1280 pixels e 1080 linhas x 1920 pixels; e para receptores portáteis em: 352 x 288 (CIF), 320 x 240 (4:3 QVGA), 320 x 180 (16:9 QVGA), 160 x 120 (4:3 SQVGA) e 160 x 90 (16:9 SQVGA).

**Definição do televisor:** É a definição que o televisor ou monitor de TV foi projetado para exibir. Caso a definição do sinal recebido seja maior ou menor que a definição nativa do televisor ou monitor de TV, será ajustado automaticamente para a definição do mesmo (Upconvert ou Downconvert).

**Definição Horizontal:** É o nível de detalhamento que a imagem pode possuir na horizontal. É medida em número de pontos (pixels) em cada uma das linhas, padronizada no sistema ISDB em 720, 1280 e 1920 colunas ou pontos por linha.

**Definição Vertical:** É o nível de detalhamento que a imagem pode possuir na vertical. É medida em número de linhas horizontais, padronizada no sistema ISDB em 480, 720 e 1080 linhas.

**Display:** É a tela da TV. Pode ser usado independente do sistema, bastando que seja acoplado ao receptor do padrão local de transmissão digital. Conceito parecido com o de monitor, utilizado em computadores. Mesmo os televisores do padrão americano ATSC, se possuírem entrada de áudio e vídeo externa, poderão ser utilizados se acoplados a um conversor do padrão brasileiro. Nesse caso, a entrada para antena do televisor ficará desligada e esse televisor será utilizado como display, ou monitor, pois a recepção do sinal será feita pelo conversor. A TV mostrará somente a imagem recebida pelo conversor.

**Downconvert:** Quando um televisor recebe sinal cuja definição é superior a da definição nativa, este é automaticamente ajustado, reduzindo a definição original do sinal.

**DVB:** (*Digital Video Broadcasting*) Padrão europeu de TV Digital. Foi desenvolvido depois do americano ATSC com o intuito de ampliar a competitividade entre a TV aberta e a TV por assinatura – TV via satélite (DTH) e TV a cabo. Sua principal aplicação é a transmissão de múltiplos programas em um só canal. Utiliza a modulação COFDM, também usada no sistema ISDB. Não é compatível com o sistema brasileiro de TV terrestre.

**EPG:** (*Electronic Program Guide* - Guia Eletrônico de Programação) Com ele, o telespectador vai ver a grade de programação da emissora de TV, na tela. Para usufruir do recurso o receptor precisa ter o suporte. No sinal HDTV a programação de uma semana pode ser consultada, no One-seg a programação é diária.

**Formato da Imagem:** O formato atual de imagem mais comum é o 4:3. Porém, como algumas gravações de DVDs, por exemplo, são feitas na proporção 16:9 (wide screen), criou-se também o 4:3 (letter box), que cria faixas pretas nas extremidades superior e inferior da tela, reduzindo a imagem sem distorcê-la.

**Full HD:** Maior resolução de tela para a TV digital, 1080 linhas horizontais por 1920 verticais.

**Ginga:** O Ginga é um middleware, software que tem o papel de garantir que as aplicações interativas irão funcionar nos televisores e conversores de todos os fabricantes. Parte do código é aberto, o que significa que as pessoas poderão estudá-lo, usá-lo e modificá-lo sem o pagamento de licenças, como acontece com o sistema operacional Linux.

**Ginga-J:** Subsistema do middleware Ginga com suporte a especificação JavaDTV.

**Ginga-NCL:** Subsistema do middleware Ginga com suporte a linguagem NCL.

**HD Ready:** Modelo de televisão de plasma ou LCD que reproduz imagem em alta definição, com resolução de 720 linhas horizontais.

**HDMI:** Este tipo de conexão permite que informações digitais de imagem e som sejam transmitidas juntas sem perda de dados. Já disponível nas TVs e DVD players, de alta definição.

**HDTV:** (*High definition TV* ou TV de alta definição) Existem hoje dois tipos de HDTV - telas de Plasma e LCD (cristal líquido) -. Se a resolução for a mesma, ambas apresentarão qualidade similar. A imagem será melhor, quanto maior a resolução da tela. Atualmente as TVs com maior resolução são as Full HD 1080i, as HD Ready possuem 720 linhas. Modelos de HDTV com conversor digital integrado também começam a ser lançados, nesse caso elas já estão prontas para receber o sinal da TV Digital. Geralmente as HDTVs são encontradas no mercado nos seguintes tamanhos: LCD - 19, 22, 26, 32, 37, 40, 42, 46, 47, 52 polegadas; plasma - 19, 32, 42, 50, 63 polegadas.

**Interlaced Scan:** (i) (imagem entrelaçada) Embora as imagens que vemos nos televisores aparentem estar preenchendo toda a tela de uma só vez, são formadas em linhas. A imagem é chamada de entrelaçada porque são exibidas primeiramente todas as linhas ímpares (1, 3, 5, 7...), e, somente após o preenchimento de toda a tela, é iniciada a reprodução das linhas pares (2, 4, 6, 8...). A maioria dos televisores de cinescópio (tubo) disponíveis no mercado utiliza esta tecnologia.

**ISDB:** (*Integrated Services Digital Broadcasting*) É o padrão japonês de TV Digital. Dos três sistemas existentes, foi o último a ser desenvolvido e integra em suas aplicações a TV de alta definição (HDTV), múltiplas programações, TV móvel e portátil e datacasting. Foi desenvolvido para a convergência com outros aparelhos, como celulares 3G e computadores de mão, que, para tanto, deverão ser equipados com o chip receptor. Utiliza a modulação COFDM, também usada no sistema DVB, com algumas melhorias que tornam os dois sistemas incompatíveis.

**ISDB-Tb:** Recentemente, o SBTVD (Sistema Brasileiro de Televisão Digital) foi chamado de ISDTB-tb, pois a TV Digital no Brasil incorpora diversas tecnologias já são utilizadas na TV Digital do Japão, o ISDB-T (*Integrated Services Digital Broadcasting Terrestrial* ou Serviço Integrado de Transmissão Digital Terrestre).

**JavaDTV:** O JavaDTV é uma especificação baseada na linguagem Java, concebida para ser flexível e adaptada para os receptores digitais no padrão ISDB-T. Como se trata de uma linguagem de ampla utilização internacional e com o respaldo de uma gigante do setor de

software, a Oracle, o Java é uma tecnologia “à prova de futuro”, com muitos anos de bons serviços à frente.

**Linguagem NCL:** A linguagem NCL (*Nested Context Language*) e sua plataforma de execução - Ginga-NCL - é considerada uma das maiores inovações nacionais dos últimos anos e foi desenvolvida pela Pontifícia Universidade Católica do Rio de Janeiro- PUC-RJ.

**Middleware:** Middleware é um software que oferece suporte para o desenvolvimento de conteúdo e aplicativos para TV Digital, entre elas a possibilidade desses conteúdos serem exibidos nos mais diferentes sistemas de recepção, independente do fabricante e tipo de receptor (TV, celular, PDAs etc.).

**Modulação:** É o processo em que certas características de uma onda eletromagnética (também chamada de portadora) variam de acordo com uma mensagem que se deseja transmitir. Ou seja, o termo técnico que descreve a maneira como a informação é “empacotada para viagem”. Quando se faz uma transmissão, o sinal pode sofrer uma série de interferências e degradações. A modulação é responsável pela “proteção” do sinal, de modo que a informação originalmente transmitida possa ser reconstruída da maneira mais fiel possível.

**Multicasting:** (ou Múltipla programação) É a possibilidade de transmitir mais de um programa ao mesmo tempo em um mesmo canal. Permite ao telespectador escolher entre vários programas ou ângulos de câmera. Com o multicasting, uma dona de casa poderá escolher entre ver um programa de receitas pela manhã ou deixar seus filhos assistirem a outra programação, que está sendo transmitida ao mesmo tempo e no mesmo canal. Em outra aplicação de multicasting, o espectador de uma partida de futebol poderá escolher o ângulo / câmera que deseja acompanhar, pois os sinais de várias câmeras serão transmitidos no mesmo canal.

**One-Seg:** (ou 1-SEG. Receptor de 1 segmento) A modulação do sistema japonês adotada no Brasil (BST-ODFM) divide por canal de TV 13 segmentos: pode-se utilizar 12 segmentos para enviar sinais de áudio e de vídeo em alta definição, dados ou multiprogramação, e utilizar o segmento central para enviar áudio e vídeo em baixa definição e dados a dispositivos móveis, como celulares e TVs portáteis. One-Seg é o receptor capaz de receber apenas o sinal destinado aos receptores portáteis.

**Padrão de transmissão de TV Digital:** Sistema digital que será usado na transmissão da TV Digital. Ao contrário da TV analógica, na qual o sistema escolhido tinha que ser usado desde a produção nos estúdios de TV até o aparelho na casa do telespectador. A maioria das emissoras já usa equipamentos digitais em seus estúdios, que independem do sistema a ser escolhido e dos displays digitais já existentes.

**Pilar Boxes:** Barras laterais, geralmente pretas, que aparecem quando uma imagem 4:3 é exibida em telas 16:9.

**Pixel:** É a aglutinação de picture element, sendo que picture, do inglês, é abreviado por pix. O pixel é o menor ponto em uma imagem. Possui três pontos de cores (vermelho, verde e azul). Assim consegue reproduzir 256 tonalidades de cores (equivalente a 8 bits). A combinação de

muitos pixels gera uma imagem. Quanto maior o número de pixels, mais definida é esta imagem. O pixel está diretamente ligado à definição. Quando falamos em “1024x768”, estamos nos referindo ao número vertical de pixels vezes o número horizontal em uma linha. Essa definição geram uma imagem de 786.432 pixels.

**Prazo de transição:** A transição para a TV Digital será gradual. Os dois sistemas coexistirão, ou seja, as emissoras terão uma transmissão analógica e outra digital durante vários anos, de forma que o telespectador tenha tempo de se adaptar ao novo sistema digital, adquirindo um novo receptor ou TV compatível.

**Progressive Scan:** (p) (imagem progressiva) Alguns televisores ou geradores de imagem possuem este sistema de processamento de sinal. Neste caso, ao contrário do sistema entrelaçado, cada quadro de imagem é formado seqüencialmente, gerando melhoria considerável na qualidade de imagem. Grande parte das TV's de LCD e Plasma disponíveis no mercado já incorporam esta tecnologia.

**Resolução:** Número de linhas horizontais que uma TV é capaz de reproduzir, sendo a imagem produzida tanto na forma progressiva quanto na forma entrelaçada. É medida em linhas de resolução horizontais, padronizada no sistema ISDB em 480, 720 e 1080 linhas.

**Resolução Nativa:** É a resolução que o televisor ou monitor de TV foi projetado para exibir. Caso a resolução do sinal recebido seja maior ou menor que a resolução nativa do televisor ou monitor de TV, este será ajustado automaticamente para a resolução do mesmo (Upconvert ou Downconvert)

**S-Video:** (*Separated Video*) Oferece melhor qualidade de imagem que o vídeo composto. São usados três fios que percorrem pelo interior de um único cabo: um para transmitir a imagem em preto e branco, outro para transmitir as informações de cor e um terceiro, que é o terra. Lembrando que esse tipo de conexão é inferior ao Vídeo Componente (Componet Video).

**SBTVD:** Sigla para “Sistema Brasileiro de TV Digital”.

**Set-Top Box:** Equipamento capaz de se conectar a uma fonte externa de sinal e transformá-lo em informações que geram imagem ao televisor. Com uma caixinha conversora, semelhante às usadas na TV a cabo ou via satélite o telespectador poderá usufruir de quase todas as vantagens da TV Digital em seu aparelho atual, como imagem perfeita, sem fantasmas e ruídos, e datacasting. Mas, para assistir à TV de alta definição, o telespectador terá que comprar um novo aparelho. Caso ele já tenha a caixinha conversora, poderá comprar somente o monitor de HDTV. Caso contrário, poderá comprar o receptor integrado.

**T-Commerce:** (ou Television-Commerce) Com a TV Digital, será possível comprar determinados produtos através da TV. No futuro, alguém interessado na trilha sonora de um filme ou de uma novela, por exemplo, poderá comprar o CD sem sair de casa ou da frente da televisão.

**Televisor Analógico:** Possui um sintonizador interno que permite receber as transmissões analógicas, mas não recebe transmissões digitais, necessitando, para isso, de um conversor digital (Set-Top Box).

**Televisor Digital:** Possui sintonizador um sintonizador interno que permite receber as transmissões digitais sem necessidade de um conversor digital. Também pode receber transmissões analógicas.

**Televisor HD Ready:** Possui sintonizador analógico, mas é capaz de reproduzir imagens com definição de 720 ou 1080 linhas horizontais. Com um conversor digital ISDB-T, poderá exibir imagens de alta definição transmitidas pelas emissoras de TV Digital no Brasil.

**Televisor HDTV:** (*High Definition TV* ou TV de alta definição) Capaz de reproduzir imagens com definição de 720 ou 1080 linhas horizontais. Os modelos cuja definição nativa é de 1080 linhas, se possuírem a função progressive scan, podendo exibir imagens com 1080 linhas de definição horizontal progressiva (1080p), são conhecidos como Full HD. Quando utilizados em fontes de sinal 1080i (ex: transmissões em HD) ou 1080p (ex: DVD de alta definição, HD-DVD ou Blue-Ray Disc), podem exibir a melhor definição disponível em alta definição.

**Televisor DTV Integrado:** (conversor digital integrado) Independente de sua tecnologia (CRT, Plasma, LCD ou Projeção), é aquele que possui o conversor digital integrado. Isso significa que pode receber sinais de TV Digital no padrão ISDB-T (padrão de TV Digital adotado no Brasil), diretamente da antena, sem necessidade de outro equipamento para converter o sinal (conversor).

**Televisor SDTV:** (*Standard Definition TV*) Possui definição nativa de 480 linhas horizontais. A maior parte dos televisores presentes no mercado pode reproduzir sinais com 480 linhas entrelaçadas (480i). Com a transmissão digital, a qualidade de imagem destes televisores será a mesma que eles apresentam quando conectados a um DVD. O conceito SDTV tem relação com a qualidade de imagem e não com o fato do produto ser digital ou analógico. Um produto SDTV pode ter um sintonizador digital.

**Transmissão terrestre:** Através de ondas de rádio-frequência, os sinais analógicos ou digitais são transmitidos pelo ar a partir das antenas terrestres (diferentemente dos satélites, que ficam no espaço) e necessitam de antenas e receptores apropriados para a sua recepção.

**TV a cabo:** Sistema que distribui conteúdo audiovisual via cabos. É transmitido por uma operadora, que recebe este conteúdo, nacional ou internacional, e o distribui às casas que pagam mensalmente pelo serviço. Normalmente tem um número significativo de canais disponibilizados.

**TV aberta:** Sistema que emite livremente conteúdo audiovisual, sem encargos e taxas para o telespectador. Para ter acesso a este conteúdo, basta que a TV esteja conectada à rede elétrica (tomada, gerador ou bateria) e situada dentro da área de cobertura de alguma emissora aberta.



**TV analógica – Transmissão (sinal aberto):** É o sinal de TV terrestre transmitido de forma analógica. É comum ocorrer perda de qualidade no processo de transmissão / recepção, ocasionando ruídos e interferências na imagem recebida.

**TV de tubo:** (CRT) CRT é um acrônimo para a expressão inglesa cathode ray tube, que em português significa “tubo de raios catódicos”. Também conhecidos como Cinescópio. Dentro de sua TV existe um tubo onde encontramos duas placas: uma positiva e outra negativa. Quando a tensão entre as placas é muito alta, gera elétrons, e quando esses atingem a placa positiva, a diferença de energia gera um feixe de luz que atravessa o tubo e para na parte de trás do vidro da televisão, formando a imagem.

**TV Digital – Transmissão (sinal aberto):** Sinal de TV terrestre transmitido de forma digital. O grande benefício deste sistema é que não há perda de qualidade no processo de transmissão. Imagem e áudio permanecem 100% com a qualidade do sinal original, eliminando ruídos e interferências características do sistema analógico.

**TV LCD:** Um feixe de luz passa por pequenas células que contém cristal líquido (daí o nome Liquid Crystal Display) controlado por uma corrente elétrica. Assim são geradas as três cores básicas para a formação de imagens vermelho, verde e azul.

**TV Móvel:** É a possibilidade de captar os sinais de TV em dispositivos em movimento: ônibus, trens, metrô, carros, barcos, etc.

**TV Plasma:** No painel de plasma, encontramos pequeninas células que contém uma mistura de gases. Quando uma corrente elétrica passa por essas células, excita o gás, que passa para o estado plasma, gerando luz.

**TV Portátil:** É a recepção em equipamentos portáteis, que podem ou não estar em movimento. Exemplo: televisores e computadores de mão equipados com receptor de TV ou telefones celulares equipados com chip receptor, nos quais o espectador pode assistir à programação deslocando-se ou não.

**TV Via Satélite:** Com o avanço da tecnologia foi possível receber o sinal diretamente via satélite nos domicílios. Um satélite recebe a transmissão de outros satélites ou de uma central terrestre, e retransmite para as casas que possuem uma antena específica apontada para ele. É um serviço pago.

**UHF:** (*Ultra High Frequency*) Frequência Ultra Alta, faixa de rádio frequência usada para propagação do sinal de TV digital, em alta definição (canal 14 ao 69).

**Upconvert:** Quando o televisor recebe o sinal cuja definição é inferior à sua definição nativa, este é automaticamente ajustado. Por exemplo: se o sinal tiver 480 linhas e a TV tiver definição nativa de 1080 linhas, este acrescentará linhas intermediárias e exibirá 1080 linhas.

**VHF:** (*Very High Frequency*) Frequência Muito Alta, usada na propagação do sinal de televisão analógica (canal 2 ao 13).

**Vídeo Componente:** (*Component Video*) Para transmissões de imagem é necessário que informações sobre cor e imagem cheguem ao aparelho em que vai ocorrer a reprodução. Há várias opções. Uma delas é o vídeo componente, onde são usados três conectores, chamados Y (conector verde), Pb (Cb ou ainda B-Y, conector azul) e Pr (Luminância, Cr ou ainda R-Y, conector vermelho). No conector Y são transmitidas as informações de imagem em preto e branco, enquanto nos outros dois conectores são transmitidas as informações de cor. Assim, é reproduzida uma imagem superior às que se conseguiria usando outras conexões, com S-Video e vídeo composto.

**Vídeo Composto:** (*Composite Video* ou RCA) Nesse padrão, o sinal de vídeo é transmitido apenas por um fio, misturando informações de imagem e cor no mesmo sinal. Por este motivo, tem a pior qualidade de imagem para transmissões de vídeo usando o cabo. É um dos tipos mais populares de conexão de vídeo e utiliza conector RCA (Radio Corporation of América, empresa que introduziu esse tipo de conector no mercado, em meados dos anos 40). A informação percorre por um fio interno, depois vem a blindagem e, em seguida, o isolamento de borracha.

**Wide Screen:** (Tela Larga ou 16:9) Aspecto de tela da TV digital.

# Sumário

<b>1. Introdução .....</b>	<b>1</b>
<b>2. Metadados na TV digital.....</b>	<b>5</b>
2.1 Metadados para TV digital: rígidos versus flexíveis .....	7
2.2 Alternativas para descrição unificada: ontologias .....	14
2.3 Ontologias na Web Semântica.....	16
2.4 Estado da arte no uso de ontologias no domínio da TV digital .....	17
2.5 Problema da heterogeneidade dos metadados .....	21
2.6 Contexto do trabalho no estado da arte .....	25
<b>3. Visão conceitual .....</b>	<b>27</b>
3.1 Visão geral dos componentes do processo proposto .....	27
3.2 Construção da ontologia de alinhamento de metadados.....	31
3.3 Extensão da ontologia OWLTV e integração com outros domínios .....	36
3.4 Considerações finais do capítulo .....	40
<b>4. Aplicação do processo de alinhamento e extensão de metadados da TV digital.....</b>	<b>43</b>
4.1 Especificações de metadados dos sistemas de TV digital .....	44
4.2 Bases de conhecimento da Web Semântica.....	45
4.3 Construção da ontologia para alinhamento de metadados .....	46
4.3.1 Etapa I – Obtenção de um arquivo base em formato XMLTV .....	47
4.3.2 Etapa II – Conversão dos metadados da tabela EIT para XMLTV .....	49
4.3.3 Etapa III – Aplicação da metodologia proposta por Thuy et al. (2009) para geração da ontologia OWLTV a partir do arquivo XMLTV .....	50
4.3.4 Etapa IV – Inclusão de classes “sameAsQuery” .....	56
4.3.5 Etapa V – Obtenção das propriedades owl:sameAs .....	58
4.4 Considerações finais do capítulo .....	60
<b>5. Validação do processo de alinhamento e extensão de metadados da TV digital .....</b>	<b>63</b>
5.1 Ontologia para alinhamento de metadados.....	63
5.2 Cenários de utilização do receptor .....	64
5.2.1 Cenário do programa “Chocolate com Pimenta” .....	64
5.2.2 Cenário do programa “Táxi” .....	65
5.3 Sistema desenvolvido em Ginga-J.....	66
5.4 Aplicação para o usuário final .....	68
5.4.1 Exemplo de pesquisa da base de conhecimento Freebase.....	70
5.4.2 Exemplo de pesquisa da base de conhecimento DBpedia.....	71
5.4.3 Exemplo de pesquisa Freebase no cenário do programa “Táxi” .....	73
5.5 Aumentando a abrangência da aplicação da ontologia.....	76
5.6 Considerações finais do capítulo .....	77
<b>6. Conclusões .....</b>	<b>79</b>
6.1 Principais contribuições.....	80
6.2 Trabalhos futuros.....	81
<b>Referências Bibliográficas .....</b>	<b>83</b>
<b>Apêndice A .....</b>	<b>89</b>



# Capítulo 1

## Introdução

A evolução da tecnologia de *broadcast* multimídia nos anos recentes tem provocado mudanças significativas no acesso das pessoas à mídia de massa. De acordo com um relatório publicado pelo instituto independente Informa Telecoms & Media (INFORMA TELECOMS & MEDIA, 2012), os moradores de vinte e cinco milhões de residências na América Latina assistem a TV digital diariamente. Esta audiência cresce a cada dia à medida que governos e operadores de TV latino americanos percebem os benefícios da transmissão digital e o número de lares equipados com a tecnologia deve atingir noventa milhões até 2016 (INFORMA TELECOMS & MEDIA, 2012).

Há alguns anos as aplicações de vídeo eram restritas, uma vez que formatos analógicos eram utilizados por televisores, vídeo cassetes e filmadoras. Contudo, após o advento do padrão MPEG-2 (HASKELL *et al.*, 1998) para compressão de vídeo digital, as aplicações de vídeo têm cada vez mais convergido para os formatos digitais. A integração da Internet e das tecnologias de comunicação móveis com as plataformas de televisão têm provido aos telespectadores novos serviços interativos de conteúdo digital. Devido a estes fatores, os equipamentos para o consumidor têm se tornado cada vez mais sofisticados, suportando uma variedade de conteúdos e conectividade com outras redes e dispositivos.

A TV digital é uma plataforma híbrida que combina elementos da televisão tradicional com a Internet, provendo ao usuário o acesso a uma diversidade de conteúdos de mídia interativa. Com o crescimento do volume e da diversidade de serviços e conteúdos multimídia, a televisão está enfrentando os mesmos desafios de complexidade e excesso de informações que já vinham sendo encarados por outras mídias digitais relacionadas com a Internet.

A tecnologia de metadados pode ser uma alternativa para lidar com esta complexidade de serviços e conteúdos digitais de forma prática e eficiente. Metadados são dados que complementam as informações digitais dos conteúdos multimídia com o objetivo de melhor descrevê-los, facilitando a estruturação e o gerenciamento de grandes volumes de informação. O uso de metadados aplicados à tecnologia de *broadcast* multimídia não se restringe a construção de um ferramental de busca e indexação de conteúdos, e abre oportunidade para o desenvolvimento de uma gama de serviços inovadores.

Atualmente existem diversas especificações de metadados utilizadas pela indústria de *broadcast* multimídia em redes de TV digital. Além disso, existem na Internet diversos repositórios de informação baseados em metadados que complementam as informações de metadados da TV digital. Contudo, como os padrões de metadados da TV digital e da Internet são

baseados em diferentes especificações não relacionadas, surge o problema de como integrar estas informações, visando criar novos serviços para telespectador que utilizem tanto informações de metadados da TV digital como informações de metadados da Internet.

Para satisfazer uma variedade de necessidades na modelagem de informações de metadados, desenvolvimento e integração de software, bem como o gerenciamento e reuso de conhecimento, diversos grupos na indústria, academia e governo vêm desenvolvendo e implantando padrões compartilhados e reutilizáveis conhecidos como ontologias (REDONDO-GARCIA *et al.*, 2010; KIM *et al.*, 2011; YOOWON *et al.*, 2011). As ontologias permitem criar modelos conceituais abstratos mais simplificados e menos formais que as especificações dos padrões de metadados usadas atualmente na indústria de *broadcast* multimídia.

Este trabalho propõe a representação das informações de metadados existentes em um ambiente de transmissão e recepção de TV digital aberta em redes de *broadcast* terrestre na forma de ontologias visando uma maior integração e organização destas informações. Nesta proposta será descrito um processo para alinhar as informações de metadados dos padrões da indústria de *broadcast* multimídia com ontologias voltadas para a descrição de conhecimento de domínios específicos existentes em repositórios da Internet.

Como a TV digital também terá acesso a Internet, a proposta também contempla o acesso a repositórios *on-line* para obter informações de ontologias complementares que estendam os metadados da TV digital, para que o usuário obtenha um maior nível de conhecimento e maiores possibilidades de interação sobre os conteúdos que são transmitidos pelas emissoras.

As ontologias propostas neste trabalho utilizam o padrão de estruturação de informações da Web Semântica (BERNERS-LEE *et al.*, 2001) conhecido como *Linked Data* criado por Tim Berners-Lee (BERNERS-LEE, 2006). Esta forma de representação propõe alguns princípios para estruturação e interligação de dados, através de tecnologias padrões da Web, como HTTP e URI de forma que a informação seja compartilhada de forma a ser lida automaticamente por máquinas e computadores. Isto permite que informações de diferentes fontes sejam facilmente interconectadas e pesquisadas. A criação de ontologias estruturadas no formato *Linked Data* facilita a interligação e complementação das informações de metadados da TV digital com repositórios de informações de metadados já existentes na Web Semântica.

A proposta do trabalho envolve a criação de ontologias para a representação das informações de metadados já existentes no contexto da TV digital, estabelecendo um alinhamento entre os padrões de metadados *broadcast* multimídia para possibilitar a comparação semântica das informações destes formatos por meio das ontologias e a extensão destes formatos com a utilização de ontologias para domínios específicos existentes na Internet.

O processo proposto permitirá que o usuário da TV digital possa facilmente pesquisar conteúdos de interesse a partir da grade de programação dos canais existentes e dos conteúdos já gravados em seu receptor; receber sugestões de conteúdos para exibição ou gravação conforme o

seu perfil e interesse; enriquecer sua experiência de assistir televisão acessando informações complementares sobre os programas transmitidos como sinopses, críticas especializadas, histórico do elenco e direção, premiações recebidas, fotos, vídeos e conteúdos relacionados disponíveis para livre acesso via Internet; entre outras funções.

A presente proposta possui uma viabilidade real de implementação sem a necessidade de impactar os padrões já existentes e que são utilizados em redes digitais de *broadcast* multimídia (TV digital aberta terrestre), como as que estão em operação atualmente no Brasil e em outros países.

O trabalho está organizado da seguinte forma:

No Capítulo 2 é realizada uma apresentação sobre o emprego atual de metadados na área de *broadcast* multimídia, alternativas para a descrição unificada usando ontologias, o estado da arte na utilização de ontologias para integração de informação em formatos distintos no domínio da TV digital e o contexto do presente trabalho no estado da arte.

No Capítulo 3 é apresentada uma visão conceitual do trabalho com uma apresentação detalhada do processo que foi criado para alinhar as informações de metadados dos padrões da indústria de *broadcast* multimídia com a possibilidade de extensão destes padrões com descrições de conhecimento de domínios específicos existentes em bases de dados da Internet.

O Capítulo 4 descreve a aplicação do processo apresentado no Capítulo 3 com o objetivo de demonstrar como é realizado o alinhamento dos metadados do padrão da TV digital terrestre ISDB-T com metadados de domínios de conhecimento específicos da Web Semântica, disponíveis nos repositórios de conteúdo das bases de conhecimento LinkedMDB, Freebase e DBpedia.

No Capítulo 5 é apresentada a validação da aplicação do Capítulo 4 com uma prova de conceito implementada em um sistema de TV digital em operação.

O Capítulo 6 contém as conclusões da pesquisa e sugestões de trabalhos futuros.





## Capítulo 2

### Metadados na TV digital

A TV digital é uma tecnologia de transmissão broadcast de conteúdos multimídia que utiliza técnicas de codificação digital para transportar informações de áudio e vídeo, bem como sinais de dados, para receptores como aparelhos de televisão e set-top boxes. Diversos países na Europa, na América do Norte e na Ásia há alguns anos já vem passando por um processo de digitalização de suas redes de transmissão de televisão. O Brasil também já iniciou este processo com o lançamento do Sistema Brasileiro de TV Digital (SBTVD) que foi realizado em dezembro de 2007 na cidade de São Paulo e atualmente já está em operação em centenas de cidades brasileiras. O SBTVD é um sistema que tem como base o padrão japonês de TV digital terrestre, denominado ISDB-T (*Integrated Services Digital Broadcasting – Terrestrial*).

Além da transmissão via redes de transmissão terrestre, também conhecida como radiodifusão, diversos canais de televisão brasileiros utilizam a transmissão via satélite para levar o sinal da TV digital aberta às residências, devido principalmente ao tamanho e dificuldade de cobertura do território brasileiro. Neste caso, o padrão utilizado pelas redes tem como base o padrão europeu, denominado DVB-S (*Digital Video Broadcasting – Satellite*).

A TV digital é uma tecnologia de transmissão de TV que usa um sistema de modulação e compressão digital para enviar vídeo, áudio e sinais de dados aos aparelhos compatíveis com a tecnologia, proporcionando assim transmissão e recepção de maior quantidade de conteúdo por uma mesma faixa de frequência (canal).

Um sistema completo de televisão em geral é composto por três componentes:

1. O "estúdio" que envolve as atividades de produção (gravação das cenas), pós-produção (edição e acabamento) e transmissão de sinais entre diferentes setores da emissora;
2. A transmissão dos programas da emissora para o usuário final (telespectador), mais conhecida como "processo de radiodifusão" (broadcast);
3. O "sistema de recepção" (antena e acessórios) e o receptor (televisor) que ficam na casa do usuário.

O primeiro componente (estúdio) já vem sendo gradualmente digitalizado, com a utilização de equipamentos de filmagem e armazenamento digitais em padrões proprietários, ilhas de edição, sistemas computacionais e transmissão entre setores da emissora com sinais totalmente digitais.

Do lado do usuário, o aparelho de televisão já possui diversos componentes digitais como o controle remoto e o sistema de sintonia com a utilização de filtros digitais.

A introdução da tecnologia digital no serviço de televisão – TV digital – se refere, portanto, a transmissão dos sinais de televisão na forma digital da emissora para a casa do usuário. Entre os benefícios proporcionados por esta nova tecnologia, pode-se citar:

1. Melhor qualidade da imagem e do som – fim dos fantasmas e chuviscos provocados pelas interferências da transmissão analógica;
2. A imagem poderá ser mais larga que a atual (*widescreen* – 16:9) contendo um maior grau de resolução (alta definição – HDTV) e um som estéreo mais envolvente (*surround* – 5.1 canais);
3. Possibilidade de se captar vários programas utilizando a largura de banda de um único programa transmitido de forma analógica;
4. Possibilidade de transmissão de sinal de TV para recepção em receptores móveis como telefones celulares e aparelhos de TV portáteis;
5. Seleção de programas através de um guia contendo toda a grade de programação;
6. Transmissão de programas com vários idiomas de áudio e legendas;
7. Personalização da TV com recursos de VOD (*video on-demand*), PPV (*pay-per-view*) e PVRs (*personal video recorders*);
8. Possibilidade de oferecimento de uma ampla gama de novos serviços e aplicações interativas com acesso a Internet através de um canal de interatividade.

A tecnologia de TV digital, por envolver a manipulação de diversos tipos de conteúdos multimídia, como elementos de áudio, vídeo e dados, utiliza metadados com o objetivo de descrever os serviços que são oferecidos pelos canais de televisão. Os metadados utilizados na TV digital, além de descrever as características dos serviços – por exemplo: resolução, formato de codificação e idioma do conteúdo de áudio e vídeo – também transmitem as informações dos eventos que compõem o guia eletrônico de programação – nome do evento, horário de início, duração, sinopse, classificação indicativa, gênero, entre outras informações.

A introdução de metadados no contexto da TV digital pode proporcionar inúmeras vantagens como interoperabilidade entre plataformas dos parceiros da cadeia de valor de *broadcasting*, interfaces padronizadas abertas para compartilhamento semi-automatizado de conteúdos multimídia e intercâmbio de dados harmonizado e uniformizado através da cadeia de valor.

Os metadados fazem a representação do conteúdo de forma escalável no tempo e no espaço, servindo com um elo virtual para a criação, entrega e intercâmbio de ativos multimídia através de um espaço virtual. A utilização de modelos unificados de metadados simplifica a identificação do contexto em uma representação de conteúdo e pode facilitar a apresentação da interatividade. Metadados e conteúdos podem ser facilmente adaptados, transcodificados, transformados e encapsulados.

## 2.1 Metadados para TV digital: rígidos versus flexíveis

Segundo Lugmayr *et al.* (2004), a TV digital possibilita a utilização de dois tipos de metadados: metadados rígidos – que são destinados a um propósito específico e garantem alto desempenho; e metadados flexíveis ou granulares – que são mais complexos e permitem alta granularidade e elevado nível de detalhamento. Nos padrões atuais de TV digital a transmissão de informações está puramente baseada nas definições de metadados das tabelas de informações de serviços (*SI – Service Information*). As tabelas SI estendem as tabelas PSI (*Program Specific Information*) do padrão MPEG-2 definindo um conjunto de estruturas que possuem dados descritivos que carregam informações de serviços específicas do domínio da TV digital. O uso de tais tabelas facilita a criação, o tratamento e a rápida extração das informações.

No entanto, as tabelas SI são consideradas metadados rígidos. Conforme Lugmayr *et al.* (2004), metadados rígidos são aqueles destinados a um propósito específico e que não passam por uma padronização de metadados passível de ser personalizável. Muitos serviços necessitam de informações mais detalhadas e que não conseguem ser definidos de maneira satisfatória dentro das tabelas SI. Assim, surgiu a necessidade de se padronizar metadados flexíveis, ou seja, metadados que apresentem estruturas personalizáveis para atender os novos serviços em cenários mais complexos.

Os padrões de TV digital usados atualmente no mundo se focam puramente na definição de metadados rígidos. Eles são mais fáceis de manipular, mais rápidos para processar, mais fáceis de criar, podem ser transformados em streaming e são menos complicados que os metadados flexíveis. Contudo, serviços mais complexos do que simples guias de programação baseados em metadados de PSI/SI estão surgindo. Estes serviços mais avançados como, por exemplo, sistemas de recomendação de conteúdo baseados no perfil do usuário, seriam muito difíceis de viabilizar com metadados rígidos. Devido a isso existe uma necessidade direta por estruturas de metadados mais flexíveis do que esses para cenários de uso mais complexos.

Um formato de metadados mais flexível que vem ganhando espaço atualmente na indústria de TV é o padrão XMLTV (XMLTV, 2012). Conforme pode ser consultado no site do Projeto XMLTV (XMLTV PROJECT, 2012) existem diversos utilitários disponíveis para manipular metadados em XMLTV. Além disso, diversos sites da Internet mantêm listagens de programação de TV baseadas em XMLTV, contendo informações atualizadas dos principais canais de TV do país e dos programas transmitidos pelos canais, conforme o fragmento de arquivo XMLTV

exibido na Figura 2.1, que foi extraído do site Revista Eletrônica (REVISTA ELETRÔNICA, 2012).

```

<?xml version="1.0" encoding="ISO-8859-1"?>
<tv generator-info-name="Revista Eletronica - Unidade Lorenz Ltda" generator-info-
url="http://www.revistaeletronica.com.br">
<channel id="GRC">
<display-name lang="pt">EPTV Campinas</display-name>
<icon src="grc.gif" />
</channel>
<programme start="20120607203500 -0300" stop="20120607211500 -0300" channel="GRC" program_id="0000008302">
<title lang="pt">Jornal Nacional</title>
<title lang="en">Jornal Nacional</title>
<desc>As principais notícias do dia no Brasil e no mundo, com apresentação de William Bonner e Patrícia Poeta.
</desc>
<category lang="pt">Jornalismo</category>
<category lang="pt">Informativo</category>
<country>Brasil</country>
<video>
<colour>yes</colour>
</video>
<rating system="Advisory">
<value>Programa livre para todas as idades</value>
</rating>
</programme>
<programme start="20120607211500 -0300" stop="20120607222500 -0300" channel="GRC" program_id="0000274345">
<title lang="pt">Avenida Brasil</title>
<title lang="en">'</title>
<desc>Jorginho revela que Carminha é sua mãe biológica, mas Tufão não acredita. Jimmy paga a fiança de Cadinho.
Noêmia corta as roupas de Cadinho e o expulsa de casa. Monalisa afirma que não adiará o casamento com Silas.
</desc>
<category lang="pt">Variedades</category>
<category lang="pt">Novela</category>
<video>
<colour>yes</colour>
</video>
<rating system="Advisory">
<value>Programa impróprio para menores de 14 anos</value>
</rating>
</programme>
</tv>

```

**Figura 2.1 – Fragmento de arquivo XMLTV publicado pelo site Revista Eletrônica**

O formato XMLTV possui informações detalhadas sobre a programação da TV digital. Estas informações estão presentes nos padrões de metadados rígidos utilizados pelas emissoras de

TV digital em operação no Brasil, como o padrão ISDB-T utilizado na transmissão de TV digital terrestre e o padrão DVB-S utilizado na TV digital via satélite. Desta forma é simples criar um mapeamento entre as informações de metadados dos padrões de TV digital e o padrão XMLTV, de forma a possibilitar um maior nível de detalhamento e complexidade das informações que as tabelas SI.

Apesar de padrões abertos não ser algo novo para a indústria de TV, tanto emissoras de televisão aberta como operadoras de televisão por assinatura ainda estão em busca de padrões abertos consolidados para a geração e manipulação de metadados flexíveis para conteúdos multimídia. Conforme Alves *et al.* (2006), atualmente existem três padrões abertos de metadados flexíveis que podem ser aplicados para o contexto da TV digital: MPEG-7, MPEG-21 e TV-Anytime. Estas especificações abordam requisitos específicos, são complementares em diversos níveis, e ainda são pouco utilizadas por parte da indústria de *broadcast*.

Em 1988, o consórcio MPEG (*Moving Picture Experts Group*) foi fundado e iniciado para definir um conjunto de padrões cobrindo diversos problemas relacionados com multimídia. Os primeiros esforços resultaram em padrões de compressão de áudio e vídeo digital, como os padrões MPEG-1, MPEG-2 e MPEG-4. MPEG seguiu dois outros trilhos com a definição do MPEG-7 como um padrão de descrição para conteúdos de audiovisuais e o MPEG-21 como um padrão de empacotamento e proteção para conteúdos multimídia.

O MPEG-7 foi introduzido inicialmente em 1997. Ele define um conjunto de descrições para conteúdos audiovisuais na forma de esquemas de descrição multimídia (MDS – *Multimedia Description Schemes*) para caracterização de um conteúdo audiovisual. O nível de contexto de aplicação para o emprego do MPEG-7 em TV digital engloba: educação, jornalismo, informações de turismo, serviços culturais, entretenimento, arquivamento de áudio e vídeo, comércio e produção de conteúdo audiovisual. A Figura 2.2 exibe um fragmento de arquivo no padrão MPEG-7.

```
:image01 rdf:type mpeg7:Image
:goal01 rdf:type s:Goal
:scoring01 rdf:type s:Scoring
:image01 mpeg7:depicts :goal01
:goal01 abc:hasAction :scoring01
:scoring01 abc:hasAgent s:_b1
:_b1 :hasName 'Zinedine Zidane'

mpeg7:Image rdfs:subclass mpeg7:MultimediaContent
mpeg7:MultimediaContent rdfs:subclass abc:Manifestation
s:Scoring rdfs:subclass abc:Action
s:Goal rdfs:subclass abc:Event
```

Figura 2.2 – Fragmento de arquivo no padrão MPEG-7

O padrão MPEG-7 possui oito partes diferentes, onde as primeiras cinco partes são normativas e as partes 6 a 8 são para suporte:

- Parte 1 (Sistemas): padroniza a transmissão binária, sincronização e modos de armazenamento (ISO/MPEG N4285, 2001);
- Parte 2 (DDL – *Description Definition Language*): definição da linguagem de metadados (ISO/MPEG N4288, 2001);
- Parte 3 (Visual): definição dos esquemas de descrição visual (ISO/MPEG N4358, 2001);
- Parte 4 (Áudio): definição dos esquemas de descrição de áudio (ISO/MPEG N4224, 2001);
- Parte 5 (MDS): descrição de descritores de conteúdos multimídia diferentes do áudio e do vídeo (ISO/MPEG N4242, 2001);
- Parte 6 (Software de Referência): define o software genérico para suporte às diferentes partes padronizadas (ISO/MPEG N4206, 2001);
- Parte 7 (Testes de Conformidade): enfoque nos processos para testes de implementações de hardware e software em conformidade com MPEG-7;
- Parte 8 (Extração e Uso de Descrições MPEG-7): define a interface de descrição de conteúdos multimídia e os procedimentos para uso das ferramentas MPEG-7 e a implementação do software de referência.

Outro padrão multimídia publicado pelo consórcio MPEG é o MPEG-21. O propósito do MPEG-21 é o uso mais transparente, protegido e interoperável de conteúdos multimídia através da cadeia de distribuição. O padrão MPEG-21 define uma “Linguagem de Expressão de Direitos” como meio de compartilhamento de direitos, permissões ou restrições de conteúdos digitais para distribuição dos mesmos dos criadores para os consumidores.

O MPEG-21 é uma especificação de metadados baseada em XML que estabelece dois conceitos principais (HASKELL *et al.*, 1998): a definição de uma unidade fundamental de distribuição e transação, que é o Item Digital, e a concepção dos usuários que interagem com o mesmo. A Figura 2.3 exibe um fragmento de arquivo no padrão MPEG-21.

```
<didl:DIDL xmlns:didl="urn:mpeg:mpeg21:2002:01-DIDL-NS">
  <didl:Container>
    ...
    <!-- Item containing content -->
    <didl:Item>
      ...
      <!-- ObjectType of Item -->
    </didl:Item>
  </didl:Container>
</didl:DIDL>
```

```

        <didl:Descriptor>
            <didl:Statement mimeType="text/xml; charset=UTF-8">
                <dip:ObjectType xmlns:dip="urn:mpeg:mpeg21:2002:01-DIP-NS">
                    urn:my:Argument</dip:ObjectType>
                </didl:Statement>
            </didl:Descriptor>
            ...
        </didl:Item>
        ...
        <!-- Processing Item -->
        <didl:Item>
            <!-- Qualification of the Item as Processing Item -->
            <didl:Descriptor>
                <didl:Statement mimeType="text/xml; charset=UTF-8">
                    <dii:Type xmlns:dii="urn:mpeg:mpeg21:2002:01-DII-NS">
                        urn:mpeg:mpeg21:2002:01-DIP-NS:PI</dii:Type>
                    </didl:Statement>
                </didl:Descriptor>
                <didl:Descriptor>
                    <didl:Statement mimeType="text/xml; charset=UTF-8">
                        <r:license xmlns:r="urn:mpeg:mpeg21:2003:01-REL-R-NS">
                            <!-- optionally, specific rights can be added here.-->
                            <r:otherInfo>
                                <dc:rights xmlns:dc="http://purl.org/dc/elements/1.1/">
                                    Copyright 2003; American Physical Society</dc:rights>
                                </r:otherInfo>
                            </r:license>
                        </didl:Statement>
                    </didl:Descriptor>
                </didl:Item>
            </didl:Container>
        </didl:DIDL>

```

**Figura 2.3 – Fragmento de arquivo no padrão MPEG-21**

Os Itens Digitais podem ser considerados o núcleo do *Framework* Multimídia definido por MPEG-21 e os usuários são os entes que interagem dentro deste *Framework* (ISO/IEC 21000-1, 2002). No seu nível mais básico, MPEG-21 provê um *framework* no qual um usuário interage com outro e o objeto desta interação é o Item Digital. Devido a isto, pode-se afirmar que o objetivo principal do MPEG-21 é definir as tecnologias necessárias para suportar o intercâmbio, acesso, consumo, comércio ou manipulação de Itens Digitais de forma eficiente e transparente.

O padrão MPEG-21 é composto por 16 partes, das quais as mais relevantes são:

- Declaração de Item Digital (DID – *Digital Item Declaration*) – especifica um conjunto de termos e conceitos para a elaboração de um modelo de representação de Itens Digitais (ISO/IEC 21000-2, 2003);

- Informação sobre DII (*Digital Item Identification and Description*) – define um modelo para permitir a identificação única do Item Digital independente da sua natureza, tipo ou granularidade (ISO/IEC 21000-3, 2003);
- Gerenciamento e proteção de propriedade intelectual (IPMP – *Intellectual Property Management and Protection*) – fornece os mecanismos necessários para a proteção dos conteúdos digitais (ISO/IEC 21000-4, 2003);
- Linguagem de expressão de direitos (REL – *Rights Expression Language*) – define uma linguagem para declarar direitos e permissões do conteúdo (ISO/IEC 21000-5, 2003) (ISO/IEC 21000-6, 2003);
- Adaptação de Item Digital (DIA – *Digital Item Adaptation*) – define mecanismos de descrição de informações sobre requisitos de um ambiente e as informações para a adaptação dos Itens Digitais para permitir acessá-los de forma transparente (ISO/IEC 21000-7, 2003).

Além dos padrões MPEG-7 e MPEG-21, outro padrão de metadados flexíveis que pode ser aplicado no contexto da TV digital é o TV-Anytime (TVA). O TV-Anytime é um conjunto completo e integrado de especificações de metadados estabelecido pelo Fórum TV-Anytime (MCPARLAND, 2002). Os recursos de TVA permitem a busca, seleção, aquisição e o uso legítimo de conteúdos em sistemas de armazenamento local e/ou remoto tanto para serviços *broadcast* como para serviços on-line.

Um sistema broadcast de TVA é composto por três elementos principais: um provedor de serviços provendo os metadados TVA, um provedor de transporte que transporta os metadados e um equipamento nas casas que processa os metadados, armazena conteúdos e os executa novamente com base na solicitação do consumidor. De acordo com Evain *et al.* (2003), os principais objetivos do TVA são:

- Assegurar que os usuários tenham acesso a conteúdo personalizado (isto é, de acordo com o seu interesse específico) proveniente de diversos provedores de conteúdo;
- Agregar valor ao conteúdo, permitindo que os usuários acessem e utilizem este conteúdo quando e onde eles desejarem, sem que haja restrição;
- Independência de tecnologia no que se refere aos meios de transmissão (camada de transporte) e recepção (recepção) de conteúdo ao consumidor;
- Desenvolver especificações para permitir a interoperabilidade e integração entre sistemas;
- Especificar as estruturas de segurança necessárias para proteger os interesses de todas as partes envolvidas.



Para melhor separação no tratamento desses objetivos, o TVA é dividido em quatro linhas de trabalho: modelo de negócios; sistema, interface de transporte e referência ao conteúdo; metadados; e gerenciamento de direitos autorais.

O padrão TVA adota XML como formato para representação dos metadados e a definição formal da estrutura e da sintaxe dos metadados é realizada por XML *Schema*. O conjunto de metadados do TVA também reutiliza alguns esquemas especificados no padrão MPEG-7 (linguagem DDL – *Description Definition Language*), principalmente no que se refere à descrição de mídias e preferências de usuário. A Figura 2.4 exibe um fragmento de arquivo no padrão TVA.

```
<TVAMain xmlns='urn:tva:metadata:2005' xmlns:mpeg7='urn:mpeg:mpeg7:schema:2005'
  xmlns:xsi='http://www.w3.org/2001/XMLSchema-instance' xml:lang='en'>
  <ProgramDescription>
    <ServiceInformationTable>
      <ServiceInformation serviceId='BBCOne'>
        <Name>BBC One</Name>
        <Owner>BBC</Owner>
        <ServiceURL>dvb://233a.1004.1044</ServiceURL>
        <ServiceGenre href='urn:tva:metadata:cs:MediaType:2005:7.1.3'>
          <Name><![CDATA[Audio and video]]></Name>
        </ServiceGenre>
      </ServiceInformation>
    </ServiceInformationTable>
  </ProgramDescription>
</TVAMain>
```

**Figura 2.4 – Fragmento de arquivo no padrão TVA**

No TVA há quatro categorias de metadados: descrição de conteúdo, descrição de instância, usuário e segmentação. A primeira trata de descrição de programas por completo (gêneros de programas, informações de áudio e vídeo) e descritores específicos (solução para os serviços de busca, localização, seleção e personalização de serviços de TV digital). A segunda define informações para dar suporte a mecanismos de localização e anúncio de serviços (no guia de programação, por exemplo). A terceira categoria determina estruturas para identificação, agrupamento, perfil e histórico de uso de usuários. Estes metadados possibilitam a criação de uma grande quantidade de cenários de aplicações como, por exemplo, a personalização e o envio de histórico para provedores. Por fim, a última categoria tem a finalidade de descrever fluxos de áudio e vídeo, para permitir o acesso e manipulação, em intervalos temporais e de forma aleatória, de segmentos dos mesmos.

Apesar do TVA não definir tecnologias ou mecanismos de entrega e recepção, atualmente o padrão europeu DVB-S, por meio da elaboração de extensões ou novas especificações, realiza esforços para incorporar o suporte a TVA no seu modelo de referência (LEE *et al.*, 2002).

Fazendo uma análise comparativa entre as três especificações de metadados flexíveis para conteúdos multimídia apresentadas (MPEG-7, MPEG-21 e TVA), pode-se verificar que cada padrão apresenta um enfoque em requisitos específicos, de forma que os mesmos são complementares em diversos níveis.

O MPEG-7 tem como enfoque a descrição de conteúdo multimídia, sendo a especificação mais completa para tal finalidade, tendo sido inclusive incorporada pela especificação TVA na categoria de metadados que trata da descrição de conteúdos.

Os diferenciais do MPEG-21 são os aspectos relacionados ao gerenciamento de direitos digitais (DRM – *Digital Rights Management*) – que é um domínio coberto por três partes da especificação – e a ênfase que é dada ao controle da qualidade dos serviços para os usuários finais (QoS – *Quality of Service*) – com a existência de metadados voltados para a descrição da capacidade do receptor e dos recursos da rede de entrega. Alguns destes diferenciais de MPEG-21 foram também incorporados pela especificação TVA.

A especificação TVA, por sua vez, além de ser a única que foi concebida originalmente para operação no ambiente de *broadcast*, tem um maior enfoque na descrição de serviços interativos para TV, como a entrega de conteúdo personalizado conforme o perfil do usuário e o acesso a serviços on-line.

Devido a esta complementaridade existente entre as especificações de metadados flexíveis para conteúdos multimídia, muitas vezes existem situações onde é necessário utilizar informações de metadados de diferentes padrões para satisfazer os requisitos específicos de um projeto ou aplicação. Por exemplo, um projeto que deseje fazer a entrega de conteúdo personalizado conforme o perfil do usuário com um maior controle da qualidade dos serviços poderia fazer uma integração entre metadados TVA com metadados MPEG-21 para melhor atender a estes requisitos.

Desta forma, nessas situações faz-se necessária a integração de informações de metadados de diferentes formatos em um modelo unificado. A utilização de ontologias é uma alternativa para atender essas necessidades de unificação de descrições de diferentes domínios de conhecimento.

## **2.2 Alternativas para descrição unificada: ontologias**

Ontologia é um modelo de dados que representa um conjunto de conceitos dentro de um domínio e os relacionamentos entre estes. Segundo Gruber *et al.* (1993), uma ontologia é definida como uma especificação explícita e formal de uma conceituação compartilhada.

Conforme Uschold *et al.* (2004), existem diversas discussões sobre o que exatamente pode ser classificado como uma ontologia. Contudo, existe um senso comum que abrange todas as abordagens:

- Uma ontologia é um vocabulário de termos que se referem a elementos de interesse em um determinado domínio;
- Uma ontologia é uma especificação de significado para os termos, fundamentada em alguma forma de lógica.

Ontologias podem representar diversos tipos de elementos em uma área de conhecimento. Estes elementos são visualizados na ontologia como classes (também chamadas de conceitos) e são tipicamente organizados em uma taxonomia de classes e subclasses. Cada classe pode possuir propriedades (também chamadas de papéis) descrevendo suas características e atributos, bem como várias restrições sobre as propriedades (também chamadas de facetas). Uma ontologia em conjunto com as instâncias concretas de uma classe (também chamadas de indivíduos) constitui uma base de conhecimento.

Tarefas que envolvem a utilização de sistemas distribuídos e heterogêneos demandam o suporte de mais de uma ontologia. Diversas ontologias necessitam de ser acessadas a partir de diferentes sistemas. Conforme Choi *et al.* (2006), a natureza distribuída do desenvolvimento de ontologias levou ao surgimento de ontologias desiguais para domínios iguais ou similares. Devido a isso, diversas partes com diferentes ontologias não interagem entre si. Para resolver estes problemas, é necessário o uso de técnicas de mapeamento de ontologias voltadas para a interoperabilidade.

O mapeamento de ontologias, segundo Choi *et al.* (2006), pode ser classificado em três categorias:

1. Mapeamento entre uma ontologia integrada global e ontologias locais;
2. Mapeamento entre ontologias locais;
3. Mapeamento com fusão e alinhamento de ontologias.

A fusão de ontologias, segundo Pinto *et al.* (1999), é o processo de geração de uma ontologia única e coerente a partir de duas ou mais ontologias diferentes existentes que são relacionadas com o mesmo assunto. Uma ontologia única e coerente resultante da fusão inclui informações de todas as ontologias originais, mas praticamente sem realizar modificações.

O alinhamento de ontologias é a tarefa de criação de ligações entre duas ontologias originais. Segundo Noy *et al.* (2000), o alinhamento de ontologias é realizado se as ontologias originais estiverem consistentes entre si e foram mantidas separadas. O alinhamento geralmente é realizado quando as ontologias possuem domínios complementares.

Conforme Pinto *et al.* (1999), a integração de ontologias é o processo de geração de uma única ontologia em um assunto a partir de duas ou mais ontologias existentes e que tratam de diferentes assuntos. Os diferentes assuntos das ontologias distintas podem ser relacionados. Alguma modificação é esperada na ontologia integrada resultante.

## 2.3 Ontologias na Web Semântica

Grande parte das ontologias publicadas na Internet atualmente utilizam o padrão de estruturação de informações da Web Semântica (BERNERS-LEE *et al.*, 2001) conhecido como *Linked Data*, criado por Berners-Lee (2006). O desenvolvimento da WWW (*World Wide Web*) possibilitou a criação de um espaço de informação global composto por páginas interligadas que constituem os websites. Apesar da maior parte dos websites possuírem algum grau de estruturação, a linguagem na qual eles são criados, a HTML (*Hypertext Markup Language*), é orientada para a estruturação de documentos de texto e não para a estruturação de dados para serem processados por outras aplicações.

O formato *Linked Data* provê um paradigma de publicação no qual não apenas documentos, mas também dados possam ser estruturados e interligados na Web visando a extensão da Web em um espaço global de dados baseado em padrões abertos, denominado Web Semântica. O propósito principal da Web Semântica é direcionar a evolução da Web para permitir que ferramentas computacionais encontrem, compartilhem e combinem informações de forma mais fácil.

A estrutura tradicional da Web baseada em páginas HTML é projetada para uso por pessoas e não por máquinas, dificultando a realização de tarefas por ferramentas computacionais sem a intervenção humana. A Web Semântica propõe a estruturação das informações da Web em um formato que possa ser prontamente interpretado por ferramentas computacionais, de forma que máquinas possam realizar tarefas mais tediosas que envolvem a busca, combinação e processamento de dados na Web.

A Web Semântica, conforme foi vislumbrada originalmente por Berners-Lee *et al.* (2001), é um sistema que possibilita que máquinas entendam e respondam a solicitações humanas complexas baseados no seu significado. Este entendimento requer que fontes relevantes de informação possam ser estruturadas semanticamente.

Várias tecnologias e linguagens são utilizadas para estruturação de dados da Web Semântica como RDF (*Resource Description Framework*), OWL (*Web Ontology Language*) e XML (*Extensible Markup Language*). Enquanto a linguagem HTML é voltada para descrição de documentos e ligações entre eles, as linguagens RDF, OWL e XML, por sua vez, são voltadas para descrição de elementos como pessoas, animais, lugares, objetos, entre outros que podem ser estruturados na forma de ontologias.

Estas tecnologias são combinadas com o objetivo de prover descrições que suplementam ou substituem o conteúdo das páginas Web. Na Web Semântica o conteúdo é publicado na forma de ontologias armazenadas em repositórios da Internet. As ontologias são lidas por máquinas que adicionam significado ao conteúdo, ou seja, descrevem a estrutura do conhecimento a partir da informação semântica atrelada ao conteúdo. Desta forma, as ferramentas computacionais podem processar o conhecimento, usando processos similares ao processo de inferência e raciocínio humano, para obter resultados mais significativos e permitir que computadores realizem automaticamente pesquisas e associações complexas de informações na Internet.

## **2.4 Estado da arte no uso de ontologias no domínio da TV digital**

Diversos trabalhos acadêmicos fizeram uso de ontologias e metadados para realizar a integração de informações em diferentes domínios. Em particular, no domínio das informações relacionadas com a área de TV digital, diversos trabalhos propuseram a criação e o mapeamento de ontologias.

Yoowon *et al.* (2011) propuseram um novo modelo de *t-commerce* baseado em interações semânticas. O propósito do trabalho foi o projeto e implementação de um sistema que empregava regras de interação espacial, ontologias e anotações com descrições dos objetos disponíveis para interação nos conteúdos de vídeo, para coletar informações sobre as preferências de compra do telespectador com base em sua interação sobre a programação de TV. A partir de uma seleção realizada pelo telespectador sobre um objeto de interesse, o sistema baseado em interações semânticas exibe uma listagem catalogada de produtos relacionados ao objeto de interesse escolhido, e direciona o usuário para sites na Internet onde os produtos podem ser adquiridos.

No trabalho apresentado por Kim *et al.* (2011) foi projetado e implementado um sistema de anúncios personalizados pela TV digital interativa em redes IP (IPTV) utilizando ontologias para criar relações semânticas entre os conteúdos de TV e cada telespectador e seu respectivo estereótipo. A proposta dos autores foi a de construir primeiramente uma ontologia do telespectador baseada nas relações semânticas entre o mesmo e os conteúdos de TV de sua preferência. Com base nesta ontologia, o sistema proposto classifica cada telespectador em um determinado estereótipo, através da comparação semântica entre a ontologia do telespectador e a ontologia do estereótipo. Após esta etapa de classificação, o sistema passa a exibir para o telespectador apenas os anúncios comerciais que são direcionados para o estereótipo relacionado.

A pesquisa dos autores Redondo-Garcia *et al.* (2010) definiu uma especificação de um sistema de conteúdos baseado em ontologias que otimiza a forma que os usuários acessam os metadados da TV digital. Primeiramente, o sistema lê as informações de metadados rígidos contidos no fluxo de transporte do sinal de TV digital no padrão DVB-SI. Então, o sistema organiza e armazena os metadados utilizando ontologias, adicionando informações de outras fontes quando necessário. Finalmente, o sistema permite que o usuário realize buscas e receba recomendações de conteúdos conforme suas preferências.

Rey-López *et al.* (2010) propuseram a incorporação de conteúdos adicionais relacionados a segmentos de programas de TV pela criação de relações semânticas obtidas utilizando ontologias e informações de segmentação do padrão de metadados MPEG-7 e TVA. Como aplicações práticas do trabalho, os autores utilizaram os estudos de casos: *t-learning* – com o objetivo de utilizar os programas de TV para engajar o usuário em novas experiências educacionais; e propaganda personalizada – cujo objetivo é oferecer aos telespectadores produtos e serviços de seu interesse, maximizando o retorno dos comerciais de televisão.

O projeto POLYSEMA (PAPADIMITRIOU *et al.*, 2007) propôs um framework de receptor de TV digital inteligente com capacidade de coletar, estender e processar metadados semânticos relacionados ao conteúdo broadcast multimídia, de modo a oferecer serviços de informação diferenciados para os telespectadores. Neste contexto, os provedores de serviços e emissoras de TV devem prover descrições de metadados no padrão MPEG-7 a serem transmitidas juntamente com os conteúdos multimídia. Os receptores decodificam as informações de metadados MPEG-7 a partir do fluxo de transporte do sinal da TV digital, ou alternativamente, podem fazer o download destes metadados via Internet a partir dos web sites dos provedores de serviços.

Tsinaraki *et al.* (2004) descreveram uma metodologia sistemática para extensão de padrões de descrição de conteúdo audiovisual (MPEG-7 e TVA) com descrições de conhecimento de domínios específicos por meio do uso de ontologias. Nesse trabalho foi demonstrada a viabilidade de fazer uma representação dos metadados voltados para descrição de conteúdo audiovisual de MPEG-7 e TVA na forma de ontologias descritas na linguagem OWL (GUINNESS *et al.*, 2009). O emprego de ontologias, segundo os autores, foi fundamental para possibilitar uma extensão dos padrões MPEG-7 e TVA com descrições de conhecimentos de domínios específicos, de forma transparente para as aplicações e ferramentas já existentes, sem a necessidade de fazer modificações no software em operação.

Em outro trabalho, Tsinaraki *et al.* (2005) descreveram um *framework* desenvolvido para suporte de busca e indexação semântica de conteúdos audiovisuais baseada em ontologias seguindo os padrões MPEG-7 e TVA para descrições de metadados. O objetivo dos autores foi prover uma metodologia para aprimorar a eficiência da busca e recuperação de conteúdos audiovisuais, mantendo ao mesmo tempo uma compatibilidade com padrões internacionais. Nesse *framework*, ontologias de domínios específicos guiaram a definição tanto de metadados específicos da aplicação como de metadados de descrição de instância, que descrevem os conteúdos dos programas audiovisuais e/ou seus segmentos. As ontologias também foram usadas para prover descrições compatíveis em ambos os padrões MPEG-7 e TVA para o mesmo conteúdo. A metodologia desenvolvida nesse trabalho abriu a possibilidade de oferecimento de serviços baseados em MPEG-7 e TVA com a integração de ontologias de domínios específicos visando aprimorar os mecanismos de busca e recuperação de conteúdos audiovisuais.

No projeto AVATAR (FERNÁNDEZ *et al.*, 2006), os autores propuseram um sistema para recomendação automática de conteúdos, voltado para facilitar o acesso aos programas transmitidos por canais de TV digital. O sistema proposto fez uso de tecnologias de Web Semântica e ontologias descritas na linguagem OWL. A ontologia desenvolvida para o projeto armazena classes, instâncias e propriedades organizadas hierarquicamente, que identificam os recursos e relacionamentos comumente usados no domínio da TV, tais como categorias de programas, elenco, gêneros e sinopses. Os autores usaram o padrão de metadados TVA como base para criação da ontologia em OWL, de forma a tornar possível um mapeamento entre as informações de conteúdos TVA e as instâncias específicas existentes na ontologia que foi criada.

No projeto IndexTV (ROVIRA *et al.*, 2004) o objetivo foi o desenvolvimento de ferramentas de descrição de conteúdo e os mecanismos necessários para transformar os receptores de TV digital em sistemas que permitam a personalização e interação mais eficiente dos usuários com os conteúdos transmitidos. No projeto foi implementado um sistema completo de personalização de TV digital baseado na aplicação do padrão MPEG-7 para a descrição de conteúdos, sincronização de metadados transmitidos via *broadcasting* e aplicação de mecanismos de inteligência artificial para suportar a personalização de conteúdo para o usuário conforme seu perfil e preferência pessoais.

No trabalho *myTV* (MCPARLAND *et al.*, 2001) foi realizada uma implementação do padrão TVA em um sistema de TV digital compatível com o padrão europeu DVB-S que conta também com acesso a Internet. O projeto apresentou diversas soluções técnicas para o transporte de metadados TVA em uma rede de *broadcast* de TV digital DVB-S. Além disso, o sistema proposto previu o acesso a repositório de informações de metadados TVA em servidores da Internet usando o canal de interatividade do receptor de TV digital. Esse projeto comprovou o funcionamento do padrão TVA em uma implementação real e elencou diversas dificuldades a serem superadas durante esse processo.

O projeto *Entercaction* (REY-LÓPEZ *et al.*, 2006) apresentou um novo conceito para o aprendizado através da TV digital interativa, em que os programas são usados como gancho para atrair os usuários para uma experiência educacional. A metodologia do projeto identificou as características dos programas que poderiam aguçar a curiosidade do telespectador e selecionou os conteúdos educacionais (objetos de aprendizado) apropriados para satisfazer esta curiosidade. Uma ontologia foi usada para classificar os programas de TV e armazenar suas características mais relevantes. Outra ontologia, baseada nas especificações SCORM (ADL SCORM, 2009), foi criada para armazenar informações sobre os objetos de aprendizado e para estabelecer relacionamentos semânticos entre eles e os programas de TV. A metodologia do projeto contemplou também a implementação do mecanismo de seleção dos conteúdos educacionais para complementar os programas de TV, usando as informações contidas nas ontologias que foram criadas.

A Tabela 2.1 apresenta um quadro comparativo dos trabalhos apresentados no que concerne à utilização de ontologias OWL, ao padrão de metadados utilizado, se o trabalho propõe uma integração entre diferentes padrões de metadados, se é proposta a integração das informações da TV digital com informações da Internet e se a implementação prática da pesquisa exige o envolvimento do provedor de serviços ou emissora de TV digital.

TRABALHO	ONTOLOGIAS OWL	PADRÃO DE METADADOS	INTEGRAÇÃO DE METADADOS	INTEGRAÇÃO COM INFORMAÇÕES DA INTERNET	ENVOLVIMENTO DE PROVEDOR DE SERVIÇOS / EMISSORA TV
Yoowon <i>et al.</i> (2011)	X	MPEG-7		X	X
Kim <i>et al.</i> (2011)	X	IPTV EPG		X	X
Redondo-Garcia <i>et al.</i> (2010)	X	DVB-SI		X	X
Rey-López <i>et al.</i> (2010)	X	MPEG-7 / TVA	X		X
Papadimitriou <i>et al.</i> (2007)		MPEG-7		X	X
Tsinaraki <i>et al.</i> (2004)	X	MPEG-7 / TVA	X		X
Tsinaraki <i>et al.</i> (2005)	X	MPEG-7 / TVA	X		X
Fernández <i>et al.</i> (2006)	X	TVA			X
Rovira <i>et al.</i> (2004)		MPEG-7			X
Mcparland <i>et al.</i> (2001)		TVA		X	X
Rey-López <i>et al.</i> (2006)	X	SCORM		X	X

**Tabela 2.1 – Comparativo dos trabalhos apresentados**

Ao analisar os trabalhos apresentados verifica-se que um requisito necessário para viabilizar a implementação de todos eles era a necessidade de envolvimento do provedor de serviços ou da emissora de TV para gerar e transportar os metadados e ontologias relacionados às soluções propostas. Devido a este requisito, as propostas possuem dificuldade de implementação real em redes de TV digital já em operação, pois necessitam que as emissoras e provedores de serviços tenham que realizar mudanças operacionais e adaptações no parque de equipamentos de transmissão.

O modelo de negócios atual da televisão aberta brasileira é baseado na comercialização de intervalos comerciais entre os programas. Apesar de existirem outras formas de renda, como venda de conteúdos para outros países, basicamente a viabilidade econômica de uma emissora de TV aberta e gratuita vem da publicidade e propaganda. Devido a este modelo vigente é muito pouco provável que as emissoras privadas de TV aberta brasileiras tivessem o interesse em realizar mudanças operacionais e adaptações em seus equipamentos de transmissão para viabilizar a implementação da presente proposta. O principal motivo desta dificuldade de envolvimento das emissoras seria o fato de que os novos serviços interativos proporcionados para o telespectador possibilitariam que o mesmo deixasse de assistir os intervalos comerciais enviados pela emissora para acessar informações e conteúdos complementares à programação do canal em repositórios da Web Semântica. Ou seja, a emissora passaria a perder audiência durante os intervalos comerciais, o que seria um risco potencial para sua principal fonte de receita.

Os fabricantes de receptores, por outro lado, provavelmente veriam na implementação de presente proposta em seus produtos uma excelente oportunidade para oferecerem uma vantagem



para o seu consumidor, tornando seus produtos diferenciados em relação à concorrência. Devido a estes fatores, a eliminação da exigência de envolvimento das emissoras e provedores de serviços que existe nos trabalhos acadêmicos que fizeram o uso de ontologias e metadados no domínio de *broadcast* multimídia será a principal inovação do presente trabalho. O objetivo é fazer com que os benefícios e novos recursos interativos proporcionados para o telespectador pelo sistema proposto possam ser implementados na prática diretamente no receptor de televisão, sem a necessidade de envolvimento operacional das emissoras e provedores de serviços de TV digital.

## 2.5 Problema da heterogeneidade dos metadados

Na medida em que um grande número de bases de conhecimento independentes vai sendo introduzido na Web Semântica o problema da heterogeneidade entre diferentes ontologias está cada vez mais presente em projetos que buscam a integração de informações de fontes distintas e relacionadas a um mesmo domínio de conhecimento. No domínio dos conteúdos transmitidos pelas redes digitais de *broadcasting*, atualmente existem diversas bases de conhecimento que possuem informações que podem ser relacionadas com os metadados rígidos transmitidos nas tabelas SI no sinal da TV digital.

Um repositório da Web Semântica que pode complementar e enriquecer os metadados da TV digital é o Freebase (FREEBASE, 2012) que possui dados estruturados de aproximadamente 23 milhões de entidades relacionadas a diversos tópicos, como pessoas, lugares, livros, filmes, programas de TV, músicas, entre outros. Apenas nos tópicos relacionados à TVD, o repositório Freebase possui 192 mil filmes e 44 mil programas de TV cadastrados e estruturados em uma ontologia bastante completa e detalhada, conforme pode ser visualizada na Figura 2.5 contendo a página da entidade relacionada ao filme “Pulp Fiction”.

The screenshot shows the Freebase interface for the entity 'Pulp Fiction'. The page includes a navigation bar with 'Data', 'Schema', 'Apps', and 'Docs' tabs. The main content area features a movie poster for 'Pulp Fiction' and a detailed description: 'Pulp Fiction is a 1994 American crime film directed by Quentin Tarantino, who co-wrote its screenplay with Roger Avary. The film is known for its rich, eclectic dialogue, ironic mix of humor and violence, nonlinear storyline, and host of cinematic allusions and pop culture references. The film was nominated for seven Oscars, including Best Picture; Tarantino and Avary won for Best Original Screenplay. It was also awarded the Palme d'Or at the 199... More W Read article at Wikipedia'. Below the description, key information is listed: 'Initial release date: Oct 14, 1994', 'Directed by: Quentin Tarantino', 'Rating: R', 'Runtime: 2 h 34 min', 'Produced by: Lawrence Bender', and 'Screenplay by: Quentin Tarantino'. A section titled 'Awards' is partially visible at the bottom. On the right side, there are sections for 'These people have edited this topic:' (with user avatars and an 'Edit this topic' button), 'Related Topics' (listing 'American Beauty', 'The Silence of the Lambs', 'Fargo', and 'L.A. Confidential'), and 'Pulp Fiction elsewhere on the web'.

Figura 2.5 – Página de visualização da entidade “Pulp Fiction” do repositório Freebase

Outro repositório da Web Semântica, o DBpedia (DBPEDIA, 2012) extrai e estrutura em uma base de dados as informações das páginas do Wikipedia (WIKIPEDIA, 2012) em arquivos RDF que podem ser acessados e consultados a partir de browsers e linguagens de pesquisa de dados da Web Semântica, como a linguagem SPARQL (SPARQL, 2012). O repositório descreve de forma estruturada em ontologias cerca de 1,83 milhões de entidades, incluindo pessoas, lugares, álbuns musicais, filmes, vídeos games, organizações, espécies e doenças. Apenas no que concerne a categoria Filmes, o repositório DBpedia possui cerca de 60 mil entidades cadastradas e estruturadas em ontologias, conforme pode ser visualizado no documento RDF relativo aos filmes da série “Star Wars” exibido na Figura 2.6.

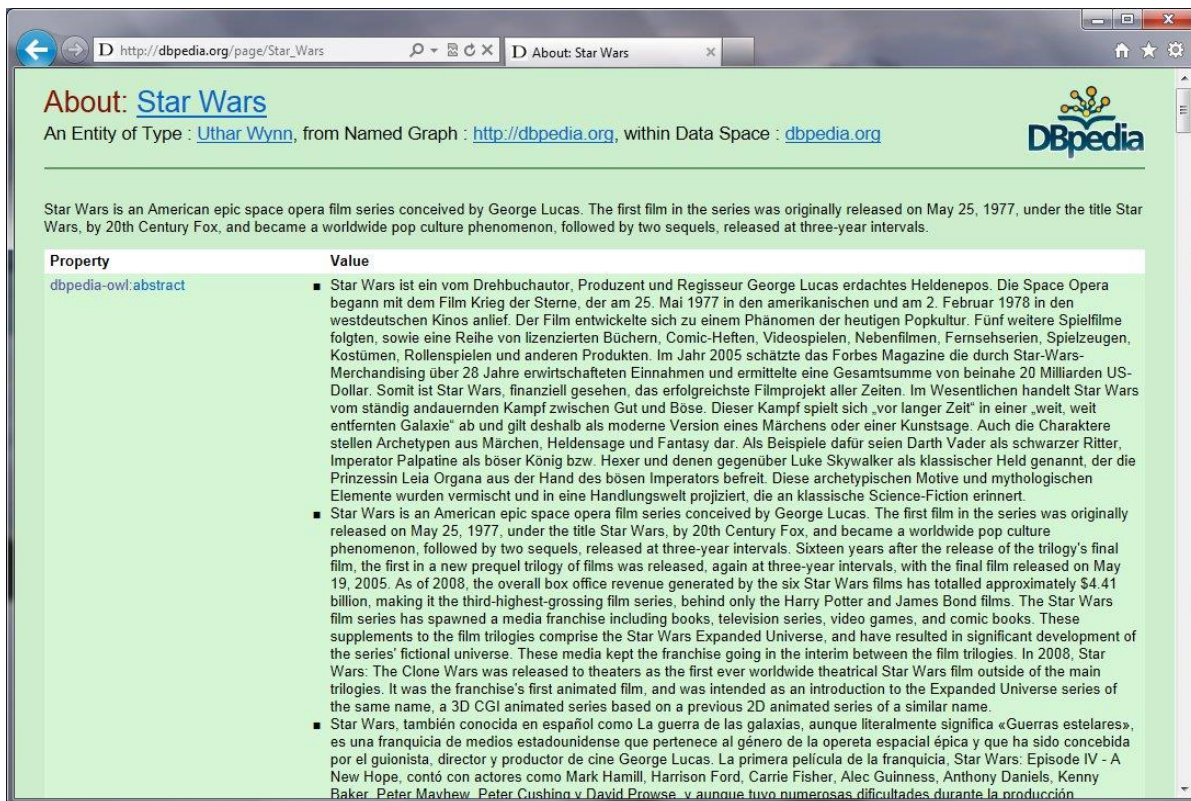
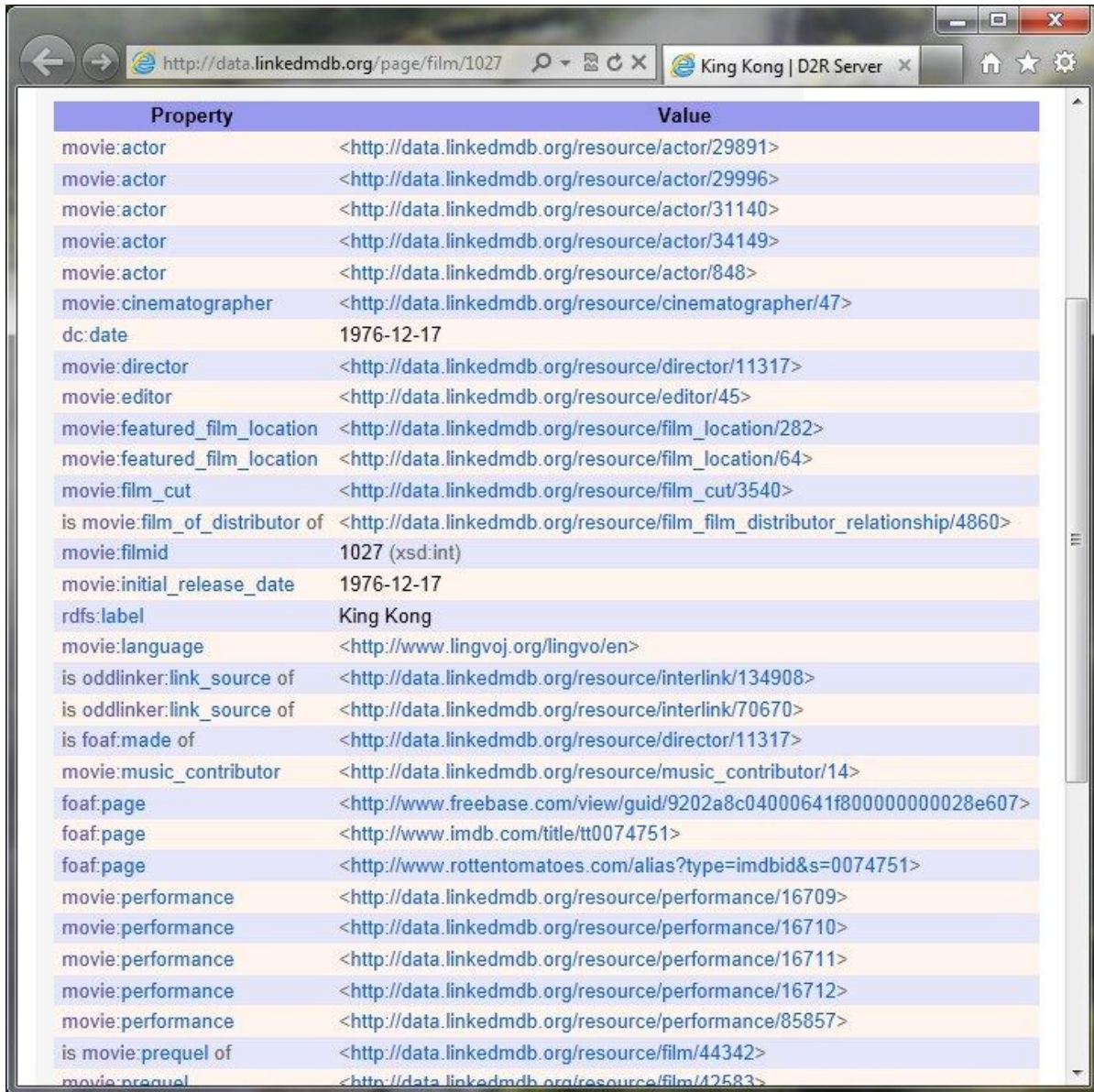


Figura 2.6 – Página de visualização da entidade “Star Wars” do repositório DBpedia

A base de dados do DBpedia possui cerca de 1 bilhão de triplas RDF, sendo aproximadamente 385 milhões extraídas da edição em inglês do Wikipedia e 665 milhões extraídas de edições em outros idiomas. Toda esta base de dados está disponível livremente para acesso via Web Semântica no padrão conhecido como *Linked Data* (BERNERS-LEE, 2006).

Outro exemplo interessante é o projeto LinkedMDB (LINKEDMDB, 2012). Este projeto objetiva criar a primeira base de conhecimento da Web Semântica dedicada à informação relacionada a filmes. O repositório possui milhões de triplas RDF com centenas de milhares de interligações com ontologias, bases de conhecimento e documentos existentes na Web.

A Figura 2.7 apresenta um exemplo de uma entidade estruturada da ontologia LinkedMDB relacionada ao filme “King Kong” com suas respectivas propriedades e valores.



Property	Value
movie:actor	<http://data.linkedmdb.org/resource/actor/29891>
movie:actor	<http://data.linkedmdb.org/resource/actor/29996>
movie:actor	<http://data.linkedmdb.org/resource/actor/31140>
movie:actor	<http://data.linkedmdb.org/resource/actor/34149>
movie:actor	<http://data.linkedmdb.org/resource/actor/848>
movie:cinematographer	<http://data.linkedmdb.org/resource/cinematographer/47>
dc:date	1976-12-17
movie:director	<http://data.linkedmdb.org/resource/director/11317>
movie:editor	<http://data.linkedmdb.org/resource/editor/45>
movie:featured_film_location	<http://data.linkedmdb.org/resource/film_location/282>
movie:featured_film_location	<http://data.linkedmdb.org/resource/film_location/64>
movie:film_cut	<http://data.linkedmdb.org/resource/film_cut/3540>
is movie:film_of_distributor of	<http://data.linkedmdb.org/resource/film_film_distributor_relationship/4860>
movie:filmid	1027 (xsd:int)
movie:initial_release_date	1976-12-17
rdfs:label	King Kong
movie:language	<http://www.lingvoj.org/lingvo/en>
is oddlinker:link_source of	<http://data.linkedmdb.org/resource/interlink/134908>
is oddlinker:link_source of	<http://data.linkedmdb.org/resource/interlink/70670>
is foaf:made of	<http://data.linkedmdb.org/resource/director/11317>
movie:music_contributor	<http://data.linkedmdb.org/resource/music_contributor/14>
foaf:page	<http://www.freebase.com/view/guid/9202a8c04000641f80000000028e607>
foaf:page	<http://www.imdb.com/title/tt0074751>
foaf:page	<http://www.rottentomatoes.com/alias?type=imdbid&s=0074751>
movie:performance	<http://data.linkedmdb.org/resource/performance/16709>
movie:performance	<http://data.linkedmdb.org/resource/performance/16710>
movie:performance	<http://data.linkedmdb.org/resource/performance/16711>
movie:performance	<http://data.linkedmdb.org/resource/performance/16712>
movie:performance	<http://data.linkedmdb.org/resource/performance/85857>
is movie:prequel of	<http://data.linkedmdb.org/resource/film/44342>
movie:prequel	<http://data.linkedmdb.org/resource/film/42583>

**Figura 2.7 – Entidade relacionada ao filme “King Kong” da ontologia LinkedMDB**

Como foi exemplificado, existem diversas bases de conhecimento na Web Semântica que possuem ontologias relacionadas ao domínio dos conteúdos transmitidos pelas redes digitais de TV digital. Estas ontologias, por estarem ligadas ao mesmo domínio de conhecimento, possuem entidades duplicadas, ou seja, entidades que possuem documentos RDF em duas ou mais bases de conhecimento na Web Semântica, e que se referem ao mesmo conteúdo. Como exemplo pode-se citar as entidades relacionadas ao filme “Titanic”, que possuem os seguintes URIs nas bases de conhecimento Freebase, DBpedia e LinkedMDB, respectivamente:

- <http://rdf.freebase.com/ns/en.titanic>
- [http://dbpedia.org/resource/Titanic\\_\(1997\\_film\)](http://dbpedia.org/resource/Titanic_(1997_film))
- <http://data.linkedmdb.org/resource/film/72>

Conforme o guia oficial da linguagem OWL (OWL, 2012), uma forma de indicar que duas entidades URI se referem à mesma coisa, ou seja, que os indivíduos possuem a mesma “identidade”, é utilizar a propriedade nativa *owl:sameAs*. No caso do exemplo do filme “Titanic” a propriedade `<owl:sameAs rdf:resource="http://rdf.freebase.com/ns/en.titanic"/>` poderia ser adicionada no documento RDF indicado pelo URI “[http://dbpedia.org/resource/Titanic\\_\(1997\\_film\)](http://dbpedia.org/resource/Titanic_(1997_film))” da base de conhecimento DBpedia, visando indicar que a entidade DBpedia se refere ao mesmo conteúdo que a entidade do repositório Freebase indicada pela propriedade *owl:sameAs*.

Segundo Halpin *et al.* (2010) a propriedade *owl:sameAs* pode ser considerada apenas um tipo de “ligação de identidade” (“*identity link*”), uma ligação que indica que duas entidades são idênticas em algum aspecto, ou seja, as entidades interligadas pela propriedade *owl:sameAs* não necessitam de ser exatamente iguais em suas propriedades e atributos, mas devem ser bastante similares ou representar a mesma coisa em contextos diferentes.

A abordagem proposta neste trabalho, assim como as propostas de Yoowon *et al.* (2011), Kim *et al.* (2011), Redondo-Garcia *et al.* (2010), Rey-López *et al.* (2010), Tsinaraki *et al.* (2004, 2005), Fernández *et al.* (2006) e Rey-López *et al.* (2006), envolve a criação de ontologias baseadas em OWL (GUINNESS *et al.*, 2009) devido a larga utilização desta linguagem na pesquisa e desenvolvimento da Web Semântica. Projetada para ser usada por aplicações que necessitam processar o conteúdo da informação ao invés de simplesmente apresentar a informação para o usuário, a OWL pode ser usada para reproduzir ontologias através da representação explícita do significado de termos em vocabulários e relacionamentos entre estes termos.

A OWL tem maior facilidade para expressar significados e semântica que as linguagens XML e RDF, que são utilizadas pelos padrões de metadados para *broadcast* multimídia como o MPEG-7, o MPEG-21 e o TVA. Com o uso da propriedade nativa *owl:sameAs*, a OWL possibilita a inter-relacionamento de entidades URI pertencentes a diferentes bases de conhecimento da Web Semântica que se referem a mesma coisa, tornando mais fácil a criação destas ligações entre entidades similares de ontologias distintas.

A OWL se tornou nos últimos anos o esforço dominante de padronização da comunidade de pesquisa internacional para a interoperabilidade entre ontologias, como pode ser verificado nas dezenas de ontologias de domínios específicos de conhecimento expressas nessa linguagem que estão publicadas no site *Protege Ontology Library* (PROTEGE ONTOLOGY LIBRARY, 2012). A utilização de uma linguagem padrão para expressar ontologias facilita a interoperabilidade entre aplicações, incluindo aquelas voltadas para a manipulação de conteúdos *broadcast*

multimídia. A OWL possibilita realizar extração de partes, união, mapeamento e importação de ontologias para diferentes aplicações.

Devido aos benefícios da utilização da OWL, neste trabalho são definidas ontologias em OWL compatíveis com os padrões de metadados da TV digital. Um mapeamento entre as ontologias OWL e os padrões foi criado para possibilitar a integração semântica das informações desses formatos e a extensão dos mesmos com a utilização de ontologias para domínios específicos existentes na Internet.

## 2.6 Contexto do trabalho no estado da arte

Este trabalho propõe a integração de metadados existentes em um ambiente de transmissão e recepção de TV digital aberta em redes de transmissão terrestre. Uma vez que o dispositivo receptor de TV digital aberta terá acesso a um canal de interatividade via Internet, será apresentado um processo para integrar as informações de metadados dos padrões da indústria de *broadcast* multimídia com ontologias voltadas para a descrição de conhecimento de domínios específicos existentes em repositórios de conteúdo da Web Semântica.

A presente proposta pode ser contextualizada no estado da arte por também fazer o uso de ontologias e metadados para realizar a integração de informações em diferentes domínios. O projeto se assemelha bastante aos objetivos do projeto descrito nos artigos de Rey-López *et al.* (2010) e Tsinaraki *et al.* (2004, 2005), mas se diferencia por focar na integração de informações de metadados dos domínios de *broadcast* multimídia fazendo extensões destes padrões com informações de metadados de domínios específicos disponíveis em repositórios da Web Semântica.

Uma grande desafio a ser superado é a eliminação da necessidade de envolvimento do provedor de serviços e da emissora de TV para gerar e transportar os metadados relacionados às ontologias propostas. A solução a ser apresentada para este problema deverá prever um mapeamento de ontologias para que as informações necessárias para gerar os metadados que são armazenados, processados e executados no receptor na casa do usuário sejam obtidas a partir dos metadados já transportados pelas redes de TV digital em operação, e também a partir de repositórios de informações de domínios específicos da Web Semântica, que serão acessados através do canal de interatividade do receptor. Somente com superação desta limitação da necessidade de envolvimento do provedor de serviços e da emissora de TV, a proposta possuirá viabilidade real de implementação, sem a necessidade de impactar os padrões abertos adotados nas redes de transmissão de TV digital terrestre, como as que estão em operação atualmente no Brasil e em outros países.

As aplicações que farão uso das ontologias geradas pelo trabalho deverão permitir que o usuário da TV digital possa enriquecer sua experiência de assistir TV com o acesso a uma gama de serviços interativos. Por meio destes serviços o telespectador poderá obter informações mais

detalhadas sobre os conteúdos transmitidos, como consultar informações sobre o histórico do diretor, dos atores principais e do elenco de um filme. O usuário deve poder também realizar gravações de conteúdos desejados e fazer buscas a partir da grade de programação dos canais existentes e dos conteúdos já gravados em seu receptor e, além disso, receber sugestões de conteúdos para exibição ou gravação conforme o seu perfil e interesse. Os novos serviços e aplicações interativas baseadas no processo proposto deverão possibilitar também o acesso a fotos, vídeos e conteúdos relacionados aos programas transmitidos disponíveis para livre acesso via Internet, entre outras funções.

# Capítulo 3

## Visão conceitual

O objetivo deste capítulo é apresentar a uma visão conceitual do trabalho com a descrição detalhada do processo que foi criado para estruturar as informações de metadados dos padrões da indústria de TV digital aberta em uma ontologia com a possibilidade de extensão destes padrões com descrições de conhecimento de domínios específicos existentes em bases de dados da Internet.

Este processo engloba mecanismos de interoperabilidade entre os metadados da indústria de *broadcast* multimídia e abre oportunidade para reuso das ontologias e modelos no desenvolvimento de um grande número de aplicações e novos serviços interativos no contexto da TV digital.

Na seção 3.1 é apresentada uma visão geral dos componentes do processo proposto, com uma descrição de cada um deles. O processo de alinhamento de metadados da indústria de *broadcast* multimídia é detalhado na seção 3.2. Na seção 3.3 o foco é a apresentação do processo para extensão dos padrões da TV digital com descrições de conhecimento de domínios específicos.

### 3.1 Visão geral dos componentes do processo proposto

Nesta seção é apresentada uma descrição dos principais componentes do processo proposto para suportar a integração de ontologias existentes na indústria de TV digital e a extensão destas ontologias com ontologias de domínios específicos existentes em repositórios de conhecimento da Internet.

As abordagens já existentes para fusão e integração de metadados e ontologias não se aplicam ao contexto do presente trabalho, devido a algumas limitações de capacidade de memória e processamento do receptor de TV digital e a necessidade do não envolvimento do provedor de serviços e da emissora de TV para gerar e transportar os metadados relacionados às ontologias propostas.

Caso fosse feita a fusão de ontologias, seria necessário o envolvimento dos provedores de serviços responsáveis pela manutenção das bases de conhecimento a serem integradas na Web Semântica e também das emissoras de TV para realizar a geração de uma ontologia única e coerente a partir das diferentes ontologias existentes que são relacionadas ao domínio da TV digital.

A abordagem de integração de ontologias, por sua vez, é um processo de geração de uma única ontologia em um assunto a partir de duas ou mais ontologias existentes. Esta abordagem exigiria a introdução de um provedor de serviços específico para hospedagem, alimentação e manutenção desta nova ontologia, que possuiria uma grande complexidade e envolveria um elevado investimento para sua operação.

Devido à não aplicação das abordagens de fusão e integração de ontologias ao contexto do presente trabalho, verifica-se que a melhor abordagem a ser aplicada é o alinhamento de ontologias, que consiste na criação de ligações entre duas ou mais ontologias originais. No caso em particular da TV digital, será feita um alinhamento dos metadados já transmitidos no sinal da TV digital com ontologias relacionadas a este domínio existentes em repositórios da Web Semântica.

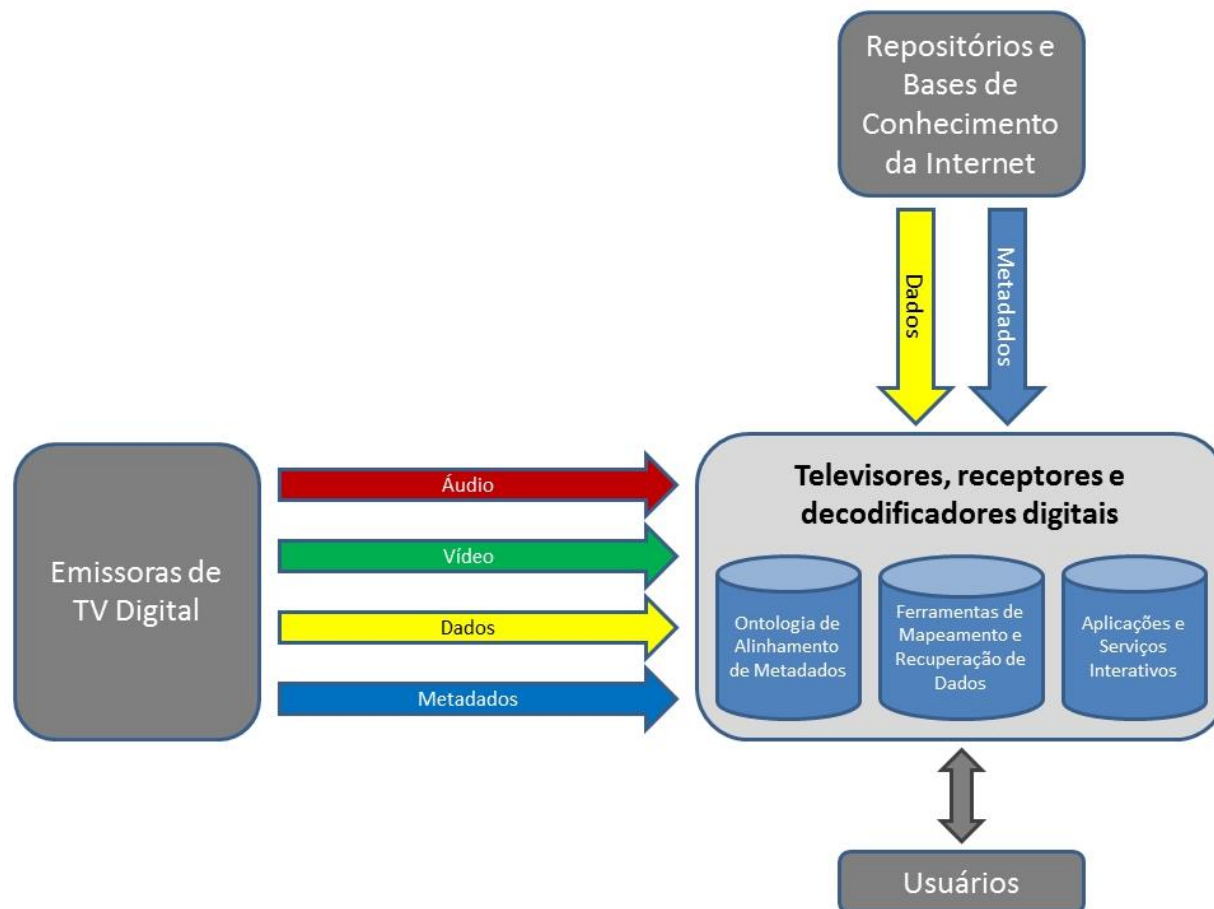
Para a realização deste alinhamento será necessário fazer a criação de uma nova ontologia a partir dos metadados transmitidos no sinal da TV digital em um processo que será descrito na seção seguinte. Esta nova ontologia de alinhamento dos metadados da TV digital será armazenada e processada localmente no receptor da TV digital e, portanto, deverá levar em consideração as limitações de capacidade de memória e processamento do mesmo.

Para que o telespectador possa ter acesso a TV digital, devem existir emissoras responsáveis pela geração e transmissão do conteúdo multimídia, que no caso do sinal digital é composto por elementos de áudio, vídeo, dados e metadados. Este sinal digital pode ser transmitido da emissora para os usuários por vários canais de transmissão como redes de radiodifusão terrestre, cabo, satélite, entre outros (MORRIS *et al.*, 2005).

Na Figura 3.1 é exibido um diagrama com os principais componentes de um sistema tradicional de transmissão e recepção de TV digital que foram selecionados para serem utilizados neste trabalho. Um dos principais componentes selecionados é o equipamento responsável pela recepção do sinal digital e apresentação dos conteúdos e serviços para o usuário ou telespectador. Este componente é representado no diagrama pelos aparelhos televisores, receptores e decodificadores digitais.

Os aparelhos televisores, receptores e decodificadores digitais recebem o sinal da TV digital das emissoras, que por este motivo, também foram selecionadas para integrar o processo proposto. Além das emissoras, outro componente fundamental do processo são os usuários que irão assistir os conteúdos e terão acesso às novas aplicações e serviços interativos que serão proporcionadas pela presente proposta.





**Figura 3.1 – Componentes do processo proposto**

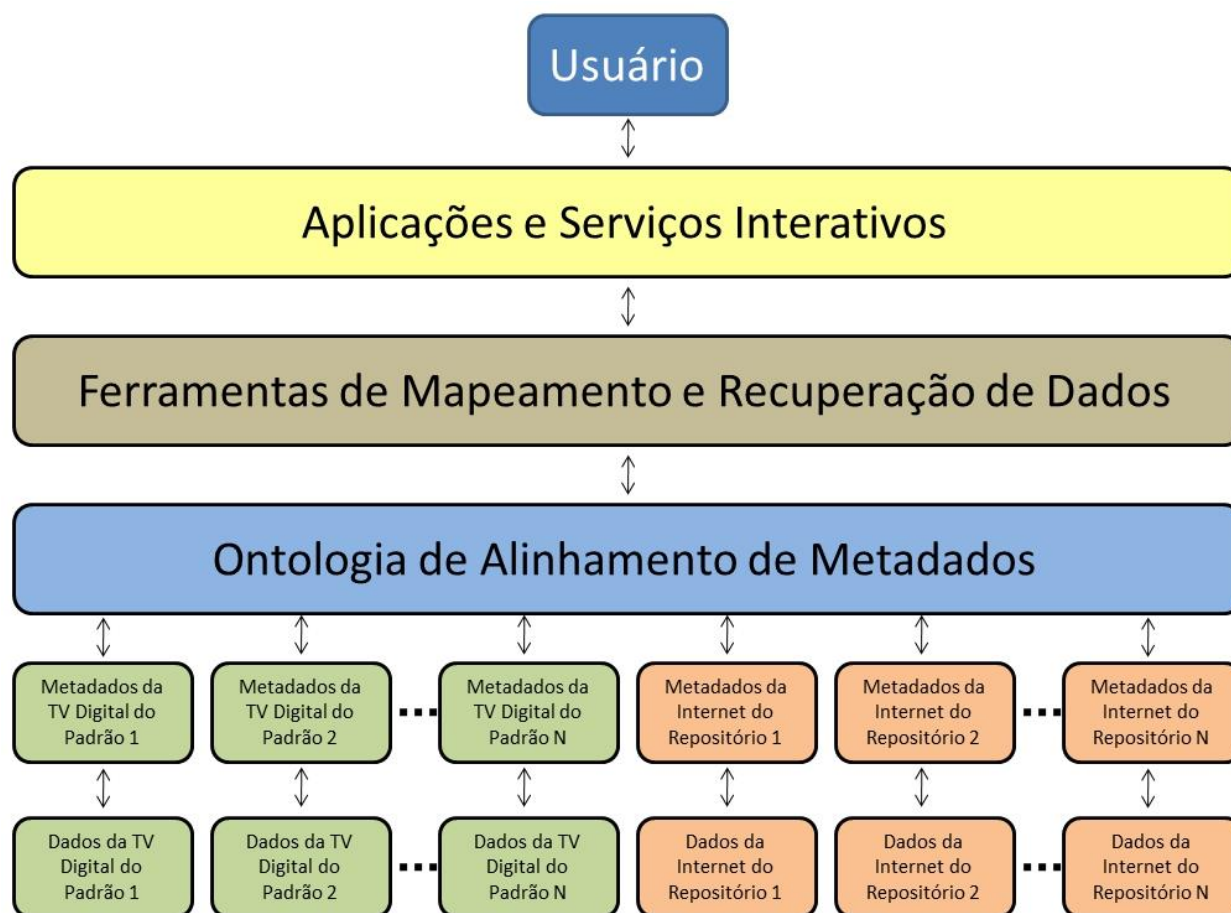
Os metadados transmitidos pelas emissoras de TV digital possuem diferentes formatos conforme o padrão de TV digital que está sendo utilizado. Para fazer uma integração entre estes diferentes padrões de metadados é proposta a criação de uma ontologia para alinhamento de metadados que permite realizar um mapeamento entre as informações de descrição de conteúdo disponíveis nos diferentes sistemas de TV digital.

Esta ontologia de alinhamento de metadados também é estendida para fazer um mapeamento entre os metadados da TV digital e metadados com descrições de conteúdos de domínios específicos existentes em repositórios e bases de conhecimento da Internet. Assume-se portanto que os aparelhos televisores, receptores e decodificadores de sinal digital possuem acesso a Internet.

Uma vez que a ontologia de alinhamento de metadados esteja disponível, a mesma pode ser armazenada internamente na memória do equipamento de recepção de sinal digital. Além da ontologia, são implementadas ferramentas que fazem o uso da ontologia de alinhamento de metadados para fazer o mapeamento e a recuperação de dados transmitidos pelos diferentes padrões de TV digital e também de dados existentes nos repositórios da Internet.

Desta forma, é possível estender e complementar as informações disponíveis na TV digital com informações de domínios de conhecimento específicos existentes na Internet, possibilitando o desenvolvimento de uma gama de aplicações e serviços interativos para os usuários. Essas aplicações e serviços interativos podem utilizar as ferramentas de mapeamento e recuperação de dados e serem executadas pelos telespectadores utilizando os recursos de memória e processamento dos equipamentos de recepção de sinal digital.

No diagrama da Figura 3.2 é ilustrada a interdependência entre os componentes do processo desenvolvido para representar a interação dos usuários com as aplicações e serviços interativos. Assume-se que o telespectador da TV digital fará interação com as aplicações e serviços interativos através de uma interface gráfica exibida na tela do televisor sobre o conteúdo de áudio e vídeo da programação. O dispositivo de interação utilizado pelos usuários normalmente é o controle remoto, mas outros dispositivos, como teclado e mouse, também são suportados pelo modelo.



**Figura 3.2 – Interdependência entre os componentes do processo desenvolvido**

O usuário, ao interagir com uma aplicação ou serviço interativo, poderá obter informações detalhadas sobre o elenco de um filme, ou o histórico de resultados no campeonato de uma equipe que participa de um evento esportivo. Estas informações que o usuário necessita em geral

não são transmitidas pelos dados e metadados gerados pelas emissoras no sinal de TV digital. Por este motivo, faz-se necessário recuperar estas informações de outros repositórios e bases de conhecimento existentes na Internet. Para isso, o código programável das aplicações e serviços interativos utiliza as ferramentas disponíveis no receptor digital para mapeamento e recuperação de dados. Assim, por meio de métodos e funções de software disponibilizadas pelas ferramentas é possível realizar consultas sobre informações complementares às que estão sendo transmitidas pelas emissoras no sinal de TV digital.

Os métodos e funções de software das ferramentas de mapeamento e recuperação de dados utilizam como base de inteligência a ontologia de alinhamento de metadados para realizar o mapeamento entre as informações de descrição de conhecimento da TV digital e os metadados de domínios específicos dos repositórios da Internet.

Portanto, a ontologia de alinhamento de metadados é a componente chave do sistema, pois é ela que possibilita fazer um relacionamento entre dados existentes em um repositório da Internet com um conteúdo transmitido na TV digital. Através deste mapeamento entre as informações de metadados dos diferentes domínios, as ferramentas disponíveis no receptor digital conseguem recuperar os dados que serão utilizados pelas aplicações e serviços interativos.

### **3.2 Construção da ontologia de alinhamento de metadados**

Conforme foi apresentado na seção anterior, um sinal de TV digital contém quatro tipos de conteúdos multimídia principais: áudio, vídeo, dados e metadados. As informações de metadados usadas para descrever os dados transmitidos pelo sinal são diferentes conforme o padrão de TV digital utilizado.

Para fins de ilustração da aplicação do processo de construção da ontologia de alinhamento de metadados da TV digital, assume-se que o receptor de TV digital da arquitetura proposta é um receptor conectado. Os receptores conectados possuem acesso ao sinal de TV digital transmitido via rede terrestre (ISDB-T) e contam também com acesso à Internet via redes cabeadas ou sem fio.

No Brasil e em vários outros países da América Latina o padrão de TV digital aberta utilizado para transmissão terrestre é baseado no padrão japonês ISDB-T com algumas modificações visando a adaptação às necessidades específicas de cada país. Contudo, pode-se observar que em todos os países que adotaram o ISDB-T, assim como na maior parte dos diferentes padrões abertos de TV digital que existem atualmente no mundo, o sistema utiliza metadados relacionados à descrição dos conteúdos que são transmitidos pelas emissoras.

Em geral, a informação de metadados mais comum que é transmitida pelo sistema de TV digital para descrição do conteúdo é o título. Além desta informação, diversas outras, como classificação parental, gênero, palavras chave, sinopse e duração do conteúdo podem ser encontradas nos metadados e dados transmitidos pelos diferentes sistemas.

O primeiro passo da aplicação do processo para extensão dos metadados da TV digital é criar uma ontologia que seja um modelo unificado para representação do conhecimento e mapeamento das informações de metadados entre os diferentes padrões. Esta ontologia é composta por uma hierarquia de classes e subclasses, com suas respectivas propriedades, restrições e instâncias.

Como base para a criação desta ontologia foi utilizado o padrão XMLTV (XMLTV, 2012). Este formato foi escolhido devido a sua crescente utilização pela indústria de TV no Brasil e em outros países, e também pelo fato do mesmo contemplar praticamente todas as informações de metadados que são transmitidas nas tabelas de informações de serviços (*SI – Service Information*) do padrão de TV digital aberta utilizado no Brasil em redes de transmissão terrestre (ISDB-T).

A aplicação do processo de alinhamento e extensão de metadados se inicia com a geração de uma ontologia baseada no padrão XMLTV para ser armazenada na memória do receptor de TV digital. Para geração desta ontologia, foi utilizada a metodologia proposta por Thuy *et al.* (2009), que toma como base o esquema XML definido em um formato DTD (*Document Type Definition*) e aplica transformações para convertê-lo em uma ontologia baseada em OWL, denominada OWLTV. O esquema XML do padrão XMLTV está disponível para download na Internet (XMLTV DTD, 2012).

A tabela 3.1 exemplifica a aplicação da metodologia proposta por Thuy *et al.* (2009) para geração da ontologia OWLTV a partir do esquema XMLTV em DTD. Conforme pode ser observado no exemplo, na definição do elemento DTD `<!ELEMENT channel (display-name+, icon*, url*) >`, como o elemento *channel* possui elementos filhos (*display-name+*, *icon\**, *url\**), este elemento é considerado uma classe OWL *owl:Class* representada pelo identificador único *rdf:ID="channel"*.

Trecho do Esquema DTD do Formato XMLTV	Ontologia OWLTV Relacionada
<pre> &lt;!ELEMENT channel (display-name+, icon*, url*) &gt; &lt;!ATTLIST channel id CDATA #REQUIRED &gt;  &lt;!ELEMENT display-name (#PCDATA)&gt; &lt;!ATTLIST display-name lang CDATA #IMPLIED&gt;  &lt;!ELEMENT icon EMPTY&gt; &lt;!ATTLIST icon src          CDATA #REQUIRED               width         CDATA #IMPLIED               height        CDATA #IMPLIED&gt;  &lt;!ELEMENT url (#PCDATA)&gt; </pre>	<pre> &lt;owl:Class rdf:ID="channel"&gt;   &lt;owl:minCardinality&gt;0&lt;/owl:minCardinality&gt;   &lt;owl:maxCardinality&gt;unbounded&lt;/owl:maxCardinality&gt;   &lt;owl:disjointWith&gt;     &lt;owl:Class rdf:ID="display-name"/&gt;&lt;/owl:disjointWith&gt;   &lt;owl:disjointWith&gt;     &lt;owl:Class rdf:ID="icon"/&gt;&lt;/owl:disjointWith&gt;   &lt;owl:disjointWith&gt;     &lt;owl:Class rdf:ID="url"/&gt;&lt;/owl:disjointWith&gt; &lt;/owl:Class&gt; &lt;owl:Class rdf:ID="display-name"&gt;   &lt;owl:minCardinality&gt;1&lt;/owl:minCardinality&gt;   &lt;rdfs:range rdf:resource="&amp;xsd:string"/&gt; &lt;/owl:Class&gt; &lt;owl:Class rdf:ID="icon"&gt;   &lt;owl:minCardinality&gt;0&lt;/owl:minCardinality&gt;   &lt;owl:maxCardinality&gt;unbounded&lt;/owl:maxCardinality&gt;   &lt;rdfs:range rdf:resource="&amp;xsd:string"/&gt; &lt;/owl:Class&gt; &lt;owl:Class rdf:ID="url"&gt;   &lt;owl:minCardinality&gt;0&lt;/owl:minCardinality&gt;   &lt;owl:maxCardinality&gt;unbounded&lt;/owl:maxCardinality&gt;   &lt;rdfs:range rdf:resource="&amp;xsd:string"/&gt; &lt;/owl:Class&gt; &lt;owl:ObjectProperty rdf:ID="has_display-name"&gt; </pre>

	<pre> &lt;rdfs:domain rdf:resource="#channel"/&gt; &lt;rdfs:range rdf:resource="#display-name"/&gt; &lt;/owl:ObjectProperty&gt; &lt;owl:ObjectProperty rdf:ID="has_icon"&gt; &lt;rdfs:domain rdf:resource="#channel"/&gt; &lt;rdfs:range rdf:resource="#icon"/&gt; &lt;/owl:ObjectProperty&gt; &lt;owl:ObjectProperty rdf:ID="has_url"&gt; &lt;rdfs:domain rdf:resource="#channel"/&gt; &lt;rdfs:range rdf:resource="#url"/&gt; &lt;/owl:ObjectProperty&gt; &lt;owl:DatatypeProperty rdf:ID="channel_id"&gt; &lt;rdfs:domain rdf:resource="#channel"/&gt; &lt;rdfs:range rdf:resource="&amp;xsd:string"/&gt; &lt;owl:minCardinality&gt;1&lt;/owl:minCardinality&gt; &lt;/owl:DatatypeProperty&gt; &lt;owl:DatatypeProperty rdf:ID="display-name_value"&gt; &lt;rdfs:domain rdf:resource="#display-name"/&gt; &lt;rdfs:range rdf:resource="&amp;xsd:string"/&gt; &lt;owl:Cardinality&gt;0&lt;/owl:Cardinality&gt; &lt;/owl:DatatypeProperty&gt; &lt;owl:DatatypeProperty rdf:ID="display-name_lang"&gt; &lt;rdfs:domain rdf:resource="#display-name"/&gt; &lt;rdfs:range rdf:resource="&amp;xsd:string"/&gt; &lt;owl:Cardinality&gt;0&lt;/owl:Cardinality&gt; &lt;/owl:DatatypeProperty&gt; &lt;owl:DatatypeProperty rdf:ID="icon_src"&gt; &lt;rdfs:domain rdf:resource="#icon"/&gt; &lt;rdfs:range rdf:resource="&amp;xsd:string"/&gt; &lt;owl:minCardinality&gt;1&lt;/owl:minCardinality&gt; &lt;/owl:DatatypeProperty&gt; &lt;owl:DatatypeProperty rdf:ID="icon_width"&gt; &lt;rdfs:domain rdf:resource="#icon"/&gt; &lt;rdfs:range rdf:resource="&amp;xsd:string"/&gt; &lt;owl:Cardinality&gt;0&lt;/owl:Cardinality&gt; &lt;/owl:DatatypeProperty&gt; &lt;owl:DatatypeProperty rdf:ID="icon_height"&gt; &lt;rdfs:domain rdf:resource="#icon"/&gt; &lt;rdfs:range rdf:resource="&amp;xsd:string"/&gt; &lt;owl:Cardinality&gt;0&lt;/owl:Cardinality&gt; &lt;/owl:DatatypeProperty&gt; </pre>
--	-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

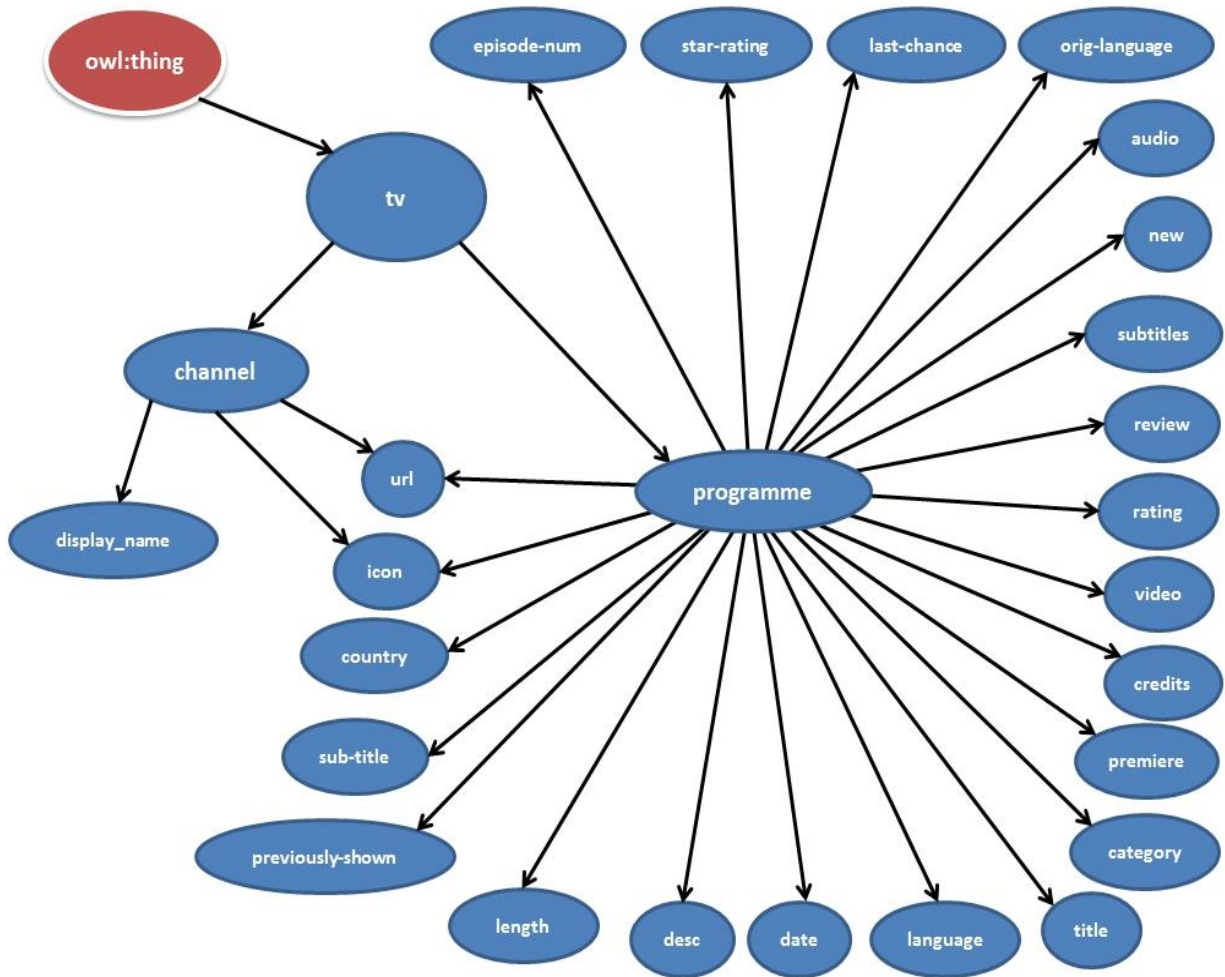
**Tabela 3.1 – Exemplificação da aplicação da metodologia proposta por Thuy *et al.* (2009)**

Para elementos aninhados, a metodologia proposta por Thuy *et al.* (2009) não utiliza a sintaxe OWL *rdfs:subClassOf*, devido ao fato de muitos elementos aninhados em DTD não serem necessariamente subclasses do elemento pai. Neste caso é adicionada na ontologia OWLTV uma nova propriedade de objeto descrita por *owl:ObjectProperty* para estabelecer o relacionamento entre o elemento pai e o elemento filho. O nome da propriedade de objeto é derivado pela concatenação para palavra “has” com o símbolo “\_” e o nome do elemento filho, como pode ser observado no exemplo da Tabela 3.1, nas propriedades de objeto identificadas por “has\_display\_name”, “has\_icon” e “has\_url”. A especificação dos elementos pai e filho é representada por *rdfs:domain* e *rdfs:range* respectivamente.

No caso dos atributos dos elementos DTD, descritos pela tag *<!ATTLIST>* que contém apenas o tipo de dado (*PCDATA* ou *CDATA*) e suas respectivas restrições (*#FIXED*, *#REQUIRED*, *#IMPLIED* entre outras), os mesmos são mapeados para propriedades OWL definidas por *owl:DatatypeProperty*, como no caso do atributo *id* do elemento *channel* e também do atributo *lang* do elemento *display-name*. O tratamento semântico do tipo de dado e das

restrições dos elementos e atributos DTD no processo de conversão para a ontologia OWLTV também é realizado conforme a metodologia proposta por Thuy *et al.* (2009).

Como pode ser observado na representação gráfica parcial de classes apresentada na Figura 3.3, que exibe as classes da ontologia até o terceiro nível hierárquico, a ontologia OWLTV resultante da aplicação da metodologia preserva toda a estrutura hierárquica do documento XMLTV original. Porém a representação em OWL agrega informação semântica, e prepara os metadados XMLTV para serem utilizados por aplicações da Web Semântica.



**Figura 3.3 – Representação gráfica parcial da ontologia OWLTV obtida a partir do padrão XMLTV**

Na Figura 3.4 é apresentado um diagrama de classes que exibe a hierarquia de classes da ontologia OWLTV gerada a partir da aplicação da metodologia proposta por Thuy *et al.* (2009) sobre as entidades do esquema XMLTV em DTD.

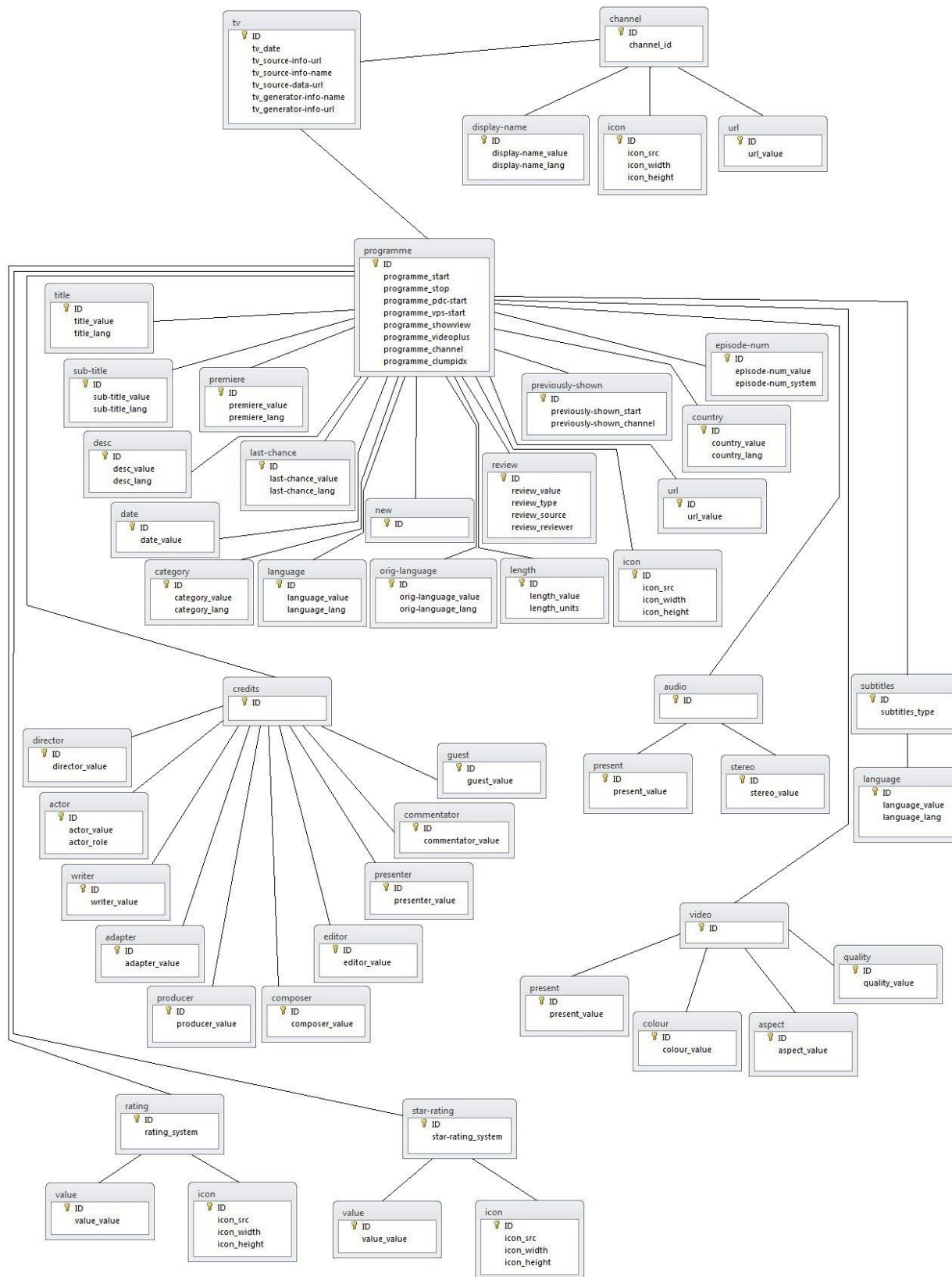


Figura 3.4 – Diagrama de classes da ontologia OWLTV obtida a partir do padrão XMLTV





Para fins de comparação, a pesquisa da entidade “Jornal Nacional” no repositório FreeBase retorna apenas uma única instância da classe *Freebase.type.object.name* com o atributo *xml:lang* com o valor “en”, conforme exibido na Figura 3.6.

```
<fb:type.object.name xml:lang="en">Jornal Nacional</fb:type.object.name>
```

**Figura 3.6 – Instância da classe *Freebase.type.object.name* para a entidade “Jornal Nacional”**

Como observado no exemplo anterior, na análise de duas classes de ontologias distintas que poderiam ser relacionadas, verifica-se que as instâncias das classes das ontologias OWLTV (*OWLTV.tv.programme.title*) e Freebase (*Freebase.type.object.name*) não são precisamente as mesmas. Desta forma, não será possível utilizar a declaração *owl:equivalentClass* para fazer este relacionamento, pois além das instâncias terem de ser exatamente as mesmas, as propriedades das classes também teriam de ser exatamente iguais.

Para resolver esta limitação de uso da declaração *owl:equivalentClass*, de forma a permitir a extensão e integração da ontologia OWLTV com outras ontologias, uma nova abordagem é proposta neste trabalho. Esta abordagem objetiva fazer com que os indivíduos da ontologia OWLTV a ser gerada dinamicamente pelo receptor de TV digital possam ser relacionados com indivíduos de ontologias de outros domínios da Web Semântica através da propriedade nativa *owl:sameAs*. No caso do exemplo utilizado anteriormente relativo ao programa “Jornal Nacional”, a propriedade `<owl:sameAs rdf:resource="http://rdf.freebase.com/ns/en.jornal_nacional"/>` poderia ser adicionada à entidade OWLTV correspondente ao programa “Jornal Nacional” exibido no dia 07 de junho de 2012 às 20:35 UTC-3, para estender as informações da ontologia OWLTV com a entidade correspondente na ontologia Freebase.

Na abordagem aqui proposta é adicionada na ontologia OWLTV uma nova classe descrita por *owl:Class* para estabelecer o relacionamento entre entidades da ontologia OWLTV e entidades de ontologias distintas. O nome desta nova classe é *sameAsQuery* e ela possui três propriedades OWL definidas por *owl:DatatypeProperty*, como pode ser observado na definição da classe em OWL exibida na Figura 3.7. A propriedade *owltv\_entity* armazena o nome da entidade relacionada na ontologia OWLTV, a propriedade *linked\_ontology* armazena nome da ontologia a ser relacionada com a entidade definida pela propriedade *owltv\_entity*, a terceira propriedade, *query\_code*, armazena o código da consulta a ser realizada na base de conhecimento definida na propriedade *linked\_ontology* para extração das entidades relacionadas à entidade da ontologia OWLTV definidas na propriedade *owltv\_entity*.

```
<owl:Class rdf:ID="sameAsQuery">
  <owl:minCardinality>0</owl:minCardinality>
  <owl:maxCardinality>unbounded</owl:maxCardinality>
</owl:Class>
<owl:DatatypeProperty rdf:ID="owltv_entity">
  <rdfs:domain rdf:resource="#sameAsQuery"/>
```

```

<owl:minCardinality>1</owl:minCardinality>
<rdfs:range rdf:resource="#xsd:string"/>
</owl:DatatypeProperty>
<owl:DatatypeProperty rdf:ID="linked_ontology">
  <rdfs:domain rdf:resource="#sameAsQuery"/>
  <owl:minCardinality>1</owl:minCardinality>
  <rdfs:range rdf:resource="#xsd:string"/>
</owl:DatatypeProperty>
<owl:DatatypeProperty rdf:ID="query_code">
  <rdfs:domain rdf:resource="#sameAsQuery"/>
  <owl:minCardinality>1</owl:minCardinality>
  <rdfs:range rdf:resource="#xsd:string"/>
</owl:DatatypeProperty>

```

**Figura 3.7 – Classe *sameAsQuery* que alinha entidades OWLTV com entidades de ontologias distintas**

Tomando como exemplo a entidade OWLTV correspondente ao programa “Jornal Nacional” exibido no dia 07 de junho de 2012 às 20:35 UTC-3, um exemplo de instância da classe *sameAsQuery* está exibido na Figura 3.8, cujo código da consulta atribuído à propriedade *query\_code* é baseado na linguagem *Metaweb Query Language* (MQL, 2012) que é utilizada para realizar consultas na base de conhecimento Freebase. O valor da propriedade *owltv\_entity* é o identificador da entidade relacionada ao programa “Jornal Nacional”, por exemplo, “programme\_1”, e o valor da propriedade *linked\_ontology* é o nome da base de conhecimento a ser consultada, no caso “Freebase”.

```

<sameAsQuery rdf:ID="sameAsQuery_1">
  <owltv_entity rdf:datatype="#xsd:string">programme_1</owltv_entity>
  <linked_ontology rdf:datatype="#xsd:string">Freebase</linked_ontology>
  <query_code rdf:datatype="#xsd:string">
    [{
      "id": null,
      "name|=": ["Jornal Nacional"],
      "type|=": ["/tv/tv_program", "/film/film"]
    }]
  </query_code>
</sameAsQuery>

```

**Figura 3.8 – Exemplo de instância da classe *sameAsQuery***

Os valores das instâncias da classe *sameAsQuery* irão variar conforme a entidade da ontologia OWLTV definida na propriedade *owltv\_entity* e a base de conhecimento definida na propriedade *linked\_ontology*. Para realizar os ajustes necessários nos valores das instâncias da classe *sameAsQuery*, o processo proposto neste trabalho prevê a existência de funções de software específicas para consulta em cada nova ontologia a ser integrada à OWLTV. A utilização destas funções de software específicas será necessária pelo fato de cada repositório de conhecimento da Web Semântica utilizar linguagens e ferramentas específicas para a consulta às

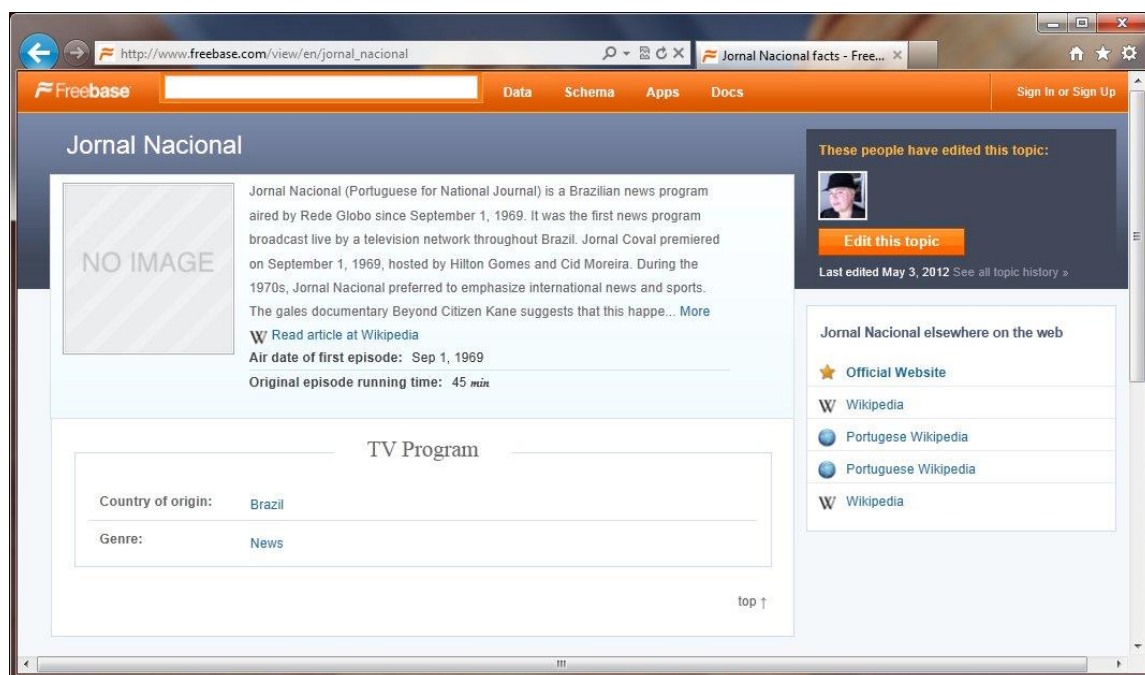
suas informações. No exemplo da Figura 3.8, o valor da instância da classe *sameAsQuery* será gerado dinamicamente pela função de software específica para geração de consultas à base de conhecimento Freebase.

O resultado da execução do código MQL da Figura 3.8 na ferramenta de consulta da base de conhecimento Freebase está exibido na Figura 3.9. Como pode ser observado, a consulta retorna apenas uma única instância do repositório Freebase com identificador (*id*) igual a *"/en/jornal\_nacional"*. Desta forma, a função de software de geração de consultas à base de conhecimento Freebase do processo proposto poderá inferir que a entidade OWLTV correspondente ao programa “Jornal Nacional” exibido no dia 07 de junho de 2012 às 20:35 UTC-3 é relacionada à entidade Freebase identificada por *"/en/jornal\_nacional"*.

```
{ {
  "id": "/en/jornal_nacional"
}}
```

**Figura 3.9 – Resultado da consulta MQL da Figura 3.7**

A Figura 3.10 exibe a página do site Freebase relacionada à entidade *"/en/jornal\_nacional"*. Com base no resultado da execução desta consulta pela função de software específica da base de conhecimento Freebase o processo proposto poderá acrescentar a propriedade `<owl:sameAs rdf:resource="http://rdf.freebase.com/ns/en.jornal_nacional"/>` à entidade OWLTV correspondente ao programa “Jornal Nacional” exibido no dia 07 de junho de 2012 às 20:35 UTC-3, para estender as informações da ontologia OWLTV com a entidade correspondente na ontologia Freebase.



**Figura 3.10 - Página de visualização da entidade *"/en/jornal\_nacional"* do repositório Freebase**

O mesmo processo de alinhamento que foi realizado com a entidade relativa ao programa “Jornal Nacional” que pertence à classe *programme*, poderia ser realizado com outras entidades pertencentes a outras classes da ontologia OWLTV como as classes *country*, *director*, *actor*, entre outras, como veremos no exemplo do último capítulo do presente trabalho. Além disso, diversas ontologias podem ser integradas a uma mesma entidade, com a inserção de novas instâncias da classe *sameAsQuery* e utilizando funções de software específicas para realizar as buscas em cada ontologia a ser integrada. Por exemplo, a entidade “programme\_1” poderia ser integrada com as ontologias LinkedMDB e DBPedia com a adição das novas instâncias da classe *sameAsQuery* cujas propriedades *linked\_ontology* fossem iguais a “LinkedMDB” e “DBPedia”, respectivamente, conforme exemplificado na Figura 3.11.

```
<sameAsQuery rdf:ID="sameAsQuery_2">
  <owltv_entity rdf:datatype="&xsd:string">programme_1</owltv_entity>
  <linked_ontology rdf:datatype="&xsd:string">LinkedMDB</linked_ontology>
  <query_code rdf:datatype="&xsd:string">
    ...
  </query_code>
</sameAsQuery>
<sameAsQuery rdf:ID="sameAsQuery_3">
  <owltv_entity rdf:datatype="&xsd:string">programme_1</owltv_entity>
  <linked_ontology rdf:datatype="&xsd:string">DBPedia</linked_ontology>
  <query_code rdf:datatype="&xsd:string">
    ...
  </query_code>
</sameAsQuery>
```

Figura 3.11 – Exemplos de instâncias da classe *sameAsQuery*

### 3.4 Considerações finais do capítulo

Os novos desafios impostos pelo crescimento da complexidade da TV digital e sua integração com a Internet fazem com que os *broadcasters* e fabricantes de receptores tenham que lidar com uma diversidade de padrões de metadados para prover aos usuários serviços interativos de alto valor agregado. Este capítulo propõe um processo para integração de diferentes especificações de metadados existentes no contexto da TV digital, objetivando facilitar o desenvolvimento destes novos serviços interativos pelos *broadcasters* e fabricantes de receptores.

A aplicação desta abordagem proposta torna possível a criação de uma ontologia que faz um alinhamento dos padrões de metadados da TV digital e também uma extensão destes padrões com informações de metadados de domínios de conhecimento específicos disponíveis em repositórios da Internet.

Uma contribuição importante desta proposta é o fato de sua implementação não impactar em mudanças operacionais nos *broadcasters* e provedores de serviços. A ontologia de alinhamento de metadados gerada a partir da aplicação do processo, assim como as ferramentas de manipulação e recuperação de dados, bem como as novas aplicações e serviços interativos são totalmente processados no equipamento receptor de sinal de TV digital. Devido a isso, a contribuição pode efetivamente ser adotada sem impacto nas redes de transmissão digitais já em operação.

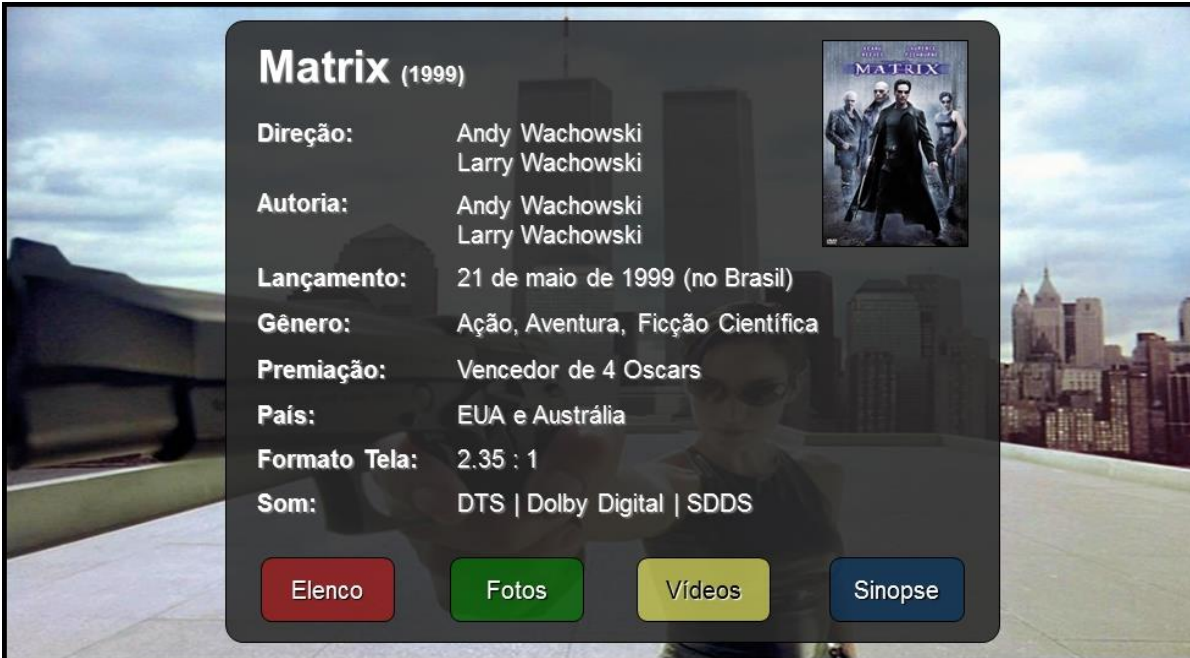
No próximo capítulo será discutida a aplicação desta abordagem para alinhar e estender os padrões da TV digital com domínios de conhecimento específicos da Internet, através de uma integração entre os padrões de metadados da TV digital utilizados no Brasil em redes terrestres (ISDB-T) com as especificações de metadados definidas pelas bases de conhecimento da Web Semântica existentes nos repositórios de conteúdo dos sites LinkedMDB, Freebase e DBpedia.

A ontologia de alinhamento de metadados a ser gerada irá permitir que um usuário da TV digital no Brasil, ao assistir um conteúdo, tenha acesso a novas funções interativas em complemento às informações já providas nos metadados do guia eletrônico de programação (EPG), exibido na Figura 3.12.



Figura 3.12 – Guia eletrônico de programação (EPG) da TV digital

A ontologia gerada pela aplicação do processo proposto irá complementar as informações do guia eletrônico de programação da TV digital, com informações de metadados existentes nos repositórios da Internet, permitindo que o usuário tenha acesso a dados complementares do programa assistido como: direção, autoria, data de lançamento, gênero, premiações, país, formato de tela, mixagem de som, elenco, sinopse, fotos e vídeos relacionados, conforme exibido na Figura 3.13.



**Matrix** (1999)

**Direção:** Andy Wachowski  
Larry Wachowski

**Autoria:** Andy Wachowski  
Larry Wachowski

**Lançamento:** 21 de maio de 1999 (no Brasil)

**Gênero:** Ação, Aventura, Ficção Científica

**Premiação:** Vencedor de 4 Oscars

**País:** EUA e Austrália

**Formato Tela:** 2.35 : 1

**Som:** DTS | Dolby Digital | SDDS

[Elenco](#) [Fotos](#) [Vídeos](#) [Sinopse](#)

Figura 3.13 – Informações adicionais dos repositórios de conteúdo da Web Semântica

## Capítulo 4

# Aplicação do processo de alinhamento e extensão de metadados da TV digital

A Figura 4.1 apresenta uma visão geral do processo de alinhamento e extensão de metadados na TV digital. O processo consiste basicamente na integração das informações de metadados dos objetos multimídia existentes no contexto da TV digital com informações de metadados de repositórios existentes na Web Semântica, de forma que os metadados integrados relacionados aos objetos multimídia sejam disponibilizados para usuário final em um formato comum baseado em OWL.

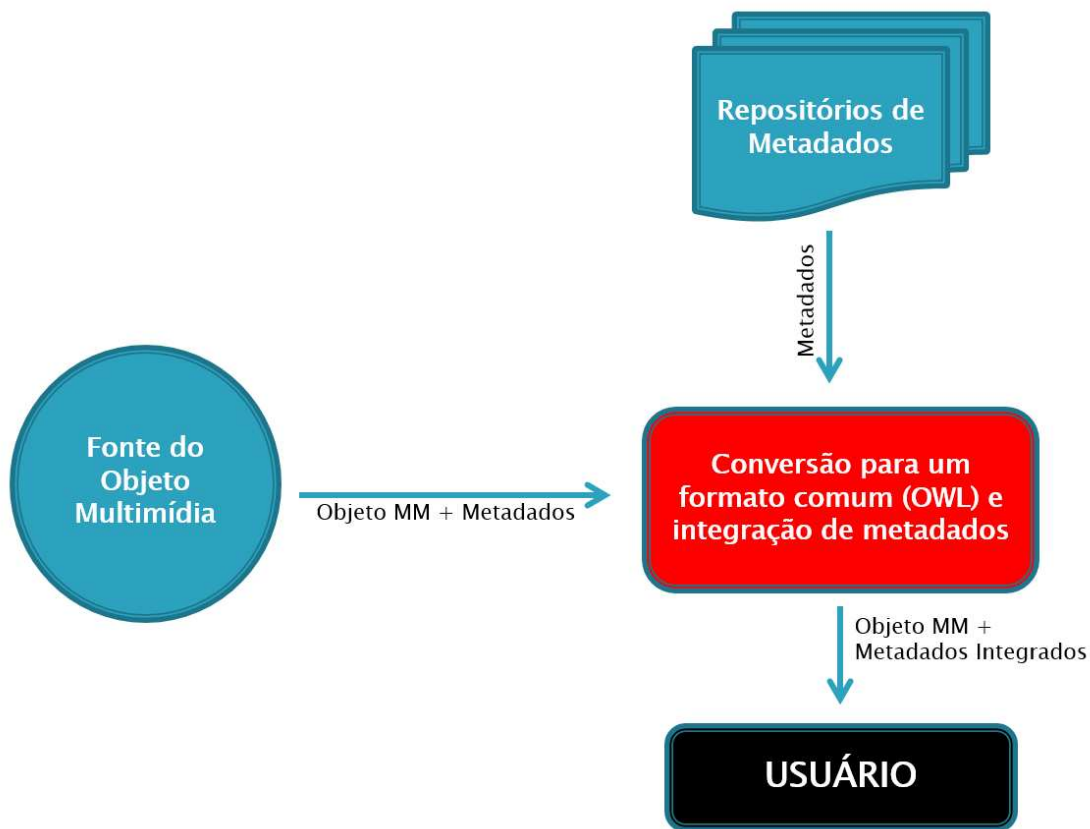


Figura 4.1 – Visão geral do processo

Neste capítulo será apresentada a aplicação do processo de alinhamento e extensão de metadados da TV digital. Primeiramente será feita a representação das informações de metadados voltadas para a classificação de conteúdos já existentes no sistema de TV digital usado no Brasil em redes terrestres (ISDB-T) e uma conversão destas informações para o formato da ontologia OWLTV definida no Capítulo 3.

A ontologia OWLTV será então estendida para fazer uma integração com as informações de metadados utilizadas pelos repositórios de conteúdo das bases de conhecimento LinkedMDB, Freebase e DBpedia, visando possibilitar uma extensão dos sistemas de TV digital com as informações de metadados de bases de conhecimento da Web Semântica.

A seguir será descrito detalhadamente o processo de alinhamento e extensão de metadados em sistemas de TV digital.

## 4.1 Especificações de metadados dos sistemas de TV digital

Antes de ser iniciada a aplicação do processo de alinhamento de metadados, é necessário realizar um levantamento das informações de metadados voltadas para classificação de conteúdos já existentes no sistema de TV digital terrestre (ISDB-T) usado no Brasil. Neste levantamento pode ser constatada a existência de diversas similaridades nas informações de metadados entre o sistema ISDB-T e o padrão XMLTV, de forma a facilitar o processo de conversão destas informações para o formato da ontologia OWLTV.

As especificações de metadados do sistema brasileiro de TV digital terrestre, o ISDB-T (ABNT, 2007), bem como do sistema de TV digital usado para transmissão digital aberta via satélite no Brasil, o DVB-S (ETSI, 2003), definem uma série de tabelas que são transmitidas juntamente com o conteúdo de áudio e vídeo do canal de televisão. Entre estas tabelas, existe uma tabela, a EIT (*Event Information Table*), que contém informações voltadas para a classificação de conteúdos dos eventos da programação que está sendo transmitida.

Entre os metadados da tabela EIT voltados para a classificação de conteúdos, alguns exemplos que podem ser convertidos para o formato da ontologia OWLTV são:

- Horário de início: este é o campo chamado *start\_time* da tabela EIT.
- Duração do conteúdo: este é o campo chamado *duration* da tabela EIT.
- Título do conteúdo: este campo, chamado de *event\_name*, está contido em um descritor da tabela EIT chamado *short\_event\_descriptor*.
- Descrição do conteúdo: existem dois campos da EIT que contêm informações relacionadas à descrição do conteúdo. O primeiro é o campo que contém uma descrição curta do programa, chamado de *text*, e também contido no descritor *short\_event\_descriptor*. O segundo campo contém uma descrição mais longa do programa, também chamado de *text*, mas contido no descritor *extended\_event\_descriptor* da tabela EIT.
- Classificação parental: este campo, chamado de *rating*, está contido em um descritor da tabela EIT chamado *parental\_rating\_descriptor*.



Além dos exemplos de metadados acima, existem outras informações que são transmitidas na tabela EIT que podem ser convertidas para o formato da ontologia OWLTV, como os campos *item* e *item\_descriptor*, contidos no descritor *extended\_event\_descriptor*, que armazenam informações adicionais e palavras chaves de cada programa, e também o campo *content\_nibble\_level1*, contido em um descritor da tabela EIT chamado *content\_descriptor*, que armazena o gênero do programa.

## 4.2 Bases de conhecimento da Web Semântica

Para a aplicação do processo de alinhamento e extensão dos padrões de TV digital com domínios de conhecimento específicos da Web Semântica, assim como foi feito com os metadados do sistema de TV digital na seção anterior, deve ser realizado um levantamento das informações de metadados voltadas para classificação de conteúdos utilizadas pelas bases de conhecimento a serem integradas, que no caso de uso apresentado neste capítulo será constituído pelos repositórios LinkedMDB, Freebase e DBpedia.

Neste levantamento pode ser constatada a existência de diversas similaridades nas informações de metadados entre as bases de conhecimento LinkedMDB, Freebase e DBpedia com as informações de metadados dos sistemas de TV digital, de forma a facilitar o processo de alinhamento com a ontologia OWLTV.

Entre as entidades da ontologia LinkedMDB (LINKEDMDB, 2012) que podem ser integradas com a ontologia OWLTV, a seguir são relacionadas algumas:

- Título do conteúdo: a base de conhecimento LinkedMDB possui uma entidade que é usada para armazenar títulos de conteúdos, a entidade *dc:title*.
- Classificação parental: a base de conhecimento LinkedMDB possui uma entidade que é usada para armazenar a classificação parental dos conteúdos, a entidade *movie:content\_rating*.
- Gênero: esta informação é armazenada na entidade *movie:film\_genre*.
- Ator: esta informação é armazenada na entidade *movie:actor*.
- Diretor: esta informação é armazenada na entidade *movie:director*.
- Editor: esta informação é armazenada na entidade *movie:editor*.
- Produtor: esta informação é armazenada na entidade *movie:producer*.
- Escritor: esta informação é armazenada na entidade *movie:writer*.

Da mesma forma que na base de conhecimento LinkedMDB, na ontologia Freebase (FREEBASE, 2012) também existem algumas entidades que podem ser integradas com a ontologia OWLTV, por exemplo:

- Título do conteúdo: a base de conhecimento Freebase possui uma entidade que é usada para armazenar títulos de conteúdos, a entidade *type/object/name*.
- Classificação parental: a base de conhecimento Freebase possui uma entidade que é usada para armazenar a classificação parental dos conteúdos, a entidade *film/film/rating*.
- Gênero: esta informação é armazenada na entidade *film/film/genre*.
- Ator: esta informação é armazenada na entidade */film/actor*.
- Diretor: esta informação é armazenada na entidade *film/director*.
- Editor: esta informação é armazenada na entidade */film/editor*.
- Produtor: esta informação é armazenada na entidade */film/producer*.
- Autor: esta informação é armazenada na entidade */film/writer*.

Fazendo a mesma verificação na ontologia DBpedia (DBPEDIA, 2012) também são encontradas entidades que podem ser integradas com a ontologia OWLTV, como as duas entidades a seguir:

- Título do conteúdo: a base de conhecimento DBpedia possui três entidades que podem ser usadas para armazenar o título do conteúdo, as entidades *dbpprop:title*, *foaf:name* e *rdfs:label*.
- Descrição do conteúdo: a base de conhecimento DBpedia possui duas entidades que são usadas para armazenar a descrição do conteúdo em diversos idiomas, as entidades *dbpedia-owl:abstract* e *rdfs:comment*.

Além das entidades apresentadas acima, existem outras que podem ser integradas com a ontologia OWLTV, como as entidades voltadas para armazenamento de informações relacionadas ao país de origem, data de lançamento e endereço web (URL) dos filmes e programas de televisão.

### 4.3 Construção da ontologia para alinhamento de metadados

Nesta seção é descrita detalhadamente a aplicação do processo utilizado para a construção da ontologia para alinhamento de metadados em sistemas de TV digital.

O objetivo da criação desta ontologia no caso de uso apresentado neste capítulo é conceber um modelo conceitual para alinhar e estender as informações de metadados do sistema de TV digital terrestre (ISDB-T) com as informações de metadados das bases de conhecimento LinkedMDB, Freebase e DBpedia.

Neste caso, a ontologia OWLTV é gerada com base nas informações de metadados para classificação de conteúdos já existentes nos sistemas de TV digital (tabela EIT) e também nas informações em formato XMLTV disponíveis para download no site Revista Eletrônica (REVISTA ELETRÔNICA, 2012), onde foi constada a existência de informações mais completas para os programas como ator, diretor, país, gênero, data de criação, entre outras.

Estas informações de metadados são mapeadas, agrupadas e integradas na ontologia OWLTV em uma hierarquia de classes, propriedades, restrições e instâncias. Para o desenvolvimento do sistema responsável pela geração da ontologia OWLTV, que é realizada em tempo real pelo receptor de TV digital, foi utilizada a linguagem Java e o padrão Ginga-J (KULESZA *et al.*, 2010) que é suportado pelos receptores de TV digital compatíveis com a norma brasileira de televisão digital terrestre (ABNT, 2010).

A seguir serão descritas as etapas do processo de geração da ontologia OWLTV que é realizado pelo sistema desenvolvido em Ginga-J com execução em tempo real nos receptores de TV digital.

#### **4.3.1 Etapa I – Obtenção de um arquivo base em formato XMLTV**

Para iniciar a geração da ontologia OWLTV inicialmente pode ser realizado o download de um arquivo em formato XMLTV contendo a informação atualizada da programação do canal de TV digital que está sintonizado no receptor, caso este arquivo esteja disponível para download na Internet. Na prova de conceito implementada no presente trabalho, este download é feito a partir do site Revista Eletrônica (REVISTA ELETRÔNICA, 2012), que mantém listagens atualizadas das programações dos canais de TV do Brasil em formato XMLTV. Além do site Revista Eletrônica, qualquer outra fonte de informação de metadados da programação dos canais de TV do Brasil que disponibilize suas informações no formato XMLTV poderia ser utilizada como base para geração da ontologia OWLTV.

A Figura 4.2 apresenta um fragmento de arquivo XMLTV publicado pelo site Revista Eletrônica.

```

<?xml version="1.0" encoding="ISO-8859-1"?>
<tv generator-info-name="Revista Eletronica - Unidade Lorenz Ltda" generator-info-
url="http://www.revistaeletronica.com.br">
<channel id="GRC">
<display-name lang="pt">EPTV Campinas</display-name>
<icon src="grc.gif" />
</channel>
<programme start="20120621144000 -0300" stop="20120621155500 -0300" channel="GRC" program_id="0000109098">
<title lang="pt">Chocolate com Pimenta</title>
<desc>Sinopse Não Disponível</desc>
<category lang="pt">Variedades</category>
<category lang="pt">Novela</category>
<rating system="Advisory">
<value>Programa livre para todas as idades</value>
</rating>
</programme>
<programme start="20120621155500 -0300" stop="20120621174500 -0300" channel="GRC" program_id="0000092443">
<title lang="pt">Táxi</title>
<title lang="en">Taxi</title>
<desc>Atrapalhado policial quer recuperar seu prestígio. Com a ajuda da taxista mais rápida de New York, ele
tenta deter a gangue liderada pela bela e perigosa Vanessa.</desc>
<credits>
<director>Tim Story</director>
<actor>Queen Latifah</actor>
<actor>Jimmy Fallon</actor>
<actor>Henry Simmons</actor>
<actor>Jennifer Esposito</actor>
<actor>Gisele Bündchen</actor>
<actor>Magali Amadei</actor>
<actor>Ann-Margret</actor>
<actor>Boris McGiver</actor>
<actor>Bryna Weiss</actor>
</credits>
<date>2004</date>
<category lang="pt">Filme</category>
<category lang="pt">Comédia</category>
<country>SUA/França</country>
<video>
<colour>yes</colour>
</video>
<rating system="Advisory">
<value>Programa impróprio para menores de 14 anos</value>
</rating>
<star-rating>
<value>2/5</value>
</star-rating>
</programme>
</tv>

```

Figura 4.2 – Fragmento de arquivo XMLTV publicado pelo site Revista Eletrônica

Caso um arquivo base em formato XMLTV não esteja disponível para download na Internet, para atendimento das necessidades da Etapa I do processo de geração da ontologia

OWLTV poderia ser utilizado um modelo padrão de arquivo XMLTV que ficaria armazenado localmente na memória do receptor como, por exemplo, o modelo exibido na Figura 4.3.

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<channel id="#" />
<programme start="#" stop="#" channel="#" program_id="#">
<title lang="pt">#</title>
<desc>#</desc>
<rating system="#">
<value>#</value>
</rating>
</programme>
</tv>
```

**Figura 4.3 – Modelo de arquivo XMLTV a ser armazenado localmente na memória do receptor de TV digital**

O modelo padrão de arquivo XMLTV exibido na Figura 4.3 possui alguns elementos e atributos com o valor “#”. Este valor deve ser substituído pelo valor do metadado correspondente transmitido na tabela EIT no processo de conversão a ser realizado na Etapa II a ser descrito na seção seguinte.

### 4.3.2 Etapa II – Conversão dos metadados da tabela EIT para XMLTV

A Etapa II do processo de geração da ontologia OWLTV é a conversão dos metadados transmitidos no sinal da TV digital terrestre (ISDB-T), em particular as informações voltadas para a classificação de conteúdos disponíveis na tabela EIT, para o formato XMLTV.

Para fazer esta conversão, o sistema desenvolvido em Gingga-J utiliza-se de métodos e interfaces do pacote *javax.tv.service.guide* da API JavaTV (JavaTV, 2012) para ter acesso as informações dos metadados EIT e convertê-los para os elementos XMLTV correspondentes, conforme exibido na Tabela 4.1.

Metadado EIT	Método e Interface do pacote <i>javax.tv.service.guide</i>	Elemento XMLTV
Horário de início: <i>start_time</i>	<i>ProgramEvent.getStartTime()</i>	Atributo <i>start</i> do elemento <i>programme</i>
Duração: <i>duration</i>	<i>ProgramEvent.getDuration()</i>	Elemento <i>length</i>
Título: <i>event_name</i>	<i>ProgramEvent.getName()</i>	Elemento <i>title</i>
Descrição: <i>text</i>	<i>ProgramEvent.retrieveDescription()</i>	Elemento <i>desc</i>
Classificação parental: <i>rating</i>	<i>ProgramEvent.getRating()</i>	Elemento <i>value</i> aninhado ao elemento <i>rating</i>

**Tabela 4.1 – Mapeamento entre metadados EIT e XMLTV utilizando o pacote *javax.tv.service.guide***

A Figura 4.4 exhibe as informações correspondentes aos programas “Chocolate com Pimenta” e “Táxi” do arquivo XMLTV exibido na Figura 4.2, só que extraídas conforme transmitidas na tabela EIT do sinal da TV digital transmitido pela emissora EPTV no dia 21 de Junho de 2012. Para a extração das informações foi utilizado o pacote *javax.tv.service.guide*.

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<channel id="1" />
<programme start="20120621144000 -0300" stop="20120621155000 -0300" channel="1" program_id="1">
<title lang="pt">VALE A PENA VER DE NOVO</title>
<desc>CHOCOLATE COM PIMENTA</desc>
<rating system="Advisory">
<value>1</value>
</rating>
</programme>
<programme start="20120621155000 -0300" stop="20120621174000 -0300" channel="1" program_id="2">
<title lang="pt">SESSÃO DA TARDE - HDTV</title>
<desc>TAXI - Elenco: Queen Latifah, Jimmy Fallon, Henry Simmons, Jennifer Esposito, Gisele Bundchen, Ana
Cristina De Oliveira - Comédia - HDTV</desc>
<rating system="Advisory">
<value>10</value>
</rating>
</programme>
</tv>
```

Figura 4.4 – Arquivo XMLTV gerado a partir dos metadados da tabela EIT

O valores dos atributos *id* do elemento *channel* e *program\_id* do elemento *programme* devem ser preenchidos de forma numérica sequencial, iniciando do valor 1 e sendo incrementados automaticamente conforme o número de elementos.

O valor do atributo *stop* do elemento *programme*, que representa o horário de término de transmissão do programa, é obtido pelo sistema desenvolvido em Ginga-J através do horário de início – método *ProgramEvent.getStartTime()*, acrescido do tempo de duração do programa – método *ProgramEvent.getDuration()*.

### 4.3.3 Etapa III – Aplicação da metodologia proposta por Thuy *et al.* (2009) para geração da ontologia OWLTV a partir do arquivo XMLTV

A terceira etapa do processo de geração da ontologia OWLTV é a aplicação da metodologia proposta por Thuy *et al.* (2009) descrita no Capítulo 3 para conversão de um esquema XML para uma ontologia baseada em OWL. A Figura 4.5 apresenta a ontologia OWLTV resultante da aplicação da metodologia de Thuy *et al.* (2009) sobre o fragmento de arquivo XMLTV obtido do site Revista Eletrônica exibido na Figura 4.2.

```
<?xml version="1.0"?>
<rdf:RDF
xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-s#"
xmlns:owl="http://www.w3.org/2002/07/owl#"
xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">
<tv rdf:ID="tv_1">
<has_channel>
<channel rdf:ID="channel_1">
<channel_id rdf:datatype="xsd:string">GRC</id>
<has_display_name>
```

```

    <display-name rdf:ID="display-name_1">
      <display-name_value rdf:datatype="&xsd:string">EPTV Campinas</display-name_value>
      <display-name_lang rdf:datatype="&xsd:string">pt</lang>
    </display-name>
  </has_display_name>
  <has_icon>
    <icon rdf:ID="icon_1">
      <icon_src rdf:datatype="&xsd:string">grc.gif</src>
    </icon>
  </has_icon>
</channel>
</has_channel>
<has_programme>
<programme rdf:ID="programme_1">
  <programme_start rdf:datatype="&xsd:string">20120621144000 -0300</start>
  <programme_stop rdf:datatype="&xsd:string">20120621155500 -0300</stop>
  <programme_channel rdf:datatype="&xsd:string">GRC</channel>
  <programme_program_id rdf:datatype="&xsd:string">0000109098</program_id>
  <has_title>
    <title rdf:ID="title_1">
      <title_value rdf:datatype="&xsd:string">Chocolate com Pimenta</title_value>
      <title_lang rdf:datatype="&xsd:string">pt</title_lang>
    </title>
  </has_title>
  <has_desc>
    <desc rdf:ID="desc_1">
      <desc_value rdf:datatype="&xsd:string">Sinopse Não Disponível</desc_value>
    </desc>
  </has_desc>
  <has_category>
    <category rdf:ID="category_1">
      <category_value rdf:datatype="&xsd:string">Variedades</category_value>
      <category_lang rdf:datatype="&xsd:string">pt</category_lang>
    </category>
    <category rdf:ID="category_2">
      <category_value rdf:datatype="&xsd:string">Novela</category_value>
      <category_lang rdf:datatype="&xsd:string">pt</category_lang>
    </category>
  </has_category>
  <has_rating>
    <rating rdf:ID="rating_1">
      <rating_system rdf:datatype="&xsd:string">Advisory</rating_system>
      <has_rating_value>
        <rating_value rdf:ID="rating_value_1">
          <rating_value_value rdf:datatype="&xsd:string">Programa livre para todas as
idades</rating_value_value>
        </rating_value>
      </has_rating_value>
    </rating>
  </has_rating>
</programme>
<programme rdf:ID="programme_2">
  <programme_start rdf:datatype="&xsd:string">20120621155500 -0300</start>
  <programme_stop rdf:datatype="&xsd:string">20120621174500 -0300</stop>
  <programme_channel rdf:datatype="&xsd:string">GRC</channel>
  <programme_program_id rdf:datatype="&xsd:string">0000092443</program_id>
  <has_title>
    <title rdf:ID="title_2">

```

```

<title_value rdf:datatype="&xsd:string">Taxi</title_value>
<title_lang rdf:datatype="&xsd:string">pt</title_lang>
</title>
<title rdf:ID="title_2">
  <title_value rdf:datatype="&xsd:string">Taxi</title_value>
  <title_lang rdf:datatype="&xsd:string">en</title_lang>
</title>
</has_title>
<has_desc>
  <desc rdf:ID="desc_2">
    <desc_value rdf:datatype="&xsd:string">Atrapalhado policial quer recuperar seu prestígio. Com a ajuda da
    taxista mais rápida de New York, ele tenta deter a gangue liderada pela bela e perigosa Vanessa.</desc_value>
  </desc>
</has_desc>
<has_credits>
  <credits rdf:ID="credits_1">
    <has_director>
      <director rdf:ID="director_1">
        <director_value rdf:datatype="&xsd:string">Tim Story</director_value>
      </director>
    </has_director>
    <has_actor>
      <actor rdf:ID="actor_1">
        <actor_value rdf:datatype="&xsd:string">Queen Latifah</actor_value>
      </actor>
      <actor rdf:ID="actor_2">
        <actor_value rdf:datatype="&xsd:string">Jimmy Fallon</actor_value>
      </actor>
      <actor rdf:ID="actor_3">
        <actor_value rdf:datatype="&xsd:string">Henry Simmons</actor_value>
      </actor>
      <actor rdf:ID="actor_4">
        <actor_value rdf:datatype="&xsd:string">Jennifer Esposito</actor_value>
      </actor>
      <actor rdf:ID="actor_5">
        <actor_value rdf:datatype="&xsd:string">Gisele Bündchen</actor_value>
      </actor>
      <actor rdf:ID="actor_6">
        <actor_value rdf:datatype="&xsd:string">Magali Amadei</actor_value>
      </actor>
      <actor rdf:ID="actor_7">
        <actor_value rdf:datatype="&xsd:string">Ann-Margret</actor_value>
      </actor>
      <actor rdf:ID="actor_8">
        <actor_value rdf:datatype="&xsd:string">Boris McGiver</actor_value>
      </actor>
      <actor rdf:ID="actor_9">
        <actor_value rdf:datatype="&xsd:string">Bryna Weiss</actor_value>
      </actor>
    </has_actor>
  </credits>
</has_credits>
<has_date>
  <date rdf:ID="date_1">
    <date_value rdf:datatype="&xsd:string">2004</date_value>
  </date>
</has_date>

```



```

</date>
</has_date>
<has_category>
  <category rdf:ID="category_3">
    <category_value rdf:datatype="&xsd:string">Filme</category_value>
    <category_lang rdf:datatype="&xsd:string">pt</category_lang>
  </category>
  <category rdf:ID="category_4">
    <category_value rdf:datatype="&xsd:string">Comédia</category_value>
    <category_lang rdf:datatype="&xsd:string">pt</category_lang>
  </category>
</has_category>
<has_country>
  <country rdf:ID="country_1">
    <country_value rdf:datatype="&xsd:string">EUA/França</country_value>
  </country>
</has_country>
<has_video>
  <video rdf:ID="video_1">
    <has_colour>
      <colour rdf:ID="colour_1">
        <colour_value rdf:datatype="&xsd:string">yes</colour_value>
      </colour>
    </has_colour>
  </video>
</has_video>
<has_rating>
  <rating rdf:ID="rating_2">
    <rating_system rdf:datatype="&xsd:string">Advisory</rating_system>
    <has_rating_value>
      <rating_value rdf:ID="rating_value_2">
        <rating_value_value rdf:datatype="&xsd:string">Programa impróprio para menores de 14
anos</rating_value_value>
      </rating_value>
    </has_rating_value>
  </rating>
</has_rating>
<has_star-rating>
  <star-rating rdf:ID="star-rating_1">
    <has_star-rating_value>
      <star-rating_value rdf:ID="star-rating_value_1">
        <star-rating_value_value rdf:datatype="&xsd:string">2/5</star-rating_value_value>
      </star-rating_value>
    </has_star-rating_value>
  </star-rating>
</has_star-rating>
</programme>
</has_programme>
</tv>

```

**Figura 4.5 – Ontologia OWLTV obtida a partir do arquivo XMLTV da Figura 4.2**

A Figura 4.6 apresenta a ontologia OWLTV resultante da aplicação da metodologia de Thuy *et al.* (2009) sobre o arquivo XMLTV obtido a partir da tabela EIT exibido na Figura 4.4.

```

<?xml version="1.0"?>
<rdf:RDF
xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-s#"
xmlns:owl="http://www.w3.org/2002/07/owl#"
xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">

<tv rdf:ID="tv_1">
<has_channel>
<channel rdf:ID="channel_1">
  <channel_id rdf:datatype="&xsd:string">1</id>
</channel>
</has_channel>
<has_programme>
<programme rdf:ID="programme_1">
  <programme_start rdf:datatype="&xsd:string">20120621144000 -0300</start>
  <programme_stop rdf:datatype="&xsd:string">20120621155000 -0300</stop>
  <programme_channel rdf:datatype="&xsd:string">1</channel>
  <programme_program_id rdf:datatype="&xsd:string">1</program_id>
  <has_title>
    <title rdf:ID="title_1">
      <title_value rdf:datatype="&xsd:string">VALE A PENA VER DE NOVO</title_value>
      <title_lang rdf:datatype="&xsd:string">pt</title_lang>
    </title>
  </has_title>
  <has_desc>
    <desc rdf:ID="desc_1">
      <desc_value rdf:datatype="&xsd:string">CHOCOLATE COM PIMENTA</desc_value>
    </desc>
  </has_desc>
  <has_rating>
    <rating rdf:ID="rating_1">
      <rating_system rdf:datatype="&xsd:string">Advisory</rating_system>
      <has_rating_value>
        <rating_value rdf:ID="rating_value_1">
          <rating_value_value rdf:datatype="&xsd:string">1</rating_value_value>
        </rating_value>
      </has_rating_value>
    </rating>
  </has_rating>
</programme>
<programme rdf:ID="programme_2">
  <programme_start rdf:datatype="&xsd:string">20120621155000 -0300</start>
  <programme_stop rdf:datatype="&xsd:string">20120621174000 -0300</stop>
  <programme_channel rdf:datatype="&xsd:string">1</channel>
  <programme_program_id rdf:datatype="&xsd:string">2</program_id>
  <has_title>
    <title rdf:ID="title_2">
      <title_value rdf:datatype="&xsd:string">SESSÃO DA TARDE - HDTV</title_value>
      <title_lang rdf:datatype="&xsd:string">pt</title_lang>
    </title>
  </has_title>
  <has_desc>
    <desc rdf:ID="desc_2">
      <desc value rdf:datatype="&xsd:string">TAXI - Elenco: Queen Latifah, Jimmy Fallon, Henry Simmons,
Jennifer Esposito, Gisele Bundchen, Ana Cristina De Oliveira - Comédia - HDTV</desc_value>
    </desc>
  </has_desc>
</programme>

```

```

</has_desc>
<has_rating>
  <rating rdf:ID="rating_2">
    <rating_system rdf:datatype="&xsd:string">Advisory</rating_system>
    <has_rating_value>
      <rating_value rdf:ID="rating_value_2">
        <rating_value_value rdf:datatype="&xsd:string">10</rating_value_value>
      </rating_value>
    </has_rating_value>
  </rating>
</has_rating>
</programme>
</has_programme>
</tv>

```

Figura 4.6 – Ontologia OWLTV obtida a partir do arquivo XMLTV da Figura 4.3

Analisando os arquivos apresentados nas Figuras 4.5 e 4.6, obtidos respectivamente através do site Revista Eletrônica e da tabela EIT do sinal da TV digital, verificamos que embora eles se refiram aos mesmos programas que foram transmitidos pela emissora EPTV no dia 21 de junho de 2012, existem diferenças entre os arquivos nas propriedades de dados que se referem ao mesmo programa conforme exibido na Tabela 4.2.

Propriedade de Dados	Arquivo OWLTV Revista Eletrônica	Arquivo OWLTV Tabela EIT
<programme_start rdf:datatype="&xsd:string">	20120621144000 -0300	20120621144000 -0300
<programme_stop rdf:datatype="&xsd:string">	20120621155500 -0300	20120621155000 -0300
<title_value rdf:datatype="&xsd:string">	Chocolate com Pimenta	VALE A PENA VER DE NOVO
<desc_value rdf:datatype="&xsd:string">	Sinopse Não Disponível	CHOCOLATE COM PIMENTA
<rating_value_value rdf:datatype="&xsd:string">	Programa livre para todas as idades	L
<programme_start rdf:datatype="&xsd:string">	20120621155500 -0300	20120621155000 -0300
<programme_stop rdf:datatype="&xsd:string">	20120621174500 -0300	20120621174000 -0300
<title_value rdf:datatype="&xsd:string">	Táxi	SESSÃO DA TARDE - HDTV
<desc_value rdf:datatype="&xsd:string">	Atrapalhado policial quer recuperar seu prestígio. Com a ajuda da taxista mais rápida de New York, ele tenta deter a gangue liderada pela bela e perigosa Vanessa. - <a href="http://www.revistaeletronica.com.br">www.revistaeletronica.com.br</a>	TAXI - Elenco: Queen Latifah, Jimmy Fallon, Henry Simmons, Jennifer Esposito, Gisele Bündchen, Ana Cristina De Oliveira - Comédia - HDTV
<rating_value_value rdf:datatype="&xsd:string">	Programa impróprio para menores de 14 anos	10

Tabela 4.2 – Comparativo dos arquivos OWLTV apresentados nas Figuras 4.4 e 4.5

Como pode ser observado na Tabela 4.2, as propriedades de dados se referem aos mesmos programas, contudo possuem valores diferentes conforme a fonte de informação. Por exemplo, apesar do horário de início do programa “Chocolate com Pimenta” ser o mesmo em ambos os arquivos (21 de junho de 2012 às 14h40 UTC -3), o horário de término é diferente em cinco minutos: segundo o site Revista Eletrônica o programa termina às 15h55, enquanto que na tabela EIT o horário de término indicado é 15h50.

Além das propriedades de dados relacionadas com horário de início e término dos programas, diversas outras propriedades como título, descrição e classificação indicativa dos programas são diferentes conforme a fonte de informação utilizada para geração do arquivo OWLTV. Devido a estas diferenças, o processo utilizado para a construção da ontologia para integração de metadados em sistemas de TV digital com as bases de conhecimento da Web Semântica deve utilizar como ponto de partida para esta integração uma única fonte de informação.

No caso da prova de conceito implementada neste trabalho, como a informação dos programas publicada pelo site Revista Eletrônica é mais completa que a informação transmitida na tabela EIT do sistema de TV digital, esta fonte de informação será utilizada como ponto de partida para a integração com as bases de conhecimento da Web Semântica, a ser realizada a partir da Etapa IV do processo de alinhamento descrita na próxima seção. Como tanto as informações de metadados do site Revista Eletrônica, como as informações de metadados da tabela EIT foram convertidas para o formato da ontologia OWLTV, qualquer uma destas fontes de informações de metadados poderia ser utilizada como ponto de partida para o processo de alinhamento a ser realizado a partir da Etapa IV.

#### 4.3.4 Etapa IV – Inclusão de classes “*sameAsQuery*”

Conforme apresentado no Capítulo 3 o relacionamento entre duas ou mais entidades de ontologias distintas é realizado pela inclusão de classes descritas por *owl:Class* contendo as consultas a serem realizadas nas bases de conhecimento a serem relacionadas com a ontologia OWLTV.

Na Figura 4.7 são adicionadas à ontologia OWLTV exibida na Figura 4.5 novas instâncias da classe *sameAsQuery* que relacionam as entidades *programme* da ontologia OWLTV identificadas por “programme\_1” e “programme\_2” com as ontologias Freebase, DBpedia e LinkedMDB, cujas propriedades *linked\_ontology* são respectivamente iguais a “Freebase”, “DBpedia” e “LinkedMDB”.

```
<?xml version="1.0"?>
...
<programme rdf:ID="programme_1">
...
  <sameAsQuery rdf:ID="sameAsQuery_1">
    <owltv_entity rdf:datatype="&xsd:string">programme_1</owltv_entity>
    <linked_ontology rdf:datatype="&xsd:string">Freebase</linked_ontology>
    <query_code rdf:datatype="&xsd:string">
      [f
        "id": null,
        "name":["Chocolate com Pimenta"],
        "type":["/tv/tv_program", "/film/film"]
```

```

    ]]
  </query_code>
</sameAsQuery>
<sameAsQuery rdf:ID="sameAsQuery_2">
  <owltv_entity rdf:datatype="&xsd:string">programme_1</owltv_entity>
  <linked_ontology rdf:datatype="&xsd:string">DBpedia</linked_ontology>
  <query_code rdf:datatype="&xsd:string">
    PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
    SELECT ?entity
    WHERE { ?entity rdfs:label "Chocolate com Pimenta"@pt }
  </query_code>
</sameAsQuery>
<sameAsQuery rdf:ID="sameAsQuery_3">
  <owltv_entity rdf:datatype="&xsd:string">programme_1</owltv_entity>
  <linked_ontology rdf:datatype="&xsd:string">LinkedMDB</linked_ontology>
  <query_code rdf:datatype="&xsd:string">
    PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
    SELECT ?entity
    WHERE { ?entity rdfs:label "Chocolate com Pimenta" }
  </query_code>
</sameAsQuery>
...
</programme>
<programme rdf:ID="programme_2">
...
<sameAsQuery rdf:ID="sameAsQuery_4">
  <owltv_entity rdf:datatype="&xsd:string">programme_2</owltv_entity>
  <linked_ontology rdf:datatype="&xsd:string">Freebase</linked_ontology>
  <query_code rdf:datatype="&xsd:string">
    [[
      "id": null,
      "name":["Táxi", "Taxi"],
      "type":["/tv/tv_program", "/film/film"]
    ]]
  </query_code>
</sameAsQuery>
<sameAsQuery rdf:ID="sameAsQuery_5">
  <owltv_entity rdf:datatype="&xsd:string">programme_2</owltv_entity>
  <linked_ontology rdf:datatype="&xsd:string">DBpedia</linked_ontology>
  <query_code rdf:datatype="&xsd:string">
    PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
    PREFIX dbpprop: <http://dbpedia.org/property/>
    PREFIX dbpedia-owl: <http://dbpedia.org/ontology/>
    SELECT ?entity
    WHERE { ?entity dbpprop:title ?title ;
             rdf:type dbpedia-owl:Film .
             FILTER (?title = "Táxi"@pt || ?title = "Taxi"@en) }
  </query_code>
</sameAsQuery>

```

```

</query_code>
</sameAsQuery>
<sameAsQuery rdf:ID="sameAsQuery_6">
  <owltv_entity rdf:datatype="&xsd:string">programme_2</owltv_entity>
  <linked_ontology rdf:datatype="&xsd:string">linkedMDB</linked_ontology>
  <query_code rdf:datatype="&xsd:string">
    PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
    SELECT ?entity
    WHERE { ?entity rdfs:label ?title .
            FILTER (?title = "Táxi" || ?title = "Taxi") }
  </query_code>
</sameAsQuery>
...
</programme>
</has_programme>
</tv>

```

Figura 4.7 – Inclusão das instâncias da classe *sameAsQuery* na ontologia da Figura 4.5

As instâncias da classe *sameAsQuery* são geradas automaticamente e em tempo real pelo sistema desenvolvido em Ginga-J que possui funções de software específicas para cada base de conhecimento da Web Semântica a ser integrada. Por exemplo, para a instância da classe *sameAsQuery* cuja propriedade *linked\_ontology* é igual a “Freebase” é gerado um código em linguagem MQL (MQL, 2012). Já para as instâncias da classe *sameAsQuery* cujas propriedades *linked\_ontology* são iguais a “DBpedia” e “LinkedMDB” os códigos gerados são escritos na linguagem SPARQL (SPARQL, 2012) com pequenas variações específicas de cada base de conhecimento.

O código das consultas geradas também varia conforme a quantidade de informações disponíveis na entidade para a pesquisa. Por exemplo, a entidade identificada por “programme\_1” que é relacionada ao programa “Chocolate com Pimenta” possui apenas uma informação de título para a pesquisa nas bases de conhecimento Freebase, DBpedia e LinkedMDB. Já a entidade identificada por “programme\_2” que é relacionada ao filme “Táxi”, possui duas informações de título para pesquisa, que são os títulos do filme nos idiomas português e inglês. Devido a esta variação de quantidade de informações disponíveis para pesquisa, as funções de software do sistema desenvolvimento em Ginga-J responsáveis pela geração do código da consulta em cada base de conhecimento se encarregam de adaptar o código das consultas conforme a quantidade de informações disponíveis nas respectivas entidades.

### 4.3.5 Etapa V – Obtenção das propriedades *owl:sameAs*

A última etapa do processo de geração da ontologia OWLTV é a execução das consultas das instâncias da classe *sameAsQuery* para obtenção das propriedades *owl:sameAs* que permitirão estender as informações da ontologia OWLTV com informações de entidades correspondentes nas bases de conhecimento Freebase, DBpedia e LinkedMDB.

A execução das consultas é realizada pelas funções de software específicas de cada base de conhecimento a ser integrada no sistema desenvolvido em Ginga-J apenas quando solicitado pelo usuário, conforme será apresentado no Capítulo 5. Na Figura 4.8 são exibidas as propriedades *owl:sameAs* obtidas a partir da execução das consultas das instâncias da classe “*sameAsQuery*” da Figura 4.7.

```

<?xml version="1.0"?>
...
<programme rdf:ID="programme_1">
...
  <sameAsQuery rdf:ID="sameAsQuery_1">
...
  </sameAsQuery>
  <owl:sameAs rdf:resource="http://rdf.freebase.com/ns/en.chocolate_com_pimenta"/>

  <sameAsQuery rdf:ID="sameAsQuery_2">
...
  </sameAsQuery>
  <owl:sameAs rdf:resource="http://dbpedia.org/resource/Chocolate_com_Pimenta"/>

  <sameAsQuery rdf:ID="sameAsQuery_3">
...
  </sameAsQuery>
...
</programme>

<programme rdf:ID="programme_2">
...
  <sameAsQuery rdf:ID="sameAsQuery_4">
...
  </sameAsQuery>
  <owl:sameAs rdf:resource="http://rdf.freebase.com/ns/en.taxi"/>
  <owl:sameAs rdf:resource="http://rdf.freebase.com/ns/en.taxi_1998"/>
  <owl:sameAs rdf:resource="http://rdf.freebase.com/ns/en.taxi_2004"/>
  <owl:sameAs rdf:resource="http://rdf.freebase.com/ns/m.04p4ds"/>
  <owl:sameAs rdf:resource="http://rdf.freebase.com/ns/m.063hb48"/>
  <owl:sameAs rdf:resource="http://rdf.freebase.com/ns/m.09rr393"/>
  <owl:sameAs rdf:resource="http://rdf.freebase.com/ns/m.0cnx5wm"/>

  <sameAsQuery rdf:ID="sameAsQuery_5">
...
  </sameAsQuery>
  <owl:sameAs rdf:resource="http://dbpedia.org/resource/Taxi_%281998_film%29"/>
  <owl:sameAs rdf:resource="http://dbpedia.org/resource/Taxi_%282004_film%29"/>
  <owl:sameAs rdf:resource="http://dbpedia.org/resource/Taxi_%281996_film%29"/>

```

```

<sameAsQuery rdf:ID="sameAsQuery_6">
...
</sameAsQuery>
<owl:sameAs rdf:resource="http://data.linkedmdb.org/resource/film/863"/>
<owl:sameAs rdf:resource="http://data.linkedmdb.org/resource/film/925"/>
<owl:sameAs rdf:resource="http://data.linkedmdb.org/resource/film/926"/>
...
</programme>
</has_programme>
</tv>

```

**Figura 4.8 – Obtenção das propriedades *owl:sameAs* a partir da execução das consultas *sameAsQuery***

Como pode ser observado na Figura 4.8 a execução das consultas das instâncias da classe *sameAsQuery* possibilita a obtenção das propriedades *owl:sameAs* que estabelecem relacionamentos entre as entidades da ontologia OWLTV e entidades das bases de conhecimento Freebase, DBpedia e LinkedMDB.

Em alguns casos a execução das consultas não retorna qualquer valor, como é o caso da consulta definida pela instância da classe *sameAsQuery* identificada por “sameAsQuery\_3” que foi realizada na base de conhecimento LinkedMDB para a entidade identificada por “programme\_1” que é relacionada ao programa “Chocolate com Pimenta”. Neste caso não foram encontradas na base de conhecimento LinkedMDB entidades relacionadas ao programa “Chocolate com Pimenta” da ontologia OWLTV.

Em outros casos, a execução das consultas retorna mais de uma entidade, como foi o caso das consultas para a entidade OWLTV relacionada ao filme “Táxi” que retornou sete entidades na consulta da base de conhecimento Freebase e três entidades nas consultas das bases de conhecimento DBpedia e LinkedMDB. Neste caso, como será apresentado no próximo capítulo, todas as entidades relacionadas que forem obtidas como resultados das consultas *sameAsQuery* serão exibidas para o usuário do sistema desenvolvido em Ginga-J. O usuário poderá então navegar nas páginas de cada entidade relacionada, para obter maiores informações sobre as mesmas.

#### 4.4 Considerações finais do capítulo

Neste capítulo foi apresentada a aplicação do processo de alinhamento e extensão de metadados da TV digital para a representação das informações de metadados voltadas para a classificação de conteúdos já existentes nos sistemas de TV digital usados no Brasil em redes terrestre (ISDB-T) na forma de uma ontologia visando uma maior integração e organização destas informações.

Esta ontologia também foi estendida para fazer um alinhamento com as informações de metadados utilizadas pelas bases de conhecimento Freebase, DBpedia e LinkedMDB visando



possibilitar uma extensão dos sistemas de TV digital com as informações de metadados destas bases de conhecimento da Web Semântica.

No próximo capítulo esta ontologia de alinhamento de metadados será validada com uma prova de conceito implementada em um sistema de TV digital em operação.



## Capítulo 5

# Validação do processo de alinhamento e extensão de metadados da TV digital

O propósito deste capítulo é apresentar a validação do processo de alinhamento e extensão de metadados com uma prova de conceito implementada em um sistema de TV digital em funcionamento.

A prova de conceito foi concebida para execução em um receptor de TV digital sem a necessidade de realizar mudanças operacionais nas redes de TV digital em operação. O receptor usado para o desenvolvimento da prova de conceito é um receptor conectado com capacidade de receber o sinal da TV digital via rede terrestre (ISDB-T) e com acesso a Internet através de um canal de retorno, possibilitando o acesso a conteúdos disponíveis em repositórios da Web Semântica como as bases de conhecimento Freebase, DBpedia e LinkedMDB. O receptor possui suporte ao padrão Ginga-J (KULESZA *et al.*, 2010) sendo plenamente compatível com a norma brasileira de televisão digital terrestre (ABNT, 2010).

### 5.1 Ontologia para alinhamento de metadados

Conforme observado no Capítulo 4, para a geração da ontologia OWLTV deve ser utilizada uma única fonte de informação como ponto de partida para o alinhamento com as bases de conhecimento da Web Semântica. Esta fonte de informação pode estar tanto disponível em um repositório da Internet, como por exemplo, o site Revista Eletrônica, como pode ser transmitida na tabela EIT do sistema de TV digital. No caso da prova de conceito implementada neste trabalho, foi utilizada como ponto de partida a informação de programação dos canais de TV brasileiros publicada pelo site Revista Eletrônica em formato XMLTV.

Para ilustrar a implementação da prova de conceito será utilizada como base a ontologia OWLTV gerada no capítulo anterior que possui as informações relacionadas aos programas “Chocolate com Pimenta” e “Táxi” transmitidos pela emissora EPTV no dia 21 de junho de 2012. O usuário da TV digital, ao assistir um programa de TV aberta que está sendo transmitido por uma emissora de TV via rede terrestre (ISDB-T) terá acesso a um aplicativo desenvolvido em Ginga-J que exibirá as informações de metadados relativas ao conteúdo obtidas a partir do site Revista Eletrônica, como o título, descrição, gênero, classificação parental, duração entre outras. Estas informações são obtidas pelo aplicativo desenvolvido em Ginga-J através dos metadados da ontologia OWLTV exibida no Apêndice A.

Devido ao fato do receptor possuir acesso a Internet, o usuário da TV digital pode utilizar este recurso para obter acesso a maiores informações sobre os conteúdos que está assistindo.

Estas informações complementares são obtidas a partir dos repositórios de conhecimento Freebase, DBpedia e LinkedMDB. Foi implementada uma ferramenta na aplicação Ginga que, utilizando a ontologia de alinhamento de metadados criada no Capítulo 4, faz o acesso aos repositórios de conhecimento da Web Semântica para obter maiores informações sobre os conteúdos que estão sendo transmitidos.

## 5.2 Cenários de utilização do receptor

Para apresentar a prova de conceito, serão demonstrados dois cenários de utilização do receptor. O primeiro cenário é voltado para a integração de informações da TV digital com as bases de conhecimento dos repositórios Freebase e DBpedia para extensão das informações de metadados do programa “Chocolate com Pimenta”. O segundo cenário é voltado para a extensão das informações de metadados do programa “Táxi” com a base de conhecimento do repositório Freebase. Além destes cenários, diversos outros podem ser vislumbrados como prova de conceito da aplicação da ontologia de alinhamento de metadados da TV digital.

### 5.2.1 Cenário do programa “Chocolate com Pimenta”

Supondo-se que o usuário estava assistindo ao programa “Chocolate com Pimenta” que foi transmitido pela emissora EPTV em 12 de junho de 2012. Uma ferramenta implementada no sistema desenvolvido em Ginga-J utiliza a ontologia de alinhamento de metadados para fazer uma extensão das informações de metadados OWLTV com as informações de metadados dos repositórios de conhecimento Freebase e DBpedia.

Desta forma, o sistema desenvolvido em Ginga-J por meio da ontologia OWLTV faz a geração das instâncias da classe *sameAsQuery* para cada base de conhecimento da Web Semântica a ser integrada. De posse das instâncias da classe *sameAsQuery*, o sistema desenvolvido em Ginga-J saberá como executar as pesquisas nas bases de conhecimento da Web Semântica para obtenção das propriedades *owl:sameAs*.

Para isso é usada uma ferramenta de busca específica para cada base de conhecimento da Web Semântica a ser integrada. Esta ferramenta recebe como parâmetros de busca as informações das instâncias da classe *sameAsQuery*. Uma vez fornecidas estas informações, a pesquisa é disparada, e ao final de sua execução é exibido para o usuário uma listagem das propriedades *owl:sameAs* encontradas nas bases de conhecimento que são relacionadas com a entidade OWLTV pesquisada. O usuário da TV digital deve então selecionar a propriedade *owl:sameAs* desejada para ser direcionado à página do repositório da Web Semântica que contém as informações detalhadas sobre a respectiva propriedade. O diagrama exibido na Figura 5.1 ilustra passo a passo a aplicação deste cenário.

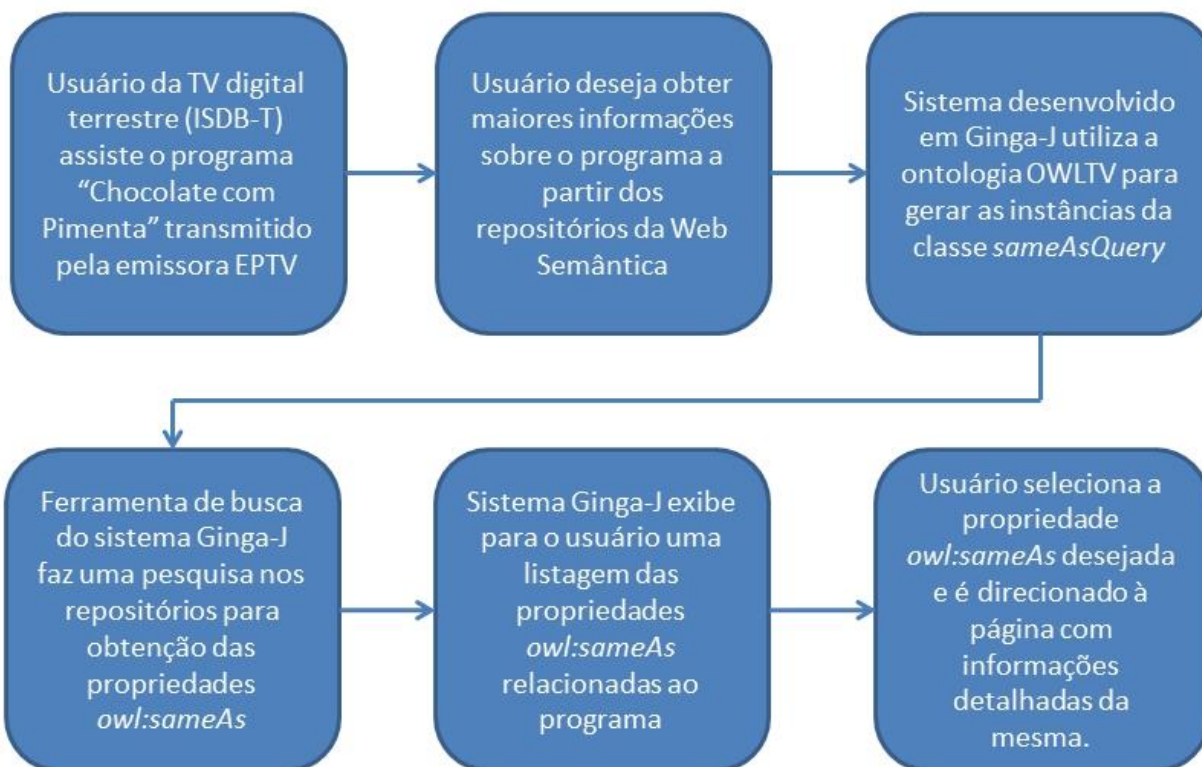


Figura 5.1 – Cenário do programa “Chocolate com Pimenta”

### 5.2.2 Cenário do programa “Táxi”

Supondo-se que o usuário estava assistindo ao programa “Táxi” que foi transmitido pela emissora EPTV em 12 de junho de 2012. Uma ferramenta implementada no sistema desenvolvido em GINGA-J utiliza a ontologia de alinhamento de metadados para fazer uma extensão das informações de metadados OWLTV com as informações de metadados do repositório de conhecimento Freebase.

Desta forma, o sistema desenvolvido em GINGA-J por meio da ontologia OWLTV faz a geração da instância da classe *sameAsQuery* com a propriedade *owltv\_entity* igual a “programme\_2” e a propriedade *linked\_ontology* igual a “Freebase”. De posse da instância da classe *sameAsQuery*, o sistema desenvolvido em GINGA-J saberá como executar a pesquisa na base de conhecimento Freebase para obtenção das propriedades *owl:sameAs*.

Para isso é usada uma ferramenta de busca específica para a base de conhecimento Freebase. Esta ferramenta recebe como parâmetros de busca as informações da instância da classe *sameAsQuery*. Uma vez fornecidas estas informações, a pesquisa é disparada, e ao final de sua execução é exibido para o usuário uma listagem das propriedades *owl:sameAs* encontradas na base de conhecimento Freebase que são relacionadas com a entidade OWLTV pesquisada, no caso o programa “Táxi”. O usuário da TV digital deve então selecionar a propriedade *owl:sameAs* desejada para ser direcionado à página do repositório Freebase que contém as

informações detalhadas sobre a respectiva propriedade. O diagrama exibido na Figura 5.2 ilustra passo a passo a aplicação deste cenário.

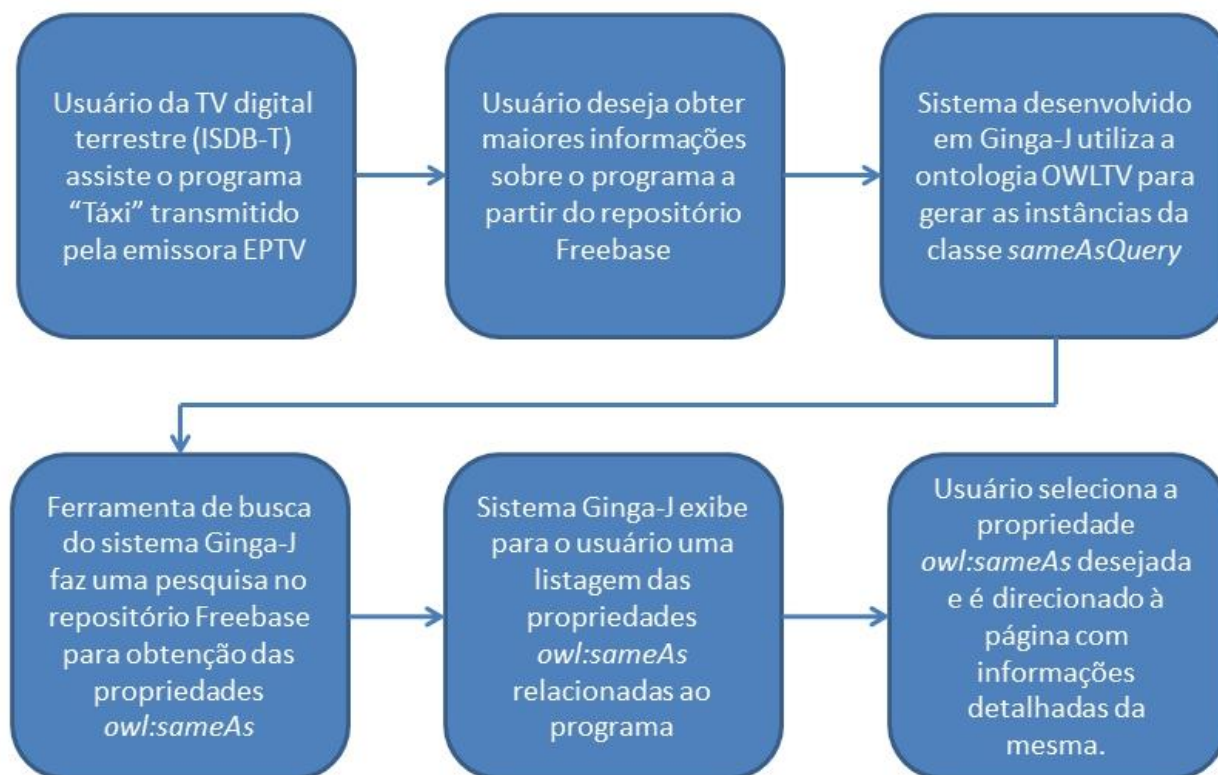


Figura 5.2 – Cenário do programa “Táxi”

### 5.3 Sistema desenvolvido em Ginga-J

Para realizar um mapeamento das informações entre os diferentes padrões de metadados da TV digital e da Internet utilizando a ontologia de alinhamento de metadados OWLTV criada no Capítulo 4, foi desenvolvido um sistema em linguagem Ginga-J.

O sistema Ginga-J é um aplicativo que pode ser transmitido pela emissora no sinal da TV digital ou ficar residente na memória no receptor. O aplicativo exibe para o usuário as informações sobre os programas que estão sendo transmitidos pela emissora via transmissão terrestre. No caso da prova de conceito implementada neste trabalho o sistema Ginga-J obtém as informações de metadados dos programas via Internet a partir do site Revista Eletrônica. Conforme foi apresentado no Capítulo 4 as informações de metadados necessárias para a geração da ontologia OWLTV também podem ser obtidas a partir da tabela EIT (*Event Information Table*) que é transmitida no sinal da TV digital e utilizada pelos receptores para a exibição do guia eletrônico de programação (EPG – *Electronic Program Guide*). O EPG possui uma apresentação específica para cada modelo de receptor. Na Figura 5.3 são apresentados exemplos de telas de aplicações EPG existentes em alguns modelos de receptores de TV digital.

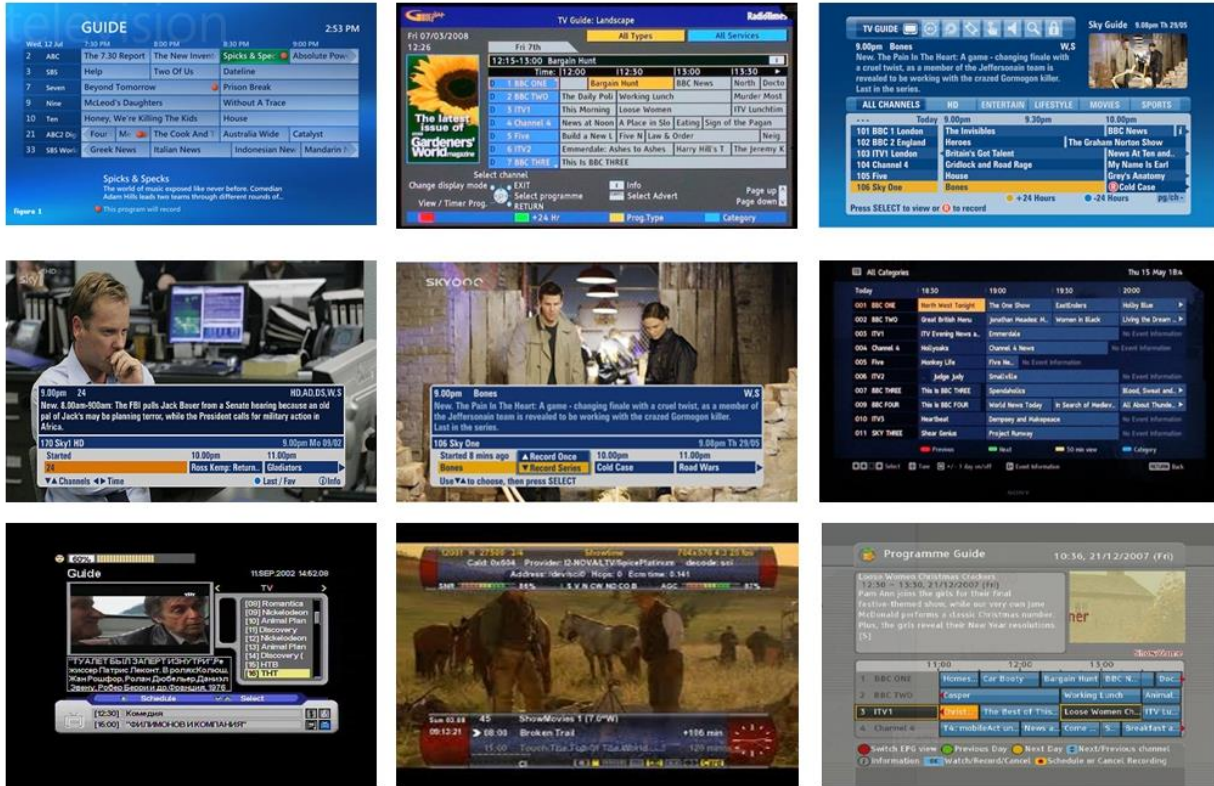


Figura 5.3 – Telas de exemplos de aplicações EPG

Apesar da apresentação da aplicação EPG ser diferente conforme o modelo de receptor, as informações exibidas pela aplicação referente aos programas transmitidos são as mesmas, independente do modelo de receptor. Estas informações sobre os programas transmitidos são obtidas pela aplicação EPG a partir da tabela EIT que é transmitida juntamente com o conteúdo de áudio e vídeo do canal de televisão. O conteúdo desta tabela é definido pelas especificações de metadados do sistema brasileiro de TV digital terrestre, o ISDB-T (ABNT, 2007).

Na Figura 5.4 é apresentado como sistema desenvolvido em Ginga-J utilizado na prova de conceito está estruturado.

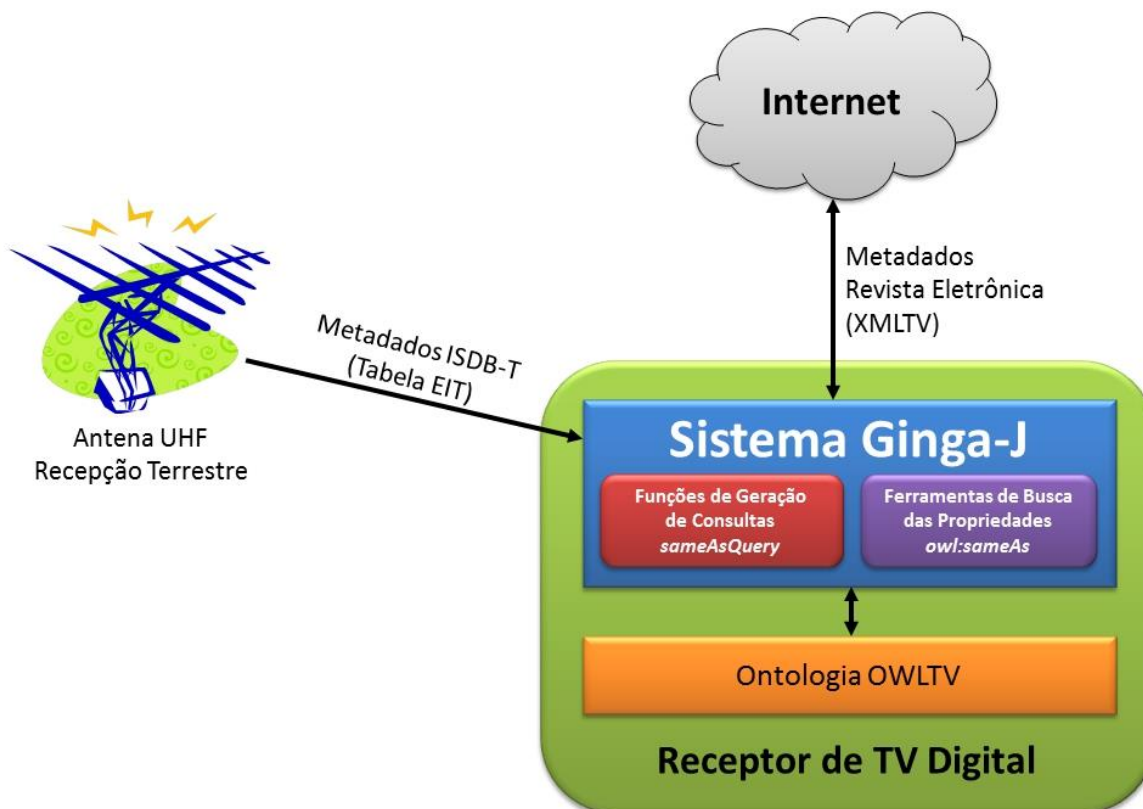


Figura 5.4 – Estrutura do sistema desenvolvido em Ginga-J

O sistema desenvolvido em Ginga-J pode ser transmitido no sinal da TV digital ou ficar residente na memória do receptor. O sistema obtém os metadados necessários para geração da ontologia OWLTV tanto através da tabela EIT transmitida no sinal da TV digital, quanto via Internet a partir do site Revista Eletrônica, conforme foi detalhado no Capítulo 4. O sistema Ginga-J possui funções específicas de software para cada repositório da Web Semântica a ser integrado. Estas funções são utilizadas tanto para a geração das consultas *sameAsQuery* como para realização de busca das propriedades *owl:sameAs* nos repositórios da Web Semântica.

## 5.4 Aplicação para o usuário final

A aplicação para o usuário final desenvolvida como prova de conceito utiliza as informações de metadados da ontologia OWLTV apresentada no Capítulo 4, bem como as funções de geração de consultas e ferramentas de busca específicas para cada repositório da Web Semântica a ser integrado. Esta aplicação foi desenvolvida para ser executada em qualquer receptor do mercado com suporte padrão Ginga-J (KULESZA *et al.*, 2010). A tela inicial da aplicação é similar a uma aplicação de guia eletrônico de programação (EPG), e está exibida na Figura 5.5.



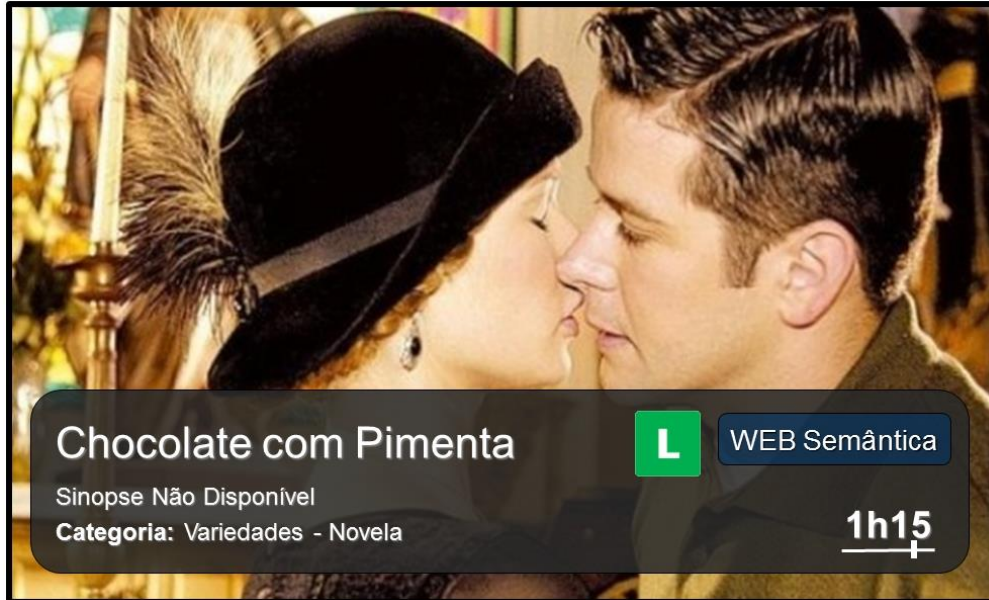


Figura 5.5 – Tela da inicial da aplicação para o usuário final

A aplicação exibe para o usuário as informações de metadados disponíveis na ontologia OWLTV para o conteúdo que está sendo transmitido, no caso do cenário da Figura 5.6, a novela “Chocolate com Pimenta”. Entre as informações exibidas, podemos destacar: título do programa, sinopse, categorias, classificação indicativa e duração. A diferença principal em relação a uma aplicação tradicional de EPG é a existência do botão “WEB Semântica”. Este botão deve ser acionado pelo usuário para obter maiores informações sobre o conteúdo que está assistindo a partir de repositórios da Web Semântica. Ao ser acionado a tela exibida na Figura 5.6 é executada.

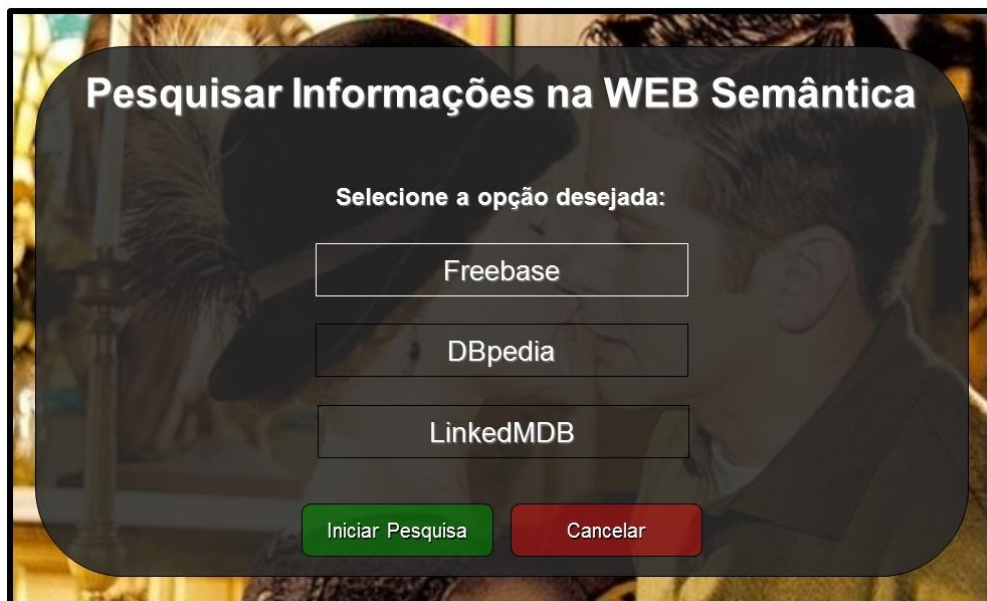


Figura 5.6 – Tela inicial da pesquisa de informações na Web Semântica

A partir da tela inicial de pesquisa de informações na Web Semântica, o usuário pode selecionar o repositório no qual será realizada a pesquisa, a partir da relação de repositórios da cujos metadados estão integrados com os metadados da TV digital por meio da ontologia de alinhamento de metadados OWLTV. No caso da prova de conceito que foi implementada, os repositórios da Web Semântica são as bases de conhecimento Freebase, DBpedia e LinkedMDB.

#### 5.4.1 Exemplo de pesquisa da base de conhecimento Freebase

Caso o usuário selecione a opção “Freebase” para iniciar a pesquisa, o sistema desenvolvido em Ginga-J irá realizar a pesquisa segundo o procedimento descrito na seção 5.2.1 e a tela de resultados com a listagem das propriedades *owl:sameAs* relacionadas ao programa “Chocolate com Pimenta” é exibida conforme a Figura 5.7.



Figura 5.7 – Tela de resultados da pesquisa ao repositório Freebase

Este exemplo tem como resultado um total de apenas 1 conteúdo existente no repositório Freebase que é relacionado ao programa “Chocolate com Pimenta”. O usuário pode então exibir informações detalhadas deste conteúdo através da opção “Detalhar Conteúdo”. A aplicação Ginga-J irá então direcionar o usuário para a página do repositório Freebase contendo informações detalhadas sobre o conteúdo selecionado, conforme a Figura 5.8.



Figura 5.8 – Página do conteúdo no repositório Freebase

#### 5.4.2 Exemplo de pesquisa da base de conhecimento DBpedia

Caso o usuário selecione a opção “DBpedia” para iniciar a pesquisa conforme exibido na Figura 5.9, o sistema desenvolvido em Ginga-J irá realizar a pesquisa segundo o procedimento descrito na seção 5.2.1.

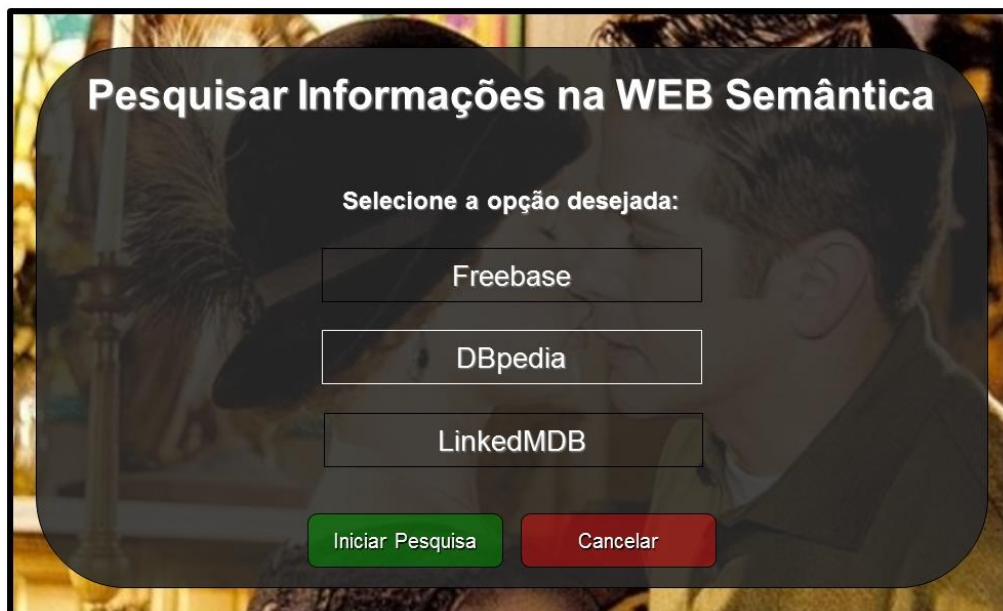


Figura 5.9 – Tela inicial da pesquisa de informações na Web Semântica

A tela de resultados com a listagem das propriedades *owl:sameAs* relacionadas ao programa “Chocolate com Pimenta” existentes no repositório DBpedia é exibida conforme a Figura 5.10.



Figura 5.10 – Tela de resultados da pesquisa ao repositório DBpedia

Este exemplo tem como resultado um total de apenas 1 conteúdo existente no repositório DBpedia que é relacionado ao programa “Chocolate com Pimenta”. O usuário pode então exibir informações detalhadas deste conteúdo através da opção “Detalhar Conteúdo”. A aplicação Ginga-J irá então direcionar o usuário para a página do repositório DBpedia contendo informações detalhadas sobre o conteúdo selecionado, conforme a Figura 5.11.

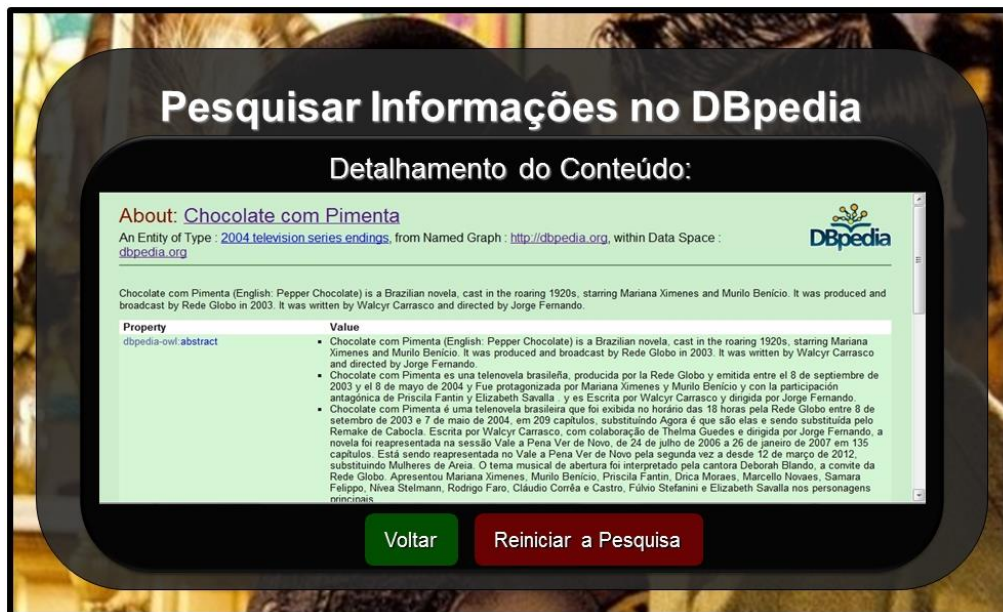


Figura 5.11 – Página do conteúdo no repositório DBpedia

### 5.4.3 Exemplo de pesquisa Freebase no cenário do programa “Táxi”

Quando utilizada no cenário do programa “Táxi” a aplicação exibe para o usuário as informações de metadados disponíveis na ontologia OWLTV para o programa, conforme exibido na Figura 5.12. Entre as informações exibidas, podemos destacar: título do programa, sinopse, título em inglês, direção, elenco, ano de lançamento, categorias, país, classificação indicativa e duração. A diferença principal em relação a uma aplicação tradicional de EPG é a existência do botão “WEB Semântica”. Este botão deve ser acionado pelo usuário para obter maiores informações sobre o conteúdo que está assistindo a partir de repositórios da Web Semântica. Ao ser acionado a tela exibida na Figura 5.6 é executada.

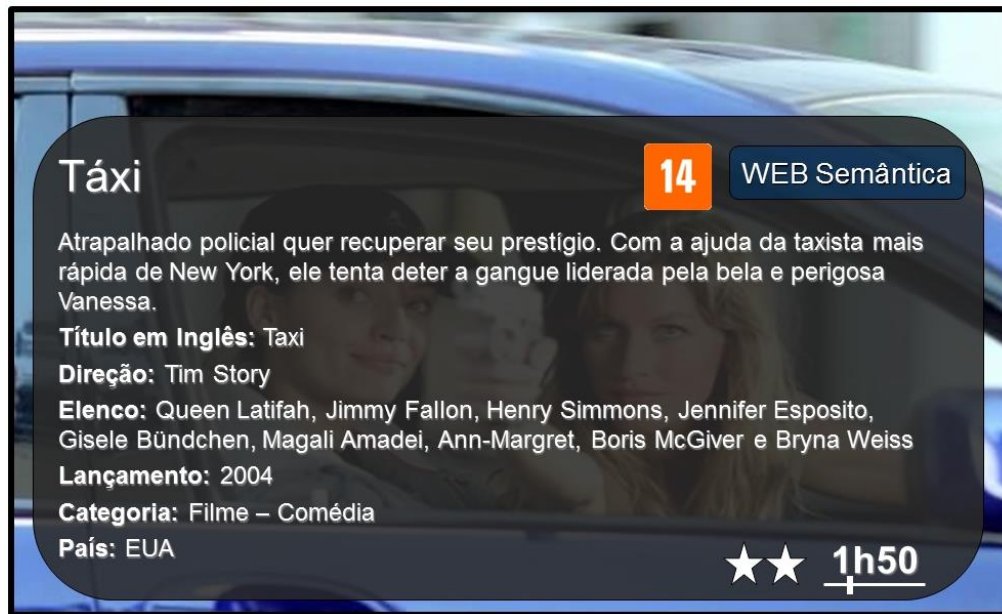


Figura 5.12 – Tela inicial da aplicação no cenário do programa “Táxi”

Caso o usuário selecione a opção “Freebase” para iniciar a pesquisa, o sistema desenvolvido em Ginga-J irá realizar a pesquisa segundo o procedimento descrito na seção 5.2.2 e a tela de resultados com a listagem das propriedades *owl:sameAs* relacionadas ao programa “Táxi” é exibida conforme a Figura 5.13.

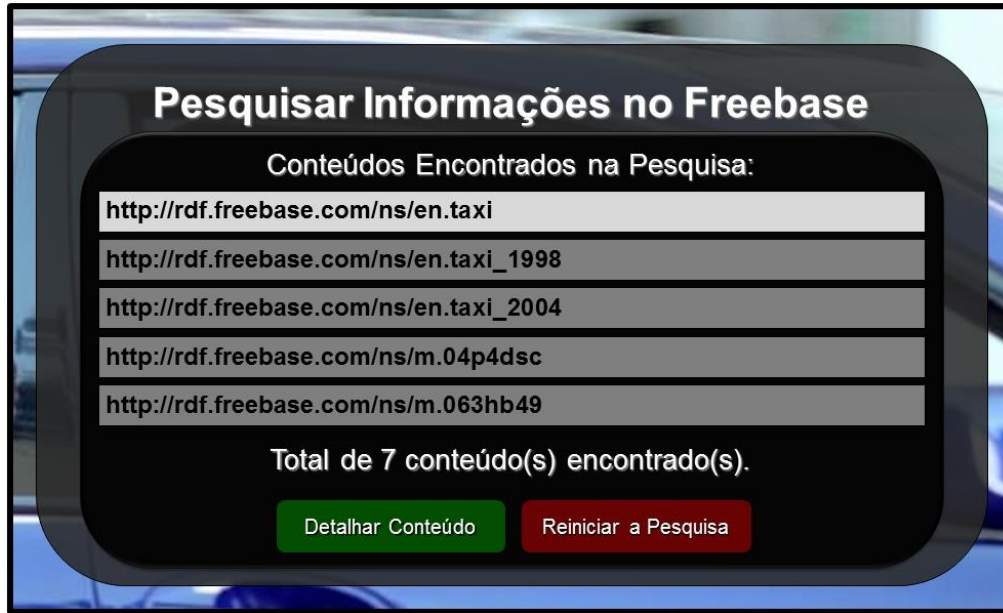


Figura 5.13 – Tela de resultados da pesquisa ao repositório Freebase

Este exemplo tem como resultado um total de sete conteúdos existentes no repositório Freebase que são relacionados ao programa “Táxi”. O usuário pode então exibir informações detalhadas de cada conteúdo através da opção “Detalhar Conteúdo”. A aplicação Ginga-J irá então direcionar o usuário para a página do repositório Freebase contendo informações detalhadas sobre o conteúdo selecionado, conforme a Figura 5.14.

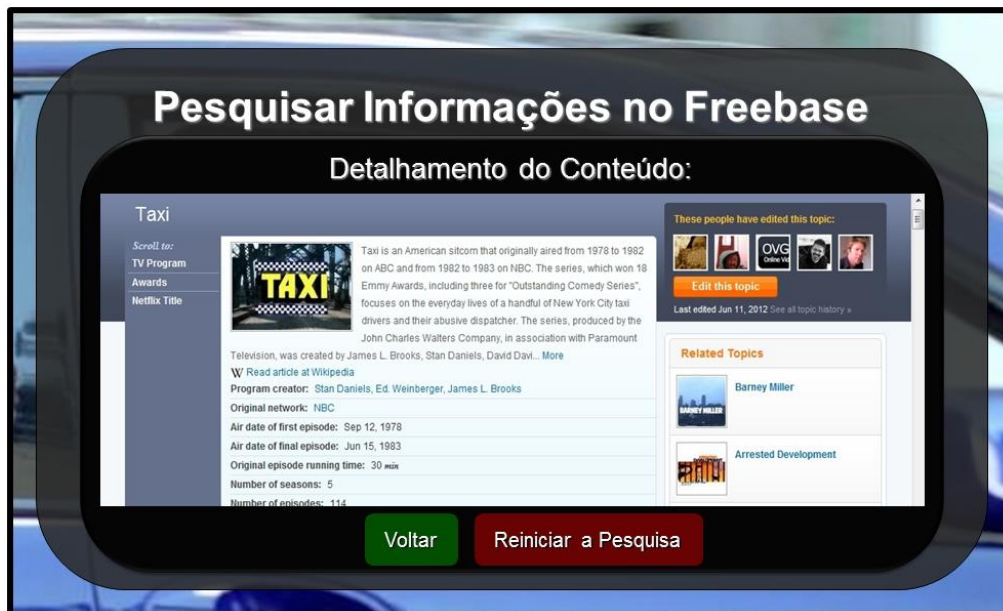


Figura 5.14 – Página do conteúdo no repositório Freebase

Conforme pode ser observado na Figura 5.14, o conteúdo exibido não se refere ao mesmo programa com o título “Táxi” que foi transmitido pela emissora EPTV no dia 12 de junho de

2012, e sim a outro programa cujo título em inglês também é “Taxi”. Neste caso, o usuário deverá retornar à tela de resultados da pesquisa exibida na Figura 5.15 e selecionar outro conteúdo da lista.

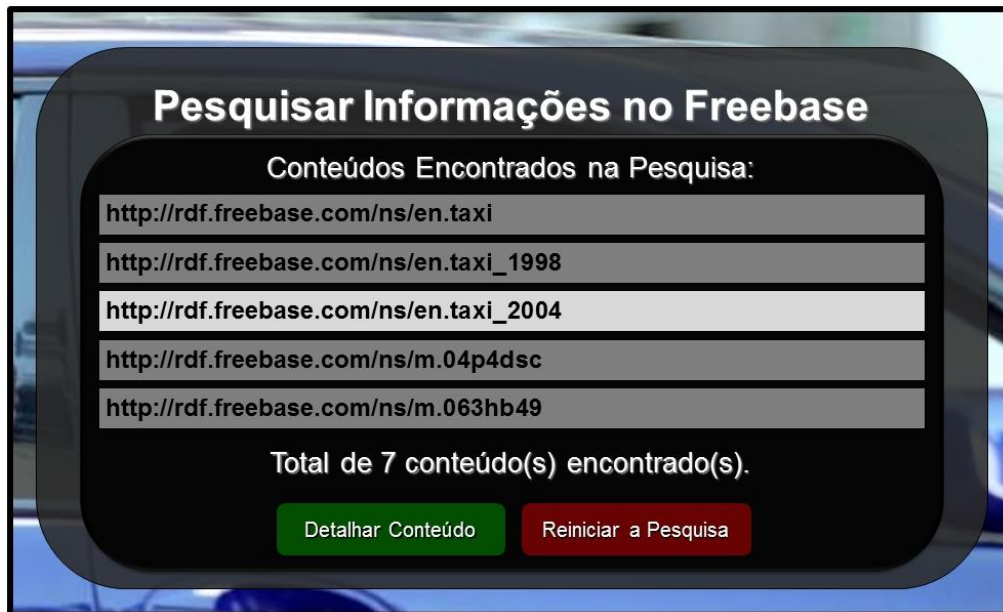


Figura 5.15 – Tela de resultados da pesquisa ao repositório Freebase

Caso o usuário selecione o terceiro conteúdo da lista de resultados da pesquisa, a aplicação Ginga-J irá então direcionar o usuário para a página do repositório Freebase contendo informações detalhadas sobre o conteúdo selecionado, que conforme exibido na Figura 5.16, se refere ao mesmo programa com o título “Táxi” que foi transmitido pela emissora EPTV no dia 12 de junho de 2012.



Figura 5.16 – Página do conteúdo no repositório Freebase

## 5.5 Aumentando a abrangência da aplicação da ontologia

Uma evolução da prova de conceito apresentada seria aumentar a abrangência da aplicação da ontologia com a integração de outras informações de metadados da TV digital com as informações de metadados dos domínios de conhecimento da Web Semântica. Além do título do conteúdo, dois tipos de informação que poderiam ser integrados e que trariam grande benefício para o usuário são as informações relacionadas à direção e ao elenco dos programas. Desta forma, ao assistir um programa o usuário poderia selecionar um indivíduo na relação de diretores e atores para ter acesso a informações detalhadas sobre o mesmo a partir de bases de conhecimento da Web Semântica. A Figura 5.17 exibe informações detalhadas da entidade relacionada ao diretor “Quentin Tarantino” no repositório Freebase.

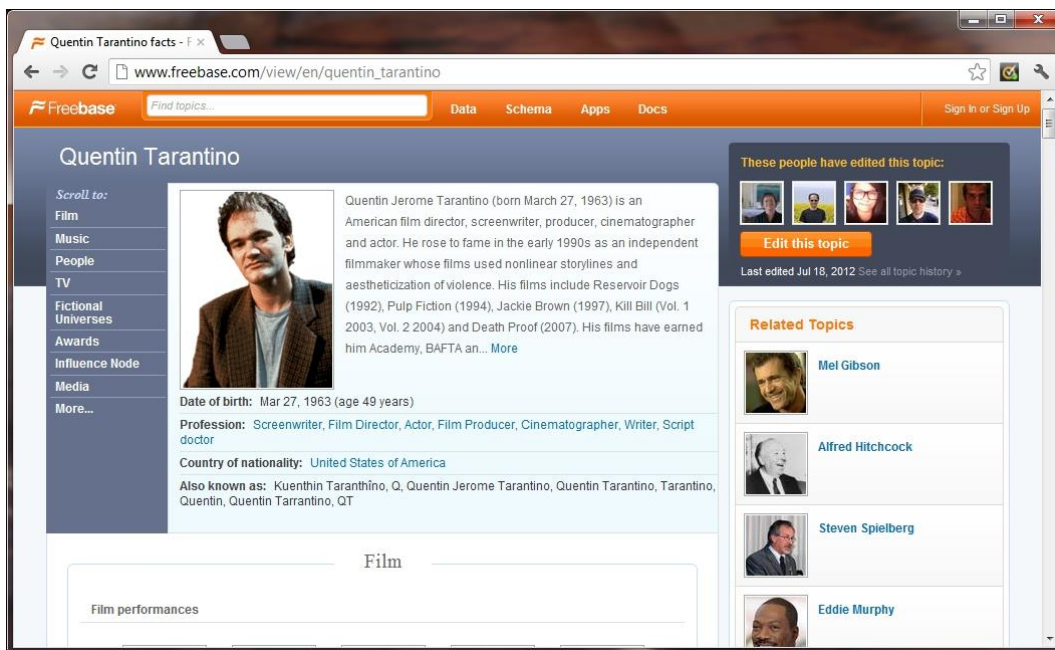


Figura 5.17 – Página da entidade relacionada ao diretor “Quentin Tarantino” no repositório Freebase

Como pode ser observado na Figura 5.17 o repositório Freebase possui informações detalhadas sobre o diretor “Quentin Tarantino” como informações pessoais, currículo completo, filmes nos quais trabalhou, premiações, entre outras informações. Da mesma forma que o repositório Freebase, outros repositórios como DBpedia e LinkedMDB também trazem informações de metadados detalhadas relacionadas à entidade “Quentin Tarantino”.

Para realizar esta integração um primeiro passo seria criar uma nova instância da classe *sameAsQuery* na ontologia OWLTV para fazer um relacionamento entre a entidade *director* da ontologia OWLTV e os repositórios da Web Semântica a serem integrados com esta entidade. No caso dos repositórios Freebase, DBpedia e LinkedMDB, poderiam ser adicionadas à ontologia OWLTV instâncias da classe *sameAsQuery* com as propriedades *linked\_ontology* com os valores



“Freebase”, “DBpedia” e “LinkedMDB”. Estas instâncias conteriam as consultas a serem realizadas nas bases de conhecimento a serem relacionadas com a ontologia OWLTV.

Por exemplo, na Figura 5.18 está exibido o conteúdo da instância da classe *sameAsQuery* para o cenário de integração de informações de metadados relacionadas a entidade OWLTV identificada por “director\_1” que é relacionada ao diretor “Quentin Tarantino”.

```
<sameAsQuery rdf:ID="sameAsQuery_1">
  <owltv_entity rdf:datatype="&xsd:string">director_1</owltv_entity>
  <linked_ontology rdf:datatype="&xsd:string">Freebase</linked_ontology>
  <query_code rdf:datatype="&xsd:string">
    [
      "id": null,
      "name|=":[ "Quentin Tarantino"],
      "type|=":[ "/film/director"]
    ]
  </query_code>
</sameAsQuery>
```

**Figura 5.18 – Instância da classe *sameAsQuery***

A partir da instância da classe *sameAsQuery* exibida na Figura 5.18 seria possível fazer a execução da consulta utilizando a função de software específicas para a base de conhecimento Freebase para obtenção da propriedade *owl:sameAs* relacionada à entidade “Quentin Tarantino” que é representada pela URL [http://rdf.freebase.com/ns/en.quentin\\_tarantino](http://rdf.freebase.com/ns/en.quentin_tarantino).

O mesmo processo que foi exemplificado com a entidade *director* poderia ser aplicado a outras entidades da ontologia OWLTV como as entidades *actor*, *producer*, *writer*, *editor*, *composer* entre outras.

## 5.6 Considerações finais do capítulo

Neste capítulo foi apresentada a validação do processo de alinhamento e extensão de metadados com uma prova de conceito concebida para execução em um receptor de TV digital sem a necessidade de realizar mudanças operacionais nas redes de TV digital em operação. Para implementação da prova de conceito foi considerada a aplicação da ontologia de alinhamento de metadados da TV digital OWLTV com os padrões de metadados da Web Semântica relacionados como os domínios de conhecimento Freebase, DBpedia e LinkedMDB. No próximo capítulo serão apresentadas as conclusões do trabalho com sugestões de trabalhos futuros visando aumentar a abrangência da aplicação da ontologia com a integração de informações de metadados da Web Semântica com as informações de metadados OWLTV não apenas relacionadas ao título do conteúdo, mas também com os metadados relacionados ao elenco e direção dos programas. Também será proposto um desenvolvimento futuro de um sistema de geração de consultas personalizadas conforme o interesse do usuário.



# Capítulo 6

## Conclusões

Com a evolução nos anos recentes da tecnologia de *broadcast* multimídia, e em particular da TV digital, a forma de acesso das pessoas à mídia de massa vem sendo modificada significativamente. A integração da televisão com a Internet e as tecnologias de comunicação móveis têm provido aos telespectadores novos serviços interativos de conteúdo digital, fazendo com que a audiência da TV digital cresça a cada dia.

Os televisores e dispositivos de acesso à TV digital, como set-top boxes, têm se tornado cada vez mais sofisticados, suportando uma variedade de conteúdos e conectividade com outras redes e dispositivos, combinando elementos da televisão tradicional com a Internet para prover ao usuário o acesso a uma diversidade de conteúdos de mídia interativa.

Este crescimento do volume e da diversidade de serviços e conteúdos multimídia na TV digital também acarreta os mesmos problemas de complexidade e excesso de informações que já vinham sendo encarados por outras mídias digitais relacionadas com a Internet. Um meio de lidar com esta complexidade de serviços e conteúdos digitais de forma prática e eficiente é a tecnologia de metadados, que são dados que complementam as informações digitais dos conteúdos multimídia com o objetivo de melhor descrevê-los, facilitando a estruturação e o gerenciamento de grandes volumes de informação.

Atualmente existem diversos padrões de metadados que são utilizados pela indústria de *broadcast* multimídia para estruturar a transmissão de informações nas redes de TV digital. Contudo, a integração entre as informações desses padrões de metadados se tornou um desafio com advento dos receptores de TV digital para redes híbridas. Estes receptores possuem as seguintes características:

- Acesso ao sinal digital transmitido por diversas redes digitais, como por exemplo, a rede terrestre (ISDB-T) que leva o sinal da TV digital aberta brasileira;
- Conexão a repositórios de conteúdos da Internet através de um canal de retorno.

Esta tese propôs a representação das informações de metadados existentes em um ambiente de transmissão e recepção de TV digital aberta em redes de transmissão terrestre na forma de ontologias visando uma maior integração e organização destas informações. As ontologias permitiram criar modelos conceituais abstratos mais simplificados e menos formais que as especificações dos padrões de metadados usadas atualmente na indústria de *broadcast* multimídia.

Neste trabalho foi descrito um processo para integrar as informações de metadados dos padrões da indústria de *broadcast* multimídia com informações de metadados voltadas para a descrição de conhecimento de domínios específicos existentes em repositórios da Web Semântica. Esta integração de padrões de metadados foi possível através da criação de ontologias para a representação das informações de metadados já existentes no contexto da TV digital estabelecendo uma integração entre os padrões de metadados *broadcast* multimídia para possibilitar a comparação semântica das informações destes formatos por meio das ontologias e a extensão destes formatos com a utilização de ontologias para domínios específicos existentes na Web Semântica.

## 6.1 Principais contribuições

Esta tese possui duas contribuições que constituem o fator de originalidade da mesma. A primeira contribuição é o processo que foi criado para estender e integrar as informações de metadados dos padrões da indústria de TV digital aberta com a possibilidade de extensão destes padrões com descrições de conhecimento de domínios específicos existentes em domínios de conhecimento da Web Semântica. Para fazer uma integração entre estes diferentes padrões de metadados foi proposta a criação de uma ontologia para alinhamento de metadados da TV digital, denominada OWLTV, que permitiu realizar um mapeamento entre as informações de descrição de conteúdo disponíveis no sistema de TV digital e as informações relacionadas existentes em repositórios da Web Semântica.

Este processo englobou mecanismos de interoperabilidade entre os metadados da indústria de *broadcast* multimídia e abriu oportunidade para reuso das ontologias e modelos no desenvolvimento de um grande número de aplicações e novos serviços interativos no contexto da TV digital. Estas aplicações e serviços interativos utilizam métodos e funções de software das ferramentas de mapeamento e recuperação de dados que são executadas pelos telespectadores utilizando os recursos de memória e processamento dos equipamentos de recepção de sinal digital. Estes métodos e funções de software utilizam a ontologia de alinhamento de metadados como base de inteligência para realizar o mapeamento entre as informações de descrição de conhecimento da TV digital e os metadados de domínios específicos dos repositórios da Web Semântica.

A ontologia de alinhamento de metadados OWLTV é, portanto, a componente chave do sistema, pois é ela que possibilita fazer um relacionamento entre dados existentes em um repositório da Web Semântica com um conteúdo transmitido no sinal da TV digital. Através deste mapeamento entre as informações de metadados dos diferentes domínios, as ferramentas disponíveis no receptor digital conseguem recuperar os dados que serão utilizados pelas aplicações e serviços interativos.

Uma consequência da primeira contribuição é o fato da implementação do processo proposto não impactar em mudanças operacionais nos *broadcasters* e provedores de serviços. Ao

analisar diversos trabalhos realizados por outros autores, apresentados no Capítulo 2, verificou-se que um requisito necessário para viabilizar a implementação de todos eles era a necessidade de envolvimento do provedor de serviços e da emissora de TV para gerar e transportar os metadados relacionados às ontologias propostas. Devido a este requisito, estas propostas já existentes possuíam dificuldade de implementação real em redes de TV digital em operação, pois necessitavam que as emissoras e provedores de serviços tivessem que realizar mudanças operacionais e adaptações no parque de equipamentos de transmissão.

A superação desta limitação existente nos trabalhos acadêmicos que fizeram o uso de ontologias no domínio de *broadcast* multimídia foi a principal inovação do presente trabalho. O objetivo foi permitir com que os benefícios e novos recursos interativos proporcionados para o telespectador pela ontologia proposta pudessem ser implementados na prática sem a necessidade de envolvimento operacional das emissoras e provedores de serviços de TV digital. A ontologia de alinhamento de metadados gerada a partir da aplicação do processo proposto, assim como as ferramentas de manipulação e recuperação de dados, bem como as novas aplicações e serviços interativos são totalmente processados no equipamento receptor de sinal de TV digital sem qualquer impacto operacional nas redes de TV digital em operação.

A segunda contribuição foi a validação do processo de alinhamento e extensão de metadados com a implementação de uma prova de conceito em um sistema de TV digital em funcionamento. Esta prova de conceito foi desenvolvida e executada em um receptor de TV digital com capacidade de receber o sinal da TV digital via rede terrestre (ISDB-T). Além disso, o receptor também tem um canal de retorno, possibilitando o acesso a conteúdos disponíveis em repositórios da Web Semântica como as bases de conhecimento Freebase, DBpedia e LinkedMDB. O desenvolvimento da prova de conceito foi feito utilizando o padrão Ginga-J (KULESZA *et al.*, 2010) sendo plenamente compatível com a norma brasileira de televisão digital terrestre (ABNT, 2010) e suportado por diversos modelos de receptores (televisores e set-top boxes) disponíveis no mercado.

Esta contribuição do desenvolvimento da prova de conceito foi importante para comprovar a validade e aplicabilidade da proposta, e também para verificar as dificuldades técnicas que surgiram em sua implementação e a superação das mesmas. Além disso, a prova de conceito permitiu visualizar o benefício real proporcionado ao usuário final da TV digital, pela ontologia OWLTV criada para estender e alinhar as informações de metadados dos padrões da indústria de TV digital aberta com a possibilidade de extensão destes padrões com descrições de conhecimento de domínios específicos existentes em bases de conhecimento da Web Semântica.

## 6.2 Trabalhos futuros

Uma sugestão de trabalho futuro seria o desenvolvimento de um módulo gerador de consultas no sistema Ginga-J que permita que o próprio usuário da TV digital crie suas consultas aos repositórios de conhecimento da Web Semântica, utilizando as informações de metadados

existentes na ontologia OWLTV para fazer o mapeamento de informações. Desta forma, o usuário poderia criar consultas personalizadas avançadas como, por exemplo:

- Quais os diretores que já dirigiram a protagonista principal do filme “Táxi”, a atriz “Queen Latifah” em outros filmes onde ela atuou;
- Todos os filmes já estrelados pelos atores “Jimmy Fallon” e “Henry Simmons” do filme “Táxi”, exibindo para cada filme encontrado seu elenco completo;
- Todos os filmes já dirigidos pelo diretor “Tim Story” do filme “Táxi”, exibindo para cada filme encontrado o elenco, produtores e músicas da trilha sonora.

Estes são apenas alguns dos exemplos de consultas avançadas que poderiam ser criadas utilizando a flexibilidade proporcionada pela ontologia OWLTV e sua capacidade de integração com as tecnologias de busca e recuperação de informações existentes na Web Semântica.

Além dos repositórios da Web Semântica utilizados na prova de conceito do presente trabalho, este módulo gerador de consultas poderia permitir a integração de novos repositórios que possuam compatibilidade com as linguagens de consulta MQL (MQL, 2012) e SPARQL (SPARQL, 2012).

Outra sugestão de trabalho futuro seria a incorporação do processo proposto em um aplicativo comercial a ser desenvolvido no padrão Ginga que poderia ser instalado pelos usuários nos receptores de TV digital com suporte a esta tecnologia de middleware, que a partir de 2013 se tornou obrigatória em todos os televisores fabricados no Brasil.

## Referências Bibliográficas

ABNT. *NBR 15603-2 Televisão digital terrestre - Multiplexação e serviços de informação (SI) Parte 2: Sintaxes e definições da informação básica de SI*. Associação Brasileira de Normas Técnicas, 2007.

ABNT. *NBR 15606-4 Televisão digital terrestre - Codificação de dados e especificações de transmissão para radiodifusão digital Parte 4: Ginga-J - Ambiente para a execução de aplicações procedurais*. Associação Brasileira de Normas Técnicas, 2010.

ADL SCORM. *Sharable Content Object Reference Model*. Disponível em: <http://www.adlnet.org>. Acessado em: 27/07/2009.

ALVES, L.; KULESZA, R.; SILVA, F. S.; JUCÁ, P.; BRESSAN, G. *Análise Comparativa de Metadados em TV Digital*. II Workshop de TV Digital do Simpósio Brasileiro de Redes de Computadores, Curitiba, 2006.

BERNERS-LEE, T. *Linked Data*. International Journal on Semantic Web and Information Systems, vol. 4, n. 2, W3C, 2006.

BERNERS-LEE, T.; HENDLER, J.; LASSILA, O. *The Semantic Web*. Scientific American Magazine, Maio, 2001.

CHOI, N.; SONG, I.; HAN, H. *A Survey on Ontology Mapping*. SIGMOD Record, 35(3), 2006.

DBPEDIA. 2012, <http://dbpedia.org/About>, Acessado em 11/06/2012.

ETSI. EM 300 468 v1.5.1 *Digital Vídeo Broadcasting (DVB-S); Specification for Service Information (SI) in DVB-S systems*. European Telecommunications Standards Institute, 2003.

ETSI. TS 102 822-2 V1.4.1 *Broadcast and On-line Services: Search, select, and rightful use of content on personal storage systems ("TV-Anytime"); Part 2: Phase 1 - System description*. European Telecommunications Standards Institute, 2007.

EVAIN, J. P.; et. al. *TV-Anytime Phase 1 – a decisive milestone in open standards for Personal Video Recorders*. In: [broadcastpapers.com](http://broadcastpapers.com), 2003.

FERNÁNDEZ, Y.; ARIAS, J.; NORES, M.; SOLLA, A.; CABRER, M. *AVATAR: An Improved Solution for Personalized TV based on Semantic Inference*. IEEE Trans. on Consumer Electronics: Vol. 52(1), Fevereiro, 2006.

FREEBASE. 2012, <http://www.freebase.com/>, Acessado em 11/06/2012.

GRUBER, T. *A translation approach to portable ontology specifications*. Knowledge Acquisition 5:199-220; 1993

GUINNESS, D.; HARMELEN, F. *OWL Web Ontology Language Overview*. W3C Recommendation, 2004, <http://www.w3.org/TR/owl-features/>. Acessado em 23/07/2009.

HALPIN, H.; HAYES P. J.; MCCUSKER J. P.; MCGUINNESS, D. L.; THOMPSON H. S. *When owl:sameAs Isn't the Same: An Analysis of Identity in Linked Data*. In: International Semantic Web Conference (1), 2010.

HASKELL, B.; PURI, A.; NETRAVALI A. *Digital video: an introduction to MPEG-2*. Journal of Electronic Imaging 7, 1998.

INFORMA TELECOMS & MEDIA. *Americas TV - 15th edition*. Setembro de 2011. Disponível em <https://commerce.informatm.com/reports/americas-tv-15th-edition.html>. Acessado em 10/01/2012.

ISO/IEC 21000-1. *Multimedia framework – part 1: Vision, Technologies and Strategy*. 2002.

ISO/IEC 21000-2. *Information technology – multimedia framework – part 2: Digital item declaration*. 2003.

ISO/IEC 21000-3. *Information technology – multimedia framework – part 3: Digital item identification*. 2003.

ISO/IEC 21000-4. *Information technology – multimedia framework – part 4: Intellectual property management and protection*. 2003.

ISO/IEC 21000-5 (N5939). *MPEG-21 multimedia framework – part 5: Rights expression language (REL)(fds)*. 2003.

ISO/IEC 21000-6 (N5842). *MPEG-21 multimedia framework – part 6: Rights data dictionary (RDD)(fds)*. 2003.

ISO/IEC 21000-7 (N5845). *MPEG-21 multimedia framework – part 7: Digital item adaptation (fcd)*. 2003.

ISO/MPEG N4206. *Multimedia content description interface – part 6 reference software*. MPEG Systems Group, 2001.

ISO/MPEG N4224. *Multimedia content description interface – part 4 audio*. MPEG Systems Group, 2001.

ISO/MPEG N4242. *Multimedia content description interface – part 5 multimedia description schemes*. MPEG Systems Group, 2001.



ISO/MPEG N4285. *Multimedia content description interface – part 1 systems*. MPEG Systems Group, 2001.

ISO/MPEG N4288. *Multimedia content description interface – part 2 description definition language*. MPEG Systems Group, 2001.

ISO/MPEG N4358. *Multimedia content description interface – part 3 visual*. MPEG Systems Group, 2001.

JAVATV. *Overview (Java TV API 1.1)*, 2012, <http://docs.oracle.com/javame/config/cdc/opt-pkgs/api/jsr927/index.html>, Acessado em 21/6/2012.

KIM, J.; KANG, S. *An ontology-based personalized target advertisement system on interactive TV*. 2011 IEEE International Conference on Consumer Electronics, Las Vegas, EUA, pp. 9-12, Janeiro, 2011.

KULESZA, R.; FERREIRA, J.; FILHO, S.; LÍVIO, A.; BRANDÃO, R.; ARAÚJO, J.; FILHO, G. *Ginga-J: Implementação de Referência do Ambiente Imperativo do Middleware Ginga*, Belo Horizonte: Webmedia 2010 – Simpósio Brasileiro de Sistemas Multimídia e Web, 2010.

LEE, H-K.; et. al. *An Effective Architecture Of Advanced TV Set-Top And Interactive Services Using TV-Anytime Metadata*, Conference Publication of the International Broadcasting Convention (IBC), Amsterdam, 2002.

LINKEDMDB. 2012, <http://www.linkedmdb.org/>, Acessado em 12/06/2012.

LUGMAYR, A.; NIIRANEN, S.; KALLI, S. *Digital Interactive TV and Metadata: future broadcast multimedia*, p.31, Springer, 2004.

MCPARLAND, A.; MORRIS, J.; LEBAN, M.; PARNALL, S.; HICKMAN, A.; ASHLEY, A.; HAATAJA, M.; JONG, F.. *myTV: A Practical Implementation of TV-Anytime on DVB-S and the Internet*. Conference Publication of the International Broadcasting Convention (IBC), Amsterdam, 2001.

MCPARLAND, A. *TV-Anytime – using all that extra data*. R&D White Paper, BBC, 2002.

MORRIS, S; SMITH-CHAIGNEAU, A. *Interactive TV Standards: A Guide to MHP, OCAP, and JavaTV*, p.32, Focal Press, 2005.

MQL. *Metaweb Query Language – Freebase API Reference*, 2012, <http://www.freebase.com/docs/mql>, Acessado em 17/06/2012.

NOY, N.; MUSEN, M. *PROMPT: Algorithm and Tool for Automated Ontology Merging and Alignment*. In: Proc. of the 17<sup>th</sup> National Conference on Artificial Intelligence (AAAI), Austin-TX, EUA, 2000.

OWL. *OWL Web Ontology Language Reference*, 2012, <http://www.w3.org/TR/owl-ref/>, Acessado em 11/06/2012.

PAPADIMITRIOU, A.; ANAGNOSTOPOULOS, C.; TSETSOS, V.; PASKALIS, S.; HADJIEFTHYMIADES, S. *A Semantics-aware Platform for Interactive TV Services*. In: Proc. of I-MEDIA'07 Conference, Graz, Austria, Setembro, 2007.

PINTO, H.; GOMEZ-PEREZ, A.; MARTINS, J. *Some Issues on Ontology Integration*. In Proc. of IJCAI99's Workshop on Ontologies and Problem Solving Methods: Lessons Learned and Future Trends, 1999.

PROTÉGÉ. *What is Protégé?*, 2010, <http://protege.stanford.edu/overview/>, Acessado em 12/10/2010.

PROTEGE ONTOLOGY LIBRARY. 2012, [http://protegewiki.stanford.edu/wiki/Protege\\_Ontology\\_Library](http://protegewiki.stanford.edu/wiki/Protege_Ontology_Library), Acessado em 25/05/2012.

REDONDO-GARCIA, J.L.; VALIENTE-ROCHA, P.; LOZANO-TELLO, A. *Ontology-based system for content management in Digital Television: System with media content extraction and ontology-based storage functionalities*. 5th Iberian Conference on Information Systems and Technologies (CISTI), Santiago de Compostela, Espanha, pp. 16–19, Setembro, 2010.

REVISTA ELETRÔNICA. 2012, <http://xmltv.revistaeletronica.com.br/wp/>, Acessado em 05/06/2012.

REY-LÓPEZ, M.; FERNÁNDEZ-VILAS, A.; DÍAZ-REDONDO, R.; LÓPEZ-NORES M.; PAZOS-ARIAS, J.; GIL-SOLLA A.; RAMOS-CABRER M.; GARCÍA-DUQUE J. *Enhancing TV programmes with additional contents using MPEG-7 segmentation information*. International Journal Expert Systems with Applications 37 (1124–1133), 2010.

REY-LÓPEZ, M.; DÍAZ-REDONDO, R.; FERNÁNDEZ-VILAS, A.; PAZOS-ARIAS, J. *Entercation Experiences: Engaging Viewers in Education through TV Programs*. In: Proc. of 4th European Conference on Interactive Television (EuroITV 2006), Athens, Greece, 2006.

ROVIRA, M.; GONZÁLEZ, J.; LÓPEZ, A.; MAS, J.; PUIG, A.; FABREGAT, J.; FERNANDEZ, G. *IndexTV: A MPEG-7 Based Personalized Recommendation System for Digital TV*. In: Proc. of IEEE Intl. Conf. on Multimedia and Expo, 2004.

SPARQL. 2012, <http://www.w3.org/TR/rdf-sparql-query/>, Acessado em 11/06/2012.

THUY, P.; LEE, Y.; LEE, S. *DTD2OWL: Automatic Transforming XML Documents into OWL Ontology*. 2<sup>nd</sup> International Conference on Interaction Sciences: Information Technology, Culture and Human, ISBN: 978-1-60558-710-3 (16-18), Agosto, 2009.

TSINARAKI, C.; POLYDOROS, P.; CHRISTODOULAKIS, S. *Integration of OWL ontologies in MPEG-7 and TV-Anytime compliant Semantic Indexing*, In: Proc. of 16th International Conference on Advanced Information Systems Engineering, 2004.

TSINARAKI, C.; POLYDOROS, P.; KAZASIS, F.; CHRISTODOULAKIS, S. *Ontology-Based Semantic Indexing for MPEG-7 and TV-Anytime Audiovisual Content*. Special issue of Multimedia Tools and Applications Journal on Video Segmentation for Semantic Annotation and Transcoding 26 (299-325), 2005.

USCHOLD, M.; GRÜNINGER, M. *Ontologies and semantics for seamless connectivity*. SIGMOD Record, 33(3), 2004.

WIKIPEDIA. 2012, <http://www.wikipedia.org/>, Acessado em 11/06/2012.

XMLTV. 2012, [http://wiki.xmltv.org/index.php/Main\\_Page](http://wiki.xmltv.org/index.php/Main_Page), Acessado em 05/06/2012.

XMLTV DTD. 2012, <http://xmltv.cvs.sourceforge.net/viewvc/xmltv/xmltv/xmltv.dtd>, Acessado em 05/06/2012.

XMLTV PROJECT. 2012, <http://wiki.xmltv.org/index.php/XMLTVProject>, Acessado em 05/06/2012.

YOOWON K.; SEUNG-BO P.; KEE-SUNG L.; GEUN-SIK J.; JAHYUN C. *Semantic Interaction TV Based on Spatial Relations and Ontology*. 2011 International Conference on Information Science and Applications (ICISA), Jeju Island, pp. 26-29, Abril, 2011.



# Apêndice A

O conteúdo completo da ontologia OWLTV utilizada na implementação da prova de conceito está exibido na Figura A.1.

```
<?xml version="1.0"?>
<rdf:RDF
xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-s#"
xmlns:owl="http://www.w3.org/2002/07/owl#"
xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">

<tv rdf:ID="tv_1">
<has_channel>

<channel rdf:ID="channel_1">
  <channel_id rdf:datatype="&xsd:string">GRC</id>
  <has_display_name>
    <display-name rdf:ID="display-name_1">
      <display-name_value rdf:datatype="&xsd:string">EPTV Campinas</display-name_value>
      <display-name_lang rdf:datatype="&xsd:string">pt</lang>
    </display-name>
  </has_display_name>
  <has_icon>
    <icon rdf:ID="icon_1">
      <icon_src rdf:datatype="&xsd:string">grc.gif</src>
    </icon>
  </has_icon>
</channel>

</has_channel>

<has_programme>

<programme rdf:ID="programme_1">
  <programme_start rdf:datatype="&xsd:string">20120621144000 -0300</start>
  <programme_stop rdf:datatype="&xsd:string">20120621155500 -0300</stop>
  <programme_channel rdf:datatype="&xsd:string">GRC</channel>
  <programme_program_id rdf:datatype="&xsd:string">0000109098</program_id>
  <has_title>
    <title rdf:ID="title_1">
      <title_value rdf:datatype="&xsd:string">Chocolate com Pimenta</title_value>
      <title_lang rdf:datatype="&xsd:string">pt</title_lang>
    </title>
  </has_title>
  <has_desc>
    <desc rdf:ID="desc_1">
      <desc_value rdf:datatype="&xsd:string">Sinopse Não Disponível</desc_value>
    </desc>
  </has_desc>
  <has_category>
    <category rdf:ID="category_1">
      <category_value rdf:datatype="&xsd:string">Variedades</category_value>
      <category_lang rdf:datatype="&xsd:string">pt</category_lang>
    </category>
  </has_category>
</programme>
</has_programme>
</tv>
</rdf:RDF>
```

```

<category rdf:ID="category_2">
  <category_value rdf:datatype="&xsd:string">Novela</category_value>
  <category_lang rdf:datatype="&xsd:string">pt</category_lang>
</category>
</has_category>
<has_rating>
  <rating rdf:ID="rating_1">
    <rating_system rdf:datatype="&xsd:string">Advisory</rating_system>
    <has_rating_value>
      <rating_value rdf:ID="rating_value_1">
        <rating_value_value rdf:datatype="&xsd:string">Programa livre para todas as
idades</rating_value_value>
      </rating_value>
    </has_rating_value>
  </rating>
</has_rating>

<sameAsQuery rdf:ID="sameAsQuery_1">
  <owltv_entity rdf:datatype="&xsd:string">programme_1</owltv_entity>
  <linked_ontology rdf:datatype="&xsd:string">Freebase</linked_ontology>
  <query_code rdf:datatype="&xsd:string">
    [{
  "id": null,
  "name|=":[ "Chocolate com Pimenta" ],
  "type|=":[ "/tv/tv_program", "/film/film" ]
    }]
  </query_code>
</sameAsQuery>
<owl:sameAs rdf:resource="http://rdf.freebase.com/ns/en.chocolate_com_pimenta"/>

<sameAsQuery rdf:ID="sameAsQuery_2">
  <owltv_entity rdf:datatype="&xsd:string">programme_1</owltv_entity>
  <linked_ontology rdf:datatype="&xsd:string">DBpedia</linked_ontology>
  <query_code rdf:datatype="&xsd:string">
    PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
    SELECT ?entity
    WHERE { ?entity rdfs:label "Chocolate com Pimenta"@pt }
  </query_code>
</sameAsQuery>
<owl:sameAs rdf:resource="http://dbpedia.org/resource/Chocolate_com_Pimenta"/>

<sameAsQuery rdf:ID="sameAsQuery_3">
  <owltv_entity rdf:datatype="&xsd:string">programme_1</owltv_entity>
  <linked_ontology rdf:datatype="&xsd:string">LinkedMDB</linked_ontology>
  <query_code rdf:datatype="&xsd:string">
    PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
    SELECT ?entity
    WHERE { ?entity rdfs:label "Chocolate com Pimenta" }
  </query_code>
</sameAsQuery>
</programme>

<programme rdf:ID="programme_2">
  <programme_start rdf:datatype="&xsd:string">20120621155500 -0300</start>

```

```
<programme_stop rdf:datatype="&xsd:string">20120621174500 -0300</stop>
<programme_channel rdf:datatype="&xsd:string">GRC</channel>
<programme_program_id rdf:datatype="&xsd:string">0000092443</program_id>
<has_title>
  <title rdf:ID="title_2">
    <title_value rdf:datatype="&xsd:string">Táxi</title_value>
    <title_lang rdf:datatype="&xsd:string">pt</title_lang>
  </title>
  <title rdf:ID="title_2">
    <title_value rdf:datatype="&xsd:string">Taxi</title_value>
    <title_lang rdf:datatype="&xsd:string">en</title_lang>
  </title>
</has_title>
<has_desc>
  <desc rdf:ID="desc_2">
    <desc_value rdf:datatype="&xsd:string">Atrapalhado policial quer recuperar seu prestígio. Com a ajuda da taxista mais rápida de New York, ele tenta deter a gangue liderada pela bela e perigosa Vanessa.</desc_value>
  </desc>
</has_desc>
<has_credits>
  <credits rdf:ID="credits_1">
    <has_director>
      <director rdf:ID="director_1">
        <director_value rdf:datatype="&xsd:string">Tim Story</director_value>
      </director>
    </has_director>
    <has_actor>
      <actor rdf:ID="actor_1">
        <actor_value rdf:datatype="&xsd:string">Queen Latifah</actor_value>
      </actor>
      <actor rdf:ID="actor_2">
        <actor_value rdf:datatype="&xsd:string">Jimmy Fallon</actor_value>
      </actor>
      <actor rdf:ID="actor_3">
        <actor_value rdf:datatype="&xsd:string">Henry Simmons</actor_value>
      </actor>
      <actor rdf:ID="actor_4">
        <actor_value rdf:datatype="&xsd:string">Jennifer Esposito</actor_value>
      </actor>
      <actor rdf:ID="actor_5">
        <actor_value rdf:datatype="&xsd:string">Gisele Bündchen</actor_value>
      </actor>
      <actor rdf:ID="actor_6">
        <actor_value rdf:datatype="&xsd:string">Magali Amadei</actor_value>
      </actor>
      <actor rdf:ID="actor_7">
        <actor_value rdf:datatype="&xsd:string">Ann-Margret</actor_value>
      </actor>
      <actor rdf:ID="actor_8">
        <actor_value rdf:datatype="&xsd:string">Boris McGiver</actor_value>
      </actor>
      <actor rdf:ID="actor_9">
        <actor_value rdf:datatype="&xsd:string">Bryna Weiss</actor_value>
      </actor>
    </has_actor>
  </credits>
</has_credits>
```

```

</credits>
</has_credits>
<has_date>
  <date rdf:ID="date_1">
    <date_value rdf:datatype="&xsd:string">2004</date_value>
  </date>
</has_date>
<has_category>
  <category rdf:ID="category_3">
    <category_value rdf:datatype="&xsd:string">Filme</category_value>
    <category_lang rdf:datatype="&xsd:string">pt</category_lang>
  </category>
  <category rdf:ID="category_4">
    <category_value rdf:datatype="&xsd:string">Comédia</category_value>
    <category_lang rdf:datatype="&xsd:string">pt</category_lang>
  </category>
</has_category>
<has_country>
  <country rdf:ID="country_1">
    <country_value rdf:datatype="&xsd:string">EUA/França</country_value>
  </country>
</has_country>
<has_video>
  <video rdf:ID="video_1">
    <has_colour>
      <colour rdf:ID="colour_1">
        <colour_value rdf:datatype="&xsd:string">yes</colour_value>
      </colour>
    </has_colour>
  </video>
</has_video>
<has_rating>
  <rating rdf:ID="rating_2">
    <rating_system rdf:datatype="&xsd:string">Advisory</rating_system>
    <has_rating_value>
      <rating_value rdf:ID="rating_value_2">
        <rating_value_value rdf:datatype="&xsd:string">Programa impróprio para menores de 14
anos</rating_value_value>
      </rating_value>
    </has_rating_value>
  </rating>
</has_rating>
<has_star-rating>
  <star-rating rdf:ID="star-rating_1">
    <has_star-rating_value>
      <star-rating_value rdf:ID="star-rating_value_1">
        <star-rating_value_value rdf:datatype="&xsd:string">2/5</star-rating_value_value>
      </star-rating_value>
    </has_star-rating_value>
  </star-rating>
</has_star-rating>

<sameAsQuery rdf:ID="sameAsQuery_4">
  <owltv_entity rdf:datatype="&xsd:string">programme_2</owltv_entity>

```



```

<linked_ontology rdf:datatype="&xsd:string">Freebase</linked_ontology>
<query_code rdf:datatype="&xsd:string">
  [{
    "id": null,
    "name|=":["Táxi", "Taxi"],
    "type|=":["/tv/tv_program", "/film/film"]
  }]
</query_code>
</sameAsQuery>
<owl:sameAs rdf:resource="http://rdf.freebase.com/ns/en.taxi"/>
<owl:sameAs rdf:resource="http://rdf.freebase.com/ns/en.taxi_1998"/>
<owl:sameAs rdf:resource="http://rdf.freebase.com/ns/en.taxi_2004"/>
<owl:sameAs rdf:resource="http://rdf.freebase.com/ns/m.04p4dsc"/>
<owl:sameAs rdf:resource="http://rdf.freebase.com/ns/m.063hb49"/>
<owl:sameAs rdf:resource="http://rdf.freebase.com/ns/m.09rr393"/>
<owl:sameAs rdf:resource="http://rdf.freebase.com/ns/m.0cnx5wm"/>

<sameAsQuery rdf:ID="sameAsQuery_5">
  <owltv_entity rdf:datatype="&xsd:string">programme_2</owltv_entity>
  <linked_ontology rdf:datatype="&xsd:string">DBpedia</linked_ontology>
  <query_code rdf:datatype="&xsd:string">
    PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
    PREFIX dbpprop: <http://dbpedia.org/property/>
    PREFIX dbpedia-owl: <http://dbpedia.org/ontology/>
    SELECT ?entity
    WHERE { ?entity dbpprop:title ?title ;
            rdf:type dbpedia-owl:Film .
            FILTER (?title = "Táxi"@pt || ?title = "Taxi"@en) }
  </query_code>
</sameAsQuery>
<owl:sameAs rdf:resource="http://dbpedia.org/resource/Taxi_%281998_film%29"/>
<owl:sameAs rdf:resource="http://dbpedia.org/resource/Taxi_%282004_film%29"/>
<owl:sameAs rdf:resource="http://dbpedia.org/resource/Taxi_%281996_film%29"/>

<sameAsQuery rdf:ID="sameAsQuery_6">
  <owltv_entity rdf:datatype="&xsd:string">programme_2</owltv_entity>
  <linked_ontology rdf:datatype="&xsd:string">LinkedMDB</linked_ontology>
  <query_code rdf:datatype="&xsd:string">
    PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
    SELECT ?entity
    WHERE { ?entity rdfs:label ?title ;
            FILTER (?title = "Táxi" || ?title = "Taxi") }
  </query_code>
</sameAsQuery>
<owl:sameAs rdf:resource="http://data.linkedmdb.org/resource/film/863"/>
<owl:sameAs rdf:resource="http://data.linkedmdb.org/resource/film/925"/>
<owl:sameAs rdf:resource="http://data.linkedmdb.org/resource/film/926"/>

</programme>
</has_programme>
</tv>

```

**Figura A.1 – Ontologia OWLTV utilizada na implementação da prova de conceito**