Universidade do Estado do Rio de Janeiro
Centro de Tecnologia e Ciências
Escola Superior de Desenho Industrial

Leonardo Burlamaqui Lima da Rocha

Ontologia de notícias: um modelo para classificação do conteúdo dos jornais on-line brasileiros, segundo a lógica da Web Semântica

Rio de Janeiro
2012

Leonardo Burlamaqui Lima da Rocha

Ontologia de notícias: um modelo para classificação do conteúdo dos jornais on-line brasileiros, segundo a lógica da Web Semântica

Dissertação apresentada, como requisito parcial para obtenção do título de Mestre, ao Programa de Pós-graduação em Design, da Universidade do Estado do Rio de Janeiro. Área de concentração: Design e Tecnologia.

Orientador: Prof. Dr. Washington Dias Lessa

Co-orientador: Prof. Dr. André Soares Monat

Rio de Janeiro

2012

CATALOGAÇÃO NA FONTE
UERJ/REDE SIRIUS/BIBLIOTECA CTC/G

R672 Rocha, Leonardo Burlamaqui Lima da.

Ontologia de notícias : um modelo para classificação do conteúdo dos jornais online brasileiros, segundo a lógica da web semântica / Leonardo Burlamaqui Lima da Rocha. - 2012.
124 f. : il.

Orientador: Washington Dias Lessa.
Coorientador: André Soares Monat.
Dissertação (Mestrado). Universidade do Estado do Rio de Janeiro, Escola Superior de Desenho Industrial.
1. Jornalismo eletrônico - Teses. 2. Jornais eletrônicos - Teses. 3. Ontologia- Teses. 4. Web semântica I. Lessa, Washington Dias. II. Monat, André Soares. III. Universidade do Estado do Rio de Janeiro. Escola Superior de Desenho Industrial. IV. Título.

CDU 070:004.738.52

Autorizo, apenas para fins acadêmicos e científicos, a reprodução total ou parcial desta dissertação, desde que citada a fonte.

_____ _____

Assinatura Data

Leonardo Burlamaqui Lima da Rocha

Ontologia de notícias: um modelo para classificação do conteúdo dos jornais on-line brasileiros, segundo a lógica da Web Semântica

Dissertação apresentada, como requisito parcial para obtenção do grau de Mestre, ao Programa de Pós-graduação em Design, da Universidade do Estado do Rio de Janeiro. Área de concentração: Design e Tecnologia.

Aprovada em _____

Banca Examinadora:

_____
Prof. Dr. Washington Dias Lessa (Orientador)
Escola Superior de Desenho Industrial da UERJ

_____
Prof. Dr. André Soares Monat (Co-orientador)
Escola Superior de Desenho Industrial da UERJ

_____
Prof[a]. Dr[a]. Lucy Carlinda da Rocha de Niemeyer
Escola Superior de Desenho Industrial da UERJ

_____
Prof. Dr. Giancarlo Guizzardi
Universidade Federal do Espírito Santo

Rio de Janeiro

2012

DEDICATÓRIA

Dedico este trabalho a todos os designers, que, de alguma forma, estão envolvidos com tecnologias que nos levam à fronteira do que se entende por Design.

Dedico-o também a meus pais, Claudio Raimundo e Vania.

E, como não poderia deixar de ser, à minha mulher, Nádia Maria.

AGRADECIMENTOS

Agradeço à minha família e aos meus amigos pelo incansável apoio.

Ao Washington pela orientação e ao André pela co-orientação.

Ao Giancarlo e à Lucy por me darem a honra de fazerem parte da minha Banca Examinadora.

À ESDI que tanto amo, graças às pessoas que desta instituição fazem parte.

À Globo.com, que me incentivou e me permitiu dar início a este trabalho.

Aos autores dos trabalhos citados.

RESUMO

ROCHA, Leonardo Burlamaqui Lima da. **Ontologia de notícias**: um modelo para classificação do conteúdo dos jornais on-line brasileiros, segundo a lógica da Web Semântica. 2012. 124f. Dissertação (Mestrado em Design) – Escola Superior de Desenho Industrial, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 2012.

Este trabalho tem por objetivo propor um modelo de ontologia simples e generalista, capaz de descrever os conceitos mais básicos que permeiam o domínio de conhecimento dos jornais on-line brasileiros não especializados, fundamentado tanto na prática quanto conceitualmente, em conformidade com os princípios da Web Semântica. A partir de uma nova forma de classificação e organização do conteúdo, a ontologia proposta deve ter condições de atender as necessidades comuns de ambas as partes, jornal e leitor, que são, resumidamente, a busca e a recuperação das informações.

Palavras-chave: Jornalismo on-line. Notícias. Classificação. Conteúdo. Ontologia. Web Semântica.

ABSTRACT

This work aims to propose a simple and generalist ontology model, able to describe the most basic concepts that underlie the knowledge domain of general purpose Brazilian on-line newspapers, based both on practice and theory, in accordance with the principles of the Semantic Web. From a new way of classifying and organizing content, the given ontology should be able to meet the common needs of both parties, journalists and readers, which can be summarized as search and information retrieval.

Keywords: Web journalism. News. Classification. Content. Ontology. Semantic Web.

LISTA DE FIGURAS

Figura 1 - Capa dos jornais impressos *O Globo* e *O Estado de S. Paulo* 28
Figura 2 - *Screenshot* da página inicial do jornal *on-line G1*, com o menu de editorias aberto 29
Figura 3 - *Screenshot* da página inicial do jornal *on-line Último Segundo* 30
Figura 4 - *Screenshots* da página inicial dos jornais *on-line NYT*, *BBC News*, *MSNBC.COM* e *CNN* 33
Figura 5 - Capa do livro de Christian Wolff, *Philosophia Prima Sive Ontologia*, de 1730 46
Figura 6 - Primeira página da obra *Categorias*, de Aristóteles 52
Figura 7 - Vocabulário visual para leitura dos diagramas de representação das ontologias 60
Figura 8 - *UFO-A*: distinção entre Indivíduo e Universal 61
Figura 9 - *UFO-A*: Indivíduo 62
Figura 10 - *UFO-A*: Universal 63
Figura 11 - *UFO-A*: Aspecto e Substancial 64
Figura 12 - *UFO-A*: Substancial 65
Figura 13 - *UFO-A*: Substancial Universal 66
Figura 14 - *UFO-A*: Aspecto 68
Figura 15 - *UFO-A*: Relação 70
Figura 16 - *UFO-B*: Perdurante 73
Figura 17 - *C*omparação entre as relações de intervalo temporal 74
Figura 18 - *UFO-C*: distinção entre Agente e Objeto 75
Figura 19 - *UFO-C*: Aspecto Intencional 76
Figura 20 - *UFO-C*: Comprometimento 78
Figura 21 - *UFO-C*: Delegação 79
Figura 22 - *UFO-C*: Ação 80
Figura 23 - Ontologia de notícias: Universais 86
Figura 24 - Ontologia de notícias: Indivíduos 94
Figura 25 - *Screenshot* da página de uma das matérias utilizadas pelo protótipo 101
Figura 26 - Trecho do código gerado na implementação da ontologia de notícias 102

LISTA DE ABREVIATURAS E SIGLAS

BBC	*British Broadcasting Corporation*
BFO	*Basic Formal Ontology*
BWW	*Bunge-Wand-Weber*
CMS	*Content Management System*
DL	*Description Logics*
DOLCE	*Descriptive Ontology for Linguistic and Cognitive Engineering*
GFO	*General Formal Ontology*
GOL	*General Ontology Language*
GUM	*Generalised Upper Model*
IA	Inteligência Artificial
NLP	*Natural Language Processing*
NYT	*The New York Times*
OWL	*Ontology Web Language*
RDF	*Resource Description Framework*
RDFS	*RDF Schema*
SPARQL	*SPARQL Protocol and RDF Query Language*
SUMO	*Suggested Upper Merged Ontology*
TCP/IP	*Transmission Control Protocol / Internet Protocol*
Turtle	*Terse RDF Triple Language*
UFO	*Unified Foundational Ontology*
URI	*Uniform Resource Identifier*
W3C	*World Wide Web Consortium*
Web	*World Wide Web*
WWW	*World Wide Web*
XML	*Extensible Markup Language*

SUMÁRIO

	INTRODUÇÃO	13
1	**JORNALISMO ON-LINE**	17
1.1	**Do jornalismo impresso ao on-line**	17
1.2	**Cenário atual**	20
1.3	**Novas iniciativas**	32
2	**WEB SEMÂNTICA**	38
2.1	**Da Web atual à Web Semântica**	38
2.2	**A Web Semântica**	39
2.3	**Web Semântica e Inteligência Artificial**	40
3	**ONTOLOGIA**	45
3.1	**Origem do termo ontologia**	45
3.2	**Ontologia na Ciência da Computação**	46
3.3	**Objetivos e características da ontologia**	48
3.4	**Padrões de linguagem**	50
3.5	**Primeiro modelo de ontologia**	51
3.6	**Ontologias de fundamentação e de domínio**	54
3.7	**Ontologia de referência**	57
3.7.1	UFO-A	61
3.7.2	UFO-B	71
3.7.3	UFO-C	74
4	**PROPOSTA DE SOLUÇÃO**	81
4.1	**Passos iniciais**	81
4.2	**Processo de conceituação da ontologia**	83
4.3	**Universais**	84
4.4	**Lead da notícia como conjunto mínimo de questões de competência**	89
4.5	**Tipos de indivíduos**	93
4.6	**Relações entre entidades**	95
4.7	**Considerações**	98
5	**RESULTADOS ALCANÇADOS**	100
6	**CONCLUSÃO E FUTUROS TRABALHOS**	105
	REFERÊNCIAS	108

GLOSSÁRIO 114
ANEXO A - SCREENSHOT DA MATÉRIA SENADO ABRE SINDICÂNCIA PARA APURAR POSSÍVEL NOVA ‘FARRA DE PASSAGENS’, DO G1 . 119
ANEXO B - SCREENSHOT DA PÁGINA DE TÓPICOS DO NYT 120
ANEXO C - SCREENSHOT DA PÁGINA DE TÓPICOS DO ESTADÃO . 121
ANEXO D - FRAGMENTO DE UFO-A: ONTOLOGIA DE ENDURANTES 122
ANEXO E - LISTA DOS ASSUNTOS DO G1 INICIADOS COM A LETRA A 123
ANEXO F - TRECHO DO CÓDIGO, QUE TRATA DOS INDIVÍDUOS DO TIPO AÇÃO NOTICIOSA, RESULTANTE DA PROTOTIPAÇÃO DA ONTOLOGIA DE NOTÍCIAS 124

INTRODUÇÃO

Desde a sua concepção, a *Internet** se caracteriza pela capacidade de conectar redes, documentos, dados e *softwares* acessíveis em qualquer lugar do mundo, a qualquer hora. Ao longo dos anos, a quantidade de informações disponíveis na *World Wide Web - WWW** cresceu vertiginosamente, assim como o número de pessoas que nela participam. Do mesmo modo, inúmeras aplicações – de todos os tipos e propósitos – vêm surgindo, conectando os mais diversos dispositivos à rede, fornecendo, buscando e trocando informações *on-line*, sendo cada vez mais necessária a adoção de padrões que propiciem a sua interoperabilidade.

Acompanhando o crescimento da *Web**, a produção de *conteúdo digital**, ou simplesmente *conteúdo*, nos jornais *on-line* sofreu um aumento expressivo, o que acarretou os mesmos problemas – experimentados, diariamente, pelos usuários de Internet – às empresas e aos seus leitores *on-line*: dificuldades para buscar, disponibilizar e relacionar informações. Contribuiu também para isso o fato do conteúdo, de modo geral, ser organizado da mesma forma que o seu similar impresso, ou seja, em *editorias** ou cadernos, muitas vezes contrariando a provável necessidade do leitor.

Como ponto de partida para o trabalho, o atual panorama dos jornais *on-line* brasileiros será analisado, no que concerne à classificação do seu conteúdo, para que seja possível identificar os problemas mais pontuais acarretados pela utilização limitada dos recursos próprios da Web e, desta forma, evidenciar a necessidade de reorganização do conteúdo jornalístico.

Em seguida, a partir de uma indicação do cenário mundial, serão levantadas algumas das iniciativas mais representativas e bem recebidas no âmbito internacional que visam a potencialização do jornalismo *on-line* – seja através da aplicação sistemática de conceitos inovadores de classificação do conteúdo digital ou de novas técnicas de rotulação/anotação da informação – ampliando as dimensões da estrutura narrativa e, por conseguinte, propiciando ao leitor uma experiência de imersão mais rica e profunda durante o consumo do conteúdo.

---

* Os termos em itálico sucedidos de asterisco (*) estão definidos no Glossário.

Concluído o levantamento, será proposta uma solução adequada à realidade dos jornais *on-line* brasileiros, capaz de ser implantada dentro de seus produtos, trazendo os subsídios necessários para que os resultados sejam compatíveis com aqueles alcançados pelas exemplares iniciativas externas que, tanto do ponto de vista funcional quanto do tecnológico, estabeleceram padrões consoantes o paradigma da Web. Para tal, os parâmetros da *Web Semântica*[1] servirão como base tecnológica a ser adotada, pois dá as condições necessárias à otimização do fluxo de trabalho nas empresas do setor jornalístico – desde a criação do conteúdo até a sua disponibilização e possível reutilização – e torna mais fácil, rápido e preciso o processo de obtenção das informações pelos leitores, mostrando-se, assim, a escolha mais acertada.

A *Web Semântica** vem, se não para solucionar, para atenuar o caos advindo da explosão de informações na *Grande Rede**, funcionando como uma espécie de extensão da Web atual capaz de atribuir sentido aos conteúdos publicados na Internet. Ao dar significado às informações, as máquinas serão capazes de "compreendê-las" e, deste modo, ajudar no uso e gerenciamento do conteúdo disponível na Web, fazendo com que seres humanos e computadores tenham condições de trabalhar num nível ainda maior de cooperação.

Sendo a *ontologia**, na Web Semântica, o aparato de representação formal e explícita do conhecimento – que se caracteriza, entre outras coisas, pelo uso de um vocabulário controlado e estruturado onde o significado de cada uma das palavras é preservado – a sua construção é de responsabilidade não apenas dos profissionais da *Ciência da Computação** e da *Ciência da Informação**, mas também daqueles que, de alguma forma, têm de lidar com o somatório de efeitos produzidos a partir da sua utilização como, por exemplo, no âmbito da comunicação, da interação e da experiência de uso do produto ou serviço a que se destina.

Entendendo que a ontologia faz parte da camada de organização das informações, a interface, de modo complementar, faz parte da camada de distribuição e, principalmente, de visualização destas informações. Consequentemente, o designer se mostra fundamental ao longo do trabalho de conceituação e construção da ontologia, visto que o seu papel é verificar se todas as

[1] Refere-se ao arcabouço tecnológico, ou seja, ao seu entendimento como artefato, e não à Web Semântica como disciplina ou área de conhecimento.

*questões de competência** (GRÜNINGER, 1994) – que correspondem, na área de engenharia de ontologias, ao resultado da especificação de requisitos, tendo por objetivo avaliar o compromisso ontológico estabelecido na definição da ontologia – estas levantadas no plano da distribuição e visualização do conteúdo, estão sendo devidamente respondidas por este modelo de representação, sem que apresente complexidades desnecessárias. E, mais do que isso, a função do designer é entender as potencialidades que podem vir a ser exploradas pelo produto, tanto no nível da comunicação visual quanto do próprio negócio, por conta das particularidades das tecnologias envolvidas na Web Semântica.

Pelo seu caráter multidisciplinar e, principalmente, pela sua constante preocupação com o usuário, o Design representa uma área de conhecimento que pode, e muito, se envolver com questões que são objeto de estudo de outras áreas como, no caso deste trabalho, a construção de ontologias, que é de domínio das Ciências da Informação e da Computação. É importante esclarecer que a inserção do Design nessa esfera tem por objetivo tão somente colaborar no trabalho desenvolvido pelos *engenheiros de ontologia**, que são os responsáveis pela conceituação e construção de modelos ontológicos – sendo estes os profissionais que estudam os métodos e metodologias voltados à construção de ontologias, tendo como objetivo tornar explícito o conhecimento contido nas aplicações de *software*, assim como nos procedimentos organizacionais e de negócio para um domínio particular – estabelecendo um ponto de contato entre estes profissionais e os designers. Desse modo, é possível garantir que a ontologia seja o exato reflexo daquilo que se pretende estabelecer em termos de experiência de uso – ou, melhor dizendo, de experiência de consumo – das informações nela armazenadas, e vice-versa.

A título de exemplo, vale dizer que o autor deste trabalho, na condição de *arquiteto de informação** (PINHEIRO, 2007) – que, por mais de três anos, fez parte do departamento de *Busca & Semântica* da *Globo.com* e, atualmente, é responsável por um grupo de estudo, denominado *Comitê de Semântica*, dentro da companhia – tem como atribuição oferecer instrumentos, estes muitas vezes baseados em tecnologias semânticas, que promovam melhorias às tarefas de busca, recuperação e distribuição do conteúdo gerado pelos produtos jornalísticos *on-line* que integram as *Organizações Globo* – entre eles, *G1*, *Globoesporte.com*, *Ego* e *TVG* – de modo a tornar a sua operação mais rápida e eficiente, e disponibilizar informações mais

precisas e relevantes para os seus leitores, que, por conseguinte, serão capazes de obtê-las mais facilmente. Na prática, isso se traduz em apoio às equipes responsáveis pela criação e manutenção desses produtos, a partir da construção de ontologias que atendam as necessidades de classificação e organização do seu conteúdo; e, com menor frequência, na elaboração de suas interfaces, funcionalidades e interações, o que vem a constituir aquilo que se entende por *experiência de uso* ou, segundo Pinheiro (2007), *Design da Experiência**.

Assim sendo, a proposta de solução a ser apresentada consistirá na elaboração de um modelo simples e resumido de *ontologia de domínio**, cuja definição e características serão vistas adiante, capaz de formalizar os conceitos mais básicos que permeiam as áreas de conhecimento dos jornais *on-line* brasileiros não especializados, fundamentado tanto na prática quanto conceitualmente, sem perder de vista as soluções adotadas no mundo, sejam estas no mesmo campo de atuação ou não. Dentre os desafios encontrados, o mais complexo é justamente a construção desse modelo, dado que este deve ser um reflexo da nova organização do conteúdo, capaz de atender as necessidades comuns de ambas as partes, jornal e leitor, que são, resumidamente, a busca e a recuperação das informações.

Para que seja possível atestar a eficácia do modelo, um protótipo – cujos critérios utilizados estão expostos em um capítulo específico – será apresentado como parte integrante deste trabalho, no momento da defesa da dissertação. Uma vez colhidos os resultados da experiência, estes serão inseridos posteriormente no presente documento, de modo que se registre o que pela Banca Examinadora será apreciado.

Cabe aqui informar que este trabalho inclui um glossário com o intuito de clarificar o significado de alguns dos termos utilizados, estes grifados em itálico e sucedidos de asterisco, conforme já pôde ser observado nas páginas que compõem a Introdução.

1 Jornalismo on-line

1.1 Do jornalismo impresso ao on-line

No país, apesar da Internet já estar presente e disponível ao público desde 1995, após finalmente ter ultrapassado as fronteiras acadêmicas, pouco se fez em relação à classificação do conteúdo dos jornais *on-line* brasileiros. No início, muitos veículos optaram pela simples transposição de conteúdos impressos para o *on-line*, sem tratamento editorial e mercadológico; os jornais impressos, caracterizados pela composição em capa e cadernos e pela limitação de espaço (do papel), tinham suas edições transferidas para a versão *on-line* sem qualquer alteração do conteúdo. Hoje, embora essa transposição não mais ocorra, os jornais *on-line* brasileiros ainda se baseiam nos jornais impressos; suas páginas iniciais assemelham-se às capas de jornal em papel e as matérias são associadas a *editorias** que, analogamente às versões impressas, funcionam como grandes diretórios, tais quais os cadernos de seus predecessores, sendo um reflexo da divisão/especialização do trabalho dentro dos jornais.

Sobre o jornalismo impresso, Lene (2008) diz:

> [...] todo veículo de comunicação deve ter um projeto editorial, que determina os assuntos e os ângulos a serem abordados em cada número. Ele engloba o projeto gráfico, que envolve os aspectos visuais e seu formato gráfico com as alternativas de diagramação, fontes tipológicas, posição, tamanhos dos títulos e o uso de cores e imagens. Cada uma das seções que compõem a redação de um veículo jornalístico recebe o nome de editoria (Geral, Cidades ou Dia-a-dia; Esportes; Política; Polícia; Cultural...). (LENE, 2008).

Assim sendo, no jornalismo impresso, percebe-se uma relação intrínseca entre as seções do veículo jornalístico e a sua redação. Ou seja, para cada seção há um grupo de trabalho envolvendo jornalistas, redatores, repórteres, colunistas e colaboradores, todos voltados para cobrir e analisar, única e exclusivamente, os acontecimentos relacionados aos assuntos de sua pauta. E mais: de modo geral, pode-se afirmar que o sistema de classificação do conteúdo utilizado nos jornais

impressos é resultado da forma como esses grupos de trabalho, cuja totalidade equivale à sua redação, são definidos e estruturados.

No jornalismo *on-line*, entretanto, a estrutura dos grupos de trabalho tende a ser muito mais flexível e, em alguns casos – como no *G1*, por exemplo – não há grupos claramente definidos. Desse modo, os profissionais acabam por atuar de forma ainda mais ampla e colaborativa, onde a versatilidade se sobrepõe à profundidade. Como efeito, é possível notar que (a) o sistema adotado para a classificação do conteúdo nos jornais *on-line* nada mais é que uma herança dos veículos impressos, cuja orientação já não possui o mesmo respaldo técnico-operacional, considerando que não há mais um vínculo entre a redação e as seções do veículo jornalístico, estas comumente chamadas de editorias e (b) nos jornais *on-line*, as funções dos profissionais que integram a redação se distinguem daqueles inseridos em veículos impressos, por conta da generalização – ou simplificação, dependendo do ponto de vista – decorrente do alto nível de versatilidade e do novo fluxo de operação necessários ao seu trabalho, o que vem a estabelecer tão somente dois papéis, um voltado à produção do conteúdo jornalístico e outro destinado à sua gestão.

No que concerne aos papéis existentes dentro da estrutura organizacional de um *jornal on-line**, serão utilizadas apenas as denominações *produtor de conteúdo digital**, ou simplesmente *produtor de conteúdo*, que equivale ao repórter que faz o levantamento e constrói o conteúdo digital em si, ao jornalista que redige a matéria ou que, de alguma forma, dá corpo ao fato noticioso, seja ele fotógrafo, ilustrador, infografista, videografista ou editor de áudio/vídeo; e *gestor de conteúdo digital**, ou tão somente *gestor de conteúdo*, cuja atribuição é classificar o conteúdo digital, ou garantir a sua correta classificação, mantendo-o organizado da melhor forma possível, de modo a facilitar as operações de busca e recuperação de informações.

Vale esclarecer que o conceito de editoria, que por diversas vezes será abordado ao longo do trabalho, não se refere aos grupos especializados de profissionais que, dentro de uma organização jornalística, cobre somente assuntos específicos de sua competência. O termo diz respeito tão somente ao conceito de diretório – normalmente representado sob a forma de seção – sendo este parte integrante do primeiro nível da estrutura de navegação dos *websites** de notícias.

O que aqui se pretende discutir é a forma como o conteúdo resultante do jornalismo digital é até hoje, no Brasil, tratado, categorizado e inter-relacionado. No

cenário mundial, já é possível encontrar algumas iniciativas que buscam o rompimento desse paradigma, adotando uma abordagem que explora o potencial da Web – de modo a aproveitar todos os recursos tecnológicos existentes e garantir uma série de benefícios, tanto para os produtores e gestores de conteúdo digital quanto para o seu público consumidor – distanciando-se consideravelmente da lógica dos jornais impressos.

É preciso adotar um novo modelo de solução, aproveitando-se de algumas das iniciativas internacionais – que já passaram por inúmeras adaptações e melhorias, fruto da longa experiência e da busca constante por inovações no meio digital – para, finalmente, romper com o modelo de classificação do conteúdo ainda utilizado pelos jornais *on-line* brasileiros, que, se é injusto apontá-lo como incorreto ou inapropriado, poder-se-ia dizer ultrapassado, tanto em relação à prática do mercado globalizado quanto das necessidades dos leitores e, até mesmo, dos produtores e gestores de conteúdo digital, cuja dinâmica de trabalho exige uma operação cada vez mais ágil, a partir de um volume de informações em extraordinária expansão.

Ao se estabelecer uma proposta compatível com as potencialidades que se configuram na Web, intensifica-se o *fluxo* de consumo do conteúdo, entre outras coisas, permitindo ao leitor uma experiência de imersão mais rica e profunda, sendo esta o resultado da produção de conhecimento propiciada pelas novas formas de classificação e relacionamento dos fatos noticiosos em meio digital.

Segundo Csikszentmihalyi (1990), a metáfora do fluxo tem sido utilizada para descrever a sensação de ação sem esforço que muitas pessoas sentem em momentos que se sobressaem como os melhores de suas vidas. Trata-se de uma experiência de completa imersão. O conceito de fluxo é definido como a sensação holística que as pessoas sentem ao agir com total envolvimento em determinada atividade. Quando em fluxo, o indivíduo se sente tão envolvido pela atividade que nada mais parece importar. O elemento chave dessa experiência é o fato de ser autotélica, o que significa que a atividade deve ser válida por si mesma, ou seja, não deve necessitar de uma recompensa externa.

Apesar de já existirem estudos anteriores sobre o conceito de fluxo em ambiente virtual, a primeira vez em que esta ideia foi utilizada para compreensão do comportamento do consumidor *on-line* foi no trabalho de Hoffman e Novak (1996). Desde então, diversos autores têm empregado a teoria do fluxo em seus estudos

sobre compras pela Internet e, como não poderia deixar de ser, também se aplica ao presente trabalho, entendendo que o processo de leitura e consumo de notícias *on-line* é capaz de produzir o efeito em questão, do mesmo modo como é possível ocorrer em um ambiente de compras na Web.

Tais consequências podem levar ao aumento do tempo de permanência do leitor, à expansão da oferta de espaços publicitários e de patrocínio em áreas dedicadas a nichos de consumo qualificados e, por conseguinte, ao crescimento da receita dos jornais *on-line* brasileiros, isto sem contar com a percepção de valor que poderá ser agregada às suas marcas pelo próprio público, dado o avanço destes veículos como fontes de informação.

1.2 **Cenário atual**

Para diagnóstico adequado dos jornais *on-line* no Brasil, é preciso, antes de mais nada, considerar a trajetória destes veículos até o momento presente, bem como os elementos fundamentais que os caracterizam, de modo a estabelecer uma generalização que sirva de ponto de partida para o modelo de solução que será proposto mais à frente.

Sobre o referido assunto, tomam-se por base os trabalhos de Bardoel e Deuze (2001) e Palacios (2002), que estabelecem como características do jornalismo desenvolvido para a Web a *multimidialidade*, a *interatividade*, a *hipertextualidade*, a *personalização*, a *memória* e a *atualização contínua*, conforme as seguintes definições:

Multimidialidade

Para Palacios (2003), o conceito diz respeito à:

> [...] convergência dos formatos das mídias tradicionais (imagem, texto e som) na narração do fato jornalístico. A convergência torna-se possível em função do processo de digitalização da informação e sua posterior circulação e/ou disponibilização em múltiplas plataformas e suportes, numa situação de agregação e complementaridade. (PALACIOS, 2003, p. 3).

Interatividade

Entre as demais características, Bardoel e Deuze (2001) consideram que a notícia *on-line* possui:

> [...] a capacidade de fazer com que o leitor/usuário sinta-se mais diretamente parte do processo jornalístico. Isto pode acontecer de diversas maneiras: pela troca de e-mails entre leitores e jornalistas, através da disponibilização da opinião dos leitores, como é feito em sites que abrigam fóruns de discussões, através de chats com jornalistas, etc. (BARDOEL; DEUZE, 2001, p. 5).

Machado (1997), por sua vez, ressalta que:

> A interatividade ocorre também no âmbito da própria notícia, ou seja, a navegação pelo hipertexto também pode ser classificada como uma situação interativa. Adota-se o termo multi-interativo para designar o conjunto de processos que envolvem a situação do leitor de um jornal na Web. (MACHADO, 1997, p. 146).

A partir dos esclarecimentos de Bardoel e Deuze (2001) e Machado (1997), é possível identificar dois elementos distintos presentes no conceito de *interatividade* que são (a) a navegação *per se* através do hipertexto, o que já equivale a uma situação interativa e (b) as atividades de *mídia social**, que correspondem aos processos de comunicação entre usuários e seu engajamento no enriquecimento da narrativa jornalística, estejam estes na condição de leitores ou jornalistas, promovidos por ferramentas e aplicativos específicos para este fim.

Faz-se necessário observar que o termo *mídia social*, ou *social media*, é empregado para designar o uso do meio eletrônico para interação social, permitindo a criação, o compartilhamento e o intercâmbio do conteúdo gerado pelo usuário. Cada vez mais apoiada por tecnologias ubíquas e escaláveis, a mídia social traz mudanças consideráveis à forma como as organizações, comunidades e indivíduos se comunicam. Como exemplos de produtos de mídia social, temos *Facebook*, *Twitter* e *Flickr*.

Também é válido esclarecer que por narrativa jornalística, ou narrativa do fato jornalístico, entende-se não apenas a reportagem, que é o gênero privilegiado da narração jornalística, mas também a notícia pura e simples de um fato qualquer,

visto que a notícia não é um acontecimento em si, mas sim a sua narração. Desse modo, as diferenças entre notícia e reportagem se limitam às formas de estruturação da narração. Como ingredientes imprescindíveis à narrativa, há personagens, ação dramática e descrições de ambiente. No que diz respeito a essa questão, Sodré e Ferrari (1986) clarificam:

> Narrativa, sabe-se, é todo e qualquer discurso capaz de evocar um mundo concebido como real, material e espiritual, situado em um espaço determinado. Os fatos atribuídos a, por exemplo, Riobaldo e Diadorim (*Grande Sertão, Veredas*), no texto de Guimarães Rosa, pressupõem a aceitação do mundo imaginado pelo escritor como algo suscetível de evocar um espaço humano real. O romance, o conto, às vezes mesmo o poema, constituem formas diferentes de narrativa.
>
> Mas a narrativa não é privilégio da arte ficcional. Quando o jornal diário noticia um fato qualquer, como um atropelamento, já traz aí, em germe, uma narrativa. O desdobramento das clássicas perguntas a que a notícia pretende responder (quem, o quê, como, quando, onde, por quê) constituirá de pleno direito uma narrativa, não mais regida pelo imaginário, como na literatura de ficção, mas pela realidade factual do dia-a-dia, pelos pontos rítmicos do cotidiano que, discursivamente trabalhados, tornam-se *reportagem*. Esta é uma extensão da notícia e, por excelência, a forma-narrativa do veículo impresso (embora a entrevista, sobretudo o perfil, possa também, às vezes, assumir uma forma-narrativa). A reportagem constitui, assim, basicamente, um dos *gêneros* jornalísticos. (SODRÉ; FERRARI, 1986, p. 11)

Por meio do exposto, evidencia-se que não há como falar em narrativa jornalística sem considerar a estrutura organizacional das informações. Sobre esse assunto, Fidalgo (2004) explica:

> Ao princípio, a informação é dada em traços gerais, consistindo preferencialmente na resposta sumária às perguntas que tradicionalmente enformam um *lead*: quem, o quê, quando, onde, por quê. Depois vêm as notícias subsequentes completando e pormenorizando a informação. De certo modo, poder-se-ia entender essa sucessão de notícias como uma extensão da noção de pirâmide invertida, usada na feitura de uma notícia, a um conjunto de notícias sobre o mesmo evento. (FIDALGO, 2004, p. 2).

Assim, ao narrar um fato jornalístico, o produtor de conteúdo deve, em princípio, tentar responder às questões essenciais próprias do *lead* – ou seja: *o quê*, *quem*, *quando*, *onde*, *como* e *por quê* – ainda que a urgência do imediatismo, este

próprio do meio *on-line*, se sobreponha às exigências da objetividade e da verificabilidade.

Na Internet, a narrativa jornalística conta com diversos formatos de mídia. Texto, imagem, áudio e vídeo em formato digital são utilizados de forma integrada e complementar, de modo a enriquecer e pormenorizar a informação, gerando o que aqui já foi definido como *multimidialidade*.

Hipertextualidade

O hipertexto está relacionado com os conceitos e as propriedades da *hipermídia*, que, para Machado (1997), é:

> [...] uma forma combinatória, permutacional e interativa de multimídia, em que textos, sons e imagens (estáticas e em movimento) estão ligados entre si por elos probabilísticos e móveis, que podem ser configurados pelos receptores de diferentes maneiras. (MACHADO, 1997, p. 146).

A partir do exposto por Machado (1997), fica clara a relação entre os conceitos de *interatividade* e *hipertextualidade*, uma vez que a situação interativa abrange também a navegação propiciada pela difusão de *links*, cujo objetivo é a interconexão de diferentes formatos de mídia.

Vale observar que há uma distinção entre conceitos aparentemente próximos como texto, hipertexto, multimídia e hipermídia, conforme apontado por Gosciola (2003). Pierre Lévy (1993), por exemplo, assume os conceitos de multimídia, hipertexto e hipermídia como hipertexto, o que mostra que o termo hipermídia já é, de certo modo, passível de interpretações diferentes.

Para Laufer e Scavetta (1997), texto é um conjunto de parágrafos e, geralmente, lido linearmente do começo ao fim. De acordo com Landow e Delany (1991), o texto possui três atributos cruciais: linearidade, demarcação e estabilidade. Nota-se, entretanto, que tais atributos não são exclusivos do meio texto.

Segundo Laufer e Scavetta (1997), o hipertexto leva os atributos do texto mais adiante, definindo-o como um agrupamento de textos em meio digital, ligados por elos semânticos ancorados em uma palavra ou uma frase promovendo, assim, uma leitura não linear. De acordo com Kilian (2000), trata-se de um texto que faz referência a outros textos e que possibilita ir ao encontro deles. O autor

complementa afirmando que o hipertexto é um documento eletrônico – texto, gráfico, áudio, vídeo ou uma combinação destes quatro elementos – ligado a outros documentos eletrônicos; o hipertexto é arquivado no disco de um computador, que permite ao leitor saltar para outro arquivo no mesmo computador ou em computadores distantes.

Multimídia, por sua vez, é a integração de gráficos, animações, vídeo, música, fala e texto, baseada em computador, para comunicar conteúdo intelectual aos leitores por um caminho simples ou uma linha de apresentação, ou por um navegador não direcional, de acordo com Lipton (1992). Para Laufer e Scavetta (1997), multimídia é o conjunto de meios utilizados ao mesmo tempo para a comunicação de conteúdos, podendo ser navegado de maneira linear ou não linear.

Entendendo que o hipertexto trouxe para a Internet e para o CD-ROM o recurso básico responsável pela capacidade de comunicação altamente dinâmica, ao facilitar a navegação não linear entre textos e ao permitir, por exemplo, o acesso a palavras procuradas, a hipermídia pode ser interpretada como a ampliação natural e consecutiva deste conceito, dado que tem como recurso básico o acesso e a navegação não linear entre conteúdos, sejam estes textos, fotos, gráficos, animações ou vídeos.

Personalização

Conceito que está diretamente relacionado à *interatividade*, sendo fruto da relação do leitor com o produto noticioso, suas escolhas e sua linearidade dentro de uma narrativa particular. Para Zuin e Correia (2008), a personalização está:

> [...] próxima da dissertação, uma vez que o usuário tem condições de fazer e/ou modificar o processo, ou ainda construir de acordo com sua mediação. (ZUIN; CORREIA, 2008, p. 13).

Memória

O armazenamento do conteúdo em banco de dados permite que o usuário acesse qualquer conteúdo e estabeleça seus próprios fluxos de leitura, criando

assim uma relação particular com a questão da temporalidade. Sobre isto, Machado (2004) diz o seguinte:

> O uso das bases de dados permite a ininterrupção do fluxo informacional, o que torna possível e confiável o acesso a informações. [...] o texto webjornalístico, incrementado pelo banco de dados, pode viabilizar vários caminhos a determinados conteúdos e o usuário pode acessá-los a qualquer momento. (Machado, 2004, p. 5).

Atualização contínua

O jornalismo na Web permite atualizações de conteúdo em tempo real. Além disso, o seu conteúdo pode ser acessado a qualquer instante e em qualquer lugar via Internet. Sobre esse assunto, Palacios (2003) esclarece que:

> A rapidez do acesso, combinada com a facilidade de produção e de disponibilização, propiciadas pela digitalização da informação e pelas tecnologias telemáticas, permitem uma extrema agilidade de atualização do material nos jornais da Web. Isso possibilita o acompanhamento contínuo em torno do desenvolvimento dos assuntos jornalísticos de maior interesse. (PALACIOS, 2003, p. 4-5).

Uma vez levantados os elementos que caracterizam o jornalismo *on-line*, faz-se necessária uma investigação sobre o seu processo evolutivo e as mudanças que dizem respeito à narrativa jornalística. Nesse contexto, temos os trabalhos de Pavlik (2001), Mielniczuk (2003) e Schwingel (2005).

Pavlik (2001) aponta três estágios de evolução para o jornalismo digital: o primeiro, caracterizado pela transposição do conteúdo impresso para a Internet; o segundo, pela agregação de recursos e criação de conteúdos originais; e o terceiro – que, segundo ele, estaria começando a emergir, na época da elaboração do seu trabalho – marcado por um produto totalmente exclusivo para a Internet.

Já Mielniczuk (2003), vai um pouco mais adiante e propõe que a trajetória dos produtos jornalísticos desenvolvidos para a Web seja dividida em três momentos: produtos de primeira geração ou fase de transposição, produtos de segunda geração ou fase de metáfora e produtos de terceira geração ou fase do jornalismo digital. Na fase transpositiva, a disponibilização de informações jornalísticas na Web fica restrita à possibilidade de ocupar um espaço, sem explorá-lo como um meio que apresenta características específicas. Na fase metafórica, o jornal impresso funciona

como uma referência para a elaboração das interfaces dos produtos e começam a ocorrer experiências na tentativa de explorar as características oferecidas pela rede. O jornalismo digital de terceira geração é descrito pela autora como “sites jornalísticos que extrapolam a ideia de uma versão para Web de um jornal impresso já existente”.

Schwingel (2005), por sua vez, complementa o trabalho de Mielniczuk (2003) apresentando indícios de uma nova etapa do jornalismo digital – ou seja, a quarta geração de produtos jornalísticos desenvolvidos para a Web – na qual ocorre a consolidação da utilização de bancos de dados complexos através da utilização de ferramentas automatizadas e diferenciadas (sistemas de apuração, edição e veiculação de informações) na criação e manutenção de produtos jornalísticos. Tais ferramentas estabeleceriam vínculos entre diferentes plataformas (Web, e-mail etc) e ambientes distintos (*chats*, fóruns etc), utilizando-se de tecnologias também diferenciadas (*PHP*, *Pearl*, *XML* etc), de acordo com seus interesses e conveniências.

É necessário, contudo, esclarecer que a noção de banco de dados aplicada ao jornalismo digital vai além dos conceitos de depósito integrado de dados, de coleção de documentos ou de repositório para busca e recuperação de informações. Em conformidade com o princípio da transcodificação, desenvolvido por Manovich (2001) – segundo o qual todos os objetos digitais podem ser traduzidos para outros formatos – Machado (2004) demonstra como os bancos de dados são uma forma cultural com estatuto próprio no jornalismo digital. Para o autor, o banco de dados desempenha as funções simultâneas e complementares de (a) formato para a estruturação da informação, (b) suporte para modelos de narrativa multimídia e (c) memória dos conteúdos publicados, o que resulta num formato para esta modalidade.

Terminologias à parte, Pavlik (2001), Mielniczuk (2003) e Schwingel (2005) reconhecem a existência de um estágio no qual, em última análise, o jornalismo digital deixa de ser uma versão do modelo impresso, e passa a ser uma nova e distinta forma de se apresentar informações, momento este de extrema relevância para o trabalho. Assim sendo, como referência a essa fase na evolução do jornalismo *on-line*, será adotado o termo *jornalismo digital de quarta geração*, assim definido por Schwingel (2005), ou simplesmente *jornalismo digital*.

Especificamente sobre o jornalismo digital no Brasil, Palacios (2002), ao realizar um estudo que avalia as características anteriormente apontadas – que são *multimidialidade*, *interatividade*, *hipertextualidade*, *personalização*, *memória* e *atualização contínua* – conclui que poucos são os recursos oferecidos pela Web que são utilizados pelos jornais *on-line* brasileiros, apresentando fortes características das publicações em papel, não explorando de forma satisfatória as possibilidades oferecidas pelo ambiente digital para o desenvolvimento de produtos jornalísticos. Entre outras coisas, percebe-se que a *multimidialidade* não aparece como uma característica explorada e que a *hipertextualidade* ainda é utilizada apenas para organizar a publicação, não sendo explorada na própria narrativa do fato jornalístico.

Apesar da análise – que pode ser considerada ultrapassada, já que foi realizada há nove anos – é possível afirmar que, no Brasil, estamos nos aproximando cada vez mais do jornalismo digital de quarta geração. Podemos observar, por exemplo, que algumas publicações organizam a narrativa em blocos de textos com *links* que levam a outros conteúdos ou extensões desta narrativa. Os *links* disponibilizam também versões e complementos em imagem, áudio, vídeo e outros formatos de informação, que vêm a constituir a narrativa jornalística. Em um primeiro momento, esses recursos são apenas informações a mais. Num passo seguinte, os recursos multimídia tornam-se parte da narrativa. Nesse estágio, a informação a mais aparece também como uma matéria sobre o mesmo assunto que já havia sido publicada pela mesma empresa. São explorados níveis de interatividade como o uso jornalístico do chat, enquetes e, principalmente, possibilidades do usuário comentar notícias e ler comentários de outros usuários. Finalmente, observa-se a valorização do arquivo, apesar dele já existir anteriormente. A divisão em editorias ainda é visível, porém algumas transformações já são praticadas, principalmente pelos portais de serviços, que fazem uso das *breaking news* – que podem ser definidas como reportagens ou boletins especiais, que cobrem acontecimentos que acabaram de ocorrer ou que estão em andamento, cuja veiculação interrompe (ou recebe mais destaque que) as notícias do momento ou aquelas que estão previstas na programação normal, por decisão editorial – sendo esta a situação que melhor retrata o jornalismo *on-line* brasileiro na atualidade. Nesse ponto, é válida uma análise mais apurada e aprofundada do assunto para compreensão do cenário e posterior identificação dos problemas.

Hoje, as editorias utilizadas pelas versões *on-line* dos jornais impressos, a exemplo de *O Globo* e *O Estado de S. Paulo* (vide Figura 1), ainda lançam mão da

Figura 1 - Capa dos jornais impressos *O Globo* e *O Estado de S. Paulo*

solução adotada por seus irmãos mais velhos, e mesmo os jornais puramente *on-line*, a exemplo de *G1* (vide Figura 2) e *Último Segundo* (vide Figura 3), fazem referência a este modelo. Desse modo, evidenciamos conteúdos de todo tipo sendo classificados num espectro limitado de assuntos como: *Blogs e Colunas*, *Brasil*, *Carros*, *Ciência e Saúde*, *Concursos e Emprego*, *Economia*, *Especiais*, *Esporte*, *Loteria*, *Mundo*, *Natureza*, *Planeta Bizarro*, *Política*, *Pop & Arte*, *Tecnologia e Games*, *Vestibular e Educação*, sendo estas as editorias que compõem o *G1*. Alternativamente, para atender a demanda do momento, algumas categorias – muitas vezes de caráter temporário – são disponibilizadas conjuntamente às

Figura 2 - *Screenshot* da página inicial do jornal *on-line G1*, com o menu de editorias aberto

editorias, porém com exposição diferenciada, nos jornais. Eis alguns exemplos: *Caso Isabella*, *Tsunami no Pacífico*, *Tragédia em Realengo*. Ou seja, dependendo da emergência ou da procura, um assunto até então inexistente pode receber destaque nos jornais *on-line*, sendo incorporado à sua estrutura de navegação, o que é característico das *breaking news*.

A questão das editorias é um velho dilema que se estende até hoje, não apenas por causa do modelo já estabelecido pelos jornais tradicionais, mas pelo simples fato da maioria dos editores-chefe, e demais responsáveis, dos jornais *on-line* vir da área impressa, o que contribui para a perpetuação de uma estrutura de conteúdo inadequada e incapaz de tirar todo proveito da dinâmica da Internet.

O problema das editorias se reflete tanto na estrutura de navegação quanto na organização dos conteúdos e suas inter-relações dentro do jornal, a saber: no G1, ao visualizar uma matéria cujo título é *Senado abre sindicância para apurar possível nova 'farra de passagens'* (vide Anexo A) – que, resumidamente, relata uma sindicância interna no Senado, instalada a pedido do senador Heráclito Fortes, para apuração de irregularidades na emissão de passagens aéreas para senadores –

Figura 3 - *Screenshot* da página inicial do jornal *on-line Último Segundo*

nota-se que a mesma faz parte da editoria *Política*. No entanto, não há qualquer referência a assuntos que talvez sejam mais pertinentes ao leitor como, por exemplo, *Escândalo*, *Senado Federal* ou *Congresso Nacional*. São relacionadas apenas duas matérias para aprofundamento da leitura, enquanto é sabido, após realização de busca, que existem ao menos oito matérias sobre o tema. Além disso, no corpo da matéria há uma imagem de referência integrada ao texto e somente um *link*, que aponta para uma das matérias já relacionadas. Com base nesse exemplo, é possível identificar as seguintes restrições técnicas e operacionais:

a) Existe uma gama extremamente restrita de editorias ao qual o conteúdo pode se relacionar, fazendo com que o mesmo seja associado a assuntos muito genéricos e, às vezes, estranhos à natureza do tema abordado;
b) Por causa dessa restrição, não é possível estabelecer um relacionamento entre conteúdos – que tratam de um tema específico – de modo automatizado, tendo de fazê-lo manualmente;
c) Da mesma forma, é inviável estabelecer novos eixos de navegação, automática ou semiautomaticamente, a partir de termos que são

relevantes ao conteúdo ou que se repetem ao longo do texto. Exemplo: *Senado Federal*, *Mesa Diretora do Senado*, *Câmara dos Deputados* e *Heráclito Fortes*.

d) A busca e recuperação da informação – seja ela no formato de texto, imagem, áudio ou vídeo – a integrar o conteúdo publicado é limitada e demorada, fazendo com que os esforços dispendidos na operação fiquem concentrados nos temas de maior importância. Consequentemente, a narrativa jornalística acaba por ficar prejudicada na maior parte dos conteúdos ofertados.

e) Nenhuma das editorias existentes possui relação funcional, seja de hierarquia, dependência etc, entre si. Desse modo, não é possível fazer com que os conteúdos associados a uma editoria também sejam apresentados automaticamente em outra. Seria interessante, por exemplo, que a região (editoria) *DF (Distrito Federal)* fosse, na prática, um subitem da editoria *Brasil*; desse modo, todo conteúdo de *DF* seria exibido em *Brasil*.

A partir do exemplo dado e do entendimento que, teoricamente, as editorias nada mais são que categorias generalistas de classificação, é possível perceber que a adoção de uma estrutura que visa organizar seu conteúdo em categorias – que, além de poucas, são isoladas e não fazem alusão à particularidade do tema abordado – não se aproveita das potencialidades da Web e, sendo assim, deixa de aperfeiçoar a experiência de uso dos produtos jornalísticos desenvolvidos para este meio, tanto do lado do produtor de conteúdo quanto do leitor. Ao se verificar tais consequências, e levando em conta os conceitos envolvidos na classificação, algumas questões essenciais à superação dos problemas podem ser apontadas, fazendo-se antes, porém, as seguintes considerações:

- Por *conteúdo digital*, ou simplesmente *conteúdo*, entende-se toda e qualquer representação digital da informação, capaz de ser publicada na Web. Um conteúdo pode ser apresentado na forma de texto, imagem, vídeo etc;
- A *categoria* é um recurso de classificação que agrupa conteúdos similares – estes, por sua vez, abordam, retratam ou citam um tema em comum, seja ele qual for – recebendo o nome daquilo a que faz

referência. Exemplos: *Escândalo*, *Política*, *Heráclito Fortes*, *iPhone*, *Senado Federal*, *Brasil* e *Carnaval 2011*;

- A *entidade nomeada* é uma forma particular de categorização que envolve a individualização de categorias do tipo *evento*, *lugar*, *organização*, *pessoa*, *produto* e congêneres. Exemplos: *Heráclito Fortes*, *iPhone*, *Senado Federal*, *Brasil* e *Carnaval 2011*. Nota-se, com base nos exemplos, que qualquer entidade nomeada pode ser tratada como categoria.

Uma vez esclarecidos os conceitos básicos de classificação, surgem questões que dizem respeito aos seguintes parâmetros:

a) Quantidade de categorias;
b) Especialização das categorias;
c) Entidades nomeadas como categorias;
d) Associação entre categorias;
e) Associação entre conteúdos e categorias.

É interessante observar que a associação entre conteúdos e categorias está estreitamente relacionada ao processo de rotulação/anotação do conteúdo, e será explicado em breve.

Com base nos parâmetros expostos, uma orientação é claramente estabelecida, no que diz respeito tanto ao levantamento das iniciativas quanto à proposição de uma solução, conforme será possível verificar adiante.

1.3 **Novas iniciativas**

No âmbito internacional, é possível destacar, dentre uma grande variedade de jornais *on-line* de alto padrão editorial, nomes como *The New York Times – NYT*, *BBC News*, *MSNBC.COM* e *CNN* (vide Figura 4), sendo estes os produtos jornalísticos que foram utilizados como referência analítica. Todos são reconhecidos pela busca incessante por inovações, por fazerem uso do estado da arte em tecnologia, além de investirem constantemente em design, conduzindo pesquisas de qualidade e usabilidade, entre outras coisas.

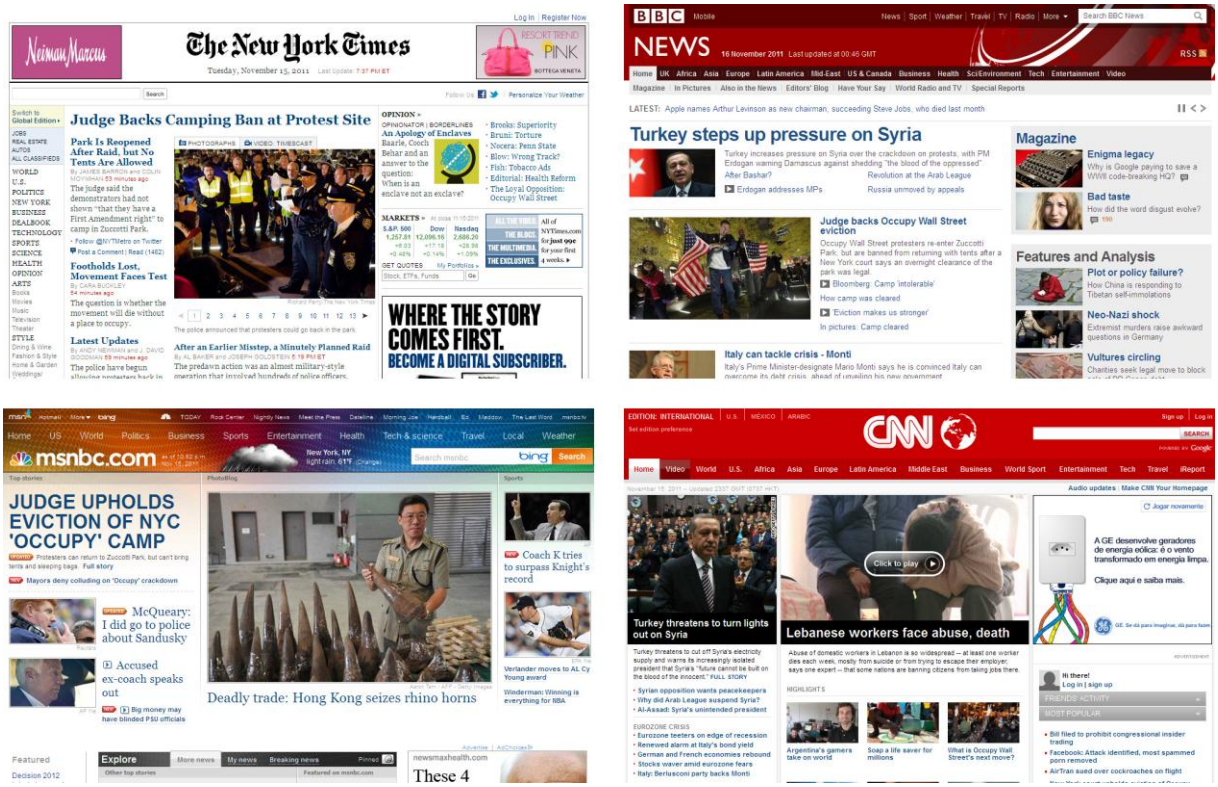


Figura 4 - *Screenshots* da página inicial dos jornais *on-line NYT*, *BBC News*, *MSNBC.COM* e *CNN*

Antes de nos aprofundarmos na questão da classificação do conteúdo, porém, é importante esclarecer que os jornais avaliados possuem um fluxo de trabalho que conta com o recurso de rotulação/anotação do conteúdo – que pode ser entendido como a atividade por trás da associação entre conteúdos e categorias – que é realizado de forma detalhada e meticulosa, demonstrando enorme preocupação com os metadados, que são a peça-chave na gestão das informações, responsáveis pelo modo como o conteúdo será indexado dentro do banco de dados. E mais: de modo geral, os metadados estão diretamente relacionados à narrativa jornalística, sendo muitas vezes o equivalente às respostas dadas para as clássicas perguntas que compõem o *lead* da notícia.

Segundo Colle (2002), os metadados são dados sobre outros dados e se constituem em ferramentas que devem guiar os usuários tanto na consulta quanto na busca de novas informações sobre os mesmos e sobre o sistema em si.

Preocupada com o correto entendimento do que é metadado, Schwingel (2005) esclarece:

> Para uma maior elucidação, podemos usar como exemplo hipotético a cobertura jornalística de algum escândalo financeiro ocorrido na década passada, no qual diversas pessoas tiveram seus nomes associados às fraudes constatadas. Caso alguma delas esteja vinculada atualmente a outro tipo de negócio que contenha irregularidades, a cobertura atual deveria contextualizar a questão com a trajetória de vida dos envolvidos. (SCHWINGEL, 2005, p. 4).

No que diz respeito ao banco de dados, e seus efeitos sobre a estruturação e a organização das informações, Fidalgo (2004) destaca a sua flexibilidade para a forma de apresentação das notícias, assim como suas vantagens para garantir objetividade às mesmas. O autor introduz o conceito de *resolução semântica*, que é explicada da seguinte forma:

> Uma primeira notícia sobre um acontecimento, que à partida surge com um determinado sentido, pode ser complementada, alterada, corrigida, à medida que outras notícias sobre o mesmo acontecimento se lhe seguem. O que de início tinha contornos indefinidos, deixando múltiplas hipóteses em aberto, vai ganhando sucessivamente formas cada vez mais definidas. Um exemplo ajudará a perceber intuitivamente o que foi dito. Uma notícia de última hora pode ter a seguinte forma: "Houve uma forte explosão em Bagdad". É evidente que se trata de uma notícia de muito baixa resolução semântica. Não se sabe que tipo de explosão foi, se acidental ou provocada, se provocada por um carro armadilhado ou se por um ataque de aviação, se houve ou não houve vítimas, em que local da cidade, etc., etc. Porém, à medida que forem chegando notícias subsequentes, a notícia do que ocorreu vai ganhando forma, ou seja, aumenta a sua resolução semântica. (FIDALGO, 2004, p. 3).

Assim sendo, o que se percebe cada vez mais entre os jornais *on-line* é a busca por uma alta resolução semântica. Para propiciá-la, faz-se necessária a construção de cenários onde, a partir de um determinado conteúdo, seja estabelecido o imediato acesso a outros que lhe digam respeito, com a finalidade de fornecer mais detalhes sobre o fato jornalístico ou de aprofundar os assuntos que deste são alvo. Isso, entretanto, só pode ser viabilizado com o uso de um banco de dados capaz de não apenas armazenar todo conteúdo produzido – e, por conseguinte, todas as informações nele contido – mas de relacioná-los de tal maneira que seja possível identificar aqueles que abordam um mesmo fato ou que, de alguma forma, são relevantes entre si.

Sobre o entendimento do que é o banco de dados, é interessante observar que Manovich (2001) o vê como a forma estruturadora dos produtos digitais na contemporaneidade, pois é um complexo de armazenagem de formas culturais que pode servir, inclusive, para criar novos gêneros e narrativas nas mais distintas áreas, enquanto que Gonçalves (2004) o encara como um formato para o jornalismo digital.

Voltando às iniciativas internacionais, uma das soluções mais recentes adotadas pelo *NYT*, *BBC News*, *MSNBC.COM* e *CNN* é a ideia de *tópico*, que corresponde ao conceito de categoria, tal qual a definição dada no presente trabalho. Desse modo, entende-se por tópico toda e qualquer categoria que, independente da forma como é representada em termos de desenho de interface – ou seja, da aparência visual de sua *webpage** – agrega conteúdos de todo tipo, relacionados ao tema a que se faz alusão. A diversidade de tópicos disponíveis é enorme, sendo permanentemente utilizado como um dos eixos da narrativa jornalística. Os tópicos são associados entre si, havendo a possibilidade de navegar por entre os conteúdos, a partir de uma cadeia de tópicos inter-relacionados.

Cada tópico tem definição e significado específicos – o que garante a relevância dos conteúdos a ele vinculados – sendo representado por uma *webpage* individualizada. Assim sendo, o tópico funciona como um importante nó de navegação e, portanto, de consumo dos conteúdos que ele tem por finalidade agregar.

As editorias continuam sendo usadas, mas desta vez com algumas mudanças. Na prática, elas correspondem a tópicos que se manifestam de modo diferenciado na interface do produto jornalístico, fazendo parte da estrutura de navegação básica do *website*. Nos demais aspectos, o seu comportamento é exatamente igual ao do tópico.

Além das editorias, que podem ser consideradas parte integrante dos tópicos, há outros tipos de tópico presentes na estrutura estabelecida. Observa-se, no entanto, abordagens distintas entre os jornais estudados. A solução mais interessante é a adotada pelo *NYT*, que divide os tópicos em assunto, lugar, organização e pessoa (vide Anexo B).

Existe também o conceito de *cobertura especial*, comumente chamado de *história* (*story*). Trata-se de uma solução utilizada pelos jornais estudados, cujo desenho da interface se distingue dos demais e que, da mesma forma que o tópico, tem por objetivo reunir conteúdos diversos sobre um determinado tema, sendo este

um acontecimento específico como, por exemplo, *Posse de Dilma*, *Réveillon 2011* e *Carnaval 2011*. Necessariamente, o tema abordado tem de ser um evento de caráter temporário, com início, meio e fim, ou seja: sua existência deve estar fortemente vinculada a um período, ainda que indeterminado, de tempo.

Embora as histórias e as *breaking news* possuam algumas características em comum – como, por exemplo, o fato de retratarem um evento e receberem destaque nos jornais *on-line* – os seus conceitos são distintos: resumidamente, enquanto que o primeiro diz respeito a um acontecimento cuja veiculação está programada pelo jornal, o segundo, ao contrário, está associado à interrupção das notícias previstas na programação normal. Isso não significa, porém, que as *breaking news* não possam ser consideradas histórias, findo o período emergencial; de certo modo, elas podem ser encaradas como uma especialização do conceito de história.

Apesar da diferenciação conceitual entre história e tópico, nos jornais avaliados os dois tipos de categoria se comportam de modo semelhante, havendo associações entre si, inclusive. Na *MSNBC.COM*, por exemplo, há uma estreita relação entre história e assunto e vice-versa, explorando um pouco mais as dimensões da estrutura narrativa.

Considerando-se os cinco parâmetros levantados anteriormente, é possível constatar que todos os jornais estrangeiros estudados incorporam cada um deles, sem exceção, devido à lógica de funcionamento dos tópicos. O conceito de história, por sua vez, serve de apoio aos tópicos e aprofunda ainda mais a narrativa do fato jornalístico.

Em relação à tecnologia, é válido dizer que no *NYT*, *BBC News* e *MSNBC.COM* a plataforma que abarca tais inovações é baseada na Web Semântica. No tocante à classificação de conteúdo, o *NYT* é o melhor exemplo de integração dessa tecnologia ao fluxo de trabalho. Segundo Sandhaus e Larson (2009), a empresa possui um complexo sistema de anotação semântica, onde o próprio *Content Management System - CMS* sugere categorias a serem relacionadas ao conteúdo, com base na análise dos dados nele contidos, a partir de uma ferramenta integrada de *Natural Language Processing - NLP**.

Além disso, tanto no *NYT* quanto na *BBC News*, observa-se que cada conceito representado na ontologia possui um identificador único chamado *Uniform Resource Identifier - URI* – que é uma cadeia de caracteres compacta usada para identificar ou denominar um *recurso** na Internet – sendo este responsável pela

definição do nome de sua *webpage*, o que é considerado uma boa prática do ponto de vista da Web Semântica. Isso vai ao encontro da ideia de tópico, dado que o sistema de classificação utilizado em ambos os jornais tem a capacidade de gerar automaticamente, sempre que necessário, uma *webpage* para cada novo conceito criado.

É importante notar que, no Brasil, embora o jornal *on-line Estadão* (vide Anexo C) se utilize do termo *tópico*, este nada mais é do que o resultado do uso de *tags* sem tratamento semântico, ou seja, da rotulação do conteúdo com termos desprovidos de significado, ao contrário do que acontece na Web Semântica – não sendo possível, consequentemente, estabelecer qualquer tipo de relação entre uma *tag* e outra; a relação é estabelecida tão somente entre conteúdos que são marcados com uma mesma *tag* – distanciando-se consideravelmente da solução que neste trabalho se pretende propor.

2 Web Semântica

2.1 Da Web atual à Web Semântica

A Web alterou radicalmente o modo como as informações são armazenadas, transmitidas e acessadas nos últimos anos. Desde a sua gênese, a Web tem crescido consideravelmente em quantidade de usuários e volume de informações. A presença cada vez mais intensa da Web em nosso cotidiano faz dela um ambiente bastante atraente para a disseminação de informações, seja entre pessoas ou organizações.

Desde que a Web foi disponibilizada como um serviço público na Internet, em 6 de agosto de 1991, sua estrutura tem desempenhado com sucesso a função para a qual foi concebida: compartilhamento de documentos em um ambiente distribuído. A partir de 1998, no entanto – com o aumento expressivo da quantidade de informações e serviços disponíveis na Web – percebeu-se a necessidade de se ter um ambiente que não apenas permita a troca de informações, mas que também possibilite às máquinas realizar tarefas com os dados existentes.

Nesse sentido, vem à tona a questão do gerenciamento de conteúdo – no que diz respeito à dificuldade na busca, organização, acesso e manutenção das informações disponíveis aos usuários – sendo necessário que as máquinas estejam aptas a “entender” a semântica das informações presentes na Web, a fim de ajudar no uso e gerenciamento deste gigantesco repositório. Tendo isso em vista, como as máquinas poderiam manipular de forma eficiente as informações descritas por, e para, seres humanos?

Primeiramente, torna-se fundamental considerar que no ambiente em discussão haja *interoperabilidade semântica**, conceito que corresponde à habilidade de sistemas de computador transmitir dados com significado compartilhado, sem ambiguidade, sendo isto um requisito necessário à realização de operações lógicas computáveis e inferências de máquina, bem como à descoberta de conhecimento e ao estabelecimento de federações de dados entre sistemas de informação.

Em suma, a interoperabilidade semântica tem a ver não apenas com o empacotamento dos dados (sintaxe), mas com a transmissão simultânea do significado (semântica) destes dados. Isso somente é possível adicionando dados sobre os dados em questão, ou seja, metadados; cada elemento de dado é ligado a um vocabulário controlado e compartilhado. O significado do dado é transmitido juntamente com o próprio dado em si, numa espécie de rótulo autodescritivo que é independente de qualquer sistema de informação. É este vocabulário compartilhado, e seus vínculos com uma ou mais ontologias, que provê a capacidade de interpretação, de inferência e de lógica da máquina.

Com a finalidade de propiciar a correta interpretação das informações por máquinas, Berners-Lee, Hendler e Lassila (2001), com a colaboração da *World Wide Web Consortium - W3C**, deram início a uma série de estudos para implantação de uma nova arquitetura capaz de estruturar as informações na Web com base em sua contextualização semântica, sendo este significado compreendido e utilizado por computadores para o processamento e intercâmbio de conteúdo: a Web Semântica.

2.2 **A Web Semântica**

Idealizada por Berners-Lee, Hendler e Lassila (2001), a Web Semântica é uma extensão da Web atual, que permitirá aos computadores e seres humanos trabalharem em cooperação. Ela interliga significados de palavras e, neste âmbito, tem como finalidade conseguir atribuir sentido aos conteúdos publicados na Internet, de modo que seja perceptível tanto pelo homem como pelo computador. A proposta da Web Semântica é estender os princípios dos documentos da Web para os dados. Os dados podem ser acessados usando a arquitetura Web, sendo relacionados uns com os outros da mesma forma que os documentos. Isso também implica a criação de uma plataforma comum que permita o compartilhamento e a reutilização dos dados por meio das fronteiras das aplicações, empresas e comunidades, podendo ser processados automática e manualmente, além de revelar novos relacionamentos possíveis entre porções de dados.

Embora muito se fale a respeito do que é a Web Semântica, uma das melhores definições é a de Berners-Lee, Hendler e Lassila, em seu artigo – que

representa um marco sobre o assunto – intitulado *The Semantic Web*, publicado na revista *Scientific American*, em 2001, que diz o seguinte:

> A Web Semântica não é uma Web em separado, mas uma extensão da atual, na qual a informação é dotada de significado bem definido, permitindo que computadores e pessoas trabalhem em cooperação. [2] (BERNERS-LEE; HENDLER; LASSILA, 2001).

A Web Semântica tem por objetivo oferecer um ambiente padronizado para o intercâmbio de informações, propondo uma nova arquitetura para representação de dados que possibilite o desenvolvimento de novas soluções nas mais diversas áreas de aplicação. Na visão de Berners-Lee, Hendler e Lassila (2001), isso se traduz num ambiente onde programas e dispositivos sejam capazes de trocar informações entre si e, por meio de *agentes de software** – que, em resumo, são programas de computador autônomos, capazes de se comunicar com outros agentes e com seres humanos, que agem em prol de uma meta – permitam a automatização de tarefas rotineiras dos usuários.

Onze anos após o lançamento da ideia, há uma tensa – e, muitas vezes, incrédula – expectativa sobre os conceitos nas quais a Web Semântica se baseia. Embora já se possam encontrar algumas iniciativas contundentes na área, os verdadeiros benefícios da Web Semântica ainda estão longe de ser alcançados. Existem inúmeros entraves a serem superados, a começar pela transposição do campo teórico, haja visto o número de estudos e proposições científicas, ainda sem comprovação prática, na qual se fundamenta.

2.3 **Web Semântica e Inteligência Artificial**

Estruturalmente, a Web Semântica pode ser vista como uma camada de definições de significado acoplada à Internet, fazendo com que o processo de busca se aproxime do conceito de guia e, consequentemente, distancie-se da noção de

[2] Texto original, na íntegra: “The Semantic Web is not a separate Web but an extension of the current one, in which information is given well-defined meaning, better enabling computers and people to work in cooperation.”

catálogo, capaz até mesmo de servir como base para sistemas cujo objetivo é raciocinar de modo semelhante ao ser humano.

A Web Semântica traz à tona a possibilidade da compreensão da linguagem humana pelas máquinas na recuperação da informação, sem que o usuário precise dominar refinadas estratégias de busca. Tal promessa, porém, não deve nos levar a ideia equivocada de que a Web Semântica é uma etapa, nível ou ramo da Inteligência Artificial - IA, pois as questões tratadas pela Web Semântica são outras.

Segundo Alesso e Smith (2009), a criação de máquinas inteligentes sempre foi objeto de interesse da humanidade, porém apenas com o surgimento do computador moderno é que a IA ganhou corpo para se estabelecer como ciência, com problemas e metodologias próprias, transformando-se numa área de pesquisa da Ciência da Computação. Inicialmente, a IA visava reproduzir o pensamento do homem, replicando faculdades humanas como criatividade, autoaperfeiçoamento e uso da linguagem. Atualmente, a IA possui outro objetivo, mas permanece sendo uma noção de múltiplas interpretações, sendo, deste modo, um conceito de difícil definição.

Considerado um dos grandes nomes da IA, Alan Turing (1950) fez grandes contribuições a este ramo de pesquisa, sendo uma delas a formulação do *Teste de Turing*, em sua publicação chamada *Computing Machinery and Intelligence*, de 1950, cujo objetivo era determinar se máquinas são capazes de exibir comportamento inteligente. No exemplo original do autor, um juiz humano conversa simultaneamente, e de forma natural, com um humano e com uma máquina, esta criada para ter desempenho indistinguível do ser humano, sem prévio conhecimento de suas identidades. Se o juiz não tiver condições de diferenciar com segurança a máquina do humano, então é considerado que a máquina passou no teste. A conversa é limitada a um canal contendo apenas texto – ou seja, um terminal de vídeo com teclado – de modo que o resultado não dependa da habilidade da máquina de converter palavras em áudio.

O Teste de Turing rapidamente se tornou o ponto de partida para as pesquisas na área. Em seu artigo, Turing (1950) fez previsões de que até o ano 2000 os computadores passariam no teste. Porém, até o momento, nenhuma máquina obteve sucesso, conseguindo somente passar em versões simplificadas desse teste.

Embora a Web Semântica se utilize de conceitos oriundos da IA – como, por exemplo, o uso de ontologias para representação do conhecimento – esta última possui um objetivo diferente da primeira. Antoniou e Harmelen (2004) dizem o seguinte a respeito disso:

> Se o objetivo da Inteligência Artificial é construir um agente de software que mostre inteligência no nível humano (e superior), o objetivo da Web Semântica é auxiliar humanos a realizarem suas tarefas diárias na rede. [3] (ANTONIOU; HARMELEN, 2004, p. 16).

E Halpin (2004) afirma o seguinte:

> A principal diferença entre a IA de vanguarda e a Web Semântica é o fato desta última, claramente, não compartilhar do objetivo da primeira, que é, como determinado pela Proposta de Dartmouth, a 'inteligência de nível humano'. [4] (HALPIN, 2004, p. 2-3).

De acordo com Halpin (2004), os primeiros movimentos da IA tinham como objetivo o que foi estabelecido por McCarthy *et al*. (1955), na Conferência de Dartmouth, em 1956:

> O estudo é proceder com base na suposição de que todos os aspectos da aprendizagem ou qualquer outra característica da inteligência podem, em princípio, ser tão precisamente descritos que uma máquina pode ser concebida para simulá-los. [5] (MCCARTHY *et al*., 1955, p. 1).

Naquela ocasião, entretanto, a "inteligência" em si não foi claramente definida. A proposta apresentada por McCarthy *et al*. (1955) deu um papel central ao "senso comum", de modo que um programa tem bom senso se automaticamente deduz, por

---

[3] Texto original, na íntegra: "If the ultimate goal of AI is to build an intelligent agent exhibiting human-level intelligence (and higher), the goal of the Semantic Web is to assist human users in their day-to-day on-line activities."

[4] Texto original, na íntegra: "The first major difference between early artificial intelligence and the Semantic Web is that the Semantic Web is clearly not pursuing the original goal of AI as stated by the Dartmouth Proposal: 'human-level intelligence'."

[5] Texto original, na íntegra: "The study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it."

si só, uma gama variada de consequências imediatas de qualquer coisa que lhe seja dita e já seja sabida.

A partir de 1956, uma enorme quantidade de esquemas de representação – de redes semânticas a *frames** – surgiu, ao ponto de se criar uma expectativa de que, nos próximos vinte anos, as máquinas seriam capazes de fazer qualquer trabalho realizável pelo homem. Entretanto, ainda que muitos desses programas conseguissem simular a inteligência em um domínio específico, como provar teoremas lógicos, ficou nítido que esta estratégia não conduziria as máquinas ao nível da inteligência geral.

Embora a IA tivesse feito um bom trabalho em domínios restritos até a década de 70, a ampliação de tal capacidade havia se mostrado uma tarefa nada simples, dado que os esquemas de representação não indicavam de forma óbvia o conhecimento que representavam, o que gerou a iniciativa de se estabelecer uma descrição formal do conhecimento utilizando uma estrutura bem compreendida, como a *Lógica de Primeira Ordem**, ou *Lógica dos Predicados*, que era equivalente à maioria dos sistemas de representação do conhecimento utilizados na época.

O passo seguinte foi, durante a década de 80, formalizar tanto conhecimento de “senso comum” quanto possível, utilizando rigorosos padrões de lógica, a fim de superar pequenas estratégias de domínio específico, estas representadas pelos sistemas especialistas, por exemplo. No entanto, essa abordagem jamais convergia em uma maneira formal e universal para se representar todo conhecimento. A unificação e formalização de todo conhecimento de “senso comum” mostrou-se, deste modo, demasiado distante. Enquanto alguns pesquisadores de IA continuavam sustentando que todo esse conhecimento necessário poderia ser codificado em pouco tempo, muitos deles acabaram deixando o campo e, consequentemente, a indústria de IA desmoronou.

Na década de 90, ficou claro que todo conhecimento útil estava situado na tarefa particular em questão e no agente, sendo improvável que qualquer representação do conhecimento tradicional ou fundamento lógico pudesse captar estes aspectos do conhecimento. Assim sendo, a alegação de que uma máquina poderia simular a inteligência no nível humano através da formalização completa de fatos e inferências acabou sendo derrubada. Por outro lado, pôde-se verificar que os frutos de tais pesquisas poderiam ser convertidos em tecnologia de grande utilidade, independentemente das reivindicações originais da IA. Ao invés de tentar formalizar

o senso comum, a IA poderia, nos anos a seguir, colocar em prática o valioso aprendizado obtido no que diz respeito à indexação e recuperação de informação em uma coleção digital global de documentos, o que mais tarde veio a se configurar como a Web.

Assim sendo, o objetivo da Web Semântica é mais modesto e está alinhado com os trabalhos mais recentes de IA, ao defender a criação de máquinas capazes de exibir um comportamento considerado “inteligente”; a sua obtenção, no entanto, é muito mais difícil de ser verificada, pois, levando-se em conta que a “inteligência” das máquinas é diferente da dos seres humanos, não há uma versão do Teste de Turing capaz de detectar apenas a inteligência em nível de máquina. De acordo com Halpin (2004), é importante observar, contudo, que alguns pesquisadores da área afirmam ter razões suficientes para acreditar que alguns dos objetivos secundários da IA podem ser atingidos através da Web Semântica, em particular a criação de ontologias utilizáveis que representam o mundo real.

Portanto, fica claro que a Web Semântica não se trata de um ramo da IA, uma vez que seu objetivo é bem distinto.

3 Ontologia

3.1 Origem do termo ontologia

Na Filosofia, Lalande (1976) define ontologia como:

> Parte da filosofia que especula sobre 'o ser enquanto ser', segundo a expressão de Aristóteles. [...] somente o nome é novo (século XVII); quanto à ciência em si, ela já existia com a mesma definição desde os escolásticos: chamava-se *transcendentia* essas determinações comuns a todos os seres. (LALANDE, 1976, p. 714-715).

Ou, segundo Bunge (2006), a ontologia é "o ramo da filosofia que estuda as feições mais universais da realidade, tais como existência real, mudança, tempo, chance, mente e vida" (BUNGE, 2006, p. 267).

De acordo com Guizzardi (2005), o termo ontologia foi cunhado no século 17 pelos filósofos Rudolf Göckel, em seu *Lexicon Philosophicum*, e Jacob Lorhard, em seu *Ogdoas Escolástica*, tornando-se popular no meio filosófico somente no século 18, a partir da publicação da obra *Philosophia Prima Sive Ontologia* (vide Figura 5), de Christian Wolff, em 1730. Smith (2003), por sua vez, esclarece que a primeira ocorrência do termo ontologia na língua inglesa aparece no *Bailey's Dictionary*, em 1721, conforme registrado pelo *Oxford English Dictionary*, com as seguintes definições: (a) ramo da Metafísica que estuda a natureza e as relações do ser e (b) teoria da natureza do ser e dos tipos de existência. Etimologicamente, *ont-* vem do verbo grego *einai* (ser) e, desta forma, a palavra latina *ontologia* (*ont-* + *logia*) pode ser traduzida como o estudo da existência.

Smith (2003), em seu trabalho *Ontology*, define o termo ontologia como um campo da metafísica que representa o estudo do ser, da sua existência, do tipo de estrutura de objetos, suas propriedades e relações de tudo que faz parte da realidade, sendo físico ou abstrato.

É interessante destacar que, segundo Guizzardi (2005), no início do século XX, o filósofo Edmund Husserl definiu o termo *Ontologia Formal** para se referir a uma parte específica da disciplina filosófica de Ontologia. Fazendo-se uma analogia com a *Lógica Formal**, enquanto que esta trata de estruturas lógicas formais – como

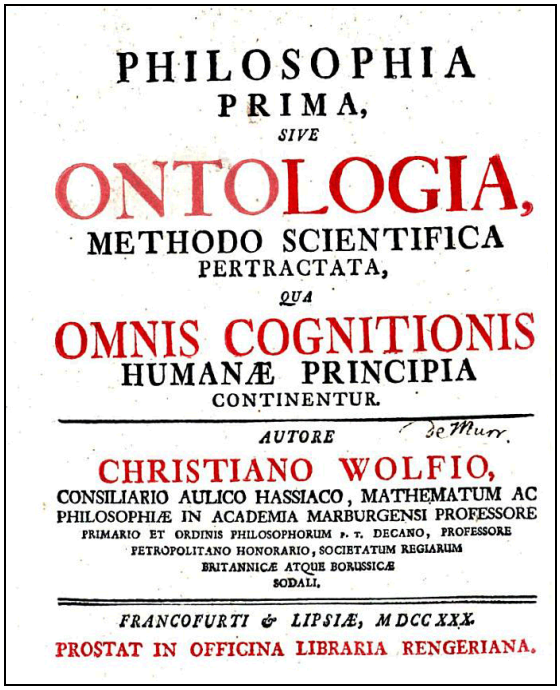

PHILOSOPHIA
PRIMA,
SIVE
ONTOLOGIA,
METHODO SCIENTIFICA
PERTRACTATA,
QUA
OMNIS COGNITIONIS
HUMANÆ PRINCIPIA
CONTINENTUR.
AUTORE
CHRISTIANO WOLFIO,
CONSILIARIO AULICO HASSIACO, MATHEMATUM AC
PHILOSOPHIÆ IN ACADEMIA MARBURGENSI PROFESSORE
PRIMARIO ET ORDINIS PHILOSOPHORUM h. t. DECANO, PROFESSORE
PETROPOLITANO HONORARIO, SOCIETATUM REGIARUM
BRITANNICÆ ATQUE BORUSSICÆ
SODALI.
FRANCOFURTI & LIPSIÆ, MDCCXXX.
PROSTAT IN OFFICINA LIBRARIA RENGERIANA.

Figura 5 - Capa do livro de Christian Wolff, *Philosophia Prima Sive Ontologia*, de 1730

verdade, *validade* e *consistência* – sem levar em consideração a sua veracidade, a Ontologia Formal lida com estruturas ontológicas formais – como *teoria das partes*, *teoria do todo*, *tipos e instanciação*, *identidade*, *dependência*, e *unidade* – isto é, com aspectos formais de objetos independentemente da sua natureza particular.

3.2 **Ontologia na Ciência da Computação**

Enquanto que, no seu sentido filosófico, o termo ontologia é considerado uma explicação da existência, ou seja, uma teoria sobre a natureza da existência, na Ciência da Computação – entendendo que esta área de estudo abrange o ramo que corresponde à IA – ontologia pode ser definida, segundo Gruber (1993), como uma especificação explícita de uma conceituação.

Conforme Neches *et al*. (1991), os estudiosos de IA se apropriaram do termo ontologia, onde o seu significado original acabou por dar lugar a uma definição mais conveniente, havendo uma homologia entre este conceito e aquele próprio da

Filosofia. Desse modo, em IA, ontologia pode ser interpretada como o conjunto de entidades com suas relações, restrições, axiomas e vocabulário.

De acordo com Guizzardi (2005), a primeira menção ao termo *ontologia* na computação data de 1967 por G. H. Mealy, em um trabalho sobre os fundamentos da modelagem de dados, na área de processamento de dados. Desde então, tem sido aplicado em diversas áreas como Sistemas de Informação, Engenharia de Software, Inteligência Artificial e Web Semântica, porém com diferentes significados e propósitos. Enquanto que nas duas primeiras áreas, ontologia é comumente utilizada em conformidade com seu significado na Filosofia, ou seja, como um sistema de categorias independente de linguagem, em outras áreas como IA e Web Semântica, tal palavra é usada, em geral, para designar um artefato concreto, projetado para um propósito específico, e representado em uma linguagem específica.

Para distinguir os dois termos, Guarino e Giaretta (1995) propuseram uma grafia diferente para ambos os significados, na qual a palavra *Ontologia* (com *o* maiúsculo) se refere ao conceito filosófico da palavra, e *ontologia* (com *o* minúsculo) alude ao sentido usado em IA. Apesar dessa pequena diferença para distinção das áreas, a proposta passou a ser adotada por justamente ter no seu sentido original a ideia de representação da existência. Em IA, o objetivo desse conceito é, em suma, representar a existência de coisas ou fenômenos do mundo real, de maneira que máquinas possam compreendê-las.

Há várias definições do termo ontologia dentro da Ciência da Computação. Seguem, abaixo, as mais conhecidas e aceitas por estudiosos da área:

> Define os termos básicos e relações associadas a uma área de conhecimento, assim como as regras para combinar esses termos e relações, a fim de definir extensões do vocabulário. [6] (NECHES *et al*., 1991, p. 40).

> Uma teoria lógica que fornece, de modo explícito, um esclarecimento parcial sobre uma conceituação. [7] (GUARINO; GIARETTA, 1995, p. 32).

---

[6] Texto original, na íntegra: “An ontology defines the basic terms and relations comprising the vocabulary of a topic area as well as the rules for combining terms and relations to define extensions to the vocabulary.”

[7] Texto original, na íntegra: “Ontology: (sense 1) a logical theory which gives an explicit, partial account of a conceptualization; (sense 2) synonym of conceptualization.”

> Um conjunto de axiomas lógicos criados para explicar o significado de um vocabulário. [8] (GUARINO, 1998, p. 4).

> Uma ontologia é uma especificação explícita de uma conceituação. [9] (GRUBER, 1993, p. 199).

De modo complementar, Gómez-Pérez, Fernández-López e Corcho (2004) esclarecem que, das definições de ontologia apresentadas, a mais utilizada entre os pesquisadores da Ciência da Computação é a de Gruber (1993), que, posteriormente, foi aperfeiçoada por Borst (1997), resultando na seguinte declaração: "ontologia é uma especificação formal de uma conceituação compartilhada" [10] (BORST, 1997, p. 12).

3.3 **Objetivos e características da ontologia**

A ontologia tem como objetivo principal fornecer um meio onde seja possível descrever o universo de um dado domínio para, então, compartilhar essa descrição. A partir dela, conceitos e relações são especificados por meio de um vocabulário de termos bem definido e de um conjunto de regras lógicas que, por sua vez, definem os tipos de relação entre os conceitos. Um de seus maiores diferenciais é o modo como as informações são rotuladas e interligadas, pois, considerando que o significado do termo é estabelecido a partir de sua relação com os demais termos em uso, o significado das palavras é preservado, o que propicia o seu processamento pelas máquinas segundo parâmetros semânticos e, desta forma, torna mais amigável o consumo destas informações.

Face às dificuldades de busca e recuperação de informações em sistemas abertos como a Internet, para a tomada de decisões a utilização de ontologias constitui a técnica adequada e imprescindível às necessidades de comunicação,

---

[8] Texto original, na íntegra: "We can now clarify the role of an ontology, considered as a set of logical axioms designed to account for the intended meaning of a vocabulary."

[9] Texto original, na íntegra: "An ontology is an explicit specification of a conceptualization."

[10] Texto original, na íntegra: "Ontologies are defined as a formal specification of a shared conceptualization."

oferecendo um melhor reuso automático de informações – e, consequentemente, de conhecimento – no atual contexto da computação.

Gradativamente, a ontologia vem ganhando espaço como padrão de conhecimento estruturado disponível para reuso em larga escala por sistemas e programas, um acontecimento sem paralelo na história da Ciência da Computação, e deve se tornar o pilar de sustentação para aplicações como comércio eletrônico, workflow e gestão de conhecimento.

A ontologia também tem servido, em certas áreas, para unificar o conhecimento e para formar consenso sobre certos conceitos, promovendo a integração de grupos de pesquisas, além de ser utilizada, em alguns casos, com propósitos educativos. O seu emprego como elemento estruturador de informações pode representar para a história das Ciências da Informação e da Computação o que, para a história humana, representou a criação de enciclopédias e bibliotecas, ou seja: o armazenamento de conhecimento. Nesse cenário que se apresenta, o conhecimento tem condições de trafegar entre computadores e sistemas que, por sua vez, podem lançar mão deles, manipulando e aplicando-os no cumprimento de suas funções.

Segundo Gruber (2009), no contexto da Ciência da Computação, uma ontologia define um conjunto de primitivas de representação com o qual se modela um domínio de conhecimento ou discurso. As primitivas de representação são tipicamente:

a) **Classes:** conjuntos ou tipos de indivíduos, os quais são organizados em uma taxonomia;
b) **Atributos:** propriedades, características ou parâmetros que classes e relações podem ter;
c) **Relações:** ligações entre os membros da classe;
d) **Axiomas**: regras pertinentes ao domínio em questão.

As definições das primitivas de representação incluem informações sobre o seu significado e as limitações na sua aplicação de lógica consistente.

Para a sua implementação dentro dos moldes da *W3C* – que é o órgão responsável pelo desenvolvimento de padrões para a criação e a interpretação dos conteúdos voltados para a Web – a ontologia se apropria da estrutura sintática da língua, sendo formada por conjuntos de triplas constituídas pelos seguintes elementos:

a) **Sujeito**;
b) **Predicado**;
c) **Objeto**.

Cada tripla pode ser interpretada como uma afirmação sobre um *recurso**, sendo que a posição dos elementos nela contidos define a semântica desta afirmação. Na posição do sujeito, está o recurso sobre o qual a afirmação está sendo feita. O predicado é outro recurso que denota uma propriedade do sujeito e o relaciona através desta propriedade com o objeto. O objeto, por sua vez, pode ser um terceiro recurso ou possuir um valor literal, como um número ou uma sequência de caracteres.

Sobre as triplas, Berners-Lee, Hendler e Lassila (2001) dizem:

> São parecidas com uma sentença linguística elementar constituída de sujeito, verbo e objeto. [...] isto é versátil o suficiente para descrever a maioria dos dados processados pelos sistemas. [11] (Berners-LEE; HENDLER; LASSILA, 2001).

3.4 **Padrões de linguagem**

Levando em consideração tão somente as atuais definições da *W3C* – ou seja, aquelas voltadas especialmente para a Web Semântica – a ontologia pode ser expressa nos seguintes padrões de linguagem:

a) **XML:** descrição de conteúdo que estabelece estruturas arbitrárias nos documentos;
b) **RDF e RDFS:** representação da informação por meio de grafo, de expressividade limitada, que dá significado às estruturas;
c) **OWL:** extensão do *RDF* que oferece mais possibilidades de expressividade, internacionalização, escalabilidade, distribuição e

[11] Texto original, na íntegra: “Meaning is expressed by RDF, which encodes it in sets of triples, each triple being rather like the subject, verb and object of an elementary sentence. These triples can be written using XML tags. In RDF, a document makes assertions that particular things (people, Web pages or whatever) have properties (such as "is a sister of," "is the author of") with certain values (another person, another Web page). This structure turns out to be a natural way to describe the vast majority of the data processed by machines.”

acessibilidade, sendo disponibilizada nas sub-linguagens *OWL Lite*, *OWL DL* e *OWL Full*, esta sem garantias computacionais.

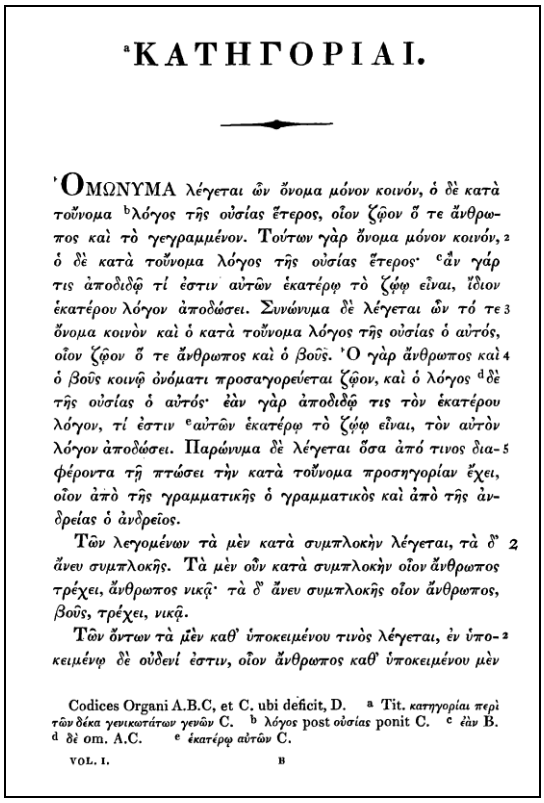
Embora existam outras linguagens similares disponíveis no mercado, os padrões apresentados acima correspondem às únicas soluções validadas, chanceladas e indicadas pela *W3C*, promovendo, assim, o uso maciço destes em escala mundial.

3.5 **Primeiro modelo de ontologia**

Nascido na cidade de Estagira, na Macedônia, onde morou até os 18 anos, Aristóteles (384 a.C. – 322 a.C.) frequentou a Academia de Platão, em Atenas, e lá estudou até a morte de seu mestre. Entre suas inúmeras obras, há um texto chamado *Categorias* (vide Figura 6) – em grego, *Κατηγοριαι*, que significa: acusação, indicação, atributo de alguma coisa, predicado – também conhecido como *Categorias Aristotélicas*, que abre não apenas o *Organon*, o conjunto de textos lógicos de Aristóteles, como também o *Corpus Aristotelicum*. Apesar de composta em apenas um livro, costuma-se dividir o conteúdo desta obra em duas partes: a primeira, que se estende do capítulo I ao IX, é chamada de *Prædicamenta* e considera-se genuinamente aristotélica; já a segunda parte, que se estende do capítulo X ao XV, é chamada de *Post-Prædicamenta* e não há certeza se a autoria é de Aristóteles ou de um de seus discípulos.

Em *Categorias*, Aristóteles faz uma análise dos vocabulários comum e filosófico, partindo das palavras, que são coisas ditas sem combinação, e das proposições, que são coisas ditas em combinação. A categoria propriamente dita refere-se às categorias ditas sem combinação, isto é, aos termos definidos enquanto coisas que servem para designar outras coisas. Assim sendo, é possível dizer, por exemplo, que a palavra *sábio* é uma categoria sem combinação, ao passo que *Platão é sábio* reúne um conjunto de categorias em combinação numa frase ou proposição.

De acordo com Chauí (2003), as categorias são coisas não combinadas com outras coisas ou palavras que servem para designar outras coisas. Elas são índices do que uma coisa é ou faz. As categorias são os predicados mais gerais de uma



[a]ΚΑΤΗΓΟΡΙΑΙ.

[1]Ὁμώνυμα λέγεται ὧν ὄνομα μόνον κοινόν, ὁ δὲ κατὰ τοὔνομα [b]λόγος τῆς οὐσίας ἕτερος, οἷον ζῷον ὅ τε ἄνθρωπος καὶ τὸ γεγραμμένον. Τούτων γὰρ ὄνομα μόνον κοινόν, [2]ὁ δὲ κατὰ τοὔνομα λόγος τῆς οὐσίας ἕτερος· [c]ἂν γάρ τις ἀποδιδῷ τί ἐστιν αὐτῶν ἑκατέρῳ τὸ ζῴῳ εἶναι, ἴδιον ἑκατέρου λόγον ἀποδώσει. Συνώνυμα δὲ λέγεται ὧν τό τε [3]ὄνομα κοινὸν καὶ ὁ κατὰ τοὔνομα λόγος τῆς οὐσίας ὁ αὐτός, οἷον ζῷον ὅ τε ἄνθρωπος καὶ ὁ βοῦς. Ὁ γὰρ ἄνθρωπος καὶ [4]ὁ βοῦς κοινῷ ὀνόματι προσαγορεύεται ζῷον, καὶ ὁ λόγος [d]δὲ τῆς οὐσίας ὁ αὐτός· ἐὰν γὰρ ἀποδιδῷ τις τὸν ἑκατέρου λόγον, τί ἐστιν [e]αὐτῶν ἑκατέρῳ τὸ ζῴῳ εἶναι, τὸν αὐτὸν [5]λόγον ἀποδώσει. Παρώνυμα δὲ λέγεται ὅσα ἀπό τινος διαφέροντα τῇ πτώσει τὴν κατὰ τοὔνομα προσηγορίαν ἔχει, οἷον ἀπὸ τῆς γραμματικῆς ὁ γραμματικὸς καὶ ἀπὸ τῆς ἀνδρείας ὁ ἀνδρεῖος.

Τῶν λεγομένων τὰ μὲν κατὰ συμπλοκὴν λέγεται, τὰ δ' [2]ἄνευ συμπλοκῆς. Τὰ μὲν οὖν κατὰ συμπλοκὴν οἷον ἄνθρωπος τρέχει, ἄνθρωπος νικᾷ· τὰ δ' ἄνευ συμπλοκῆς οἷον ἄνθρωπος, βοῦς, τρέχει, νικᾷ.

Τῶν ὄντων τὰ μὲν καθ' ὑποκειμένου τινὸς λέγεται, ἐν ὑπο-[2]κειμένῳ δὲ οὐδενί ἐστιν, οἷον ἄνθρωπος καθ' ὑποκειμένου μὲν

Codices Organi A.B.C, et C. ubi deficit, D. [a] Tit. *κατηγορίαι περὶ τῶν δέκα γενικωτάτων γενῶν* C. [b] λόγος post οὐσίας ponit C. [c] ἐὰν B. [d] δὲ om. A.C. [e] *ἑκατέρῳ αὐτῶν* C.

VOL. I. B

Figura 6 - Primeira página da obra *Categorias*, de Aristóteles

coisa ou de um sujeito, que aparecem em tudo o que dizemos ou pensamos e que revelam qual é, ou o que é, a entidade que está sendo predicada.

A categoria funciona, então, como a propriedade predicativa de um termo. As categorias são os modos do ser, de se dizer o que o ser é. Para Aristóteles, o ser se diz de muitas maneiras, o que equivale a dizer que o ser pode ser dito com inúmeras predicações ou categorias. Ao dizer *S é P*, tem-se que *P* é uma categoria de *S* porque *P* designa o que *S* é, ou seja, *P* serve como índice do que *S* é ou faz. As predicações afirmam o que as coisas são segundo sua qualidade, sua quantidade, suas relações, sua atividade ou passividade, sua espacialidade e temporalidade. As categorias, portanto, se referem a determinações – ou seja, propriedades – pertencentes a um ser e das quais o pensamento deverá servir-se para conhecê-lo e exprimi-lo.

Uma outra característica da categoria é que ela é tudo o que a nossa percepção pode captar das coisas. Por serem coisas que se dão direta e imediatamente aos sentidos, elas não necessitam de qualquer demonstração ou prova. Nesses termos, a categoria indica ou dá o sinal, numa apreensão direta e

imediata, do que uma coisa é como entidade simples. Por serem entidades simples, as categorias não podem ser desmembradas em partes para serem analisadas visando à verificação e prova. A verificação e a prova só podem acontecer quando uma categoria entra numa proposição, ou seja, quando ela se combina com outra categoria. Desse modo, o desmembramento para verificação e prova só pode acontecer com a proposição, e não com a categoria. Não é possível, por exemplo, submeter o termo *sábio* ao desmembramento analítico; por este termo ser uma entidade simples, pode-se apenas tomá-lo como um dado imediato da percepção. O mesmo não ocorre com a frase ou proposição *Platão é sábio*; com a demonstração pode-se operar a verificação e prova, havendo a possibilidade de questionar se Platão é de fato sábio; esta proposição será então submetida a uma contraprova, levantando, por exemplo, a hipótese de que Platão talvez seja sábio.

Em suma, Aristóteles enumera todos os possíveis tipos de coisa que tenham condições de assumir a função de sujeito ou predicado de uma proposição, correspondendo à primeira tentativa do homem de organizar o conhecimento e, talvez, às noções aristotélicas mais discutidas de todos os tempos. Assim sendo, essa obra pode ser entendida como a primeira ontologia já criada pelo homem e, devido às suas características generalistas, será utilizada como fonte de referência e inspiração para o modelo que será proposto mais à frente.

Ao todo, Aristóteles propõe 10 (dez) categorias que, uma vez associadas entre si e atribuídos os seus valores, são capazes de descrever – de forma estruturada – qualquer coisa. São elas:

- **Substância:** corresponde àquilo que é. Exemplo: este é o homem, esta é a mulher;
- **Quantidade:** estabelece-se quanto é, se é muito ou pouco. Exemplo: uma mulher é alta, baixa;
- **Qualidade:** pode-se dizer que é. Exemplo: é verde, azul;
- **Relação:** considera-se a relação entre uma coisa e outra. Exemplo: uma criança é menor ou igual a outra;
- **Lugar:** determina-se onde uma coisa está. Exemplo: está lá, aqui, na França;
- **Tempo:** pode-se dizer quando é, quando deixa de ser, quando foi. Exemplo: é agora, foi no século XXI;

- **Posição:** estabelece-se a postura da coisa. Exemplo: está em pé, sentada;
- **Posse:** refere-se ao que a coisa tem. Exemplo: está vestida, ou seja, tem roupa;
- **Ação:** trata-se daquilo que a coisa faz. Exemplo: ato da semente de germinar;
- **Paixão:** corresponde àquilo que a coisa sofre, se submete. Exemplo: a formiga é morta.

Nessa obra, Aristóteles diz o seguinte, de acordo com Ackrill (1963):

> Das expressões que são ditas sem qualquer combinação, cada uma significa a substância, ou então quanto, qual, em relação a quê, onde, quando, estar numa posição, ter, fazer, ser afetado. Para dar uma ideia aproximada, são exemplos de substância homem, cavalo; de quantidade: quatro pés, cinco pés; de qualidade: branco, gramatical; de relação: dobro, metade, maior; de lugar: no Liceu, no mercado; de tempo: ontem, ano passado; de posição: deitado, sentado; de posse: calçado, armado; de ação: cortando, queimando; de paixão: sendo cortado, sendo queimado. (ACKRILL, 1963, p. 5).

Considerando que *Categorias* é o equivalente a uma conceituação cuja definição, à luz da Ciência da Computação, é compatível com a de uma ontologia, pode-se designar as categorias elencadas por Aristóteles como classes, ou seja, como conceitos que representam conjuntos ou tipos de objetos.

3.6 **Ontologias de fundamentação e de domínio**

Na literatura, é possível encontrar diferentes sistemas de classificação de ontologias (GÓMEZ-PÉREZ; FERNÁNDEZ-LÓPEZ; CORCHO, 1999). Porém, aquele que parece mais adequado – por utilizar como principal critério a conceituação, uma característica fundamental das ontologias – é o sistema apresentado por Nicola Guarino, em seu artigo intitulado *Formal Ontology and Information Systems*, publicado em 1998, que sugere o desenvolvimento de diferentes tipos de ontologia de acordo com o nível de generalidade necessária.

Considerando os tipos de ontologia existentes, somente interessa explorar dois deles, que são as *ontologias de fundamentação**, ou *ontologias de alto nível*, e as ontologias de domínio.

As ontologias de fundamentação representam um tipo de ontologia que reúne resultados da Ontologia Formal em um sistema de categorias independentes de domínio – conceitos como *parte*, *todo*, *papel* e *evento* – usadas para articular conceituações de diversos domínios.

Em relação a esse tipo de ontologia, Guizzardi e Wagner (2005) dizem o seguinte:

> Uma ontologia de fundamentação, às vezes chamada também de 'ontologia de alto nível', define um conjunto de categorias ontológicas independente de domínio que formam uma base geral para ontologias específicas de domínio. [12] (GUIZZARDI; WAGNER, 2005, p. 346).

Essas categorias a que se referem os autores deixam clara a distinção entre uma *propriedade* de uma coisa, o *tipo* ao qual a coisa pertence, o *papel* desempenhado pela coisa, e assim por diante. Assim, além de características sintáticas e estruturais, pode-se considerar também a natureza do conceito, fornecendo mais parâmetros necessários ao estabelecimento de associações entre entidades similares, bem como ao descarte de possíveis associações entre entidades de naturezas diferentes.

Segundo Silva (2010), em geral, as ontologias de fundamentação são utilizadas como modelos de referência que definem os conceitos permitidos em uma linguagem de modelagem conceitual bem fundamentada, possibilitando que ela capture a semântica do mundo real. Além dos benefícios na construção de modelos conceituais de um domínio, a utilização de ontologias de fundamentação pode ser útil também no alinhamento de ontologias, pois que, identificando as meta-categorias das quais os conceitos são derivados, é possível estabelecer sua natureza, diferenciando, por exemplo, *objetos* de *processos*, ou *tipos* de coisas de *papéis* desempenhados por estas. Essa distinção pode ajudar a evitar associações incorretas no processo de alinhamento, restringindo a identificação de termos

[12] Texto original, na íntegra: "A foundational ontology, sometimes also called 'upper level ontology', defines a range of top-level domain-independent ontological categories, which form a general foundation for more elaborated domain-specific ontologies."

equivalentes àqueles derivados da mesma meta-categoria, ou seja, que possuem a mesma natureza conceitual.

Definir conceitos que devem servir de base para domínios de conhecimento de todo tipo é uma tarefa nada fácil, e as chamadas ontologias de fundamentação foram criadas para justamente lidar com esta questão, entre outras coisas. Existem ontologias de fundamentação com orientações diversas e, atualmente, há inúmeras propostas em andamento. A título de exemplo, temos: *Basic Formal Ontology - BFO*, *Cyc*, *Descriptive Ontology for Linguistic and Cognitive Engineering – DOLCE*, *General Formal Ontology - GFO*, *Generalised Upper Model - GUM*, *Gist*, *Suggested Upper Merged Ontology - SUMO* e *Unified Foundational Ontology - UFO*.

Um exemplo conhecido de ontologia de fundamentação na literatura de computação é o modelo *Bunge-Wand-Weber - BWW*, proposto por Wand e Weber (1995), baseado na teoria metafísica originalmente desenvolvida por Bunge (1977).

As ontologias de domínio, por sua vez, podem ser definidas da seguinte forma, segundo Guarino (1997, 1998):

> Ontologias de domínio descrevem o vocabulário relacionado a um domínio genérico (como medicina ou automóveis) por meio da especialização dos termos introduzidos na ontologia de fundamentação. [13] (GUARINO, 1997, p. 146, 1998, p. 10).

Uma vez esclarecidas as características das ontologias de domínio, é razoável afirmar que, pelo fato destas ontologias representarem conceitos muito específicos, não é possível garantir a compatibilidade entre dois ou mais modelos deste tipo. À medida que os sistemas passam a depender cada vez mais de ontologias de domínio, um equívoco comum é a criação de representações mais abrangentes a serem incorporadas por estas ontologias, o que acaba por gerar um grande problema para os engenheiros de ontologia; de modo geral, este é um processo realizado manualmente, que consome muito tempo e recursos. A utilização de uma ontologia de fundamentação, que forneça uma definição comum para conceitos essenciais, representa uma alternativa para se evitar tal problema.

---

[13] Texto original, na íntegra: *"Domain ontologies and task ontologies describe, respectively, the vocabulary related to a generic domain (like medicine, or automobiles) or a generic task or activity (like diagnosing or selling), by specializing the terms introduced in the top-level ontology"*.

Considerando que este trabalho tem por objetivo propor uma ontologia que trata da cobertura de fatos jornalísticos – podendo esta ser classificada como uma ontologia de domínio – é interessante que a representação deste conhecimento utilize como gabarito uma ontologia de fundamentação, pelos motivos já expostos. E, no intuito de consolidar esta abordagem, faz-se relevante a justificativa dada por Guizzardi *et al*. (2009a):

> Uma ontologia de domínio, no sentido usado pelas demais comunidades em computação, é um tipo particular de modelo conceitual. Em particular, é um modelo conceitual que deve satisfazer o requisito adicional de servir como uma representação de consenso (ou modelo de referência) de uma conceituação compartilhada por uma determinada comunidade. Portanto, se uma ontologia de domínio é, antes de qualquer coisa, um modelo conceitual, uma linguagem adequada para representação de ontologias de domínio deve satisfazer os requisitos gerais de uma linguagem adequada para modelagem conceitual, ou seja, deve ter como teoria subjacente uma ontologia de fundamentação. Em outras palavras, ontologias (no sentido adotado em filosofia e em modelagem conceitual) representam ferramentas conceituais de importância fundamental para a criação de ontologias de domínio de qualidade (no sentido adotado nas demais áreas). (GUIZZARDI *et al*., 2009a, p. 1-2).

3.7 **Ontologia de referência**

Para o presente trabalho, o modelo *Unified Foundational Ontology - UFO* (vide Anexo D) foi escolhido como base para a construção da ontologia de jornalismo, pois, considerando as suas características, existem três que atendem as necessidades básicas do projeto, que são (a) ter a capacidade de cobrir um grande espectro de agentes de software que estão por vir, bem como de situações de uso e de integração com sistemas diversos, devido ao seu alto grau de generalização, (b) dispor de uma vasta gama de definições capaz de abarcar conceitos provenientes de ontologias de domínio de todo gênero, de modo a garantir um mínimo de expressividade e interoperabilidade e (c) oferecer uma quantidade adequada – ou seja, sem excessos – de conceitos, dado que uma das grandes barreiras na utilização de ontologias, de modo geral, é a acentuada curva de aprendizado necessária até a completa assimilação do modelo pelos engenheiros de ontologia.

Grosso modo, *UFO* faz uso aprofundado de conceitos de caráter filosófico que, por sua vez, servem como alicerces para conceitos que tangem domínios específicos de conhecimento, como será possível perceber adiante.

Vale observar que, embora existam similaridades entre as *Categorias Aristotélicas* e *UFO*, já que este faz uso dos conceitos lançados no referido trabalho de Aristóteles, são perceptíveis as diferenças entre uma ontologia e outra. Enquanto que *Categorias* é extremamente simples pelo fato de considerar o conhecimento como um conjunto de proposições organizado sob a luz da Filosofia, *UFO* estabelece uma organização mais complexa e criteriosa, também baseada na Filosofia, porém complementada pela ótica da Ciência da Computação e de outras áreas do conhecimento, levando em conta a sua aplicação em sistemas de informação de toda sorte. É interessante notar, também, que o modelo *UFO* é resultado de um projeto que foi iniciado há poucos anos, sendo uma proposta vinte e três séculos mais nova que as *Categorias Aristotélicas*.

Inicialmente proposta por Giancarlo Guizzardi e Gerd Wagner, em 2004, *UFO* é uma ontologia de fundamentação que foi concebida a partir da adaptação e ampliação de uma série de teorias provenientes, principalmente, da Ontologia Formal, Lógica Filosófica, Filosofia da Linguagem, Linguística e Psicologia Cognitiva, e desenvolvida com o intuito de prover uma fundamentação ontológica para linguagens gerais de modelagem conceitual, reunindo, ao longo dos anos, teorias axiomáticas que versam sobre as principais categorias de conceitos utilizados nesta área específica.

UFO sintetiza resultados de outras ontologias de fundamentação, a *GFO*/*GOL*[14] e a *OntoClean*/*DOLCE*[15], que, embora apresentem propriedades interessantes, têm limitações consideráveis no que diz respeito à capacidade de captura dos conceitos básicos das linguagens de modelagem conceitual. Assim, a proposta de *UFO* é justamente unificar essas ontologias aproveitando suas características positivas e sanando as limitações detectadas (GUIZZARDI; WAGNER, 2005, 2010).

---

[14] *General Formal Ontology - GFO* é uma ontologia subjacente à *General Ontology Language - GOL*, que é uma linguagem desenvolvida pelo grupo de pesquisa *OntoMed*, da Universidade de Leipzig.

[15] Junção da ontologia *OntoClean* com a *Descriptive Ontology for Linguistic and Cognitive Engineering - DOLCE*, desenvolvida pelo grupo de pesquisa *ISTC-CNR-LOA*.

Em sua estruturação, a ontologia *UFO* é composta por três fragmentos denominados *UFO-A* (ontologia de *objetos*), *UFO-B* (ontologia de *eventos*) e *UFO-C* (ontologia de *entidades sociais e intencionais*).

UFO-A define o núcleo dessa ontologia, sistematizando conceitos como, por exemplo, *tipos e estruturas taxonômicas*, *relações todo-parte*, *propriedades intrínsecas e espaços de valores de atributos*, *propriedades relacionais*, entre outros. Segundo Guizzardi (2005), esse fragmento constitui uma teoria estável, formalmente caracterizada com o aparato de uma lógica modal de alta expressividade e possuindo forte suporte empírico promovido por experimentos em psicologia cognitiva.

UFO-B tem por objetivo sistematizar conceitos como *estados*, *processos*, *eventos*, *relações temporais*, entre outros. E *UFO-C*, por sua vez, construída sobre os fragmentos anteriores, visa sistematizar conceitos que incluem *agente*, *ação*, *estados intencionais*, *delegação*, *compromissos e reinvindicações sociais* etc.

É importante destacar que a ontologia *UFO* tem sido utilizada em diversos estudos de caso de construção de ontologias de domínio, bem como no desenvolvimento de aplicações baseadas nessas ontologias. Exemplos de domínios abordados incluem Eletrocardiologia, cujo trabalho foi desenvolvido por Gonçalves, Zamborlini e Guizzardi (2009), Exploração e Produção de Petróleo, por Guizzardi *et al*. (2009b) e Lopes *et al*. (2009), entre outros.

Faz-se necessário esclarecer que não existe uma ontologia de fundamentação absolutamente melhor ou pior que outra, mas aquela que melhor se adequa a uma determinada necessidade. E, como não poderia deixar de ser, *UFO* não é exceção. Além disso, como até o momento não há qualquer ferramenta capaz de comparar a habilidade dessas ontologias em responder as questões de competência de uma determinada aplicação, é preciso analisá-las caso a caso, sob um ponto de vista técnico, o que demanda uma pesquisa e análise minuciosas.

A seguir, a ontologia *UFO* será apresentada em detalhes, desconsiderando, no entanto, os conceitos e relações que não têm relevância para o presente trabalho ou que não darão suporte à construção da ontologia de notícias, proposta no capítulo seguinte, e seus possíveis desdobramentos. Para o entendimento dos diagramas de representação que serão amplamente utilizados como recurso descritivo tanto de *UFO* quanto da ontologia de notícias, um vocabulário visual foi preparado (vide Figura 7).

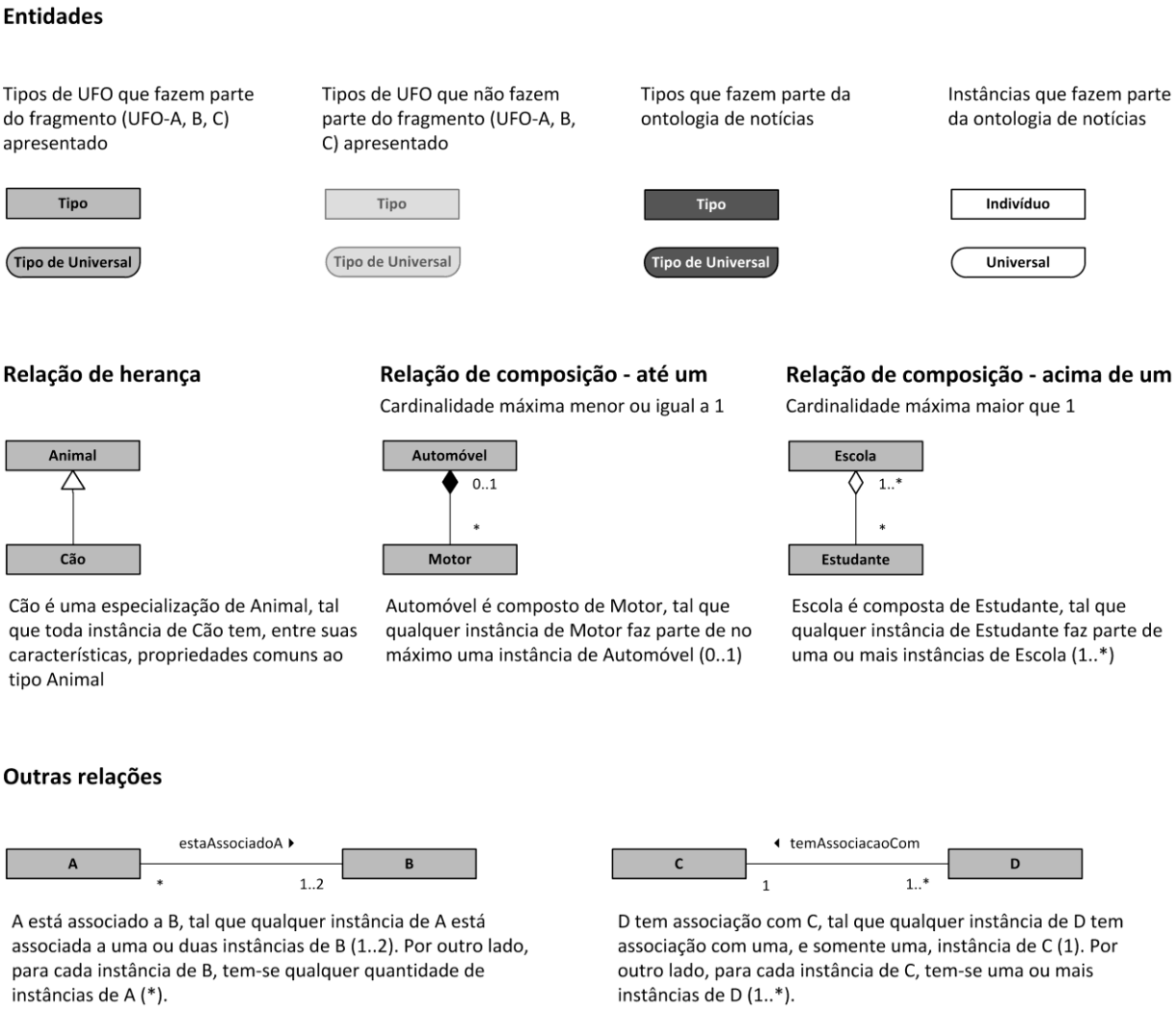


Figura 7 - Vocabulário visual para leitura dos diagramas de representação das ontologias

Entre os conceitos tratados no referido vocabulário, é importante explicar dois deles, a saber: *cardinalidade** e *instanciação**.

Cardinalidade corresponde a uma restrição de integridade que representa regras de negócio que, por sua vez, estabelecem as quantidades mínima e máxima de relacionamentos entre uma determinada entidade e outras. Resumindo, tais quantidades são expressas pela cardinalidade, através de um par de valores *A..B* onde *A* é a *cardinalidade mínima* e *B* é a *cardinalidade máxima*. Enquanto que a cardinalidade mínima estabelece a menor quantidade de relacionamentos que deve existir entre uma entidade qualquer e outras distintas, a cardinalidade máxima estabelece a maior.

Instanciação diz respeito à particularização de uma entidade ou, melhor dizendo, à individualização de um tipo ou categoria. Desse processo, tem-se como

resultado a produção de instâncias, que são objetos únicos, distintos, que representam o tipo ou categoria do qual fazem parte. Tal conceito é tratado em *UFO-A*, como poderá ser visto adiante.

3.7.1 **UFO-A**

Primeiramente, é importante esclarecer que, em *UFO*, o conceito mais genérico é chamado de *Entidade* (*Entity*)[16], que deve ser entendido como algo concebível ou perceptível. Todos os demais conceitos são especializados a partir de Entidade.

O fragmento *UFO-A* nada mais é que uma ontologia que trata particularmente de entidades chamadas *endurantes* (*endurants*), que correspondem a objetos ordinários do cotidiano que persistem no tempo. A divisão entre endurantes e *perdurantes* (*perdurants*), este último grupo tratado na ontologia *UFO-B*, pode ser intuitivamente compreendida segundo a diferença entre *objetos* e *eventos*, respectivamente.

De acordo com Guizzardi, Falbo e Guizzardi (2008), nessa ontologia há uma distinção fundamental (vide Figura 8) que ocorre entre as categorias *Indivíduo* (*Individual* ou *Particular*) e *Universal* (*Universal*). Universal é o tipo ou categoria geral que representa os padrões de características presentes em diferentes indivíduos; este tipo se aplica, por exemplo, aos conceitos ou categorias de indivíduos *Pessoa*, *Político* e *Urso*, que são padrões de características comuns presentes em certos

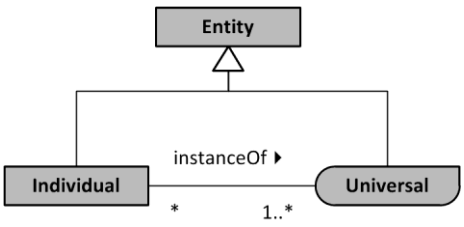


Figura 8 - *UFO-A*: distinção entre Indivíduo e Universal

[16]Os conceitos de *UFO* são apresentados em português, havendo a indicação do termo original, em inglês, entre parênteses. Nos diagramas, porém, os conceitos são apresentados somente em inglês.

indivíduos. Por outro lado, Indivíduo é a categoria ou tipo geral que se aplica aos indivíduos, que são entidades que existem na realidade e possuem uma identidade única; este tipo se aplica, por exemplo, aos indivíduos *Heráclito Fortes* e *Knut*, por exemplo. Além disso, cada indivíduo num domínio deve ser instância de pelo menos um universal. Sendo assim, com base nos exemplos dados, *Heráclito Fortes* é um indivíduo do domínio que instancia os universais *Pessoa* e *Político*, enquanto *Knut* instancia o universal *Urso*.

Em relação aos indivíduos (vide Figura 9), estes podem ser do tipo *Indivíduo Concreto* (*Concrete Particular*), que é atribuído a um objeto particular, determinado, ou *Indivíduo Abstrato* (*Abstract Particular*), que se aplica a uma qualidade ou característica separada do objeto a que pertence ou a que está ligada. Os indivíduos concretos se especializam no tipo *Endurante* (*Endurant*) – que, como já dito anteriormente, tem a ver com objetos ordinários do cotidiano que persistem no tempo – e *Perdurante* (*Perdurant*), que abrange indivíduos compostos por partes temporais e será explicado em *UFO-B*, enquanto que os indivíduos abstratos podem ser do tipo *Estrutura de Qualidade* (*Quality Structure*), *Quale* (*Quale*) ou *Proposição* (*Proposition*). Uma estrutura de qualidade pode ser entendida como uma estrutura de medição (ou um espaço de valores) onde qualidades individuais podem tomar seus valores. Ou seja, uma qualidade está associada a uma estrutura de qualidade. As qualidades *idade* e *peso*, por exemplo, estão associadas a um espaço de valores que é uma estrutura linear isomórfica ao eixo positivo dos números reais. Quale, por sua vez, corresponde a um ponto na estrutura da qualidade, que é uma percepção ou concepção de uma propriedade intrínseca.

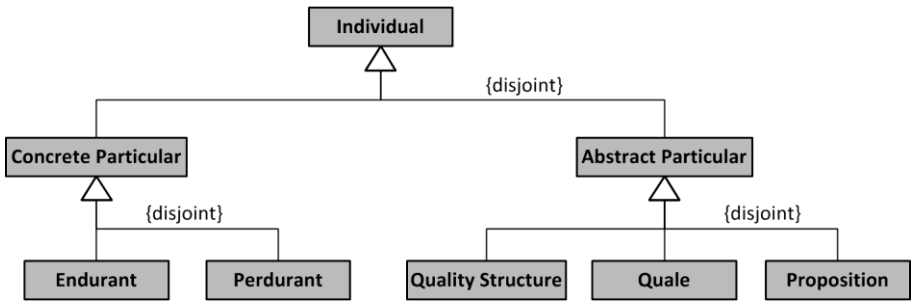


Figura 9 - *UFO-A*: Indivíduo

Os universais (vide Figura 10), por outro lado, classificam-se como *Universal de Endurante* (*Endurant Universal*) ou *Universal de Perdurante* (*Perdurant Universal*), sendo que os universais de endurante ainda podem ser do tipo *Universal Monádico* (*Monadic Universal*) ou *Relação* (*Relation Universal*). O tipo Universal Monádico é a categoria que se aplica aos conceitos que são padrões empregados em indivíduos singulares, enquanto que o tipo Relação se aplica às relações, que são padrões atribuídos a grupos de dois ou mais indivíduos.

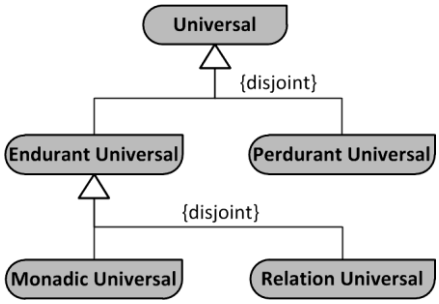


Figura 10 - *UFO-A*: Universal

Em relação à distinção entre categorias de indivíduos e universais, pode-se dizer que, de maneira geral, para cada especialização de indivíduo presente em *UFO-A*, há uma especialização de universal equivalente. Em suma, as especializações de indivíduos são instâncias das especializações de universais correspondentes. Desse modo, para os tipos Endurante e Universal Monádico, há as especializações *Substancial* (*Substantial*) e *Aspecto* (*Moment*)[17]. Substanciais são as entidades do mundo real que persistem no tempo, mantendo as suas identidades; é possível citar, por exemplo, as entidades físicas e sociais do dia-a-dia, como uma pessoa, uma cadeira e um lago. Aspectos, por sua vez, são entidades que denotam a instanciação de propriedades de outras entidades, sendo inerentes a estas últimas. Enquanto que os substanciais são existencialmente independentes, os aspectos só podem existir em outras entidades, ou seja, são existencialmente dependentes (vide Figura 11).

[17]Este conceito é derivado do termo alemão *momente*, e diz respeito às propriedades que as entidades possuem, não havendo qualquer relação com a noção de instante de tempo.

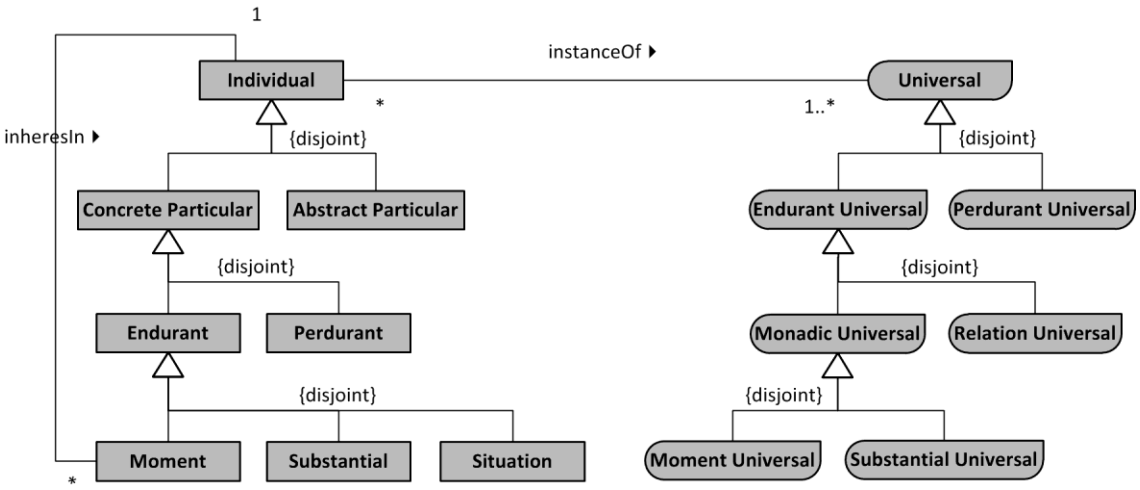


Figura 11 - *UFO-A*: Aspecto e Substancial

Assim sendo, no que concerne à classificação em questão aplicada a indivíduos, é possível entender que a *idade de Dilma* é um aspecto, ou seja, um indivíduo que objetifica sua idade como propriedade abstrata, e é inerente a ela, neste caso a *Dilma*. A *enxaqueca de Dilma*, por sua vez, pode ser interpretada da seguinte maneira: a *intensidade* é um aspecto da *enxaqueca de Dilma*, enquanto que a *enxaqueca* é um aspecto de *Dilma*; deste modo, a *intensidade* é inerente à *enxaqueca*, que é inerente a *Dilma*. Aplicando este mesmo princípio a universais monádicos, tem-se que o conceito *Pessoa* é um universal do tipo *Substancial Universal* (*Substantial Universal*) porque seus indivíduos são do tipo Substancial, como *Dilma*; e, analogamente, o conceito *enxaqueca* é um universal do tipo *Aspecto Universal* (*Moment Universal*) porque seus indivíduos são do tipo Aspecto, como a *enxaqueca de Dilma*.

Diferentemente dos universais, os indivíduos do tipo Endurante têm ainda como especialização o tipo *Situação* (*Situation*). Situações são entidades complexas constituídas, possivelmente, por vários substanciais ou situações, sendo tratadas como um sinônimo para o que a literatura chama de *estado de coisas** (*state of affairs*), ou seja, como uma porção da realidade que pode ser compreendida como um todo. Por exemplo: *Dilma está gripada e com enxaqueca*; com base nesse conceito, define-se a relação *estar presente em* entre substanciais e situações constituídas por eles; dessa forma, pode-se dizer que o substancial *Dilma* e seus aspectos *gripe* e *enxaqueca* estão presentes na situação *Dilma está gripada e com enxaqueca*.

Dentre os indivíduos classificados como Substancial (vide Figura 12), interessam neste trabalho aqueles que têm uma identidade bem definida, classificados como *Objeto* (*Object*); este tipo se divide em *Complexo Funcional* (*Functional Complex*), *Coletivo* (*Collective*) e *Quantidade* (*Quantity*). O tipo Complexo Funcional é tal que suas instâncias podem ser compostas por outros complexos funcionais, desde que as partes exerçam papéis diferentes no todo. Por exemplo, o corpo humano é um complexo funcional composto por partes com diferentes papéis, como coração e cérebro. Em contrapartida, as instâncias do tipo Coletivo possuem partes que exercem o mesmo papel funcional no todo, como, por exemplo, uma floresta, que é um conjunto de árvores, ou uma pilha de livros. Finalmente, o tipo Quantidade agrupa indivíduos que são porções maximais de uma quantidade de matéria, por exemplo, a quantidade (maximal) de água dentro de um copo.

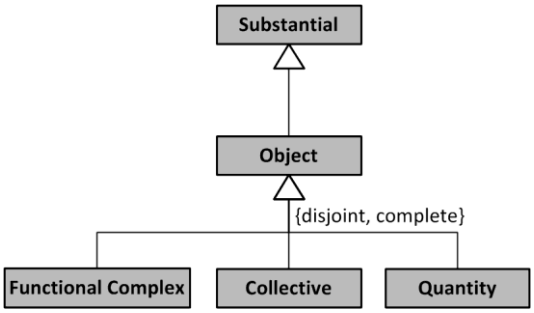


Figura 12 - *UFO-A*: Substancial

Os substanciais universais (vide Figura 13), por sua vez, podem ser do tipo *Sortal* (*Sortal Universal*) ou *Não Sortal* (*Characterizing Universal*). Os universais do tipo Sortal são tais que agregam indivíduos com o mesmo princípio de identidade. Supondo, por exemplo, que a impressão digital defina a identidade de uma pessoa, são universais do tipo Sortal os conceitos *Pessoa* e *Adulto*, uma vez que todos agregam indivíduos que possuem o mesmo princípio de identidade, como *Dilma* e *Heráclito Fortes*. Por outro lado, universais do tipo Não Sortal são tais que agregam indivíduos com princípios de identidade diferentes. Assim sendo, supondo que a identidade de uma empresa seja definida pelo CNPJ, o conceito *Item Assegurável*, que agrega pessoas e empresas, é considerado um universal do tipo Não Sortal.

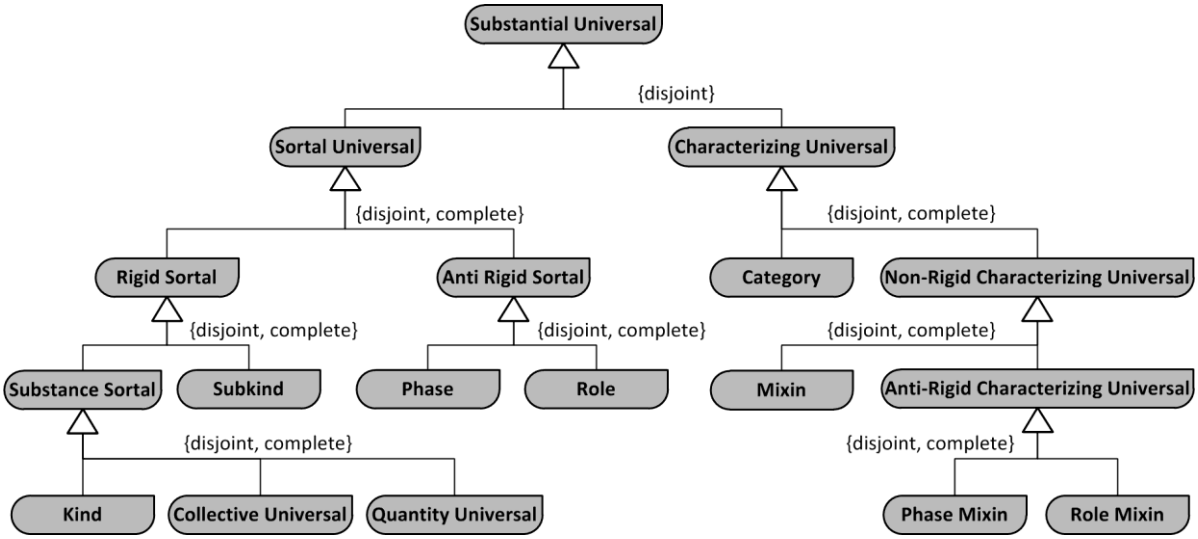


Figura 13 - *UFO-A*: Substancial Universal

Os sortais são divididos em *Sortal Rígido* (*Rigid Sortal*) e *Sortal Antirrígido* (*Anti Rigid Sortal*). Universais do tipo Sortal Rígido são conceitos cujos padrões de características são aplicados a todas as suas instâncias enquanto existirem, como *Pessoa* e *Empresa*. *Dilma*, por exemplo, é necessariamente instância de *Pessoa* enquanto existir. Universais do tipo Sortal Antirrígido, ao contrário, são conceitos cujos padrões de características podem eventualmente não se aplicar a todas as suas instâncias enquanto existirem, como *Presidente* e *Político*. Por exemplo: *Dilma* é instância do conceito *Presidente* em um determinado instante, e pode deixar de sê-lo em outro.

Os universais do tipo Sortal Rígido são classificados como *Sortal de Substância* (*Substance Sortal*) quando proveem o princípio de identidade aos seus indivíduos, ou como *Subtipo* (*Subkind*) quando apenas herdam este princípio. Considerando, por exemplo, que a impressão digital é o princípio de identidade provido a toda instância do conceito *Pessoa*, este é do tipo Sortal de Substância, enquanto os conceitos *Homem* e *Mulher* são do tipo Subtipo, pois, além de serem rígidos, especializam o conceito *Pessoa* e, portanto, herdam deste o princípio de identidade. O tipo Sortal de Substância especializa-se, ainda, em *Espécie* (*Kind*), *Universal de Coletivo* (*Collective Universal*) e *Universal de Quantidade* (*Quantity Universal*), que categorizam conceitos cujas instâncias são do tipo Complexo Funcional, Coletivo e Quantidade, respectivamente.

Sortais antirrígidos, por outro lado, são divididos em *Fase* (*Phase*) e *Papel* (*Role*). Conceitos do tipo Fase são determinados por um aspecto intrínseco do indivíduo e representam possíveis estágios na história do sortal universal correspondente, constituindo uma partição deste. *Dilma*, por exemplo, instancia a fase *Idoso*, visto que sua idade, que é um aspecto intrínseco, é superior a 60 anos. Conceitos do tipo *Papel*, como *Presidente* e *Político*, são determinados por um aspecto relacional do indivíduo e representam papéis desempenhados ao longo da história do sortal universal correlato. Por exemplo: *Dilma* instancia o papel de *Presidente*, dado que *ocupa o cargo de* – este é o aspecto relacional – *Presidente do Brasil*.

Quanto aos universais do tipo Não Sortal, estes são classificados como *Categoria* (*Category*) ou *Não Sortal Não Rígido* (*Non-Rigid Characterizing Universal*). O tipo Categoria também é chamado de *Não Sortal Rígido* (*Rigid Characterizing Universal*), e um exemplo de instância é o conceito *Item Assegurável*, que generaliza conceitos rígidos com diferentes princípios de identidade, como *Pessoa* e *Empresa*. Por sua vez, o tipo Não Sortal Não Rígido é dividido em *Misto* (*Mixin*) e *Não Sortal Antirrígido* (*Anti-Rigid Characterizing Universal*). O tipo Misto também é chamado de *Não Sortal Semi-rígido* (*Semi-Rigid Characterizing Universal*), e um exemplo de instância é o conceito *Grudável*, que é ao mesmo tempo rígido para alguns indivíduos, como *adesivo* e *cola*, e antirrígido para outros, como *arroz*. Já o tipo Não Sortal Antirrígido tem como especialização os tipos *Misto de Fase* (*Phase Mixin*) e *Misto de Papel* (*Role Mixin*), sendo exemplos de instância, respectivamente, *Adulto*, que generaliza conceitos antirrígidos com diferentes aspectos intrínsecos, como *Maior de Idade* e *Idoso*, e *Cliente,* que generaliza conceitos antirrígidos com diferentes princípios de identidade, como *Cliente Pessoa Física* e *Cliente Pessoa Jurídica*.

No que diz respeito às entidades do tipo Aspecto (vide Figura 14), a dependência existencial – que, em *UFO*, agrega o conceito que é chamado de *inerência* (*inherence*) – é utilizada para diferenciá-las, tanto entre indivíduos quanto entre universais monádicos. Os aspectos se dividem nos tipos *Aspecto Intrínseco* (*Intrinsic Moment*), que é dependente de uma única entidade, a exemplo da cor, que depende de um único objeto para existir – ou seja, diz respeito especificamente à qualidade que está presente em um dado objeto – e *Aspecto Relacional* (*Relator*), que, por outro lado, depende de várias entidades, como, por exemplo, um emprego,

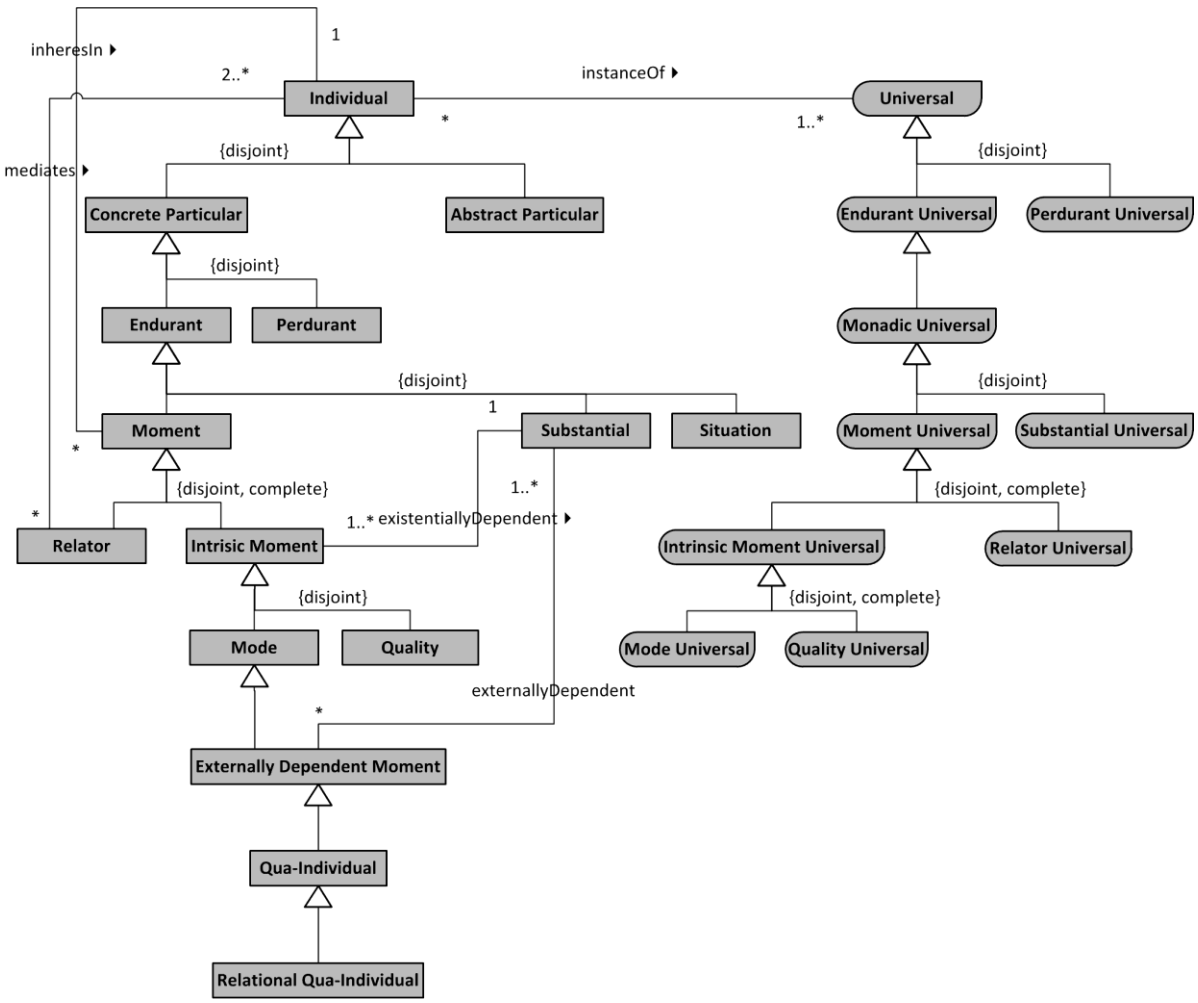


Figura 14 - *UFO-A*: Aspecto

que envolve um empregador e um empregado, ou um acordo de parceria, que agrega ao menos duas partes.

As entidades do tipo Aspecto Intrínseco denotam propriedades intrínsecas da entidade portadora. Se essas propriedades são mensuráveis, ou seja, se têm um valor em uma ou mais dimensões de qualidade, elas equivalem a aspectos do tipo *Qualidade* (*Quality*). Aspectos como *idade* e *peso*, por exemplo, têm um valor numa estrutura unidimensional – considera-se que a *idade de Dilma* é um aspecto inerente à *Dilma* e possui o valor *64 anos* – enquanto outras, como *cor* e *sabor*, têm seu valor numa estrutura multidimensional. O aspecto que não tem representação em um sistema de medida, por sua vez, é considerado como sendo do tipo *Modo* (*Mode*), como a *enxaqueca de Dilma*, por exemplo. Vale notar que essa distinção se aplica, mais uma vez, a indivíduos e universais correspondentes.

Entre os indivíduos do tipo Modo, há o tipo *Modo Externamente Dependente* (*Externally Dependent Mode*), que vale para todo e qualquer modo existencialmente dependente de um indivíduo que, por sua vez, é independente da entidade portadora. Esse tipo possui uma especialização chamada de *Indivíduo-Qua* (*Qua-Individual*) – cujo termo é proveniente da ideia de se considerar um indivíduo a respeito de certas particularidades apenas – que se refere a um aspecto que representa o conjunto de características que um indivíduo adquire em um determinado contexto. Se o contexto é dado pela instanciação de uma relação, o aspecto é considerado como sendo do tipo *Indivíduo-Qua Relacional* (*Relational Qua-Individual*), que é uma especialização do tipo anterior. Um exemplo disso é o conjunto de características que *Dilma* adquire quando da instanciação da relação de *ter parceria com Heráclito Fortes*, que corresponde a DilmaQuaParceiraDeHeráclitoFortes, assim como sua contraparte mutuamente dependente, o modo análogo de *Heráclito Fortes*, que é HeráclitoFortesQuaParceiroDeDilma.

No que concerne aos indivíduos do tipo Aspecto Relacional, é importante notar que estes só podem existir, no sentido técnico, se compostos pelos indivíduos-qua relacionais inerentes aos indivíduos mediados quando da instanciação da relação por eles representada, sendo também inerente à soma *mereológica** desses indivíduos-qua. Por exemplo, o aspecto *parceria entre Dilma e Heráclito Fortes* é composto pelos indivíduos-qua *DilmaQuaParceiraDeHeráclitoFortes* e *HeráclitoFortesQuaParceiroDeDilma*, e é inerente à soma mereológica destes.

Voltando à classificação inicial dos tipos de universais, o tipo Relação é especializado segundo a aplicação da distinção ontológica básica entre relações (vide Figura 15), podendo ser do tipo *Relação Material* (*Material Relation*) ou *Relação Formal* (*Formal Relation*). O tipo Relação Material se aplica às relações que dependem de algum interventor para valer, ou seja, um indivíduo do tipo Aspecto Relacional. A relação material *parceria com* entre *Dilma* e *Heráclito Fortes*, por exemplo, vale enquanto existir o aspecto relacional *Parceria entre Dilma e Heráclito Fortes*. Contrariamente, as relações do tipo Relação Formal são estabelecidas pela simples existência dos indivíduos relacionados. De modo geral, este tipo abrange as relações que formam a superestrutura matemática para criação da ontologia, como *dependência existencial*, *parte de*, *subconjunto de*, *instanciação*, dentre outras. Por

exemplo, a relação *parte de* entre *Dilma* e *seu Coração* vale sempre que ambos existirem.

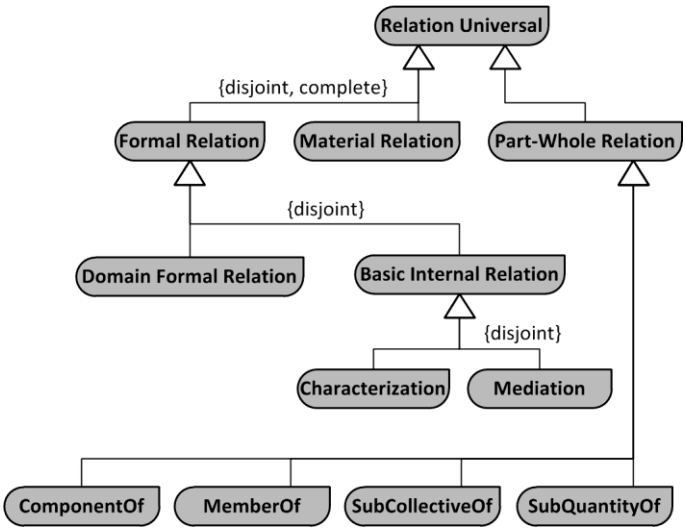


Figura 15 - *UFO-A*: Relação

As relações do tipo Formal podem, ainda, ser classificadas como *Relação Básica Interna* (*Basic Internal Relation*) ou *Relação Formal de Domínio* (*Domain Formal Relation*). O tipo Relação Básica Interna vale para relações formais internas de dependência existencial que têm representação entre as categorias de indivíduos de *UFO*. Tais relações têm como especializações o tipo *Caracterização* (*Characterization*), que se aplica à relação de *inerência* (*inheresIn*), esta responsável pela definição dos indivíduos do tipo Aspecto, e o tipo *Mediação* (*Mediation*), que se aplica à relação de *mediação* (*mediates*), sendo esta quem define os indivíduos do tipo Aspecto Relacional. O tipo Relação Formal de Domínio, por sua vez, se aplica às relações formais que são específicas de domínio, e que, por este motivo, não são representadas entre os tipos de indivíduos de *UFO*, mas entre os conceitos específicos de domínio.

Por fim, há um tipo de relação ainda não apresentado, cujo nome é *Relação Parte-todo* (*Part-Whole Relation*), que diz respeito às relações todo-parte que ocorrem entre os subtipos de substanciais. O tipo em questão se divide em *ComponenteDe* (*ComponentOf*), *MembroDe* (*MemberOf*), *SubcoletivoDe* (*SubCollectiveOf*) e *SubquantidadeDe* (*SubQuantityOf*).

ComponenteDe corresponde a todas as partes separáveis de um substancial que exercem papéis diferentes em relação ao todo, contribuindo para o funcionamento do mesmo, a exemplo do cérebro e do coração em relação ao corpo humano. *MembroDe* trata das partes de um todo que desempenham o mesmo papel, como, por exemplo, uma determinada árvore em relação à floresta do qual faz parte, ao passo que *SubcoletivoDe* apresenta a noção de partes de uma coleção que, embora possam se separar do todo, não possuem um papel funcional em relação ao mesmo, a exemplo de um conjunto de árvores que integra a Floresta Amazônica. *SubquantidadeDe*, por sua vez, se aplica às partes de uma porção de matéria que correspondem ao todo devido a uma relação de conexão topológica, a exemplo do álcool presente em uma porção de vinho.

3.7.2 **UFO-B**

Enquanto que *UFO-A* é uma ontologia que se ocupa das distinções relacionadas a endurantes, *UFO-B* trata de perdurantes – representados pelo tipo Perdurante, mencionado em *UFO-A*, que é atribuído a um indivíduo cujas partes temporais não estão presentes em sua totalidade quando este assim está – ou, mais intuitivamente, de eventos. Desse modo, os eventos – também chamados de ocorrências – diferem dos objetos, em termos de suas relações com o tempo.

Os objetos estão inteiramente presentes ou ausentes em um determinado instante do tempo, isto é, eles *são no tempo*, o que quer dizer que, se em uma circunstância um objeto possui uma determinada propriedade, em outra circunstância este mesmo objeto tem outra propriedade, possivelmente diferente da anterior. A identidade do objeto, porém, é preservada. Por exemplo, *Knut* pode pesar 50 quilos em uma circunstância e 90 quilos em outra, e ainda assim ser o mesmo indivíduo.

Ao contrário dos objetos, os eventos *acontecem no tempo*, segundo o entendimento de que se estendem no tempo acumulando partes temporais, que, por sua vez, correspondem ao seu elemento fundamental de composição. Um jogo de futebol, uma comemoração, uma entrevista e a execução de uma prova são exemplos de eventos. Em qualquer momento em que um evento está presente,

apenas algumas de suas partes temporais estão presentes de fato. Consequentemente, eventos não podem sofrer mudanças no tempo no sentido genuíno, dado que nenhuma de suas partes temporais mantém sua identidade ao longo do tempo.

Eventos são, ainda, possíveis transformações de uma situação da realidade para outra. Ou seja, eventos podem alterar o estado de coisas – que diz respeito à noção de que trata o tipo Situação, em *UFO-A* – de um *pré-estado* (*pre-state*) para um *pós-estado* (*pos-state*).

Eventos são entidades ontologicamente dependentes, ou seja, para existirem dependem de seus participantes. Por exemplo, *o acordo de parceria entre Dilma e Heráclito Fortes* é um evento que conta com a participação de *Dilma*, *Heráclito Fortes* e, possivelmente, de um *documento* que formalize tal parceria. Assim sendo, o evento é composto pela participação individual de cada uma dessas entidades, representada pelo tipo *Participação* (*Participation*). Cada uma dessas participações é por si própria um evento que pode ser do tipo *Evento Atômico* (*Atomic Event*) ou *Evento Complexo* (*Complex Event*), e é existencialmente dependente de um único substancial.

Análogo ao que foi definido em *UFO-A*, os aspectos de eventos têm seus valores – estes representados pelo tipo Quale – obtidos por sua projeção em uma estrutura de qualidade que é abrangida pelo tipo *Estrutura Temporal* (*Temporal Structure*), tendo como especialização o tipo *Intervalo Temporal* (*Time Interval*), que, por sua vez, é composto pelo tipo *Instante* (*Time Point*).

Um intervalo temporal pode possuir uma ou mais relações de intervalos temporais – que correspondem ao tipo *Relação de Intervalo Temporal* (*Time Interval Relation*), que é uma especialização do tipo Relação Formal, este originário de *UFO-A* – a partir das quais as relações correspondentes entre eventos podem ser derivadas. Nesse âmbito, cada intervalo temporal é designado como *fonte*, através da relação *source*, ou como *destino*, através da relação *target* (vide Figura 16).

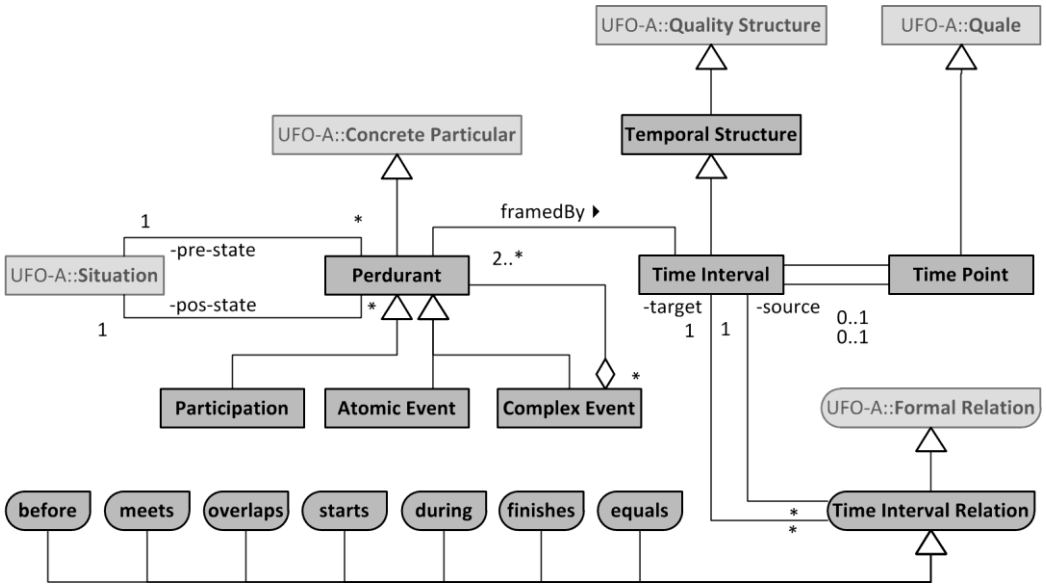


Figura 16 - *UFO-B*: Perdurante

Seguindo as definições de Allen (1983), *UFO-B* estabelece as seguintes relações de intervalo temporal entre dois eventos: *precede* (*before*), *encontra* (*meets*), *sobrepõe* (*overlaps*), *inicia* (*starts*), *durante* (*during*), *termina* (*finishes*) e *equivale* (*equals*). Enquanto as duas primeiras dão conta de eventos sequenciais, as demais tratam de eventos paralelos (vide Figura 17).

Cabe ressaltar que a distinção entre universais e indivíduos realizada em *UFO-A* permanece em *UFO-B*. Desse modo, os conceitos especializados de Indivíduo Concreto possuem especializações de Universal equivalentes.

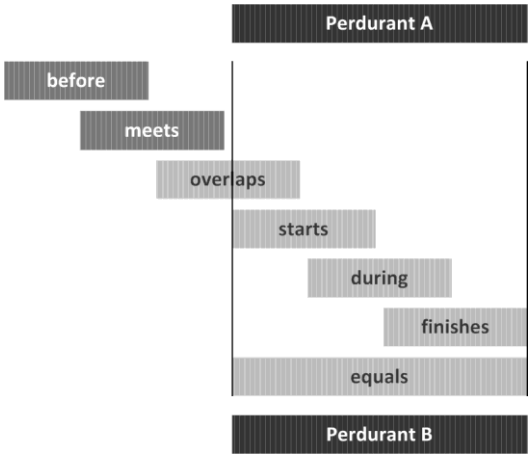


Figura 17 - Comparação entre as relações de intervalo temporal

3.7.3 **UFO-C**

A terceira ontologia que compõe *UFO* é uma ontologia de entidades sociais, construída a partir das partes A e B de *UFO*, que abrange tanto endurantes quanto perdurantes. Uma distinção fundamental em *UFO-C* (vide Figura 18) se dá entre entidades do tipo *Agente* (*Agent*) e não agentes, representados pelo tipo *Objeto* (*Object*). É importante destacar que o termo objeto em *UFO-C* tem conotação de substancial inanimado, ou seja, substancial incapaz de agir. Agentes e objetos podem ser físicos ou sociais. Por exemplo, no contexto de agentes, pessoas são do tipo *Agente Físico* (*Physical Agent*) e sociedades são do tipo *Agente Social* (*Social Agent*). Por outro lado, no que concerne a objetos, carros são do tipo *Objeto Físico* (*Physical Object*), ao passo que linguagens são do tipo *Objeto Social* (*Social Object*).

Enquanto que os agentes físicos têm como especialização o tipo *Agente Humano* (*Human Agent*), os agentes sociais podem ser classificados como *Agente Social Coletivo* (*Collective Social Agent*) ou *Organização* (*Organization*), e se aplicam, por exemplo, a grupos sociais e empresas, respectivamente.

Os objetos sociais têm como especialização o tipo *Descrição Normativa* (*Normative Description*), que estabelece regras reconhecidas por pelo menos um agente social, sendo capaz de definir entidades sociais como universais, estes exemplificados pelos tipos de comprometimentos sociais, outros objetos, como a

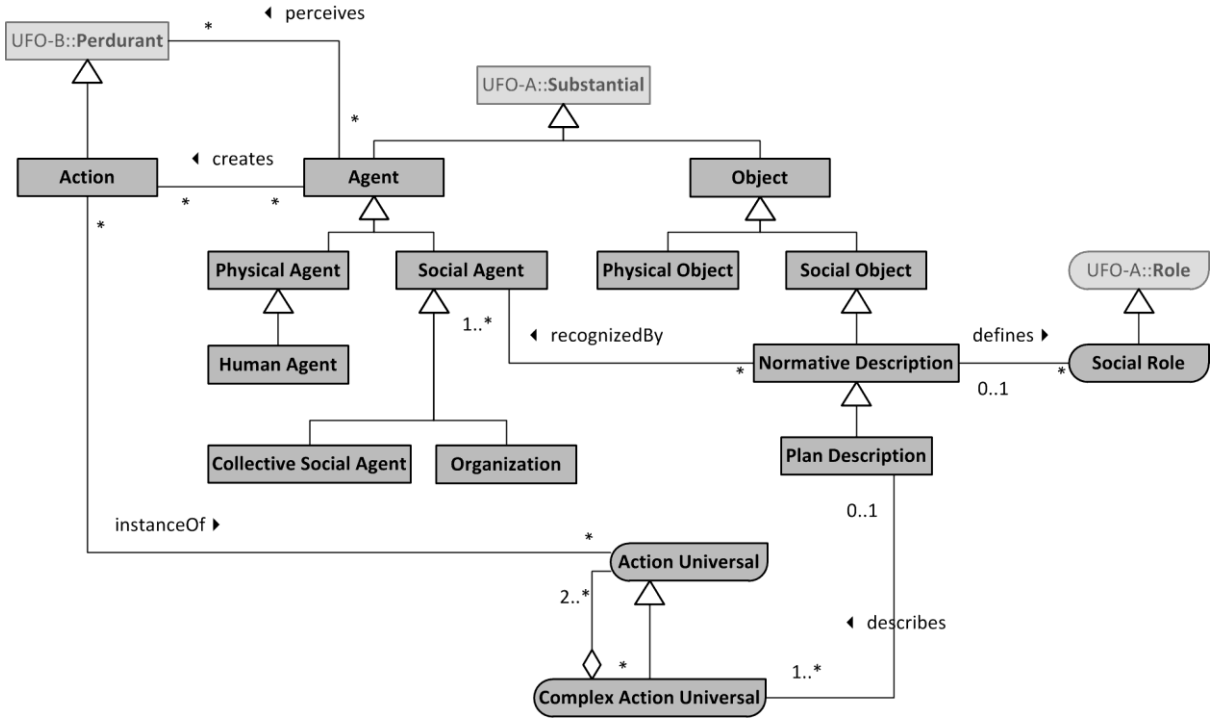


Figura 18 - *UFO-C*: distinção entre Agente e Objeto

coroa do rei da Noruega, e papéis sociais – representados pelo tipo *Papel Social* (*Social Role*), que é uma especialização de Papel, já visto em *UFO-A* – tais como presidente e cidadão. São exemplos de descrições normativas, desde que entendidos tão somente como documentos, e não como relações, a Constituição Brasileira, o Regimento Interno da Comissão de Ética em Pesquisa da UERJ e um contrato de trabalho.

As descrições normativas podem, ainda, ser do tipo *Descrição de Plano* (*Plan Description*), que tem como característica descrever universais de planos complexos, estes correspondentes ao tipo *Universal de Ação Complexa* (*Complex Action Universal*), que, por sua vez, é uma especialização de *Universal de Ação* (*Action Universal*), também chamado de *Plano*, cuja definição será vista adiante.

Conforme apresentado em *UFO-A*, aspectos são entidades que denotam particularidades e que dependem de outras entidades para existir. Em *UFO-C*, é introduzido o conceito de *Aspecto Intencional* (*Intentional Moment*), um tipo especial de aspecto que não deve ser encarado simplesmente como a noção de *alguma coisa que se intenciona*, mas, em um contexto mais amplo, como a capacidade de certas propriedades (de determinados indivíduos) de se referir a possíveis situações – estas representadas pelo tipo Situação, proveniente de *UFO-A* – na realidade.

Os aspectos intencionais (vide Figura 19) podem ser classificados como *Aspecto Mental* (*Mental Moment*) ou *Aspecto Social* (*Social Moment*), este último tipo sendo também uma especialização de Modo Externamente Dependente, definido em *UFO-A*. Enquanto que os aspectos sociais podem ser do tipo *Comprometimento Social* (*Social Commitment*) ou *Reivindicação Social* (*Claim*), os aspectos mentais têm como especialização os tipos *Crença* (*Belief*), *Desejo* (*Desire*) e *Intenção* (*Intention*), também chamado de *Comprometimento Interno* (*Internal Commitment*). É importante observar que tanto os comprometimentos sociais quanto as intenções são, também, especializações do tipo *Comprometimento* (*Commitment*).

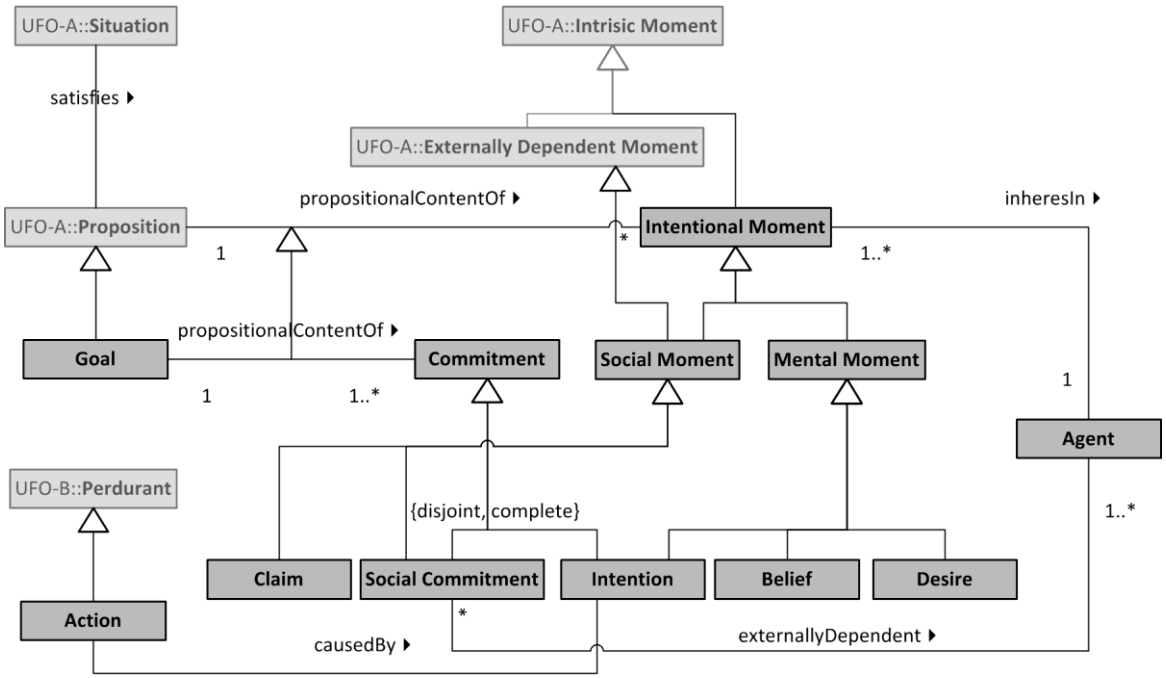


Figura 19 - *UFO-C*: Aspecto Intencional

Enquanto que o tipo Crença se aplica aos aspectos mentais que podem ser justificados por situações na realidade – um exemplo é a crença de que o planeta Terra gira em torno do Sol – os tipos Desejo e Intenção, por outro lado, abrangem aqueles que podem ser realizados ou não. O tipo Desejo é caracterizado pela vontade do agente em direção a um estado de coisas na realidade. Por exemplo, o desejo de que o Brasil seja a maior potência econômica do mundo. O tipo Intenção, por sua vez, corresponde a um estado de coisas desejado e que o agente se

compromete a perseguir; trata-se de uma intencionalidade específica cuja noção já foi mencionada anteriormente: *alguma coisa que se intenciona*.

Todo aspecto intencional possui uma representação abstrata de uma classe de situações às quais faz referência, correspondente ao tipo Proposição, presente em *UFO-A*. As proposições, por sua vez, têm como especialização o tipo *Objetivo* (*Goal*), que equivale ao conteúdo proposicional de uma intenção. Uma organização, por exemplo, pode ter a intenção de *ser bem sucedida*; neste caso, tem-se *sucesso* como objetivo.

No exemplo supracitado, um conteúdo proposicional possível seria *a organização estar entre as dez maiores do seu nicho de mercado no Brasil*. No que concerne à relação entre aspectos intencionais e situações, esta pode ser definida da seguinte maneira: uma situação no mundo real pode satisfazer o conteúdo proposicional de um aspecto intencional, isto é, satisfazer, no sentido lógico, a proposição que representa o conteúdo proposicional. Considerando o exemplo dado, a situação na qual *a referida organização é a nona maior em seu nicho de mercado no Brasil* satisfaz o conteúdo proposicional estabelecido anteriormente.

O tipo Comprometimento Social trata do comprometimento de um agente com outro. Da mesma forma que um modo externamente dependente, o comprometimento social é inerente a um agente, e é externamente dependente de outro. O comprometimento social provoca, necessariamente, o surgimento de uma intenção no primeiro agente e a reivindicação social do segundo com o primeiro. Por esse motivo, intenções levam agentes a executarem ações, estas representadas pelo tipo *Ação* (*Action*), que será explicado mais à frente.

Não obstante o fato de abarcar comprometimentos sociais e intenções, o tipo Comprometimento pode ser classificado como sendo do tipo *Comprometimento Atômico* (*Atomic Commitment*) ou *Comprometimento Complexo* (*Complex Commitment*), este composto por dois ou mais comprometimentos. *Compromisso* (*Appointment*), por sua vez, é uma outra especialização de comprometimento, cujo conteúdo proposicional – representado pelo tipo Objetivo de Compromisso (*Appointment Goal*), que é uma especialização de Objetivo – se refere, explicitamente, a um intervalo de tempo, equivalente ao tipo Intervalo de Tempo, já visto em *UFO-B* (vide Figura 20).

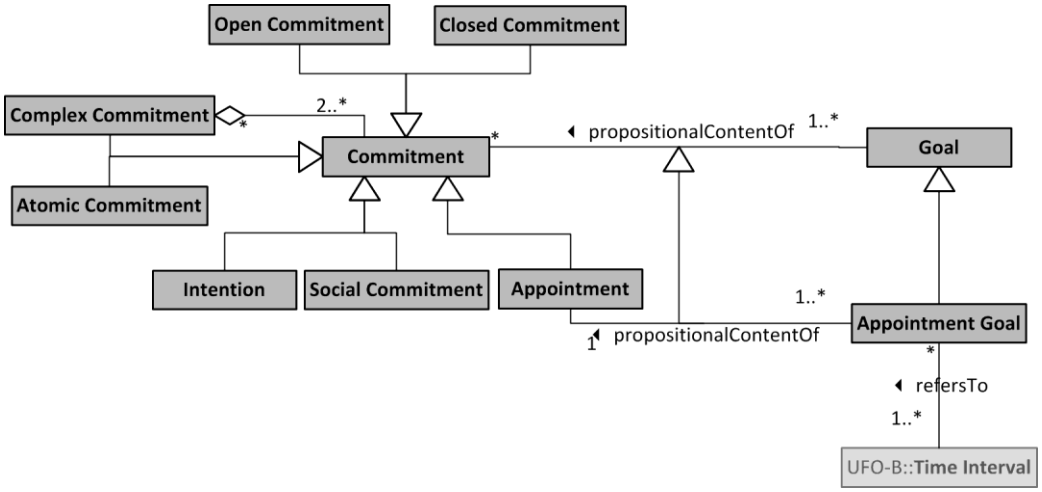


Figura 20 - *UFO-C*: Comprometimento

Além do exposto, comprometimentos podem ser do tipo *Comprometimento Aberto* (*Open Commitment*) ou *Comprometimento Fechado* (*Closed Commitment*). A diferença entre os dois se dá pelo fato de, no último, o agente ter de cumprir o comprometimento a partir da execução de uma ação específica. Diz-se, então, que um comprometimento é baseado em um plano – este representado pelo tipo Universal de Ação – sendo cumprido pelo agente, se, e somente se, este ativamente provocar uma situação que satisfaz o conteúdo proposicional do comprometimento, por meio de uma ação que, por sua vez, é uma instância do plano. Por exemplo, quando *Dilma se compromete com Heráclito Fortes a estabelecer um acordo de parceria*, este comprometimento é do tipo Comprometimento Aberto, pois não existe um plano a ser seguido para o seu cumprimento. Por outro lado, caso *Dilma se comprometa a estabelecer tal acordo através da celebração de um contrato*, este comprometimento será do tipo Comprometimento Fechado, dado que, para alcançá-lo, ela deverá, necessariamente, celebrar um contrato. Comprometimentos abertos e fechados podem explicar as noções de delegação aberta e fechada, respectivamente.

Delegação (*Delegation*) é uma especialização do tipo Relação Material, visto em *UFO-A*, derivada do tipo *Delegatum* (*Delegatum*); este, por sua vez, é uma especialização de *Aspecto Relacional Social* (*Social Relator*), que é um tipo de Aspecto Relacional, também já definido em *UFO-A*. Quando um agente (*delegator*) delega a outro (*delegatee*) um objetivo, o segundo agente se compromete – refere-se ao tipo Comprometimento Social – com o primeiro, que, por sua vez, passa a ter

o direito de reivindicar – proveniente do tipo Reivindicação Social – do segundo o cumprimento do que foi delegado. Comprometimento Social e Reivindicação Social compõem o tipo Delegatum, a partir do qual o tipo Delegação é derivado.

Comprometimentos e reivindicações sociais sempre formam um par que se refere ao mesmo conteúdo proposicional, e um aspecto relacional social é um tipo de Aspecto Relacional composto por dois ou mais pares de comprometimentos e reivindicações sociais associados (vide Figura 21).

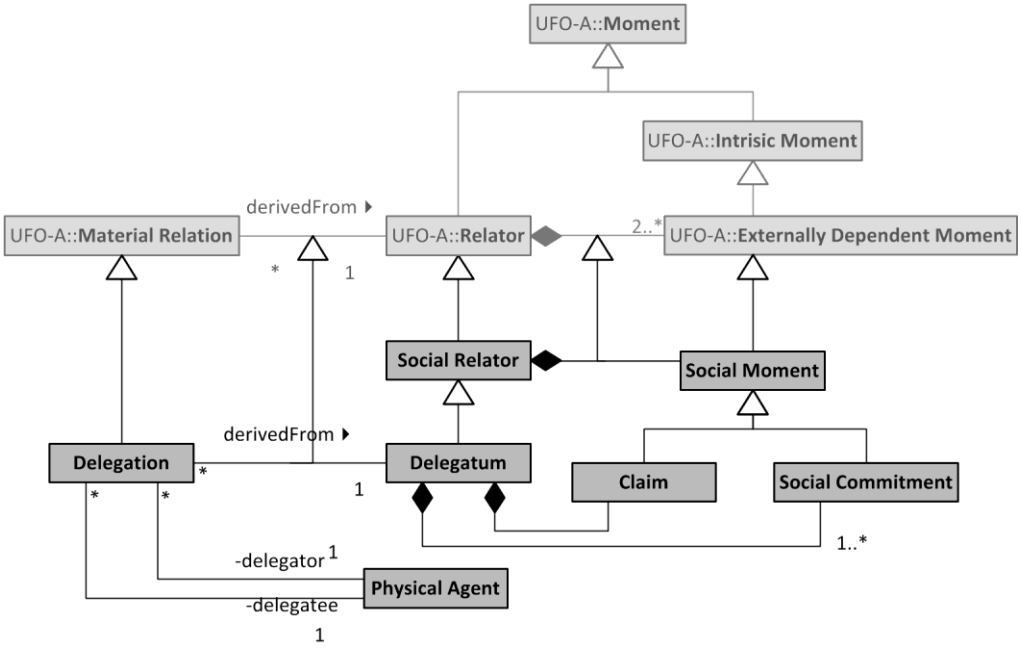


Figura 21 - *UFO-C*: Delegação

Finalmente, as ações – que correspondem ao tipo *Ação* (*Action*) – são eventos intencionais que instanciam um plano, este equivalente ao tipo Universal de Ação, com o propósito específico de satisfazer o conteúdo proposicional de alguma intenção. Por exemplo, o processo padrão de instauração de sindicância nos órgãos da administração pública é considerado um universal de ação, e suas instanciações nas comissões do Senado Federal são exemplos de ações.

Assim como eventos, ações podem ser do tipo *Ação Atômica* (*Atomic Action*) ou *Ação Complexa* (*Complex Action*). Uma ação complexa é composta por duas ou mais participações que podem, por sua vez, ser intencionais – estas equivalentes, portanto, ao tipo Ação – ou eventos não intencionais. Aplicando essa classificação ao exemplo anterior, entende-se que o processo padrão é do tipo Universal de Ação

Complexa, ao passo que as instâncias do processo nas comissões do Senado Federal são do tipo Ação Complexa.

Em *UFO*, toda participação intencional de um agente, aqui representada pelo tipo *Contribuição de Ação* (*Action Contribution*), é considerada uma ação. Ressalta-se, no entanto, que apenas agentes – que são entidades capazes de possuir aspectos intencionais – podem realizar ações. Considerando que objetos são entidades inanimadas, embora não tenham condições de realizar ações, podem participar destas por meio do tipo *Participação de Recurso* (*Resource Participation*), desde que tratados como sendo do tipo *Recurso* (*Resource*). Observa-se que tanto Contribuição de Ação quanto Participação de Recurso são especializações do tipo Participação, já apresentado em *UFO-B* (vide Figura 22)[18].

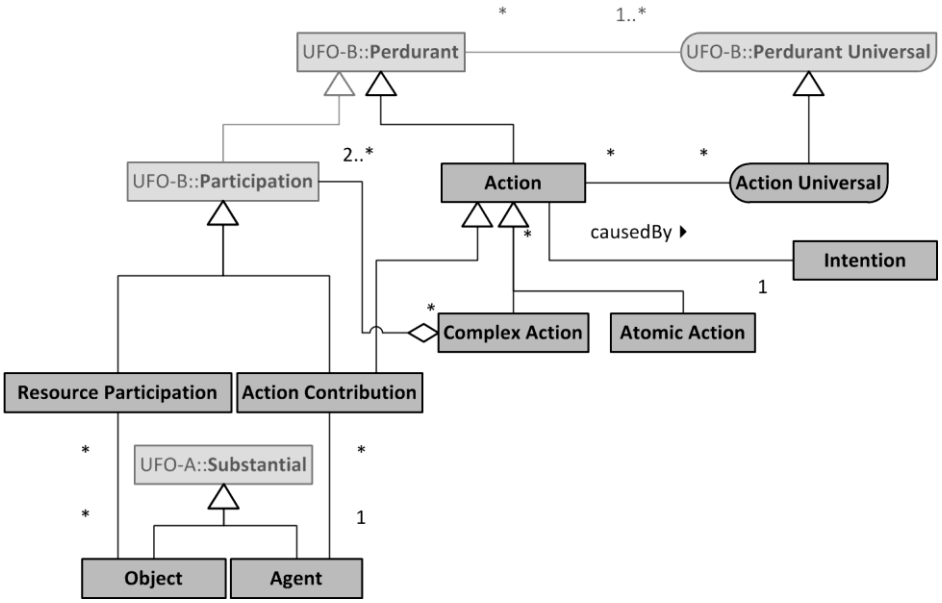


Figura 22 - *UFO-C*: Ação

[18]O tipo *Universal de Perdurante* (*Perdurant Universal*), embora não tenha sido apresentado explicitamente no trabalho, aparece na Figura 22 como o universal equivalente do tipo de indivíduo Perdurante e, desta forma, faz parte de *UFO-B*.

4 Proposta de solução

4.1 Passos iniciais

Considerando-se as iniciativas observadas, a ontologia de fundamentação utilizada como referência e os parâmetros que devem nortear as mudanças necessárias à suplantação dos problemas identificados anteriormente, uma proposta de solução pode, enfim, ser apresentada aos jornais *on-line* brasileiros.

A criação dos tópicos, que atualmente integram os jornais estrangeiros estudados, deve ser o ponto de partida dos produtos jornalísticos *on-line* brasileiros, de modo que os resultados obtidos por aqueles veículos também se façam presentes no cenário nacional.

Inicialmente, os tópicos devem ser divididos em seis tipos: assunto, evento, lugar, organização, pessoa e produto, que, além de serem categorias de fácil distinção, se assemelham aos modelos taxonômicos de cunho generalista mais comumente utilizados na Web. Sendo assim, teríamos, a título de exemplo, o seguinte resultado: *Política* como assunto, *Sindicância sobre nova 'farra de passagens'* como evento, *Distrito Federal* como lugar, *Senado Federal* como organização, *Heráclito Fortes* como pessoa e *iPhone* como produto. Posteriormente, novos tipos poderão ser estabelecidos, conforme a necessidade dos produtos jornalísticos, de modo a tornar a classificação dos tópicos mais específica.

Por definição, todo e qualquer tópico poderá ser tratado como editoria, desde que o objetivo seja destacá-lo, sendo incorporado à estrutura de navegação básica do jornal, e aplicando-se um modelo de interface diferenciado. Ou seja, a editoria não será encarada apenas como um tipo de categoria, mas como uma possível forma de apresentação dos tópicos.

Os tópicos do tipo evento assumirão o mesmo papel que tem as histórias nos produtos de jornalismo digital estrangeiros. A única diferença reside no fato de que o evento será considerado, por si só, um tópico, e não um tipo de categoria à parte.

Dado que, de modo geral, os tipos de tópico estabelecidos nesta solução têm especificidades completamente diferentes entre si, é possível que, assim como as editorias, cada tipo de tópico possua um desenho de interface diferenciado e,

provavelmente, mais adequado às suas necessidades. Por exemplo: enquanto que o tópico *Senado Federal* pode exibir vídeos, galeria de fotos, *podcasts*, matérias relacionadas, enquete e uma lista com os nomes dos senadores, o tópico *Distrito Federal* pode se limitar a apresentar algumas informações sobre a região, sua localização no mapa político do Brasil, bem como matérias e fotos associadas.

Em síntese, para os leitores a oferta de tópicos, que abrangem conceitos como assunto, história (evento) e entidade nomeada (lugar, organização, pessoa etc), se traduzirá em novos eixos dentro da narrativa, enquanto que para os produtores de conteúdo uma nova estrutura será disponibilizada, levando-se em consideração os cinco parâmetros já mencionados, onde o conteúdo será tratado, categorizado e inter-relacionado de modo consistente, flexível e sistemático.

Vale ressaltar que, a exemplo de *NYT*, *BBC News* e *MSNBC.COM*, a Web Semântica é a base tecnológica onde o modelo proposto se insere. Dessa forma, haverá a possibilidade de se associar qualquer tipo de conteúdo a uma ou mais categorias – processo este conhecido como anotação semântica – de modo simples e rápido, durante as etapas de criação e edição do conteúdo, proporcionando não apenas o relacionamento entre conteúdo e categoria, mas entre os próprios conteúdos e categorias em si, o que vai ao encontro dos parâmetros críticos levantados anteriormente.

A estrutura de categorias e suas inter-relações serão definidas adiante, na etapa de conceituação da ontologia. É preciso entender que, uma vez implementada, a ontologia deverá ser inserida num banco de triplas (*Virtuoso*, *BigOWLIM* etc) – que equivale a uma espécie de banco de dados, porém mais flexível e escalável que os modelos relacionais (*MySQL*, *SQL Server* etc) – capaz de armazenar, processar e extrair informações em larga escala, sendo este o pilar da tecnologia semântica.

Da mesma forma que um banco de dados relacional, uma vez concebido o primeiro modelo da ontologia, este pode evoluir gradativamente, conforme as necessidades de organização e manipulação das informações forem se configurando, tornando-se maior e mais detalhado. Na ontologia, porém, esse processo evolutivo é mais simples e natural, sendo a sua manutenção menos onerosa às empresas.

Considerando que, em um modelo relacional, para cada classe da ontologia exista uma tabela de entidade, e para algumas propriedades haja a necessidade de

tabelas de relacionamento, é razoável afirmar que, em casos de média e grande complexidade, para obter um modelo relacional com a mesma expressividade haja uma quantidade de tabelas tal que, para a realização de consultas ao banco de dados, a atuação de um especialista que conheça todos os relacionamentos das entidades se faça necessária.

Além disso, a atualização do modelo provavelmente exigirá uma refatoração geral de todos os agentes de software que façam consultas no modelo relacional, incluindo toda a codificação dos sistemas que utilizam esta base de dados. Quando se utiliza um modelo semântico, sendo este potencializado por consultas *SPARQL**, a responsabilidade pelo conhecimento do modelo passa a ser da máquina, o que as ferramentas de inferência já testadas fazem muito bem, especialmente o *Oracle Semantic Technologies*.

Assim sendo, com o uso de ferramentas e tecnologias adequadas aos conceitos da Web Semântica, obtém-se uma grande flexibilidade para mudanças estruturais dos dados.

4.2 **Processo de conceituação da ontologia**

A ontologia nada mais é que um reflexo da forma como se pretende distribuir e apresentar o conteúdo dos jornais *on-line* brasileiros. No entanto, durante o seu processo de conceituação, um novo conjunto de informações foi considerado: uma lista com mais de 1.600 termos – designados como *assuntos* e utilizados até hoje pelo *G1* – que carece de critérios de classificação.

A lista em questão representa uma oportunidade para se tentar entender o modo como o conteúdo é rotulado em um dos principais jornais *on-line* brasileiros, mas, não apenas isto, serve também para perceber algumas das dificuldades experimentadas pelos produtores e gestores de conteúdo, funcionando como insumo para a definição da ontologia.

Após análise criteriosa da referida lista, nota-se que a maior parte dos assuntos são, segundo a ontologia *UFO*, universais. Ou seja, os assuntos são, em sua maioria, tipos gerais que representam os padrões de características presentes em diferentes indivíduos. Assim sendo, fica claro que o detalhamento desses

universais tem papel preponderante na organização dos assuntos. A título de exemplo, a matéria *Crianças são encontradas sozinhas e mãe responderá por abandono* corresponde a um indivíduo que é instância do universal *Abandono* e, também, do universal *Abandono de menor*, sendo este uma especialização do primeiro.

Como visto no exemplo dado, a ontologia deve levar em conta que muitos dos universais são especializações de conceitos mais amplos e, neste aspecto, o objetivo da ontologia aqui proposta é oferecer universais genéricos capazes de conter a maior parte dos universais identificados entre os assuntos levantados. Nota-se que, devido à extensão da lista dos assuntos do *G1*, neste trabalho são apresentados somente os itens iniciados com a letra *A* (vide Anexo E).

4.3 **Universais**

Do mesmo modo que nos universais, a categorização dos indivíduos também deve ser abrangente. Na prática, para cada tipo de indivíduo, deve haver ao menos um universal correspondente, como se propõe em *UFO*. Consequentemente, o primeiro desafio é elencar esses universais (vide Figura 23)[19] e os seus tipos, a saber:

- *Universal de Ação Noticiosa*, que é especialização tanto de Universal de Ação, definido em *UFO-C,* quanto de Universal de Evento Noticioso;
- *Universal de Acordo*, *Universal de Lei* e *Universal de Regra*, que são especializações de Universal de Descrição Normativa, de *UFO-C*;
- *Universal de Bem de Capital*, *Universal de Construção*, *Universal de Matéria-prima*, *Universal de Obra Criativa Física* e *Universal de Produto*, que são especializações de Universal de Objeto Físico, de *UFO-C*. Universal de Obra Criativa Física tem *Universal de Escultura*, *Universal de Fotografia*, *Universal de Gravura* e *Universal de Pintura* como universais;

[19] Na ontologia de notícias, optou-se por utilizar uma nomenclatura baseada em termos em português, diferentemente do que acontece com os conceitos provenientes de *UFO*. Além disso, para fins de legibilidade, os universais definidos pela ontologia de notícias, e seus tipos, tiveram o prefixo “Universal de” suprimido.

- *Universal de Condição*, que é especialização de Aspecto Universal Intrínseco, de *UFO-A;*
- *Universal de Entidade Empresarial*, *Universal de Entidade sem Fins Lucrativos* e *Universal de Organização da Administração Pública*, que são especializações de Universal de Organização, de *UFO-C*. Universal de Entidade sem Fins Lucrativos tem *Universal de Associação Privada* e *Universal de Partido Político* como universais;
- *Universal de Equipe Esportiva*, *Universal de Grupo da Sociedade* e *Universal de Grupo Musical*, que são especializações de Universal de Agente Social Coletivo, de *UFO-C*;
- *Universal de Evento Noticioso*, que é especialização de Universal de Perdurante, de *UFO-B*;
- *Universal de Disciplina*, *Universal de Índice Financeiro*, *Universal de Moeda*, *Universal de Obra Criativa Social*, *Universal de Projeto* e *Universal de Serviço*, que são especializações de Universal de Objeto Social, de *UFO-C*. Universal de Disciplina e Universal de Serviço se especializam, respectivamente, em *Universal de Esporte* e *Universal de Serviço Financeiro*, ao passo que Universal de Obra Criativa Social se divide em *Universal de Filme*, *Universal de Livro*, *Universal de Música*, *Universal de Peça Teatral*, *Universal de Personagem* e *Universal de Programa*;
- *Universal de Lugar*, que é especialização de Substancial Universal, de *UFO-A*, e se divide em *Universal de Acidente Geográfico* e *Universal de Divisão Territorial*; este, por sua vez, tem *Universal de Logradouro*, *Universal de Bairro*, *Universal de Cidade*, *Universal de Estado*, *Universal de Região*, *Universal de País*, *Universal de Continente* e *Universal de Parte de Divisão Territorial* como universais;
- *Universal de Organismo*, que é especialização de Universal de Agente Físico, de *UFO-C*, e que se especializa em *Universal de Animal*; este, por sua vez, tem como especialização *Universal de Animal Não Humano*;
- *Universal de Pessoa*, que é universal de Universal de Agente Humano, de *UFO-C*;

- *Universal de Unidade Organizacional*, que é especialização de Universal de Agente Social, de *UFO-C*.

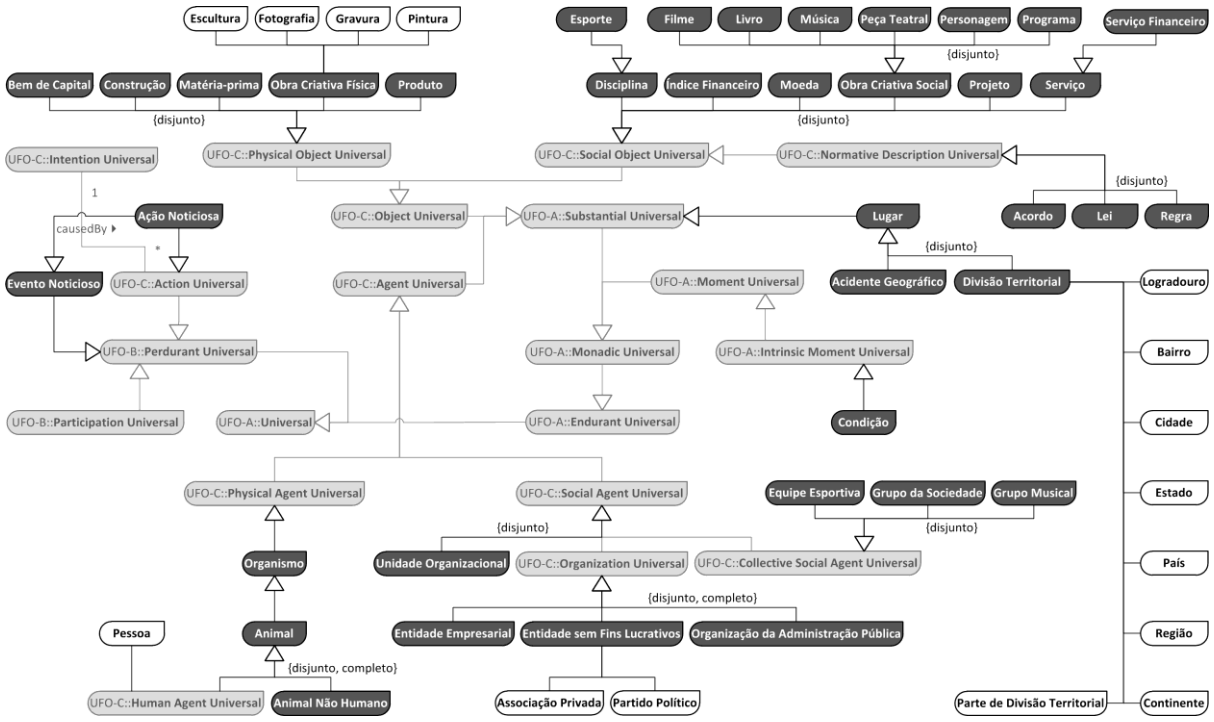


Figura 23 - Ontologia de notícias: Universais

Uma vez listados os universais e os seus tipos, resta expor suas respectivas definições e fazer algumas considerações, como será visto adiante.

Universal de Ação Noticiosa é um tipo que corresponde a planos, processos e métodos comuns a ações noticiadas com frequência – e que, portanto, são de conhecimento do veículo jornalístico – que têm o propósito específico de satisfazer o conteúdo proposicional de alguma intenção, conforme definido em *UFO-C*. Como exemplo, tem-se o inquérito policial e o processo de votação de projetos de lei.

O tipo Universal de Acordo trata das descrições normativas que têm a concordância de dois ou mais agentes. Universal de Lei é um tipo que envolve organizações da administração pública, ou pessoas a estas relacionadas, e o tipo Universal de Regra equivale a normas, instruções ou convenções destinadas à correta execução de um plano. São exemplos de Universal de Acordo, Universal de Lei e Universal de Regra, respectivamente, acordo extrajudicial, lei ordinária e convenção de edifício edilício.

O tipo Universal de Bem de Capital se aplica a objetos físicos que são entendidos como bens usados na produção de outros bens, como, por exemplo, máquinas e compostos químicos utilizados na produção agrícola. Universal de Construção é um tipo que se destina a instalações, monumentos, edificações e obras arquitetônicas, ou da engenharia civil, de todo gênero, como por exemplo, estádios e aeroportos. O tipo Universal de Matéria-prima diz respeito a materiais que sirvam de entrada para um sistema de produção qualquer, como frutas para produção de suco e minérios para fabricação de liga metálica. Universal de Obra Criativa Física tem a ver com objetos físicos elaborados por meio de um processo criativo; seus universais, Universal de Escultura, Universal de Fotografia, Universal de Gravura e Universal de Pintura são exemplos deste tipo. Universal de Produto, por sua vez, é um tipo que trata dos bens destinados à satisfação das necessidades de consumo de um indivíduo; por exemplo, televisores, roupas e alimentos.

Universal de Condição é um tipo que se refere a todos os aspectos intrínsecos que dizem respeito a problemas de saúde, tais como doenças, moléstias, distúrbios, síndromes e sintomas, sendo, portanto, existencialmente dependentes de agentes físicos do tipo Universal de Organismo.

O tipo Universal de Entidade Empresarial abarca as organizações que visam o lucro, enquanto que Entidade sem Fins Lucrativos se aplica às empresas que não têm este objetivo; seus universais, Universal de Associação Privada e Universal de Partido Político, são exemplos deste tipo e têm como propósito distinguir conceitos muito presentes no meio noticioso. Universal de Organização da Administração Pública, por sua vez, é um tipo que compreende os órgãos públicos, as autarquias e as fundações públicas.

Universal de Equipe Esportiva é um tipo que corresponde aos times e equipes voltados à prática de qualquer esporte, independente de sua categoria ou modalidade; time de futebol, dupla de tênis e equipe de nado sincronizado são alguns exemplos. O tipo Universal de Grupo da Sociedade se destina a conjuntos de pessoas ou camadas que fazem parte de uma sociedade, podendo representar uma classe econômica, profissional ou social; por exemplo, classe média, indígenas e mulheres. Universal de Grupo Musical, por outro lado, é um tipo que abrange bandas, orquestras e todo tipo de grupo que tem por objetivo executar algum tipo de música.

O tipo Universal de Evento Noticioso equivale a eventos e acontecimentos, programados ou não, que são noticiados pelo veículo jornalístico. Por exemplo, há eventos da natureza, como chuvas, furacões e tufões, e eventos esportivos, como campeonatos, competições e confrontos.

Universal de Disciplina é um tipo que abrange todos os ramos do conhecimento ou qualquer tipo de saber provindo da experiência, como, por exemplo, astrologia, biologia e economia doméstica; sua especialização, o tipo Universal de Esporte, engloba os conjuntos de exercícios físicos praticados com método, individualmente ou em equipes, como ciclismo, futebol e natação. O tipo Universal de Índice Financeiro trata dos objetos sociais que funcionam como indicadores do mercado financeiro e, portanto, têm relação com estruturas de qualidade específicas. O tipo Universal de Moeda se refere aos instrumentos oficiais para realização de transações monetárias. Universal de Obra Criativa Social tem a ver com objetos sociais elaborados por meio de um processo criativo; suas especializações, Universal de Filme, Universal de Livro, Universal de Música, Universal de Peça Teatral, Universal de Personagem e Universal de Programa, são exemplos deste tipo, sendo válido observar que Universal de Personagem é um tipo que diz respeito tão somente aos objetos tratados como agentes dentro do seu próprio universo fictício, e o tipo Universal de Programa, por sua vez, trata dos programas de televisão, rádio e outros meios de comunicação em massa transmitidos periodicamente ou não. O tipo Universal de Projeto é atribuído a todo esforço, de duração não necessariamente definida, empreendido para criar objetos. Por sua vez, Universal de Serviço é um tipo que se aplica a toda atividade econômica que resulta na criação de objetos intangíveis; por exemplo, segurança, telecomunicações e transporte.

O tipo Universal de Lugar é um substancial que abarca espaços de terra de todo gênero, posicionados em coordenadas geográficas específicas. A especialização Universal de Acidente Geográfico é um tipo destinado a manifestações contrastantes do terreno, em solo continental ou costeiro, como por exemplo, chapadas, ilhas e rios. Universal de Divisão Territorial, por outro lado, é uma especialização que vale para toda extensão de terra definida política, econômica ou socialmente; seus universais, Universal de Bairro, Universal de Cidade, Universal de Continente, Universal de Estado, Universal de Logradouro, Universal de País, Universal de Parte de Divisão Territorial e Universal de Região,

são exemplos deste tipo, sendo importante esclarecer que, enquanto Universal de Região trata de conjuntos de estados, Universal de Parte de Divisão Territorial serve aos territórios cuja divisão não se caracteriza como um dos universais previstos, como, por exemplo, Baixada Fluminense, Baixo Gávea, Triângulo Mineiro e Zona Sul.

O tipo Universal de Organismo é um agente físico que cobre tudo aquilo que pode ser considerado um ser vivo; sua especialização, Universal de Animal, é um tipo que se refere a todos os seres pertencentes ao Reino Animal. Por sua vez, Universal de Agente Humano, definido em *UFO-C*, seria uma especialização de Universal de Animal, do mesmo modo que Universal de Animal Não Humano, que é um tipo destinado aos demais animais.

Universal de Pessoa corresponde a um universal que envolve todos os agentes humanos que têm relevância para o veículo jornalístico, seja a sua presença nos noticiários perene ou não. Enfim, o tipo Universal de Unidade Organizacional se aplica aos agentes sociais que equivalem a departamentos, seções e setores específicos de uma organização, com características e funções distintas, como, por exemplo, as áreas de recursos humanos, de suporte técnico e de tecnologia da informação, estas presentes na estrutura organizacional de inúmeras companhias.

4.4 **Lead da notícia como conjunto mínimo de questões de competência**

Conforme dito anteriormente, definidos os universais, é possível se chegar aos equivalentes individuais. No entanto, antes de se adentrar nesse assunto, faz-se necessário revisitar a questão da estrutura organizacional das informações, em especial o *lead*, que, segundo Fidalgo (2004), equivale às perguntas que toda informação deve, preferencialmente, responder de forma sumária.

Dado que o *lead* de uma informação é composto, essencialmente, pelas perguntas *o quê*, *quem*, *quando*, *onde*, *como* e *por quê*, é fundamental entendermos que, na verdade, estas correspondem ao conjunto mínimo de questões de competência que a ontologia de notícias deve tentar responder. Todas as demais perguntas a serem feitas, embora não necessariamente acessórias, têm mais um

caráter enriquecedor, ou seja, uma função mais alinhada à pormenorização da informação, do que de, pura e simplesmente, trazer à tona o seu teor em linhas gerais.

Como resultado desse raciocínio, surge uma simples pergunta sobre o *lead*: afinal de contas, o que se quer dizer com *o quê*, *quem*, *quando*, *onde*, *como* e *por quê*? Sua resposta, contudo, não é tão simples e envolve um processo de decomposição e análise de cada uma das seis perguntas que dão forma ao *lead* da informação, havendo o risco de, ainda assim, não se chegar a uma resposta correta. Na prática, talvez nem exista uma, ou uma única, resposta correta, mas uma definição em consonância, por exemplo, com os princípios editoriais do veículo jornalístico, onde cada um vislumbra uma solução, de acordo com o seu ponto de vista.

Considerando o aspecto acima mencionado, a definição a que se pretende chegar neste trabalho será extremamente influenciada pela ontologia *UFO*, já que esta serve de base para a representação de todo conhecimento jornalístico, e, também, pelo levantamento preliminar dos universais que integram a ontologia de notícias, pois dará uma visão ainda mais clara e específica sobre como este conhecimento é classificado. E mais: ao se interpretar que as perguntas que compõem o *lead* são, por si só, as questões de competência de maior relevância para o modelo, chega-se à conclusão de que é exatamente este quem deve tentar responder cada uma delas com precisão. Em suma, a ontologia de notícias será responsável por dar visibilidade àquilo que se entende por *lead* da informação, tornando-a buscável e recuperável.

Iniciando-se o processo de análise das perguntas que formam o *lead*, *o quê* se mostra a mais adequada para se começar a examinar, pois parece dizer respeito à origem ou, melhor dizendo, àquilo que é representado como a própria informação em si. Portanto, ao se considerar que a maior parte do conteúdo jornalístico narra fatos e acontecimentos, é razoável afirmar que o objetivo da pergunta *o quê* é saber quais os eventos de que trata o conteúdo. De acordo com os universais levantados para a ontologia de notícias, isso equivale exatamente aos eventos noticiosos, ou seja, indivíduos que são instâncias de Universal de Evento Noticioso, e às ações noticiosas, neste caso indivíduos que são instâncias de Universal de Ação Noticiosa. Por outro lado, em algumas situações o conteúdo pode não cobrir acontecimentos, mas substanciais e aspectos, em se tratando, por exemplo, de matérias cujo teor

corresponda à avaliação de produtos recém-lançados e à análise da saúde financeira de empresas de capital aberto, respectivamente. Desse modo, embora em menor escala, a pergunta *o quê* pode se referir a indivíduos do tipo Substancial e Aspecto, ambos definidos em *UFO-A*, ou aos seus respectivos universais mesmo.

É importante esclarecer que, embora *o quê* possa dizer respeito a agentes e objetos, dado que são especializações de substanciais, na maior parte das vezes eles se farão presentes por meio da pergunta *quem*. Dentre as seis perguntas que compõem o *lead* da informação, *quem* é a que tem por objetivo saber quais são as pessoas envolvidas nos eventos e ações, uma vez que se trata de um pronome intrinsecamente ligado a pessoas. No entanto, seria limitado entender que *quem* diz respeito apenas a pessoas, dado que existem outras entidades capazes de agir. Assim sendo, para que a pergunta possa ser respondida de forma satisfatória, todo e qualquer agente deve ser considerado, incluindo os agentes sociais. Não se pode esquecer, também, que há objetos que passam por um processo de personificação, ou que já contam com esta característica, em que são atribuídas qualidades comuns às pessoas, como é o caso dos personagens. Visto que o modelo proposto contempla esse tipo de objeto, ao se falar de *quem*, este se torna mais uma opção a ser considerada. Resumindo, essa pergunta faz referência a indivíduos do tipo Agente, apresentado em *UFO-C*, bem como àqueles que são instâncias de Universal de Personagem.

A pergunta *quando*, por sua vez, não enseja maiores discussões, pois trata claramente da questão temporal; e, para isto, *UFO-B* fornece o tipo Intervalo Temporal, que, através de seus indivíduos, é capaz de dar a precisão necessária para se chegar à resposta adequada, seja fazendo referência ao passado, ao presente ou ao futuro. Porém, também seria razoável considerar que, ao se falar de tempo, o interesse não seja exatamente o período compreendido entre dois instantes, mas os eventos noticiosos e ações noticiosas que antecedem e sucedem, ou mesmo os que concorrem com, aqueles que atendem a pergunta *o quê*. Desse modo, ainda que de forma complementar, tem-se também o tipo de universal Relação de Intervalo Temporal, presente em *UFO-B*, para responder a pergunta *quando*.

Da mesma maneira que a pergunta anterior, *onde* não levanta dúvidas quanto às entidades a que faz alusão; trata-se de uma questão de competência que tem por objetivo saber em que lugar o evento noticioso ou a ação noticiosa se situa, que é o

equivalente aos indivíduos que são instâncias de Universal de Lugar, ou de suas especializações e universais.

Ao se examinar a pergunta *como*, verifica-se que há uma ideia de meio ou maneira no qual uma determinada coisa se dá. Nesse sentido, e entendendo que tal coisa é exatamente o evento noticioso ou a ação noticiosa que é alvo da narrativa jornalística, poder-se-ia considerar os indivíduos do tipo Substancial, conforme disposto em *UFO-A*, envolvidos neste cenário, para tentar se chegar a uma resposta. Entretanto, outra possibilidade, e que parece fazer mais sentido no contexto estabelecido, seria interpretar a pergunta mais como uma tentativa de saber os detalhes do acontecimento, desconstruindo-o a ponto de dar visibilidade aos eventos e ações que o compõem, se assim houver, bem como apresentando aqueles que o antecedem e o sucedem, e mesmo os que com ele concorrem, encarando-o, sempre que possível, como parte integrante de uma cadeia de acontecimentos, ao invés de algo isolado. Desse modo, além dos indivíduos que são instâncias de Universal de Evento Noticioso e de Universal de Ação Noticiosa, tem-se, novamente, a presença do tipo de universal Relação de Intervalo Temporal, definido em *UFO-B*, desta vez de relevância fundamental e objetiva, para que a pergunta *como* possa ser respondida a contento; de forma complementar, dispõe-se dos indivíduos do tipo Substancial.

Finalmente, a pergunta *por quê* tem por objetivo chegar à causa, motivo ou propósito de determinada coisa. Embora, em algumas situações, a causa possa ser desprovida de intencionalidade, como no caso de eventos de caráter não intencional, ainda assim é possível identificar relações de causalidade entre os acontecimentos veiculados pelo produto jornalístico. Por exemplo, uma inundação pode ser causada por uma forte pancada de chuva que, por sua vez, ocorreu em razão da chegada de uma frente fria. Assim sendo, os indivíduos que instanciam Universal de Evento Noticioso e Universal de Ação Noticiosa são o alvo dessa pergunta, apesar de não se resumir somente a estes. Quando a informação trata de eventos intencionais, ou seja, de ações noticiosas, cada uma delas está vinculada a uma intenção cujo propósito é satisfazer o seu conteúdo proposicional. Trocando em miúdos, quer-se dizer que, nesse âmbito, também é necessário considerar os indivíduos do tipo Intenção e Objetivo para que a pergunta *por quê* possa ser respondida adequadamente.

4.5 **Tipos de indivíduos**

Uma vez que foi possível compreender melhor as perguntas que compõem o *lead* da informação, e alinhá-las com a ontologia de notícias – segundo o entendimento de que as mesmas representam as questões de competência que o modelo deve responder – torna-se mais simples a definição dos tipos de indivíduos a serem considerados na proposta, a saber:

- *Ação Noticiosa*, que é especialização tanto de Ação, definido em *UFO-C,* quanto de Evento Noticioso;
- *Acordo*, *Lei* e *Regra*, que são especializações de Descrição Normativa, de *UFO-C*;
- *Construção* e *Obra Criativa Física*, que são especializações de Objeto Físico, de *UFO-C*. Do mesmo modo que o seu universal, Disciplina se especializa em *Esporte*. Obra Criativa Física, por outro lado, se divide em *Escultura*, *Fotografia*, *Gravura* e *Pintura*;
- *Condição*, que é especialização de Aspecto Intrínseco, de *UFO-A*;
- *Entidade Empresarial*, *Entidade sem Fins Lucrativos* e *Organização da Administração Pública*, que são especializações de Organização, de *UFO-C*. Entidade sem Fins Lucrativos se divide em *Associação Privada* e *Partido Político*;
- *Equipe Esportiva*, *Grupo da Sociedade* e *Grupo Musical*, que são especializações de Agente Social Coletivo, de *UFO-C*;
- *Evento Noticioso*, que é especialização de Perdurante, de *UFO-B*;
- *Disciplina*, *Índice Financeiro*, *Moeda*, *Obra Criativa Social*, *Projeto* e *Serviço*, que são especializações de Objeto Social, de *UFO-C*. Serviço se especializa em *Serviço Financeiro*, e Obra Criativa Social se divide em *Filme*, *Livro*, *Música*, *Peça Teatral*, *Personagem* e Programa;
- *Lugar*, que é especialização de Substancial, de *UFO-A*, e se divide em *Acidente Geográfico* e *Divisão Territorial*; este, por sua vez, se especializa em *Logradouro*, *Bairro*, *Cidade*, *Estado*, *Região*, *País*, *Continente* e *Parte de Divisão Territorial*;

- *Organismo*, que é especialização de Agente Físico, de *UFO-C*, e que se especializa em *Animal*; este, por sua vez, tem como especialização *Animal Não Humano*;
- *Pessoa*, que é especialização de Agente Humano, de *UFO-C*;
- *Unidade Organizacional*, que é especialização de Agente Social, de *UFO-C*.

É importante observar que, ao se fazer o desdobramento para indivíduos (vide Figura 24), alguns dos conceitos tratados na forma de universais foram intencionalmente descartados, como é o caso, por exemplo, de Universal de Matéria-prima. Esse tipo representa um conceito que, para as necessidades do veículo jornalístico, só se aplica ao nível dos universais, pois não há interesse em se fazer uma distinção, por exemplo, entre *algodão* e uma quantidade específica deste insumo que está nas mãos de uma pessoa.

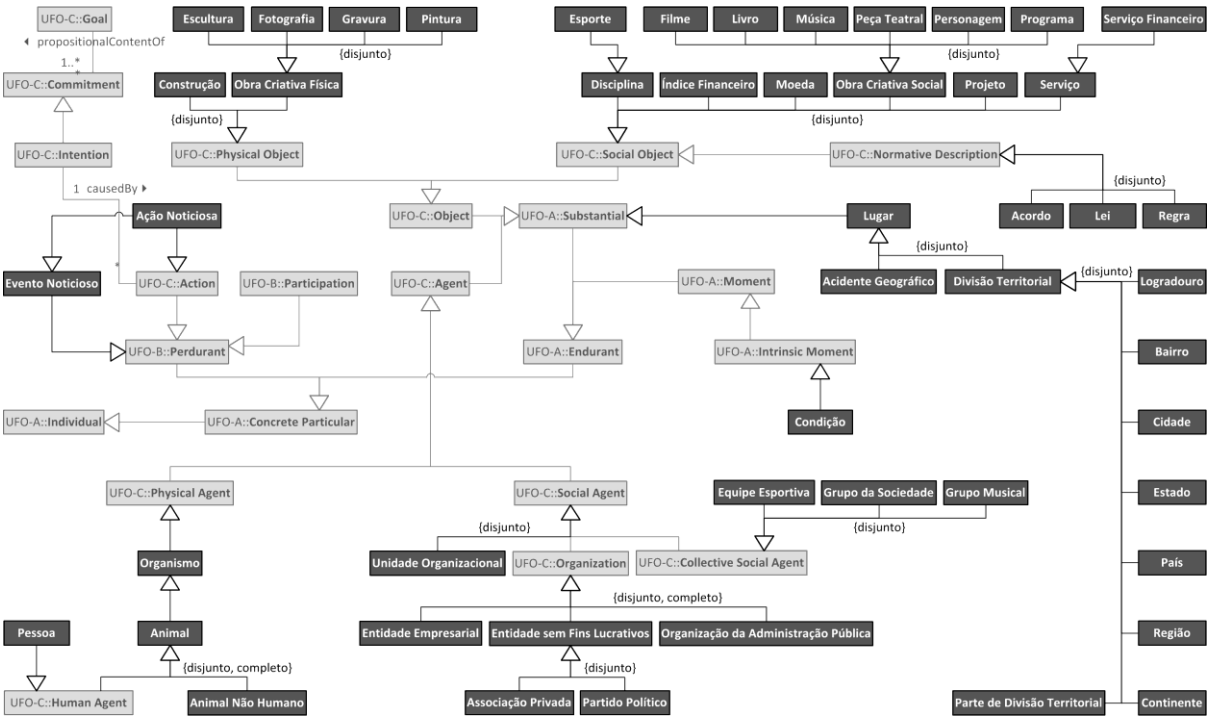


Figura 24 - Ontologia de notícias: Indivíduos

Outro ponto que é necessário ser apresentado é o fato da ontologia proposta considerar certas restrições técnicas que dizem respeito à sua implementação, uma vez que se pretende experimentá-la na prática para exposição dos resultados. Por

conseguinte, é importante esclarecer que o modelo aqui proposto não é apenas conceitual; o seu desenho reflete, de fato, a solução de implementação.

4.6 **Relações entre entidades**

No que concerne às relações entre entidades, apesar de algumas delas serem herdadas de *UFO* devido à utilização de alguns de seus conceitos, é necessário destacar aquelas que têm serventia à ontologia de notícias, bem como aquelas que podem ser consideradas exclusivas do modelo, a saber:

- *abrange*, que liga o tipo Região ao tipo Estado, nesta direção;
- *acordadoPor*, que liga o tipo Acordo ao tipo Agente, nesta direção;
- *afeta*, que liga o tipo Condição ao tipo Organismo, nesta direção;
- *atuaEm*, que liga o tipo Agente ao tipo Disciplina, nesta direção;
- *bairroDe*, que liga o tipo Bairro aos tipos Cidade e Parte de Divisão Territorial, nesta direção;
- *capitalDoEstado*, que liga o tipo Cidade ao tipo Estado, nesta direção;
- *capitalDoPais*, que liga o tipo Cidade ao tipo País, nesta direção;
- *causadaPor* (*causedBy*), de *UFO-C*, que liga o tipo Ação Noticiosa ao tipo Intenção, nesta direção;
- *cidadeDe*, que liga o tipo Cidade aos tipos Estado, País e Parte de Divisão Territorial, nesta direção;
- *colaboradoPor*, que liga o tipo Objeto ao tipo Agente, nesta direção;
- *compostoDe*, que interliga eventos do tipo Evento Noticioso;
- *concebe*, que liga o tipo Projeto ao tipo Objeto, nesta direção;
- *conteudoProposicionalDe* (*propositionalContentOf*), de *UFO-C*, que liga o tipo Objetivo ao tipo Comprometimento, nesta direção;
- *contidoPor*, que liga o tipo Acidente Geográfico ao tipo Lugar, nesta direção;
- *cria* (*creates*), de *UFO-C*, que liga o tipo Agente ao tipo Ação Noticiosa, nesta direção;

- *criadoPor*, que liga o tipo Personagem aos tipos Obra Criativa Física e Obra Criativa Social, nesta direção;
- *define*, que liga o tipo Projeto ao tipo Ação Noticiosa, nesta direção;
- *dirigidoPor*, que liga o tipo Agente Social ao tipo Pessoa, nesta direção;
- *elaboradoPor*, que liga o tipo Objeto ao tipo Agente, nesta direção;
- *envolve*, que liga o tipo Participação aos tipos Agente e Objeto, nesta direção;
- *estadoDe*, que liga o tipo Estado aos tipos País e Parte de Divisão Territorial, nesta direção;
- *estaEm*, que liga o tipo Objeto Físico ao tipo Lugar, nesta direção;
- *existencialmenteDependente* (*existentiallyDependent*), de *UFO-A*, que liga o tipo Aspecto Intrínseco ao tipo Substancial, nesta direção;
- *fazParteDe*, que liga os tipos Agente Físico e Objeto Físico ao tipo Agente Social, nesta direção;
- *inerenteA* (*inheresIn*), de *UFO-A*, que liga o tipo Aspecto ao tipo Indivíduo, nesta direção;
- *instanciaDe* (*instanceOf*), de *UFO-A*, que liga o tipo Indivíduo ao tipo Universal, nesta direção;
- *integra*, que liga o tipo Parte de Divisão Territorial ao tipo Divisão Territorial, nesta direção;
- *localizadoEm*, que liga o tipo Agente ao tipo Lugar, nesta direção;
- *logradouroDe*, que liga o tipo Logradouro aos tipos Bairro, Cidade e Parte de Divisão Territorial, nesta direção;
- *membroDe*, que liga o tipo Pessoa ao tipo Agente Social, nesta direção;
- *ocorreEm*, que liga o tipo Evento Noticioso ao tipo Lugar, nesta direção;
- *paisDe*, que liga o tipo País ao tipo Continente, nesta direção;
- *pratica*, que liga os tipos Pessoa e Equipe Esportiva ao tipo Esporte, nesta direção;
- *presta*, que liga o tipo Agente ao tipo Serviço, nesta direção;
- *produz*, que liga o tipo Bem de Capital ao tipo Produto, nesta direção;
- *provocadoPor*, que interliga eventos do tipo Evento Noticioso;

- *reconhecidoPor* (*recognizedBy*), de *UFO-C*, que liga Descrição Normativa ao tipo Agente Social, nesta direção;
- *regiaoDe*, que liga o tipo Região aos tipos País e Parte de Divisão Territorial, nesta direção;
- *regula*, que liga os tipos Acordo, Lei e Regra ao tipo Entidade, nesta direção;
- *representa*, que liga o tipo Equipe Esportiva ao tipo Organização, nesta direção;
- *sobResponsabilidadeDe*, que liga o tipo Objeto ao tipo Agente, nesta direção;
- *temAgenteCausador*, que liga o tipo Condição ao tipo Organismo, nesta direção;
- *temAgenteTransmissor*, que liga o tipo Condição ao tipo Organismo, nesta direção;
- *trabalhaPara*, que liga o tipo Pessoa ao tipo Agente Social, nesta direção;
- *unidadeDe*, que liga o tipo Unidade Organizacional ao tipo Organização, nesta direção.

A título de exemplo, as três primeiras relações apresentadas acima serão fundamentadas e detalhadas, de modo que seja possível entender como as relações sugeridas poderão interligar as entidades a que fazem referência e, desta forma, se integrar à ontologia.

A relação *abrange* equivale a uma relação de composição. Assim sendo, por meio dessa relação, as instâncias do tipo Estado associadas a uma mesma instância do tipo Região são caracterizadas como sendo parte integrante desta última instância. Trocando em miúdos, a partir dessa relação é possível saber quais são os estados que compõem uma determinada região, e vice-versa. Quanto à cardinalidade, tem-se que para cada instância do tipo Estado pode haver qualquer número de instâncias do tipo Região, ou seja, *(*)*; por outro lado, para cada instância do tipo Região deve haver ao menos duas instâncias do tipo Estado, ou seja, *(2..*)*.

Por sua vez, a relação *acordadoPor* corresponde a uma relação de mediação, no qual uma instância do tipo Acordo somente existe a partir de uma interação social, esta representada por múltiplas associações com instâncias do tipo Agente,

que atuam como mediadores da instância do tipo Acordo. Resumidamente, com base nessa relação, é possível identificar os agentes envolvidos na mediação de um determinado acordo, e vice-versa. Quanto à cardinalidade, tem-se que para cada instância do tipo Agente pode haver qualquer número de instâncias do tipo Acordo, ou seja, *(*)*, ao passo que para cada instância do tipo Acordo deve haver ao menos duas instâncias do tipo Agente, ou seja, *(2..*)*.

Finalmente, a relação *afeta* equivale a uma relação de inerência, visto que uma instância do tipo Condição nada mais é que uma propriedade abstrata que, por sua vez, se apresenta como aspecto de uma instância do tipo Organismo. Em suma, a partir dessa relação é possível descobrir as condições inerentes a um determinado organismo, e vice-versa. Quanto à cardinalidade, tem-se que para cada instância do tipo Organismo pode haver qualquer número de instâncias do tipo Condição, ou seja, *(*)*, enquanto que para cada instância do tipo Condição deve haver obrigatoriamente uma instância do tipo Organismo, ou seja, *(1)*.

No que diz respeito às relações entre entidades, é importante notar que, para se garantir a consistência dos dados, é necessária a formulação de regras de restrição, de modo que, por exemplo, um indivíduo do tipo Cidade não seja associado, simultaneamente, a indivíduos do tipo Estado e País, por meio da relação *cidadeDe*. Ainda que tais regras não façam parte do escopo do trabalho, é importante ter em vista que, para a devida operação de um sistema que se baseie no modelo proposto, estas são imprescindíveis.

4.7 **Considerações**

Como pôde ser observado, muitos dos conceitos definidos em *UFO* não foram utilizados pela ontologia aqui proposta, sequer indiretamente. Isso se deve ao fato do projeto ter como meta a obtenção de um modelo simples, conforme informado anteriormente, possuindo a capacidade de responder as questões de competência oriundas da narrativa jornalística, ainda que, em última instância, de forma minimalista. Sendo assim, optou-se por apresentar um modelo resumido, inicial, de uma ontologia cujo domínio é a cobertura de fatos jornalísticos, com o intuito de dar

um primeiro passo rumo à formalização e à explicitação deste conhecimento que é sabidamente vasto.

Entende-se que os conceitos de *UFO* definidos no presente trabalho e não cobertos pela ontologia de notícias seriam de grande valia ao detalhamento do modelo e, por conseguinte, à sua consolidação. Embora isso represente um esforço que foge do que aqui foi proposto, é exatamente esse entendimento que motivou a exposição dos referidos conceitos, que, na situação em questão, se mostram extremamente úteis.

5 Resultados alcançados

Para atestar a aplicabilidade, e ao menos alguns dos benefícios, do modelo proposto, uma prova de conceito foi realizada a partir da construção de um protótipo. Outro aspecto relevante dessa experiência é o entendimento de que se trata de um teste preliminar cujos critérios de validação estão orientados, não para se chegar a uma interface de consumo ou mesmo de gestão do conteúdo jornalístico, mas para dar visibilidade ao conhecimento gerado a partir dos relacionamentos definidos no nível dos dados e, também, no nível das informações, graças ao uso de uma ontologia capaz de lidar com tais vínculos dentro de uma dimensão semântica.

O protótipo consiste em um banco de triplas instalado em um computador cujo propósito é servir como ambiente informacional onde a ontologia, que foi implementada dentro do padrão *RDF*, na sintaxe *Turtle**, é armazenada e carregada. Completadas todas as etapas que dizem respeito à instalação do banco de triplas e a inserção da ontologia neste repositório, um conjunto não aleatório de 50 matérias foi selecionado para integrar o *corpus* de conhecimento do protótipo – obtido a partir de um processo de importação do próprio *website G1* – sendo associado aos conceitos de maior relevância abordados pelo mesmo, desde que comuns à ontologia. O conjunto de matérias escolhido especialmente para compor esse experimento aborda um evento em comum e, assim, pode ser considerado como parte de uma única cobertura jornalística que diz respeito a um caso de corrupção que teve grande repercussão no Brasil, em 2005, intitulado pelo *G1* e outros grandes jornais *on-line* como *Corrupção nos Correios* (vide Figura 25).

Acredita-se que a quantidade de matérias que foi estabelecida atenda adequadamente ao propósito do teste, servindo como uma amostra fidedigna do conteúdo produzido por um jornal *on-line*, visto que a ideia aqui é ter em mãos uma massa de informações com as quais seja possível gerar, de forma direta, alguns vínculos semânticos. A título de curiosidade, vale dizer que a quantidade definida para o protótipo equivale a aproximadamente 0,5% das matérias publicadas diariamente[20] pelo *G1*.

[20]De acordo com o G1, publica-se uma média de mil matérias por dia.

Figura 25 - *Screenshot* da página de uma das matérias utilizadas pelo protótipo

O processo de implementação da ontologia se deu de modo consideravelmente simples. Primeiramente, os conceitos de *UFO* utilizados como base para o modelo foram declarados em um arquivo em separado; em seguida, os conceitos propostos pela ontologia de notícias foram adicionados a este mesmo arquivo. Outro arquivo foi criado contendo todos os universais, enquanto que os indivíduos foram separados por tipo em arquivos distintos; deste modo, tinha-se um arquivo com indivíduos do tipo Ação Noticiosa (vide Anexo F), outro com indivíduos do tipo Lugar, e assim por diante. Optou-se por fazer dessa forma para facilitar a gestão das entidades, visto que a tarefa de implementação, e seus possíveis ajustes durante a fase de testes, foi conduzida manualmente (vide Figura 26).

Os relacionamentos existentes no protótipo em questão puderam ser visualizados através de uma *webpage* que foi gerada a cada requisição (*query**) *SPARQL* ao banco de triplas, esta realizada por meio de um *endpoint** oferecido pelo ambiente informacional. Ao todo, foram realizadas 14 requisições ao banco de dados, algumas bastante simples e outras consideradas complexas, como, por exemplo, a identificação de todas as pessoas que foram investigadas e afastadas de

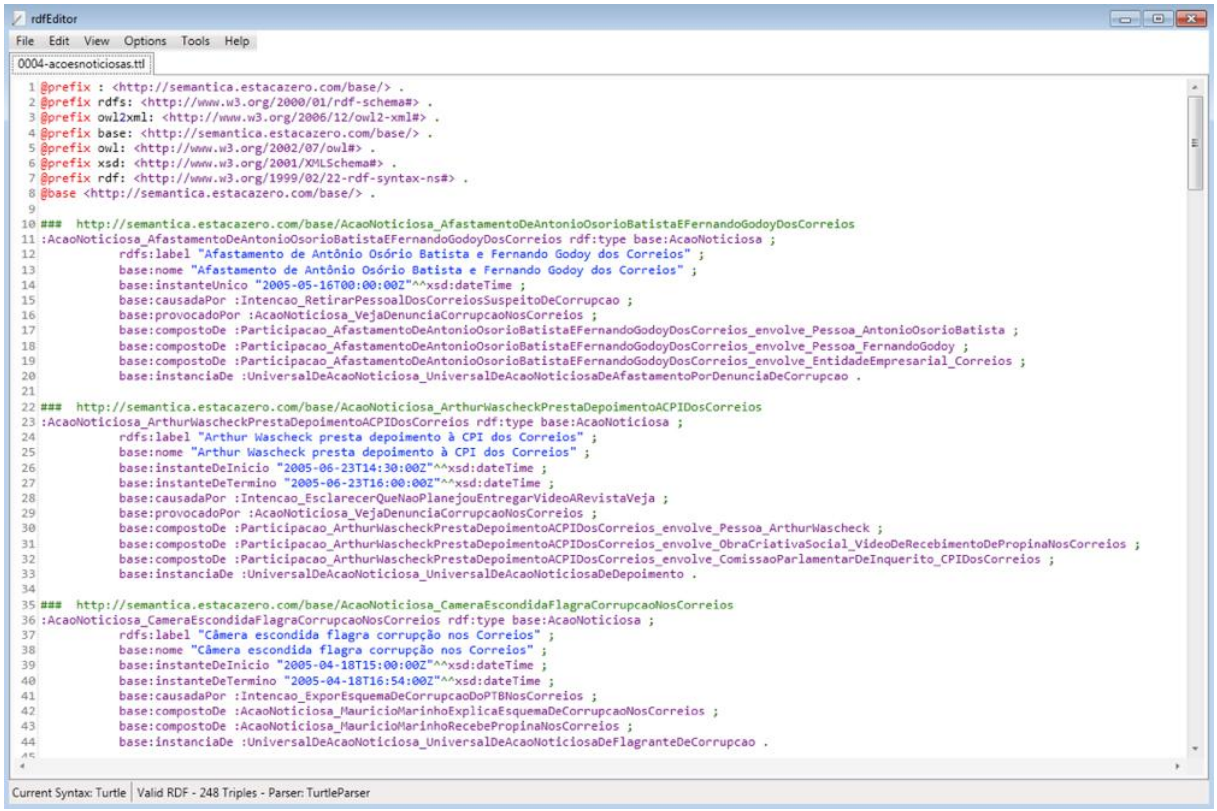


Figura 26 - Trecho do código gerado na implementação da ontologia de notícias

seus respectivos cargos devido a um possível envolvimento no caso de corrupção de que trata o *corpus* noticioso; para esta requisição, o sistema foi capaz de encontrar 6 indivíduos atendendo às condições especificadas.

Graças à análise preliminar dos conceitos que dizem respeito ao *lead* da notícia, foi possível estabelecer uma política de anotação semântica clara para as matérias integrantes do protótipo. Como efeito, tudo aquilo que não estava diretamente relacionado às perguntas que compõem o *lead* foi desconsiderado. Isso não apenas facilitou o trabalho de identificação das entidades relevantes em cada uma das matérias, como permitiu que fossem estabelecidas relações de considerável importância dentro do contexto da cobertura jornalística, estas muitas vezes implícitas no conteúdo e, portanto, nem sempre claras para os leitores ou mesmo para os produtores e gestores de conteúdo.

Entre as evidências encontradas durante as etapas de construção e verificação do protótipo, percebeu-se que a política utilizada na anotação do *corpus* noticioso pode ser encarada, claramente, como uma abordagem capaz de integrar um *framework** de anotação semântica do conteúdo jornalístico, sendo o protótipo

mesmo um passo importante nesta direção, devido à sua capacidade de demonstrar alguns dos conceitos mais fundamentais que permeiam a Web Semântica, como será visto adiante. Ao utilizar a ontologia de notícias proposta, esse *framework* poderia ser integrado ao *CMS* dos jornais *on-line*, de modo que produtores e gestores de conteúdo fossem capazes de anotar, manual ou semiautomaticamente, todo conteúdo produzido.

Entendendo que tanto a construção quanto a verificação do protótipo equivalem a meios de validação do modelo proposto neste trabalho, houve, desde a etapa de conceituação da ontologia de notícias, uma tentativa contínua de analisar a lista de tópicos do *G1*. Concluída a análise de 500 tópicos, o que representa pouco mais de 31% do total, verificou-se que cerca de 70% desta amostra poderia ser integrado ao modelo proposto sem a necessidade de criar novos universais ou tipos de indivíduos, ou seja, utilizando o modelo tal como foi concebido. Ainda que alguns desses tópicos se encaixem em conceitos dentro da ontologia considerados pouco específicos, o resultado obtido se mostrou bastante positivo e prova que o modelo tem expressividade suficiente para atender a maior parte da demanda dos jornais *on-line* brasileiros não especializados, sem muitos ajustes.

Como resultado da avaliação do protótipo, verificou-se que as matérias associadas a determinados tópicos são encontradas sem dificuldades e que, a partir de determinadas matérias, é possível encontrar outras que tratam de assuntos e eventos, ou quaisquer outros conceitos, afins. Devido à formalização de algumas relações implícitas nas matérias, foi possível encontrar informações que até então jamais seriam obtidas a partir do simples consumo do conteúdo. Isso denota que o conhecimento por trás das matérias não apenas foi explicitado e estruturado, mas também foi capaz de gerar um novo conhecimento que, conforme apresentado anteriormente, corresponde a um dos objetivos da Web Semântica.

Outro ponto que merece destaque é o fato do protótipo ter sido capaz de mostrar a facilidade com que uma ontologia de domínio pode ser estendida de uma ontologia de fundamentação; sem esta, a tarefa de implementação da ontologia de notícias seria muito mais complexa e onerosa. O processo de implementação da ontologia de notícias em si denota claramente o importante papel desempenhado pela ontologia de fundamentação; isto sem levar em conta que, em diversos momentos do processo de conceituação, *UFO* serviu como parâmetro para que equívocos de classificação não fossem cometidos.

O protótipo, que foi apresentado à Banca Examinadora durante a defesa da dissertação, é considerado parte integrante e fundamental deste trabalho, dado que os resultados obtidos através do próprio servem como confirmação da teoria aqui exposta, ou seja, como prova de que o modelo proposto de fato funciona e, portanto, tem aplicação prática.

É importante notar que, pelo fato dos jornais não especializados cobrirem um universo de conhecimento bastante abrangente, tratando de assuntos de todos os gêneros, o modelo proposto neste trabalho oferece condições de ser estendido para outros domínios de conhecimento.

6 Conclusão e futuros trabalhos

Primeiramente, é necessário esclarecer que, tal como *UFO*, a ontologia de notícias proposta neste trabalho não está terminada. Longe disso, pode-se perceber que se trata de um projeto contínuo, sem fim, dado que apenas alguns dos conceitos mais básicos foram explorados e, ainda assim, sem grande profundidade.

Por outro lado, por mais simples que seja a solução proposta, uma série de mudanças precisará ser realizada no âmbito tecnológico (banco de dados, plataforma de publicação etc), operacional (rotulação/anotação do conteúdo, extração de dados etc) e, consequentemente, humano, dado que há uma sensível alteração no fluxo de trabalho dos profissionais responsáveis pela gestão de conteúdo. Desse modo, aconselha-se que o processo de mudança seja gradativo, possibilitando a adequação do modelo mental, até então em vigor, à nova realidade estabelecida. Ainda assim, é de se esperar que focos de resistência na cadeia produtiva sejam encontrados, até que, finalmente, todos os profissionais envolvidos tenham condições de perceber as vantagens do novo modelo de trabalho, beneficiando-se do mesmo.

Para o sucesso do que aqui se propõe, é preciso que produtores e gestores de conteúdo digital reconheçam o valor da mudança, pois, tomando como base as iniciativas levantadas, os ganhos são altos tanto para estes profissionais quanto para os leitores, sendo esta percepção sustentada pelos argumentos apresentados adiante.

Levando-se em conta que o raciocínio automatizado é um dos diferenciais da Web Semântica e, portanto, um dos motivos para o uso de tecnologias afins no meio jornalístico, um passo importante seria colocar esta funcionalidade em operação, visto que, inevitavelmente, haverá uma quantidade expressiva de entidades se relacionando de forma indireta dentro do banco de dados. Tais relações implícitas podem, muitas vezes, representar informações ou porções de conhecimento úteis ao trabalho dos produtores de conteúdo.

Entre alguns dos benefícios no uso do raciocínio automatizado sobre o modelo proposto, pode-se destacar, por exemplo, a possibilidade de criação de uma plataforma de conhecimento voltada para a geração de argumentos de pauta, pois, conforme exposto anteriormente, a ontologia de notícias permite a descoberta de

novos conhecimentos, até então implícitos ou de difícil obtenção, a partir do uso de relações e inferências que promovam a análise de dados por perspectivas diferentes daquelas que são comumente lançadas sobre o modelo, o que muitas vezes é trivial para o mesmo.

Ao se utilizar a ontologia de notícias como alicerce tecnológico para a formulação do conhecimento jornalístico, é possível fornecer aos produtores de conteúdo novos fatos e pontos de vista sobre um fato noticioso, enriquecendo a sua cobertura, apontando novos assuntos relevantes, e dando respaldo técnico ao próprio trabalho investigativo do jornal. No entanto, para que isso seja alcançado, é fundamental o trabalho dos gestores de conteúdo, no sentido de formalizar e explicitar ao máximo o conhecimento por trás de cada fato jornalístico, pois, do contrário, o potencial do modelo proposto não será aproveitado em sua plenitude.

Considerando-se a proposta formalizada neste trabalho, parece razoável afirmar que os parâmetros que garantirão o seu sucesso estão sendo atendidos. A implantação deste novo modelo de classificação do conteúdo permitirá, e ensejará, a criação de uma grande quantidade de categorias cujo objetivo é reunir conteúdos afins – com um nível considerável de especialização, onde também serão abrangidas as definições que dizem respeito à pessoa, lugar, organização etc, chamadas aqui de entidades – sendo todos os tipos de conteúdo associados a um grande número de categorias, além da possibilidade destas mesmas categorias serem relacionadas entre si.

Com as categorias já estabelecidas e relacionadas, a operação dos produtores e gestores de conteúdo ganhará agilidade, tornando os processos de manutenção, armazenamento e recuperação das informações mais eficientes. Do lado do leitor, novos eixos na estrutura narrativa poderão ser apresentados, o que proporcionará uma experiência de imersão mais diversificada e profunda durante o consumo dos fatos jornalísticos, levando-o a requisitar mais conteúdos. Consequentemente, o tempo de permanência do leitor tenderá a aumentar, o que garantirá mais lucros aos jornais *on-line* brasileiros e, muito provavelmente, também se traduzirá num aumento de valor em suas marcas, devido à percepção positiva por parte do público consumidor.

Um aspecto deste trabalho que merece ser mencionado é o fato de que, até a conclusão desta dissertação, não havia uma documentação razoavelmente completa e atualizada de *UFO* disponível na língua portuguesa. Normalmente, o que se

encontra na Web são artigos acadêmicos com alguns fragmentos de *UFO*; em outros casos, quando documentos mais completos são encontrados, os mesmos estão desatualizados, com nomenclatura e conceitos obsoletos, entre outros problemas. Por conseguinte, acredita-se que este trabalho já traz em si uma contribuição significativa ao servir como documento de referência em *UFO*, sendo útil a pesquisadores e estudantes da área de Ciência da Computação.

Finalmente, é importante observar que, ao se colocar em prática a proposta em questão, os jornais *on-line* brasileiros terão condições de alcançar o que se entende por jornalismo digital de quarta geração – ao se extrapolar a ideia de uma versão para Web de um jornal impresso previamente existente – constituindo-se em produtos totalmente exclusivos para a Internet, onde o uso de bancos de dados complexos se consolida através da utilização de ferramentas automatizadas e diferenciadas na sua criação e manutenção, sem nada terem a dever aos seus concorrentes estrangeiros, nos termos aqui expostos.

REFERÊNCIAS

ACKRILL, J. L. **Aristotle**: Categories and de Interpretatione. Oxford: Clarendon Press, 1963.

ALESSO, H. P.; SMITH, C. F. **Thinking on the Web**: Berners-Lee, Godel and Turing. New Jersey: John Wiley & Sons, 2009.

ALLEN, J. F. **Maintaining Knowledge about Temporal Intervals**. Communications of the ACM, S. L., v. 26, n. 11, p. 832-843, 1983.

ANTONIOU, G.; HARMELEN, F. V. **A Semantic Web Primer**. New York: MIT, 2004.

BARDOEL, J.; DEUZE, M. Network Journalism: Converging Competences of Media Professionals and Professionalism. In: **Australian Journalism Review**, S. L., v. 2, n. 23, p. 91-103, 2001. Disponível em: <http://jclass.umd.edu/classes/jour698m/BardoelDeuze+NetworkJournalism+2001.pdf>. Acesso em: 31 ago. 2012.

BERNERS-LEE, T.; HENDLER, J.; LASSILA, O. **The Semantic Web**. Scientific American, S. L., p. 35-43, maio 2001. Disponível em: <http://www.ryerson.ca/~dgrimsha/courses/cps720_02/resources/Scientific%20American%20The%20Semantic%20Web.htm>. Acesso em: 31 ago. 2012.

BORST, W. N. **Construction of Engineering Ontologies for Knowledge Sharing and Reuse**. 1997, 227 f. Tese (Doutorado em Ciência da Computação) - University of Twente, Enschede, 1997. Disponível em: <http://purl.org/utwente/17864>. Acesso em: 31 ago. 2012.

BRINGUENTE, A. C. de O. **Reengenharia de uma Ontologia de Processo de Software e seu Uso para a Integração de Ferramentas de Apoio ao Planejamento de Projetos**. 2011. 102 f. Dissertação (Mestrado em Informática) - Universidade Federal do Espírito Santo, Vitória, 2011.

BUNGE, M. **Treatise on Basic Philosophy – Ontology I**: The Furniture of the World. New York: Reidel, 1977. v. 3.

_____. **Dicionário de Filosofia**. São Paulo: Perspectiva, 2006.

CHAUÍ, M. **Introdução à História da Filosofia**. São Paulo: Companhia das Letras, 2003.

COLLE, R. **Explotar la Información Noticiosa**. Data Minino Aplicado a la Documentación Periodística. Madri: Universidad Complutense de Madrid, 2002.

CSIKSZENTMIHALYI, M. **Flow**: The Psychology of Optimal Experience. New York: Harper and Row, 1990.

FIDALGO, A. Do Poliedro à Esfera: Os Campos de Classificação. A Resolução Semântica no Jornalismo On-line. In: ENCONTRO NACIONAL DE PESQUISADORES EM JORNALISMO, 2., 2004, Salvador. **Anais**... Brasília: SBPJor, 2004. 1 CD-ROM.

GÓMEZ-PÉREZ, A; FERNÁNDEZ-LÓPEZ, M.; CORCHO, O. **Ontological Engineering**: With Examples from the Areas of Knowledge Management, e-Commerce and the Semantic Web. Berlin: Springer, 2004.

GONÇALVES, E. O Banco de Dados como Formato no Jornalismo Digital. In: CONGRESSO DA LUSOCOM, 6., 2004, Covilhã. **Anais**... Covilhã: LUSOCOM, 2004. 1 CD-ROM.

GONÇALVES, B. N.; ZAMBORLINI, V.; GUIZZARDI, G. **An Ontological Analysis of the Electrocardiogram**. RECIIS. Electronic Journal of Communication Information and Innovation in Health, S. L., v. 3, p. 45-59, 2009.

GOSCIOLA, V. **Roteiro para as Novas Mídias**: Do Game à TV Interativa. São Paulo: Senac São Paulo, 2003.

GRUBER, T. R. **A Translation Approach to Portable Ontology Specifications**. Knowledge Acquisition, S. L., v. 5, n. 2, p. 199-220, 1993.

_____. Ontology. In: LIU, L; ÖZSU, M. T. (Eds.). **Encyclopedia of Database Systems**. Berlin: Springer, 2009. p. 1963-1965.

GRÜNINGER, M.; FOX, M. S. The Role of Competency Questions in Enterprise Engineering. In: ROLSTADÅS, A (Ed.). **Benchmarking**: Theory and Practice. London: Chapman and Hall, 1994. p. 83-95.

GUARINO, N; GIARETTA, P. Ontologies and Knowledge Bases: Towards a Terminological Clarification. In: MARS, N. (Ed.). **Towards Very Large Knowledge Bases**: Knowledge Building and Knowledge Sharing. Amsterdam: IOS Press, 1995. p. 25-32.

_____. Semantic Matching: Formal Ontological Distinction for Information Organization, Extraction, and Integration. In: PAZIENZA, M. T. (Ed.). **Information Extraction**: A Multidisciplinary Approach to an Emerging Information Technology. Berlin: Springer, 1997. p. 139-170.

_____. Formal Ontology and Information Systems. In: GUARINO, N. (Ed.). **Proceedings of the 1st International Conference on Formal Ontologies in Information Systems**. Trento: IOS Press, 1998. p. 3-15.

GUIZZARDI, G. **Ontological Foundations for Structural Conceptual Models**. 2005. 416 f. Tese (Doutorado em Ciência da Computação) - University of Twente, Enschede, 2005. Disponível em: <http://purl.org/utwente/50826>. Acesso em: 31 ago. 2012.

_____; WAGNER, G. Some Applications of a Unified Foundational Ontology in Business Modeling. In: ROSEMANN, M.; GREEN, P. (Eds.). **Ontologies and Business Systems Analysis**. Hershey: Idea Group Publishing, 2005. p. 345-367.

_____; FALBO, R. de A.; GUIZZARDI, R. S. S. **A importância de Ontologias de Fundamentação para a Engenharia de Ontologias de Domínio**: O Caso do Domínio de Processos de Software. Revista IEEE América Latina, v. 6, n. 3, p. 244-251, jul. 2008.

_____; ALMEIDA, J. P. A.; GUIZZARDI, R. S. S.; FALBO, R. de A. Ontologias de Fundamentação e Modelagem Conceitual. In: SEMINÁRIO DE PESQUISA EM ONTOLOGIA NO BRASIL, 2., 2009, Rio de Janeiro. **Anais**... ONTOBRAS, 2009a. p. 1-6.

_____; LOPES, M.; BAIÃO, F. A.; FALBO, R. de A. On the importance of Truly Ontological Distinctions for Ontology Representation Languages: An Industrial Case Study in the Domain of Oil and Gas. In: INTERNATIONAL CONFERENCE ON EXPLORING MODELING METHODS IN SYSTEMS ANALYSIS AND DESIGN, 14., 2009, Amsterdam. **Anais**... Amsterdam: EMMSAD, 2009b. p. 224-236.

_____; WAGNER, G. Using the Unified Foundational Ontology (UFO) as a Foundation for General Conceptual Modeling Languages. In: POLI, R.; HEALY, M; KAMEAS, A. (Eds.). **Theory and Applications of Ontology**: Computer Applications. Berlin: Springer, 2010. p. 175-196.

HALPIN, H. **The Semantic Web**: The Origins of Artificial Intelligence Redux. Edinburgh: University of Edinburgh, 2004. Disponível em: <http://www.ibiblio.org/hhalpin/homepage/publications/airedux.pdf>. Acesso em: 31 ago. 2012.

HOFFMAN, D. L.; NOVAK, T. P. **Marketing in Hypermedia Computer-Mediated Environments**: Conceptual Foundations. Journal of Marketing, S. L., n. 60, p. 50-68, jul. 1996.

KILIAN, C. **Writing for the Web**. Bellingham: Self-Counsel, 2000.

LALANDE, A. **Vocabulaire Technique et Critique de la Philosophie**. Paris: Presses Universitaires de France, 1976.

LANDOW, G. P.; DELANY, P. **Hypermedia and Literary Criticism**. Cambridge: MIT Press, 1991.

LAUFER, R.; SCAVETTA, D. **Texto, Hipertexto, Hipermedia**. Porto: Rés, 1997.

LENE, H. Objetividade e Subjetividade na Seleção e Classificação. **Observatório da Imprensa**. 501. ed. São Paulo, 02 set. 2008. Disponível em: <http://www.observatoriodaimprensa.com.br/news/view/objetividade-e-subjetividade-na-selecao-e-classificacao>. Acesso em: 31 ago. 2012.

LÉVY, P. **As Tecnologias da Inteligência**: O Futuro do Pensamento na Era da Informática. São Paulo: Editora 34, 1993.

LIPTON, R. **Multimedia Toolkit**: Build Your Own Solutions with DocuSource. New York: Random House, 1992.

LOPES, M.; GUIZZARDI, G.; BAIÃO, F. A.; FALBO, R. de A. Reverse Engineering a Domain Ontology to Uncover Fundamental Ontological Distinctions: An Industrial Case Study in the Domain of Oil and Gas Production and Exploration. In: INTERNATIONAL CONFERENCE ON ENTERPRISE INFORMATION SYSTEMS, 11., 2009, Milão. **Anais**... Milão: ICEIS, 2009. p. 262-267.

MACHADO, A. Hipermídia: O Labirinto como Metáfora. In: DOMINGUES, D. (Org.). **A Arte no Século XXI**: A Humanização das Tecnologias. São Paulo: Fundação Editora da UNESP, 1997. p. 144-154.

MACHADO, E. Banco de Dados como Formato no Jornalismo Digital. In: CONGRESSO LUSÓFONO DE CIÊNCIAS DA COMUNICAÇÃO, 6., 2004, Covilhã. **Anais**... Covilhã: Lusocom, 2004.

MANOVICH, L. **The Language of New Media**. Cambridge: MIT Press, 2001.

MCCARTHY, J.; MINKSY, M.; ROCHESTER, N.; SHANNON, C. **A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence**. Estados Unidos, 1955.

MEIXNER, U. **Axiomatic Formal Ontology**. Dordrecht: Kluwer Academic Publishers, 1997.

MIELNICZUK, L. **Jornalismo na Web**: Uma Contribuição para o Estudo do Formato da Notícia na Escrita Hipertextual. 2003. Tese (Doutorado em Comunicação) - Faculdade de Comunicação, Universidade Federal da Bahia, Salvador, 2003.

NECHES, R.; FIKES, R.; FININ, T.; GRUBER, T.; SENATOR, T.; SWARTOUT, W. **Enabling Technology for Knowledge Sharing**. AI Magazine, S. L., v. 12, n. 3, p. 36-56, 1991.

PALACIOS, M. **Jornalismo On-line, Informação e Memória**: Apontamentos para Debate. Covilhã: Editora Universidade da Beira Interior, 2002.

_____; MIELNICZUK, L.; BARBOSA, S.; RIBAS B.; NARITA, S. **Um Mapeamento de Características e Tendências no Jornalismo On-line Brasileiro e Português**. Comunicarte, Aveiro, v. 1, n. 2, 2002.

_____. Ruptura, Continuidade e Potencialização no Jornalismo On-line: O Lugar da Memória. In: MACHADO, E.; PALACIOS, M. (Orgs.). **Modelos do Jornalismo Digital**. Salvador: Editora Calandra, 2003.

PAVLIK, J. **Journalism and New Media**. New York: Columbia University Press, 2001.

PINHEIRO, M. **Do Design de Interface ao Design da Experiência**. Revista Design em Foco, Salvador, v. 4, n. 2, p. 9-23, jul./dez. 2007.

SANDHAUS, E.; LARSON, R. **A Century of Semantic Technology**: Semantics at The New York Times. 2009. Disponível em: <http://graphics8.nytimes.com/packages/images/rd/final.pdf>. Acesso em: 31 ago. 2012.

SCHWINGEL, C. Jornalismo Digital de Quarta Geração: A Emergência de Sistemas Automatizados para o Processo de Produção Industrial no Jornalismo Digital. In: COMPÓS, 2005, Niterói. **Anais**... Niterói: Compós, 2005, 1 CD-ROM.

SILVA, V. dos S. **Uma Abordagem para Alinhamento de Ontologias Biomédicas para Apoiar a Anotação Genômica**. 2010. 112 f. Dissertação (Mestrado em Informática) - Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2010.

SMITH, B. Ontology. In: FLORIDI, L. (Ed.). **The Blackwell Guide to the Philosophy of Computing and Information**. Malden: Blackwell, 2003. p. 155-166.

SODRÉ, M.; FERRARI, M. H. **Técnica de Reportagem**: Notas sobre a Narrativa Jornalística. São Paulo: Summus Editorial, 1986.

TURING, A. M. **Computing Machinery and Intelligence**. Mind, S. L., v. 59, n. 236, p. 433-460, out. 1950.

WAND, Y.; WEBER, R. **On the Deep Structure of Information Systems**. Information Systems Journal, S. L., v. 5, n. 3, p. 203-223, 1995.

ZAMBORLINI, V. **Estudo de Alternativas de Mapeamento de Ontologias da Linguagem OntoUML para OWL**: Abordagens para Representação de Informação Temporal. 2011. 181 f. Dissertação (Mestrado em Informática) - Universidade Federal do Espírito Santo, Vitória, 2011.

ZUIN, A. L. A.; CORREIA, C. M. de C. Jornalismo Impresso e Jornalismo On-line: A Linguagem Híbrida da Informação. In: SIMPÓSIO NACIONAL DE PESQUISADORES EM COMUNICAÇÃO E CIBERCULTURA - ABCIBER, 2., 2008, São Paulo. **Anais**... São Paulo: PUC-SP, 2008.

GLOSSÁRIO

Agentes de *software*	Programas de computador autônomos, capazes de se comunicar com outros agentes e com seres humanos, que agem em prol de uma meta
Arquiteto de informação	Aquele que estrutura o conteúdo
Cardinalidade	Corresponde a uma restrição de integridade que representa regras de negócio que, por sua vez, estabelecem as quantidades mínima e máxima de relacionamentos entre uma determinada entidade e outras
Ciência da Computação	Estudo dos algoritmos, suas aplicações e sua implementação, na forma de *software*, para execução em computadores eletrônicos
Ciência da Informação	Estudo da informação, desde a sua gênese até o processo de transformação de dados em conhecimento
Content Management System - CMS	Sistema gestor de *websites* que integra ferramentas necessárias para gerir conteúdos em tempo real, sem a necessidade de programação de código; seu objetivo é estruturar e facilitar a criação, administração, distribuição, publicação e disponibilização da infomação
Conteúdo digital	Toda e qualquer representação digital da informação, capaz de ser publicada na Web
Design da Experiência	Segundo Pinheiro (2007), é a área do Design que investiga a *experiência de uso* e utiliza conhecimentos das ciências sociais e de comportamento
Editoria	Assunto, tema ou gênero criado arbitrariamente pelo corpo

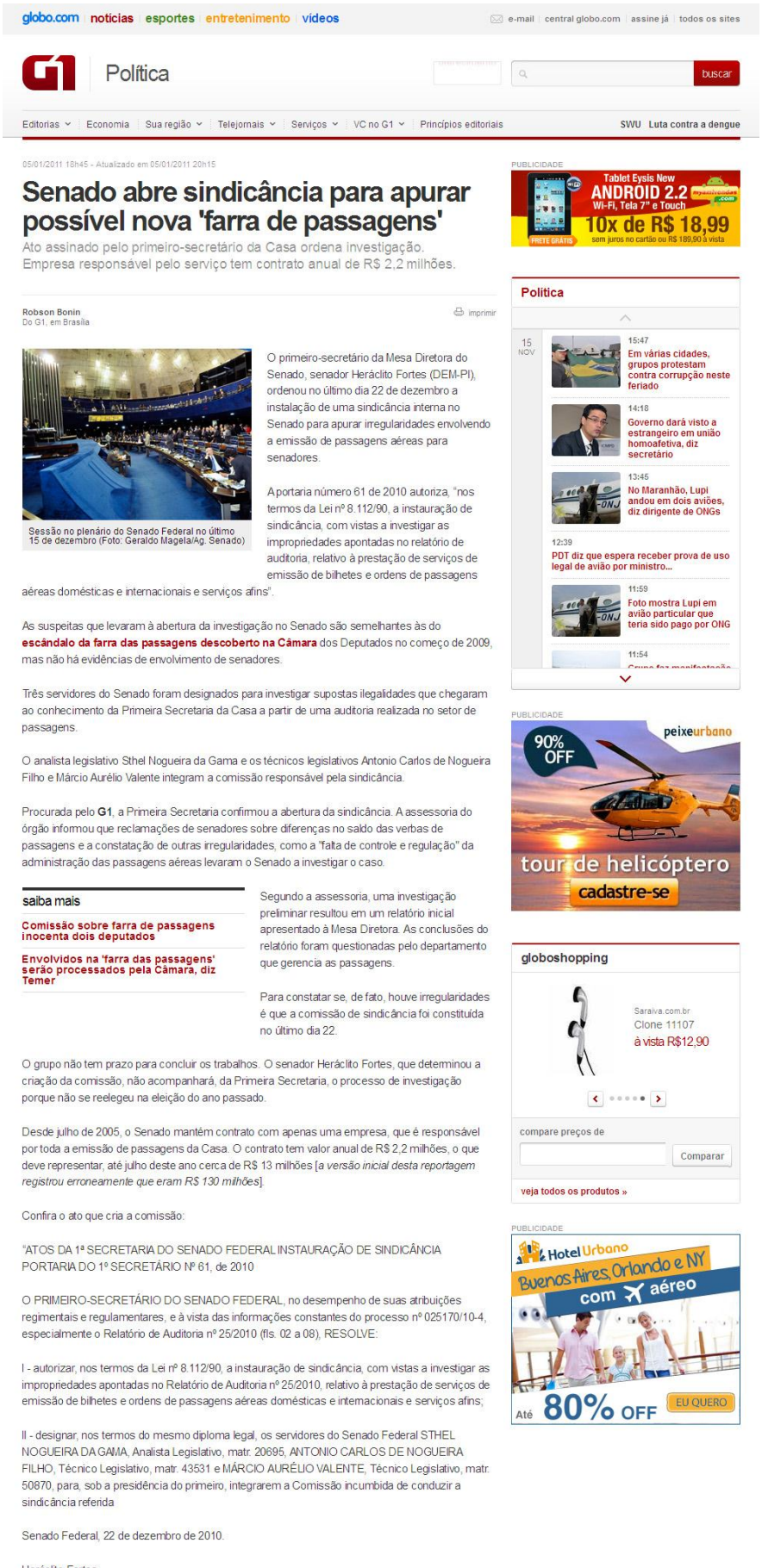
	editorial de um jornal para agrupar e organizar conteúdos afins; funciona como um diretório de informações, sendo normalmente um reflexo da divisão/especialização do trabalho dentro da redação
Endpoint	Terminal de comunicação
Engenheiro de ontologia	Aquele que é responsável pela conceituação e construção de ontologias, estudando e desenvolvendo métodos e metodologias voltados a esta atividade; seu principal objetivo é tornar explícito o conhecimento contido nas aplicações de *software*, assim como nos procedimentos organizacionais e de negócio para um domínio particular
Estado de coisas	Segundo Meixner (1997), é a intenção de um comunicado; trata-se de uma definição explicativa ou explicação, portanto, não é nem uma análise do significado estabelecido, nem uma definição nominal puramente convencional
Frames	Mecanismo de representação do conhecimento que possui uma estrutura parecida com escaninhos, nos quais são colocadas etiquetas e valores; os *frames* permitem a representação de relações hierárquicas entre conceitos e, através destas relações, é possível inferir propriedades e responder a determinadas questões
Framework	Abstração que une códigos comuns entre vários projetos de *software*, provendo uma funcionalidade genérica. Observa-se, no entanto, que um *framework* pode atingir uma funcionalidade específica, por configuração, ao se programar uma aplicação de *software*
Gestor de conteúdo digital	Aquele que é responsável pela classificação do conteúdo digital
Grande Rede	Designação dada à World Wide Web

Instanciação	Processo de individualização de um tipo ou categoria no qual se tem como resultado a produção de instâncias, que são objetos únicos, distintos, que representam o tipo ou categoria do qual fazem parte
Internet	Sistema global de redes interligadas de computadores que usam o padrão *TCP/IP** para servir bilhões de usuários no mundo inteiro
Interoperabilidade semântica	Habilidade de sistemas de computador transmitir dados com significado compartilhado, sem ambiguidade, sendo isto um requisito necessário à realização de operações lógicas computáveis e inferências de máquina, bem como à descoberta de conhecimento e ao estabelecimento de federações de dados entre sistemas de informação
Jornal *on-line*	Produto jornalístico voltado para a Web; *website* de notícias
Lógica de Primeira Ordem	Sistema lógico que estende a Lógica Proposicional; seu compromisso ontológico consiste de objetos com certas relações entre eles que são válidas ou não válidas
Lógica Formal	Disciplina da Filosofia que trata de estruturas lógicas formais sem levar em consideração a sua veracidade
Mereologia	Teoria das relações entre as partes e o todo
Mídia social	Refere-se ao uso do meio eletrônico para interação social, permitindo a criação, o compartilhamento e o intercâmbio do conteúdo gerado pelo usuário
Natural Language Processing - NLP	Campo da Ciência da Computação e da Linguística que estuda as interações entre computadores e a linguagem humana, tendo surgido como um ramo da Inteligência Artificial
Ontologia	De acordo com Borst (1997), é uma especificação formal de

	uma conceituação compartilhada
Ontologia de domínio	Tipo de ontologia que descreve o vocabulário relacionado a um domínio genérico através da especialização de conceitos introduzidos na ontologia de fundamentação
Ontologia de fundamentação	Também chamada de *ontologia de alto nível*, é um tipo de ontologia que reúne resultados da Ontologia Formal em um sistema de categorias usadas para articular conceituações de diversos domínios
Ontologia Formal	Parte específica da disciplina filosófica de Ontologia que lida com estruturas ontológicas formais, isto é, com aspectos formais de objetos independentemente da sua natureza particular
Produtor de conteúdo digital	Aquele que é responsável pela criação do conteúdo digital
Query	Requisição/consulta ao banco de dados
Questões de competência	Na área de engenharia de ontologias, correspondem ao resultado da especificação de requisitos, tendo como objetivo avaliar o compromisso ontológico estabelecido na definição da ontologia; deste modo, é possível verificar se a mesma satisfaz os requisitos especificados
Recurso	Entidade que é representada por meio de uma URI
SPARQL	Linguagem de *query** para bancos de dados, capaz de recuperar e manipular dados armazenados no formato RDF
TCP/IP	Conjunto de protocolos de comunicação entre computadores em rede
Turtle	Formatação utilizada para expressar os dados no modelo RDF

Web	Abreviação de World Wide Web
Web Semântica	Conceito que se propõe a ser uma evolução gradual da *Web* tal qual a conhecemos, por meio da adição de significado ao atual formato de representação de dados; deste modo, os dados poderão ser facilmente processados pelas aplicações, e técnicas avançadas poderão ser utilizadas a fim de tirar proveito desta representação
Webpage	Documento ou recurso de informação que é adequado para a World Wide Web, podendo ser acessado por meio de um navegador Web e exibido em um monitor ou dispositivo móvel
Website	Conjunto de *webpages* geralmente acessíveis pela Internet
World Wide Web	Sistema de documentos em hipermídia que são interligados e executados na Internet

ANEXO A - *Screenshot* da matéria *Senado abre sindicância para apurar possível nova 'farra de passagens'*, do *G1*



globo.com | notícias | esportes | entretenimento | vídeos — e-mail | central globo.com | assine já | todos os sites

G1 Política — buscar

Editorias | Economia | Sua região | Telejornais | Serviços | VC no G1 | Princípios editoriais — SWU Luta contra a dengue

05/01/2011 18h45 - Atualizado em 05/01/2011 20h15

Senado abre sindicância para apurar possível nova 'farra de passagens'

Ato assinado pelo primeiro-secretário da Casa ordena investigação. Empresa responsável pelo serviço tem contrato anual de R$ 2,2 milhões.

Robson Bonin
Do G1, em Brasília

imprimir

Sessão no plenário do Senado Federal no último 15 de dezembro (Foto: Geraldo Magela/Ag. Senado)

O primeiro-secretário da Mesa Diretora do Senado, senador Heráclito Fortes (DEM-PI), ordenou no último dia 22 de dezembro a instalação de uma sindicância interna no Senado para apurar irregularidades envolvendo a emissão de passagens aéreas para senadores.

A portaria número 61 de 2010 autoriza, "nos termos da Lei nº 8.112/90, a instauração de sindicância, com vistas a investigar as impropriedades apontadas no relatório de auditoria, relativo à prestação de serviços de emissão de bilhetes e ordens de passagens aéreas domésticas e internacionais e serviços afins".

As suspeitas que levaram à abertura da investigação no Senado são semelhantes às do **escândalo da farra das passagens descoberto na Câmara** dos Deputados no começo de 2009, mas não há evidências de envolvimento de senadores.

Três servidores do Senado foram designados para investigar supostas ilegalidades que chegaram ao conhecimento da Primeira Secretaria da Casa a partir de uma auditoria realizada no setor de passagens.

O analista legislativo Sthel Nogueira da Gama e os técnicos legislativos Antonio Carlos de Nogueira Filho e Márcio Aurélio Valente integram a comissão responsável pela sindicância.

Procurada pelo **G1**, a Primeira Secretaria confirmou a abertura da sindicância. A assessoria do órgão informou que reclamações de senadores sobre diferenças no saldo das verbas de passagens e a constatação de outras irregularidades, como a "falta de controle e regulação" da administração das passagens aéreas levaram o Senado a investigar o caso.

saiba mais

Comissão sobre farra de passagens inocenta dois deputados

Envolvidos na 'farra das passagens' serão processados pela Câmara, diz Temer

Segundo a assessoria, uma investigação preliminar resultou em um relatório inicial apresentado à Mesa Diretora. As conclusões do relatório foram questionadas pelo departamento que gerencia as passagens.

Para constatar se, de fato, houve irregularidades é que a comissão de sindicância foi constituída no último dia 22.

O grupo não tem prazo para concluir os trabalhos. O senador Heráclito Fortes, que determinou a criação da comissão, não acompanhará, da Primeira Secretaria, o processo de investigação porque não se reelegeu na eleição do ano passado.

Desde julho de 2005, o Senado mantém contrato com apenas uma empresa, que é responsável por toda a emissão de passagens da Casa. O contrato tem valor anual de R$ 2,2 milhões, o que deve representar, até julho deste ano cerca de R$ 13 milhões [*a versão inicial desta reportagem registrou erroneamente que eram R$ 130 milhões*].

Confira o ato que cria a comissão:

"ATOS DA 1ª SECRETARIA DO SENADO FEDERAL INSTAURAÇÃO DE SINDICÂNCIA PORTARIA DO 1º SECRETÁRIO Nº 61, de 2010

O PRIMEIRO-SECRETÁRIO DO SENADO FEDERAL, no desempenho de suas atribuições regimentais e regulamentares, e à vista das informações constantes do processo nº 025170/10-4, especialmente o Relatório de Auditoria nº 25/2010 (fls. 02 a 08), RESOLVE:

I - autorizar, nos termos da Lei nº 8.112/90, a instauração de sindicância, com vistas a investigar as impropriedades apontadas no Relatório de Auditoria nº 25/2010, relativo à prestação de serviços de emissão de bilhetes e ordens de passagens aéreas domésticas e internacionais e serviços afins;

II - designar, nos termos do mesmo diploma legal, os servidores do Senado Federal STHEL NOGUEIRA DA GAMA, Analista Legislativo, matr. 20695, ANTONIO CARLOS DE NOGUEIRA FILHO, Técnico Legislativo, matr. 43531 e MÁRCIO AURÉLIO VALENTE, Técnico Legislativo, matr. 50870, para, sob a presidência do primeiro, integrarem a Comissão incumbida de conduzir a sindicância referida

Senado Federal, 22 de dezembro de 2010.

PUBLICIDADE

Tablet Eysis New ANDROID 2.2 Wi-Fi, Tela 7" e Touch 10x de R$ 18,99 sem juros no cartão ou R$ 189,90 à vista

Política

15 NOV

15:47 Em várias cidades, grupos protestam contra corrupção neste feriado

14:18 Governo dará visto a estrangeiro em união homoafetiva, diz secretário

13:45 No Maranhão, Lupi andou em dois aviões, diz dirigente de ONGs

12:39 PDT diz que espera receber prova de uso legal de avião por ministro...

11:59 Foto mostra Lupi em avião particular que teria sido pago por ONG

11:54

PUBLICIDADE

peixeurbano 90% OFF tour de helicóptero cadastre-se

globoshopping

Saraiva.com.br Clone 11107 à vista R$12,90

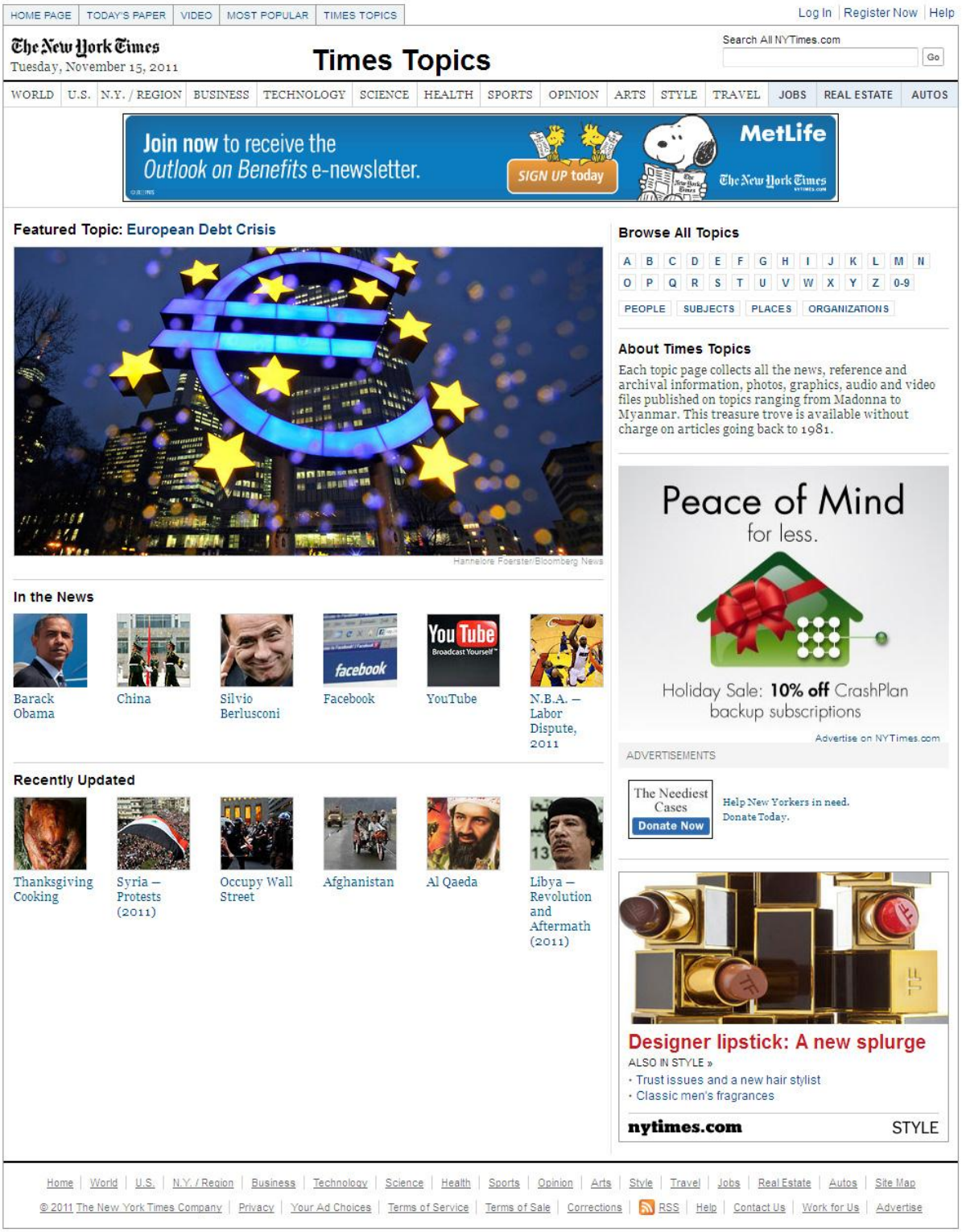
compare preços de Comparar

veja todos os produtos »

PUBLICIDADE

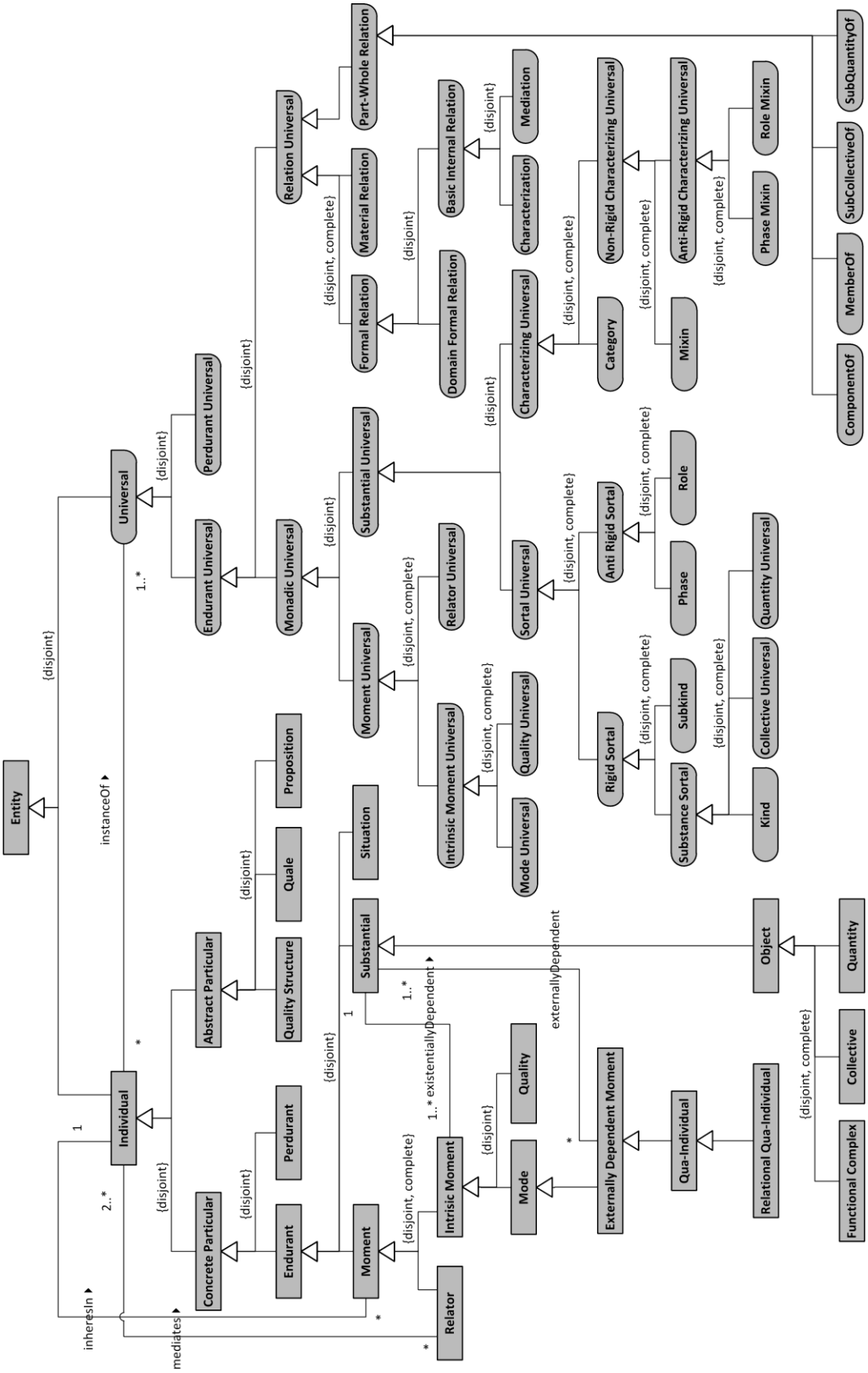
Hotel Urbano Buenos Aires, Orlando e NY com aéreo Até 80% OFF EU QUERO

ANEXO B - *Screenshot* da página de tópicos do *NYT*



ANEXO C - *Screenshot* da página de tópicos do *Estadão*

ANEXO D - Fragmento de *UFO-A*: ontologia de *endurantes*



ANEXO E - Lista dos assuntos do *G1* iniciados com a letra *A*

abandono
abandono de bebê
abandono de menor
abastecimento de água
abelha
aborto
abraço
abuso de autoridade
abuso de menores
abuso sexual
academia
açaí
ação trabalhista
acessibilidade
acessórios
acidente
acidente aéreo
acidente ambiental
acidente de carro
acidente de helicóptero
acidente de trabalho
acidente de trânsito
acidente doméstico
acidente ecológico
acidente escolar
acidente ferroviário
acidente marítimo
aço
acordo comercial
acordo internacional
adestramento de cães
administração
administração pública
adoção
adoção de animais
adolescentes
aeromodelismo
aeronáutica
aeroportos
afogamento
agentes públicos
agressão
agressão infantil
agressões contra mulheres
agricultura
agricultura familiar
agronegócio
agropecuária
agrotóxicos
água
água contaminada
AIDS
air guitar
ajuda humanitária
alagamento
albergues
álcool
alcoolismo
alegorias
alergia alimentar
alfabetização
alfândega
algodão
alianças políticas
aliciamento de menores
alimentação
alimentação saudável
alimento orgânico
alimentos
alistamento militar
alpinismo
alzheimer
amamentação
ambulantes
ameaça de bomba
ameaça de desabamento
ameaçado de extinção
amigos
amistosos
amizade
amor
anabolizantes
analfabetismo
analfabetismo funcional
anemia
animação
animais
animais silvestres
anistia
aniversário
ano novo
anorexia
antibiótiocos
anticoncepcionais
antropologia
apagão
aposentadoria
aposentados
apreensão
apreensão de alimento
apreensão de armas
apreensão de aves
apreensão de caça-níqueis
apreensão de carga
apreensão de contrabando
apreensão de drogas
apreensão de pássaros
apreensão de peixe
apreensão de remédio
apuração
aquecimento global
arbitragem
árbitros
áreas de risco
áreas públicas
armas
armas nucleares
arqueologia
arquitetura
arquivologia
arrastão
arrecadação
arte
arte contemporânea
arte de rua
arte digital
arte e ciência
arte eletrônica
arte moderna
arte sacra
artes gráficas
artes marciais
artes plásticas
artes visuais
artesanato
artilharia
árvore de natal
árvores
asilo político
assaltos
assassinatos
assédio moral
assédio sexual
assistência social
astrobiologia
astrologia
astronautas
astronomia
atacado
atacado e varejo
ataque de abelhas
ataque de cão
ataque de onça
ataque de tigre
ataque de tubarão
ataque do coração
ataque suicida
ataques
atentado terrorista
atentados
atividade econômica
atividade física
atletismo
ato público
atropelamento
atualidades
auditoria
autismo
autoajuda
autoescolas
automedicação
automobilismo
automóveis
autores
auxílio-doença
avalanche
avaliação
AVC
aventuras
aves
aves silvestres
aviação
axé

ANEXO F - Trecho do código, que trata dos indivíduos do tipo *Ação Noticiosa*, resultante da prototipação da ontologia de notícias

```
@prefix : <http://semantica.estacazero.com/base/> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix owl2xml: <http://www.w3.org/2006/12/owl2-xml#> .
@prefix base: <http://semantica.estacazero.com/base/> .
@prefix owl: <http://www.w3.org/2002/07/owl#> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@base <http://semantica.estacazero.com/base/> .

###
http://semantica.estacazero.com/base/AcaoNoticiosa_AfastamentoDeAntonioOsorioBatistaEFernandoG
odoyDosCorreios
:AcaoNoticiosa_AfastamentoDeAntonioOsorioBatistaEFernandoGodoyDosCorreios rdf:type
base:AcaoNoticiosa ;
rdfs:label "Afastamento de Antônio Osório Batista e Fernando Godoy dos Correios" ;
base:nome "Afastamento de Antônio Osório Batista e Fernando Godoy dos Correios" ;
base:instanteUnico "2005-05-16T00:00:00Z"^^xsd:dateTime ;
base:causadaPor :Intencao_RetirarPessoalDosCorreiosSuspeitoDeCorrupcao ;
base:provocadoPor :AcaoNoticiosa_VejaDenunciaCorrupcaoNosCorreios ;
base:compostoDe
:Participacao_AfastamentoDeAntonioOsorioBatistaEFernandoGodoyDosCorreios_envolve_Pessoa_Antoni
oOsorioBatista ;
base:compostoDe
:Participacao_AfastamentoDeAntonioOsorioBatistaEFernandoGodoyDosCorreios_envolve_Pessoa_Fernan
doGodoy ;
base:compostoDe
:Participacao_AfastamentoDeAntonioOsorioBatistaEFernandoGodoyDosCorreios_envolve_EntidadeEmpre
sarial_Correios ;
base:instanciaDe
:UniversalDeAcaoNoticiosa_UniversalDeAcaoNoticiosaDeAfastamentoPorDenunciaDeCorrupcao .

###
http://semantica.estacazero.com/base/AcaoNoticiosa_ArthurWascheckPrestaDepoimentoACPIDosCorrei
os
:AcaoNoticiosa_ArthurWascheckPrestaDepoimentoACPIDosCorreios rdf:type base:AcaoNoticiosa ;
rdfs:label "Arthur Wascheck presta depoimento à CPI dos Correios" ;
base:nome "Arthur Wascheck presta depoimento à CPI dos Correios" ;
base:instanteDeInicio "2005-06-23T14:30:00Z"^^xsd:dateTime ;
base:instanteDeTermino "2005-06-23T16:00:00Z"^^xsd:dateTime ;
base:causadaPor :Intencao_EsclarecerQueNaoPlanejouEntregarVideoARevistaVeja ;
base:provocadoPor :AcaoNoticiosa_VejaDenunciaCorrupcaoNosCorreios ;
base:compostoDe
:Participacao_ArthurWascheckPrestaDepoimentoACPIDosCorreios_envolve_Pessoa_ArthurWascheck ;
base:compostoDe
:Participacao_ArthurWascheckPrestaDepoimentoACPIDosCorreios_envolve_ObraCriativaSocial_VideoDe
RecebimentoDePropinaNosCorreios ;
base:compostoDe
:Participacao_ArthurWascheckPrestaDepoimentoACPIDosCorreios_envolve_ComissaoParlamentarDeInque
rito_CPIDosCorreios ;
base:instanciaDe :UniversalDeAcaoNoticiosa_UniversalDeAcaoNoticiosaDeDepoimento .

###
http://semantica.estacazero.com/base/AcaoNoticiosa_CameraEscondidaFlagraCorrupcaoNosCorreios
:AcaoNoticiosa_CameraEscondidaFlagraCorrupcaoNosCorreios rdf:type base:AcaoNoticiosa ;
rdfs:label "Câmera escondida flagra corrupção nos Correios" ;
base:nome "Câmera escondida flagra corrupção nos Correios" ;
base:instanteDeInicio "2005-04-18T15:00:00Z"^^xsd:dateTime ;
base:instanteDeTermino "2005-04-18T16:54:00Z"^^xsd:dateTime ;
base:causadaPor :Intencao_ExporEsquemaDeCorrupcaoDoPTBNosCorreios ;
base:compostoDe :AcaoNoticiosa_MauricioMarinhoExplicaEsquemaDeCorrupcaoNosCorreios ;
base:compostoDe :AcaoNoticiosa_MauricioMarinhoRecebePropinaNosCorreios ;
base:instanciaDe :UniversalDeAcaoNoticiosa_UniversalDeAcaoNoticiosaDeFlagranteDeCorrupcao .
```