



Universidade de Brasília

Instituto de Ciências Exatas
Departamento de Ciência da Computação

Aplicação de abordagens de Web Semântica,
Engenharia Web e Engenharia Semiótica a Dados
Abertos

Lucas Shimabuko Silva Rocha
Pedro Vítor Falcetti Sozin

Monografia apresentada como requisito parcial
para conclusão do Bacharelado em Ciência da Computação

Orientadora
Prof.^a Dr.^a Fernanda Lima

Brasília
2014

Universidade de Brasília — UnB
Instituto de Ciências Exatas
Departamento de Ciência da Computação
Bacharelado em Ciência da Computação

Coordenador: Prof. Dr. Homero Luiz Piccolo

Banca examinadora composta por:

Prof.^a Dr.^a Fernanda Lima (Orientadora) — CIC/UnB
Prof. Dr. Guilherme Novaes Ramos — CIC/UnB
Prof. Dr. Rodrigo Bonifacio de Almeida — CIC/UnB

CIP — Catalogação Internacional na Publicação

Rocha, Lucas Shimabuko Silva.

Aplicação de abordagens de Web Semântica, Engenharia Web e Engenharia Semiótica a Dados Abertos / Lucas Shimabuko Silva Rocha, Pedro Vítor Falcetti Sozin. Brasília : UnB, 2014.

175 p. : il. ; 29,5 cm.

Monografia (Graduação) — Universidade de Brasília, Brasília, 2014.

1. Dados Abertos, 2. Web Semântica, 3. Engenharia Semiótica,
4. Engenharia Web, 5. Método de Avaliação de Comunicabilidade

CDU 004.4

Endereço: Universidade de Brasília
Campus Universitário Darcy Ribeiro — Asa Norte
CEP 70910-900
Brasília-DF — Brasil

Resumo

A Web caminha para uma nova realidade onde dados não serão apenas disponibilizados por máquinas, incumbindo ao ser humano o processo de entendimento e reconhecimento. Atualmente o maior custo de trabalho para filtrar e adquirir informações é de responsabilidade daquele que utiliza a máquina e não da máquina em si. A Web Semântica proposta em 2001 por Tim Berners-Lee et al. busca mudar esse panorama. Valendo-se de vocabulários, dados interligados e inferências, essa abordagem muda a análise antes apenas sintática do computador para uma realidade semântica.

O Brasil se encontra em um momento de abertura de seus dados, o governo tem investido e se reunido com especialistas para tentar seguir uma tendência mundial de transparência e colaboração da população. A iniciativa já conta com um portal de dados públicos e com diversos Hackathons, encontros estimulados pelo governo, para o desenvolvimento de novas aplicações que utilizem esses dados disponibilizados. No entanto, o formato atual dos dados disponibilizados não facilita a utilização e interpretação dos mesmos por máquina, principalmente devido a falta de semântica nos dados.

O foco deste trabalho é a utilização de uma abordagem para o desenvolvimento de uma aplicação que reúne áreas comumente utilizadas de forma separada: os Dados Abertos, a Engenharia Web, a Interação Humano-Computador (IHC) e a Web Semântica.

A principal contribuição deste trabalho foi a produção de uma documentação sobre a proposta de aplicação. A utilização de um método de IHC em conjunto com tecnologias semânticas e a utilização de dados abertos é uma abordagem nova para o momento do Brasil, em que os dados abertos têm sido largamente discutidos.

Palavras-chave: Dados Abertos, Web Semântica, Engenharia Semiótica, Engenharia Web, Método de Avaliação de Comunicabilidade

Abstract

The Web moves towards a new reality where the data will no longer be just supplied by machines while the recognition and the understanding by men. Nowadays the higher amount of work to filter and get information is done by men and not by the machine used. The Semantic Web proposed by Tim Berners-Lee et al. seeks to change this reality. Using vocabularies, ontologies, linked data and reasoning, this approach changes the syntactic analysis to a semantic one.

Brazil finds itself in a phase that values transparency. The government has spent lots of money and gathered specialists to discuss and follow a worldwide tendency of Open Data and popular collaboration. This initiative already has an open government data portal, and many Hackathons, meetings encouraged by the government itself, for developing new applications based on this open data. However, the format of the available data makes it difficult to be reusable by machine, mostly because of the inexistence of tripled data.

The objective of this study was to document the model of an application development, which gather areas that are usually apart: Human-Computer Interaction (HCI), Semantic Web, Open Data, and Web Engineering.

The main contribution of this study was the production of a documentation about the proposed application. In addition, the use of a HCI method with semantic technologies and the use of open data is a new approach to Brazil's moment, wherein the open data is being widely discussed.

Keywords: Open Data, Semantic Web, Semiotic Engineering, Web Engineering, Communicability Evaluation Method

Sumário

1	Introdução	1
1.1	Contextualização	1
1.2	Problema de Pesquisa Endereçado neste Trabalho	3
1.3	Objetivo	4
1.4	Organização da Monografia	4
2	Referencial Teórico	6
2.1	Conceitos Básicos de Dados Abertos	6
2.1.1	Governo Aberto	6
2.1.2	Dados Abertos Governamentais	8
2.2	Conceitos Básicos de Web Semântica	13
2.2.1	Dados Ligados	13
2.2.2	RDF	14
2.2.3	Ontologias	15
2.2.4	SPARQL	16
2.3	Conceitos Básicos de Engenharia Web	18
2.3.1	OOHDM	18
2.3.2	SHDM	19
2.4	Conceitos Básicos de Interação Humano-Computador (IHC)	20
2.4.1	Interação Humano Computador	20
2.4.2	Projeto de IHC - DECIDE	24
2.4.3	Avaliações em IHC: Método de Avaliação de Comunicabilidade e outros	25
2.5	Conceitos Básicos sobre o Domínio: Corrupção	31
2.5.1	A Relação Entre Transparência e Corrupção	32
2.6	Aplicações de Dados Abertos	36
2.6.1	Aplicações Internacionais	36
2.6.2	Aplicações Brasileiras	39
3	Proposta de Trabalho	44
3.1	Utilização de um Conjunto dos Dados Abertos Governamentais	44
3.2	Modelagem da Aplicação	45
3.2.1	1ª Etapa do OOHDM: Levantamento de requisitos	45
3.2.2	2ª Etapa do OOHDM: Modelagem Conceitual	47
3.2.3	3ª Etapa do OOHDM: Modelagem Navegacional	48
3.3	Protótipos de Baixa Fidelidade para a Avaliação	49

3.4	Utilização de Método para a Avaliação da Aplicação	52
3.5	Conversão dos Dados para o Formato RDF	53
4	Implementação da Proposta	54
4.1	Aplicação do Método de Avaliação de Comunicabilidade (MAC)	54
4.2	Ambiente Tecnológico Utilizado	58
4.2.1	Google Refine	58
4.2.2	Virtuoso	60
4.2.3	Endpoint Olho nos Políticos	60
5	Considerações Finais	63
5.1	Resultados deste trabalho	64
5.2	Limitações	64
5.3	Trabalhos Futuros	64
A	Roteiro da Avaliação	66
B	Termo de Consentimento	67
C	Questionário Demográfico	68
D	Informações ao Participante	71
E	Tarefas	73
F	Questionário Pós-Teste	74
	Referências	76

Lista de Figuras

1.1	Formato dos dados disponibilizados 31/07/2013	3
1.2	Formato dos dados disponibilizados 17/06/2014	3
2.1	Relação entre Competição e Corrupção. (Fonte: <i>Workshop Corruption - How, and Why, To Avoid It</i> [21])	7
2.2	Visão de uma informação política. (Fonte: Bellany e Taylor[5])	8
2.3	Estrutura da Web Semântica. (Fonte Tim Berners Lee et al. [7])	13
2.4	Estrutura da Tripla RDF	14
2.5	Exemplo de Perfil FOAF. (Fonte: Wikipedia [70])	15
2.6	Nuvem de Relacionamento DBpedia. (Fonte: DBpedia [13])	16
2.7	Estrutura de SPARQL. (Fonte: W3C [68])	17
2.8	Objetivos de estudo em IHC (Fonte: Barbosa e Silva [4])	22
2.9	Interação entra usuário e sistema computacional (Fonte: Prates e Barbosa[4])	23
2.10	Protótipos de baixa fidelidade desenvolvido em papel e alta fidelidade desenvolvido com a ferramenta Justinmind [38].	26
2.11	Aspectos avaliados por meio de cada método. (Fonte: Barbosa e Silva [4])	26
2.12	Aspectos avaliados por meio de cada método. (Fonte: Barbosa e Silva [4])	27
2.13	Frequência com que cada método de avaliação é utilizado nos tipos de avaliação. (Fonte: Barbosa e Silva [4])	27
2.14	Processo de comunicação	28
2.15	Mapa de Percepção de Corrupção - Transparency International [36]	32
2.16	Tabela de Percepção da Corrupção - Países Melhores Colocados - Transparency International [36]	33
2.17	Tabela de Percepção da Corrupção - Brasil - Transparency International [36]	33
2.18	Ranking de Transparência dos Países - Open Data Index [26]	34
2.19	Ranking de Transparência Brasil - Open Data Index [26]	35
2.20	Estrutura da Data-Gov Wiki. (Fonte: Data-Gov Wiki [34])	37
2.21	Visualização do <i>Where does my money go</i> . (Fonte: <i>Where Where does my money go does</i> [27])	38
2.22	Visualização de um problema reportado na aplicação <i>Fix my street</i> . (Fonte: <i>Fix my street</i> [46])	39
2.23	Visualização das solicitações de dados no Queremos Saber. (Fonte: Queremos Saber [32])	40
2.24	Tela da Aplicação Para Onde Foi o Meu Dinheiro. (Fonte: Para Onde Foi o Meu Dinheiro [49])	40
2.25	Visualização de Dados do PAC. (Fonte: Obras do PAC [37])	41

2.26	Visualização de Dados da Empresa Claro em 2011. (Fonte: Reclamações Procon [17])	41
2.27	Dados Referentes a Aeronaves em 2011. (Fonte: Aeroportos Brasil [63]) . .	42
2.28	Semelhança entre partidos na Câmara dos Deputados. (Fonte: Radar Parlamentar [51])	43
2.29	Basômetro, apoio dos parlamentares ao governo. (Fonte: Basômetro [20]) .	43
3.1	Tarefa 1	45
3.2	Tarefa 2	46
3.3	Tarefa 3	46
3.4	Diagrama de Classes Conceituais	47
3.5	Diagrama de Classes Navegacionais	48
3.6	Diagrama de Classes Contextuais	48
3.7	Protótipo de Papel - Tela Inicial	49
3.8	Protótipo de Papel - Perfil do Político	50
3.9	Protótipo Digital - Tela Inicial	51
3.10	Protótipo Digital - Licitações de uma Pessoa Jurídica	51
3.11	Protótipo Digital - Perfil do Político	52
4.1	Resultado do MAC	55
4.2	Nova Tela - FAQ	56
4.3	Nova Tela - Tela Inicial	57
4.4	Nova Tela - Partido	57
4.5	Nova Tela - Receitas	58
4.6	Pessoa Jurídica Repetida	59
4.7	Dados Ligados Google Refine	59
4.8	Alinhamento de RDF	60
4.9	Arquitetura do Virtuoso. (Fonte: Virtuoso [61])	61
4.10	Consulta no <i>Endpoint</i>	62
4.11	Resultado da Consulta no <i>Endpoint</i>	62

Lista de Tabelas

2.1	Formato de Dados	10
2.2	Artefatos do OOHDM	19
3.1	Artefatos utilizados para desenvolver o protótipo	45
3.2	Propriedades do RDF projetado	53
4.1	Perfil dos Usuários da Avaliação	54

Capítulo 1

Introdução

Este capítulo discorre brevemente a respeito da situação atual de Dados Abertos no Brasil, explicita uma lacuna de pesquisa identificada e apresenta uma proposta de utilização de abordagens originárias de diferentes áreas para endereçar o problema explicitado.

1.1 Contextualização

Um dos diversos problemas que a sociedade brasileira enfrenta todos os dias é o da corrupção. Com frequência são noticiadas informações sobre os políticos e empresas participando de esquemas fraudulentos com o intuito de obtenção de lucro fácil e ilícito.

O conjunto de dados abertos governamentais se apresenta como uma boa oportunidade de controle social, onde qualquer cidadão pode utilizar os dados disponibilizados para monitorar a atividade governamental. Dentro dos conjuntos de dados disponibilizados pelo governo brasileiro temos, por exemplo, movimentações financeiras de políticos e partidos durante as campanhas eleitorais, informações a respeito de empresas que foram contratadas pelo governo. Todavia, o acesso a essas informações, que possibilitariam ao cidadão a verificação da idoneidade dos participantes, é de difícil acesso para toda a população.

A disponibilização de dados, sejam eles governamentais ou privados, à população, através do auxílio da Web é um fenômeno que atinge os mais diversos países do cenário mundial. Portais como o Portal de Dados Abertos Brasileiros, no Brasil, e o Data.Gov.Uk, no Reino Unido, são exemplos de plataformas para catalogação, organização e distribuição dos dados para a população. Os dados são disponibilizados com o objetivo de prestar contas à população e principalmente fornecer informação para que interessados possam utilizar esses dados para o desenvolvimento de aplicações. Os dados do portal nacional, todavia, estão disponibilizados de tal forma que uma pessoa leiga, isto é, sem noções de computação e ou do domínio específico sobre o qual aquele dado se refere, terá dificuldades de extrair alguma informação. Dados apresentados no formato PDF são muito mais difíceis de serem tratados, reutilizados e armazenados em um banco de dados do que dados no formato CSV. Atualmente, não existem no Portal Brasileiro de Dados Abertos [22] e no Portal da Transparência [23] dados com significado semântico e com metadados unificados aos dados, estes dados foram primeiramente disponibilizados em formatos mais simples e mais práticos para os fornecedores e agora estão começando a ser apresentados de forma mais adequada para os consumidores dos dados. O RDF [67], especificado pela

principal organização de padronização da Web, o W3C, tem como objetivo padronizar a disponibilização da informação na Web através de triplas, uma estrutura de sujeito -> predicado -> objeto, e assim permitir a associação de um objeto ou recurso com outras bases de dados, desde que sejam utilizados de um vocabulário ou uma ontologia, uma descrição de conceitos e relacionamentos, já conhecidos.

Apesar dos incentivos para a produção de aplicações que utilizem dados abertos, não existe um padrão ou mesmo um guia que auxilie os interessados a desenvolver tais aplicações. Portanto, para permitir que a informação dos dados abertos seja melhor utilizada pelo projetista, serão aplicadas abordagens de Engenharia Web, Web Semântica e Interação Humano-Computador, mais especificamente, a Engenharia Semiótica, aos dados governamentais.

A Engenharia Web e a área de Interação Humano-Computador contribuem para a criação de sistemas de informação que sejam de mais fácil utilização, além de permitir navegações mais inteligentes e baseadas em contexto. A Web Semântica permite incluir significado aos dados, tanto quando são somente disponibilizados quanto quando são utilizados em uma aplicação.

A Engenharia Web busca utilizar métodos para construir aplicações, como o *Object-Oriented Hypermedia Design Method* (OOHDM)[57]. O OOHDM permite projetar uma aplicação seguindo etapas bem determinadas que otimizam as formas de navegação que estarão disponíveis para o usuário. Este tema também tem ganhado importância no ramo da computação, pois cada vez mais se faz necessário o desenvolvimento de aplicações e sistemas Web mais complexos e que permitam ao usuário navegar em busca de informações de forma mais direta.

A Interação Humano-Computador (IHC) aborda a maneira com a qual o usuário interage e percebe um sistema de computador [4]. Os métodos de avaliação de IHC auxiliam os desenvolvedores na criação de interfaces mais intuitivas e agradáveis ao uso. O foco nas avaliações de sistemas computacionais, desde o seu projeto até sua fase final, é algo relativamente novo e que vem ganhando espaço principalmente pela inclusão digital.

A combinação de abordagens da Engenharia Web junto com a IHC, permite que o projetista modele uma aplicação seguindo métodos pré-determinados e desenvolva um protótipo para avaliar a comunicabilidade do que foi proposto, para conferir se a mensagem é transmitida adequadamente ao usuário.

A Web Semântica, introduzida por Tim Berners Lee et al [7], contribui para a ideia de que a informação apresentada ao usuário pode possuir significado e além disso, ser relacionada com informações semelhantes em toda a Web, propiciando que as máquinas possam realizar tarefas para o usuário, ao invés de que este as faça manualmente. Através de ontologias e vocabulários, bases de dados de diferentes organizações que tratam de um mesmo domínio podem ser acessadas e tratadas de maneira semelhante.

O crescimento vertiginoso da Web fez com que a informação fosse criada e disponibilizada rapidamente, sem, no entanto, seguir padrões estabelecidos. O que se busca na Web atualmente é a utilização de padrões, ontologias e metadados para conectar a informação permitindo a integração de seu conteúdo, diminuindo o trabalho humano no consumo de informação e permitindo a criação e desenvolvimento de funcionalidades superiores no apoio da navegação humana.

O objetivo deste trabalho é utilizar dos conhecimentos dessas abordagens para auxiliar o projetista de uma aplicação de dados abertos.

1.2 Problema de Pesquisa Endereçado neste Trabalho

Atualmente, os dados que o governo tem a obrigação de disponibilizar para a população através da Lei de Acesso à Informação [11] não se encontram em um formato que possa ser facilmente trabalhado por máquinas. As Figuras 1.1 e 1.2 retratam os panoramas dos formatos nos quais os dados estão disponibilizados no site www.dados.gov.br nos dias 31 de julho de 2013, no momento do início do trabalho, e dia 17 de junho de 2014.

http://dados.gov.br/dataset Go

31 captures
10 Mai 12 - 7 Dez 13

JUN JUL 31 SET
2012 2013 2014

Pesquisa - Portal Brasileiro de Dados Abertos

Q Pesquisa ...
Pesquisar

Comissões e Participações de Deputados em Comissões da ALMG
html xml json xml json xml json xml json xml json xml json xml
json xml json xml xml json xml json xml json xml json xml json
xml json xml json xml json xml json xml json xml json xml json
Comissões e participações de Deputados em comissões, consultáveis pelos webservices de dados abertos da Assembleia Legislativa de Minas Gerais (ALMG).
Não Abertamente Licenciado

Anuário Estatístico de Acidentes de Trabalho - AEAT
csv json xml
csv json xml csv json xml csv json xml
Anuário Estatístico de Acidentes de Trabalho - AEAT, documento público, com edição e publicação anuais em meio papel e versão digital acessível por meio do site do Ministério da...
OPEN DATA

Atendimentos de Consumidores nos Procons (Sindec)
pdf pdf pdf zip+csv zip+csv zip+csv zip+csv zip+csv zip+csv zip+csv zip+csv
zip+csv pdf
O Sistema Nacional de Informações de Defesa do Consumidor - Sindec é um sistema informatizado que integra

Tags
INEP (15)
SP (13)
IBGE (13)
avaliação educação (12)
desempenho escolar (11)
censo escolar (10)
federal (9)
educação básica (9)
São Paulo (9)
transparência (6)

Resource formats
xls (20)
pdf (18)
xml (15)
csv (15)
zip+csv (14)
html (12)
zip+txt (10)
zip+sas (8)
zip+xls (6)
json (5)

Figura 1.1: Formato dos dados disponibilizados 31/07/2013

Q Pesquisa ...
Pesquisar

Sistema BR-Brasil - Boletins de Ocorrências em Rodovias Federais
pdf pdf zip+csv zip+csv zip+csv zip+csv zip+csv zip+csv zip+csv zip+csv
zip+csv zip+csv zip+csv zip+csv zip+csv zip+csv zip+csv zip+csv
O Sistema BR-Brasil é um sistema desenvolvido pelo Departamento de Polícia Rodoviária Federal (DPRF) que visa suprir todas as deficiências operacionais em termos de informatização e...
OPEN DATA

Adequação de acesso rodoviário
text/html xml
Adequação de acesso rodoviário. Empreendimentos pertencentes a carteira do Anexo III do PPA 2012-2015. Programa 2075. Iniciativas 00AL e 00AM - escala 1:2.500.000. Sistema de referência:...
Não Abertamente Licenciado

Adequação de Rodovias
text/html xml
Adequação de Rodovias. Empreendimento pertencente a carteira do Anexo III do PPA 2012-2015. Programa 2075. Iniciativas 00AB, 009Z, 00AH, 009X, 00A6, 00A9, 009V, 00A3, 00A0, 00A1, 009W,....
Não Abertamente Licenciado

Agências da Previdência Social - APS
csv odt html
A rede de atendimento ao público da Previdência Social é composta por mais de 1.500 Agências da Previdência Social - APS, entre essas temos as APS Atendimento de Acordos Internacionais,....
OPEN DATA

Tags
INDE (135)
Equipamentos públicos (17)
INEP (15)
IBGE (13)
avaliação educação (12)
desempenho escolar (11)
Vegetação (11)
Mato Grosso (11)
educação básica (10)

Resource formats
text/html (135)
xml (54)
csv (36)
xls (20)
zip+csv (18)
html (18)
pdf (15)
zip+sas (8)
json (6)
zip+xls (5)

Figura 1.2: Formato dos dados disponibilizados 17/06/2014

No Portal, não existe no catálogo algum dado no formato RDF ou outro formato triplificado, o que permitiria interpretação e uso de metadados. O dado atualmente precisa ser minerado e trabalhado antes de efetivamente ser utilizado em alguma aplicação. Apesar disso, pode-se perceber um aumento do conjunto de dados disponibilizados de 87 para 238 e a redução de dados apresentados no formato PDF.

Esse trabalho pretende responder à seguinte pergunta de pesquisa: Como abordagens de Engenharia Web, Web Semântica, Interação Humano Computador e Engenharia Semiótica podem auxiliar o projetista de uma aplicação de dados abertos?

1.3 Objetivo

Esta seção descreve o objetivo geral e os objetivos específicos deste trabalho.

O objetivo geral do trabalho é utilizar abordagens de diversas áreas ligadas à computação para auxiliar o projetista de uma aplicação que envolva dados abertos em sua modelagem e utilização dos dados.

Os objetivos específicos deste trabalho são, divididos por área:

- Dados Abertos: utilizar conjuntos de dados abertos governamentais separados que possa trazer informações relevantes, não triviais, no contexto político.
- Engenharia Web: utilizar o método OOHDM para o desenvolvimento de artefatos que auxiliem a navegação e o desenvolvimento do protótipo da aplicação.
- Web Semântica: introduzir semântica em um conjunto de dados inicialmente sintático e ligá-lo com outras bases de dados, permitindo consultas SPARQL independente do propósito original dos dados.
- Interação Humano-Computador: utilizar o Método de Avaliação de Comunicabilidade (MAC) para avaliar a comunicabilidade do protótipo inicial para desenvolver versões aprimoradas.

Demais objetivos específicos:

- Utilizar abordagens de diferentes áreas da computação para endereçar a questão atual de Dados Abertos.
- Documentar a abordagem utilizada para auxiliar pessoas a desenvolverem trabalhos relacionados a este.

1.4 Organização da Monografia

A monografia está organizada da seguinte forma: No capítulo 2, as bases teóricas nas quais este trabalho se sustenta são apresentadas com o intuito de preparar e contextualizar para que se compreenda o ambiente no qual este trabalho está inserido. Áreas como Dados Abertos, Engenharia Web, Web Semântica e Interação Humano-Computador são apresentadas neste capítulo. No capítulo 3, a proposta para a solução do problema é apresentada, juntamente com as modelagens e protótipos construídos que auxiliaram a implementação da proposta. No capítulo 4, a implementação da proposta é apresentada

ao leitor, explicando as abordagens de implementação utilizadas, ambiente tecnológico, ferramentas e procedimentos utilizados para atingir o objetivo da monografia. No capítulo 5, os resultados são apresentados com as contribuições deste trabalho, as limitações e dificuldades encontradas e os trabalhos futuros.

Capítulo 2

Referencial Teórico

O referencial teórico abordado neste capítulo consiste de conceitos e conhecimentos básicos para auxiliar na compreensão deste trabalho. As quatro áreas (Dados Abertos, Web Semântica, Engenharia Web, Interação Humano-Computador) abordadas com uma proposta de união através do projeto de uma aplicação utilizando-se de dados abertos são descritas a seguir juntamente com uma explicação do domínio dos dados abertos escolhido para o projeto.

2.1 Conceitos Básicos de Dados Abertos

A ideia central de Dados Abertos é que a informação deve estar disponível para que qualquer pessoa a utilize, redistribuída como bem entender, sem qualquer restrição de patentes, propriedade intelectual ou outro mecanismo de controle [2]. No máximo deve-se creditar a autoria e redistribuir sobre a mesma licença daquelas previstas em estatuto.

A importância de uma clara definição do significado de aberto é a interoperabilidade. Interoperabilidade é a capacidade de diversos sistemas e organizações trabalharem juntos, interoperar ou combinar diferentes conjuntos de dados [25].

Com relação aos Dados Abertos, a ideia de interoperabilidade se torna mais evidente. A capacidade de permitir que uma parte “aberta” de um material seja livremente combinada com partes “abertas” de outros materiais é fundamental para que os objetivos práticos sejam alcançados, como a capacidade de combinar diferentes conjuntos de dados e assim desenvolver produtos e serviços melhores para a população. Dados abertos devem ser usados para promover a transparência e a prestação de contas, desta forma reduzindo danos à economia por corrupção. Em um relatório do *Workshop Corruption - How, and Why, To Avoid It*[21], é mostrado que a competitividade dos países é correlacionada com o menor índice de corrupção. Esta relação é apresentada na Figura 2.1.

2.1.1 Governo Aberto

É importante que seja clara a noção do que é Governo Aberto, mais comumente encontrado em inglês: *Open Government*. Basicamente é a noção de que as pessoas têm o direito de acessar documentos e informações do governo *Open Government*. Mas esse significado continua evoluindo, o conceito de Governo Aberto é influenciado pelo movimento de software livre, e alterou seu foco para permitir a participação mais efetiva

COMPETITIVENESS AND CORRUPTION

© V.J. ROTT Zurich

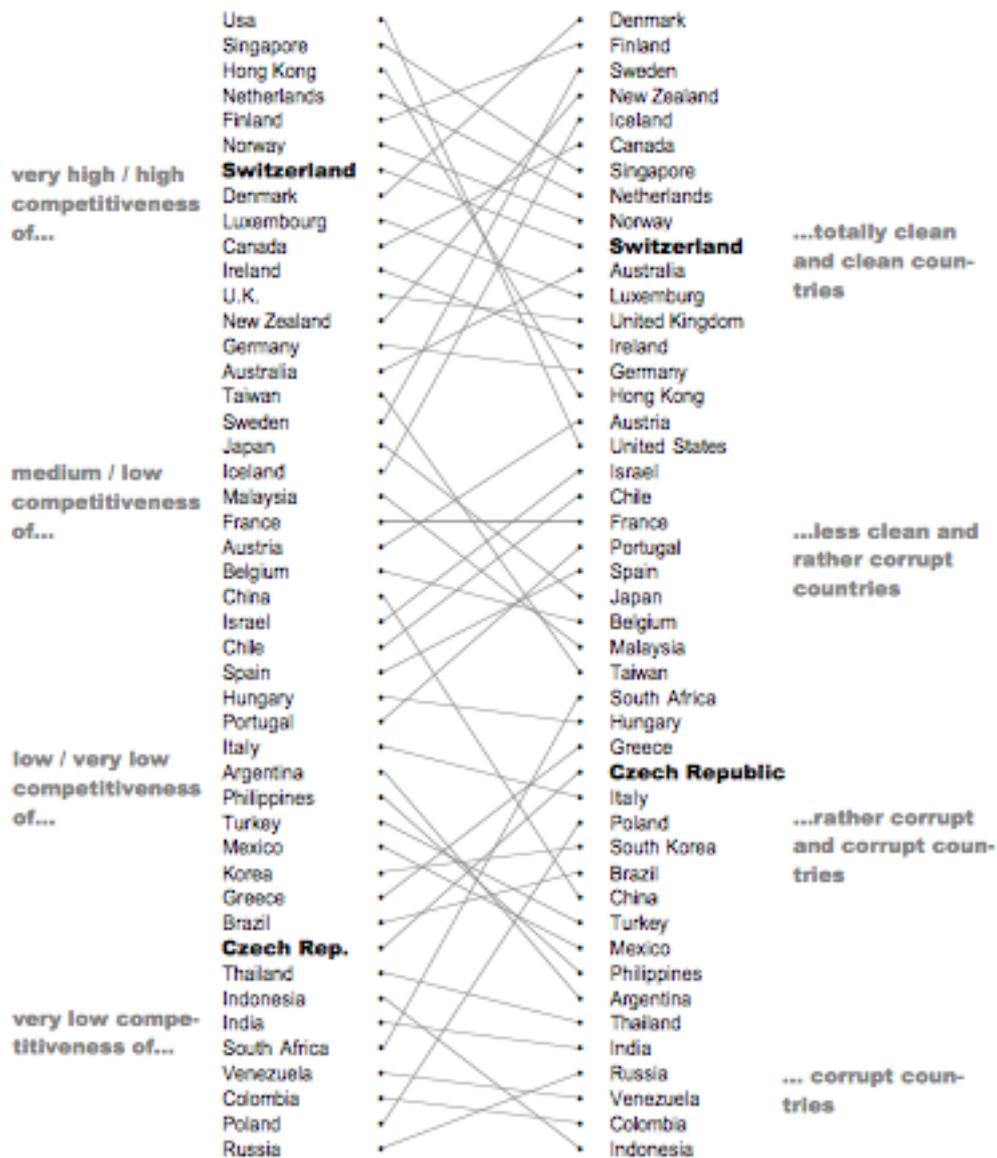


Figura 2.1: Relação entre Competição e Corrupção. (Fonte: *Workshop Corruption - How, and Why, To Avoid It*[21])

no modo de governar. Assim como no software livre se permite que usuários alterem e contribuam com o código, o governo aberto agora significa que os cidadãos não se limitam apenas a acessar a informação, mas participam de um modo mais significativo: gerando aplicações que podem mudar o cenário em que vivem. Governo aberto ainda significa maior comunicação e interoperabilidade entre os ramos do governo.

2.1.2 Dados Abertos Governamentais

O governo tem papel fundamental em Dados Abertos Governamentais, tanto por causa da quantidade e centralidade de informações que coleta, como por sua característica pública. Um dos objetivos de disponibilizar dados governamentais é melhorar o governo, expandindo informação e acesso de uma forma que novos interessados influenciem nas decisões do governo. O conceito de informação política é utilizado como forma de entender o rumo que a mistura de diferentes interesses estão tomando. Uma informação política é um conjunto de públicos estratégicos, fontes de dados, diagrama de informação de dados, e relações envolvidas entre as partes governamentais e as não governamentais [5]. O grande ponto em pensar a iniciativa de dados abertos em termos de informação política é para a identificação de fontes de dados importantes, caminhos e públicos estratégicos. É um ponto importante para entender como a utilização desses dados pode afetar o interesse de vários públicos estratégicos e a criação de valor para os dados e para a informação. A Figura 2.2 ilustra a o conceito de informação política.

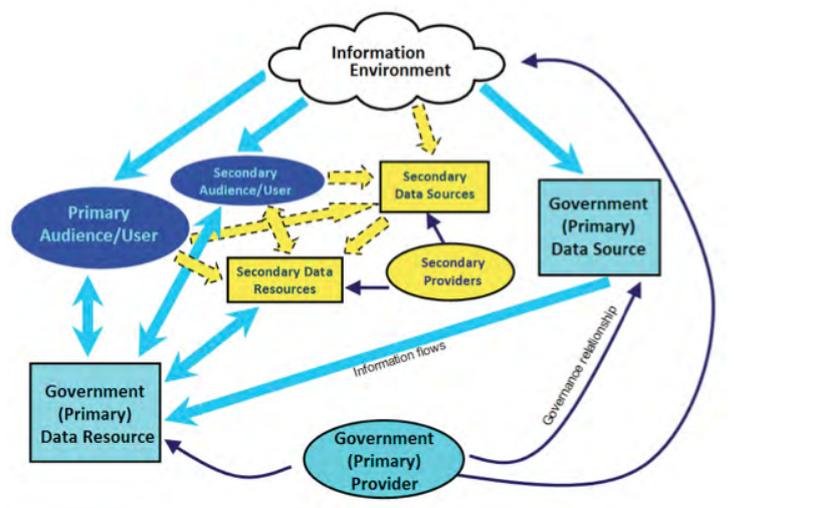


Figura 2.2: Visão de uma informação política. (Fonte: Bellany e Taylor[5])

Dentro dessa realidade de dados abertos, existe um movimento em torno de dados abertos governamentais que permite a qualquer cidadão o acesso aos dados das atividades governamentais em formato aberto para o uso e redistribuição. De acordo com David Eaves [19] existem três leis para os dados abertos governamentais:

- Se o dado não pode ser encontrado ou indexado, ele não existe.
- Se o dado não estiver em formato aberto e legível por máquina, ele não pode ser aproveitado.

- Se algum dispositivo legal impedir a sua replicação, ele não é útil.

Os dados só poderão ser considerados abertos se estiverem disponibilizados de acordo com os 8 princípios de Dados Abertos Governamentais, que foram definidos por um grupo de trabalho do *Open Government Data* [12]:

1. Dados devem estar completos:

- Todos os dados públicos devem estar disponíveis. Os dados são informações armazenadas ou gravadas, incluindo, mas não limitada a documentos, base de dados, transcrições, e gravações de áudio/vídeo. Dados públicos são dados que não estão sujeitos a limitações válidas de privacidade, segurança ou controle de acesso, reguladas por estatutos. A forma mais rápida de recorrer a alguma informação é através da rede mundial de computadores. Caso a informação desejada não esteja presente nela e a sua obtenção só se torne possível através de requisições pessoais, ela se torna indisponível para a maioria dos cidadãos. Qualquer esforço ou custo para acessar alguma informação prejudica enormemente a sua disponibilização.

2. Dados precisam ser primários:

- Os dados são publicados como foram coletados, com o maior nível possível de granularidade, de forma não agregada ou modificada. Este princípio reflete uma mudança na ideia de disponibilizar as informações diretamente para os cidadãos que a desejam para a ideia de disponibilizá-la para mediadores. Os mediadores seriam capazes de construir aplicações com suas próprias ideias e mostrar informações diferentes das que são obtidas na fonte. Através disso, ao invés de informações pré-analisadas e simplificadas pelo governo, é possível a descoberta de soluções inovadoras com os mais variados tipos de dados abertos disponibilizados pelo governo.

3. Dados precisam ser atuais:

- Os dados devem ser disponibilizados o mais rápido possível para preservar o seu valor. O dado não pode mais ser considerado aberto caso a sua disponibilização se dê em um momento em que este não é mais útil para o público. A definição de atual para um dado não se dá somente pelo momento em que é disponibilizado, mas também sobre o quão rápido os usuários desses dados conseguem atualizações sobre eles.

4. Dados precisam ser acessíveis:

- Os dados devem estar disponíveis para a maior quantidade de usuários e para os propósitos mais diversos. O princípio da acessibilidade diz respeito tanto sobre as necessidades do usuário do dado ser capaz de localizá-lo, interpretá-lo e compreendê-lo quanto a de uma aplicação ser capaz de adquiri-lo ou de decodificá-lo.

5. Dados devem ser processáveis por máquina:

- Os dados são razoavelmente estruturados para permitir o seu processamento. A partir do momento em que o volume dos dados começa a aumentar, as mais inovadoras e interessantes aplicações no meio de dados abertos governamentais necessitam de um computador melhor para pesquisar, classificar ou transformar esses dados. A Tabela 2.1 ilustra as recomendações para a disponibilização de cada tipo de dado.

Media	Formato de Dado Recomendado
Planilhas	CSV(UTF-8) ou ODS, ou XML
Documentos	XHTML/XML (para estrutura) PDF/A-1b (para paginação)
Imagens	JPEG, SVG, or PNG
Audio	descompactado WAV ou Ogg Vorbis
Video	descompactado AVI ou Ogg Theora

6. Dados devem ser não discriminatórios:

- Os dados são disponibilizados para qualquer um sem a necessidade de registro. O acesso anônimo ao dado deve ser permitido.

7. Dados devem estar num formato não proprietário:

- Os dados são disponibilizados em um formato que não é controlado exclusivamente por uma entidade. Formatos de dado proprietários adicionam restrições desnecessárias sobre quem poderá usar o dado e como essa pessoa irá fazer isso.

8. Dados devem ser livres de licenças:

- Os dados não estão sujeitos a nenhum direito autoral, patente, marca registrada, ou segredo industrial. Restrições razoáveis de privacidade, segurança e controle de acesso podem ser permitidas na forma regulada por estatutos.

Além dos oito princípios, Tauberer [64] cita mais recomendações sobre a disponibilização dos dados abertos.

1. Dados devem ser permanentes:

- Os dados devem estar disponíveis em local estável por tempo indeterminado. Prover documentos em locais fixos ajuda o público a compartilhá-los através de direcionamentos diretos para a sua fonte ao invés de disponibilizar instruções sobre como encontrá-los e como distribuí-los.

2. Dados devem promover a análise:

- Os dados disponibilizados pelo governo devem estar em formatos que facilitem a sua análise e conseqüentemente o seu reuso. Este princípio reforça a ideia de que o maior valor e o melhor aproveitamento de um dado vêm da análise pública do mesmo ao invés de uma análise prévia realizada pelo governo.
3. Dados devem ser disponibilizados em formatos seguros:
 - O governo deve disponibilizar seus dados para o público em formatos que não possuam conteúdo executável. O conteúdo executável cria um risco de segurança para o usuário já que o mesmo pode conter um *malware*. As macros utilizadas no *Microsoft Office* são um exemplo de dados com conteúdo executável.
 4. Os dados devem possuir precedência e serem confiáveis:
 - Os dados publicados devem possuir assinatura digital ou incluir informações de data de criação e publicação, autenticidade e integridade. A assinatura digital ajuda os usuários na validação da fonte do dado requerido, pois através dela é possível ter a certeza de que o mesmo não foi modificado depois que foi publicado.
 5. Participação popular:
 - O público ou os que necessitam da informação estão na melhor posição para determinar qual informação é necessária e importante para o desenvolvimento de suas aplicações.
 6. Transparência pública:
 - Este princípio está relacionado ao fato de quem não somente os dados devem ser abertos, mas também que o processo para a criação do dado também deve ser transparente.
 7. Interoperabilidade:
 - Este princípio reflete o poder da interoperabilidade sobre um dado, tornando-o mais valioso e possibilitando uma mais fácil combinação dos dados. Caso dois conjunto de dados se refiram a um mesmo objeto, seus criadores deveriam torná-los interoperáveis, desenvolvendo padrões para o compartilhamento de dados ou adotando um já existente possivelmente em conjunto com o próprio governo. A busca e o esforço para atingir a interoperabilidade são contrastantes com a necessidade de disponibilizar os dados rapidamente para o público, sendo assim, a rapidez e a agilidade são tidos como mais importantes.
 8. Escolhas tecnológicas podem causar apoio unilateral:
 - O uso de tecnologias criadas por setores privados pode gerar conflitos de interesses na regulação desse setor, e criar um incentivo para as corporações apoiadoras se envolverem na criação de políticas para isso. Escolhas tecnológicas devem ser evitadas, pois essencialmente apoiam uma única entidade.

9. Priorização:

- Apesar de todos os princípios de dados abertos governamentais serem escritos como definições, eles não podem ser tratados como valores binários. O governo tem uma quantidade limitada de recursos para o preparo dos dados para a população, e seria impossível para todos os dados abertos governamentais atingirem todos esses princípios instantaneamente e simultaneamente. Catalogar, documentar e preparar uma infraestrutura para a disponibilização de atualizações e uma possível revisão disso requer um esforço significativo e conseqüentemente uma priorização. A velocidade para a disponibilização de um dado pode contrastar com o desejo da precisão do dado. Atualmente com a limitação de recursos financeiros e tecnológicos, um planejamento para atingir todos os princípios de forma incremental se faz necessário.

Em 27 de maio de 2009, o Presidente em exercício Luiz Inácio Lula da Silva, sancionou a Lei da Transparência [10], que obriga a divulgação dos gastos previstos e realizados nos orçamentos da União, Estados, Distrito Federal e Municípios. Esse marco foi importante para o início da divulgação dos dados governamentais e abriu caminho para a Lei de Acesso a Informação (LAI) [11].

O Brasil aderiu ao movimento de dados abertos governamentais e em 18 de novembro de 2011, a Presidenta Dilma Vana Rousseff sancionou a Lei no 12.527 [11], que tem o propósito de regulamentar o direito de acesso dos cidadãos às informações públicas e seus dispositivos são aplicáveis aos três Poderes da União, Estados, Distrito Federal e Municípios. Esta lei garante ao cidadão uma das ferramentas mais poderosas no combate a corrupção: a informação. Um governo aberto é um pilar fundamental para o progresso econômico, político e social.

Ricardo Matheus, em seu artigo *New Perspectives for Electronic Government in Brazil: The Adoption of Open Government Data in National and Subnational Governments of Brazil* [44], aborda o estado da arte referente a dados abertos no Brasil no período inicial do ano de 2012 e ressalta a importância do tema para a transparência, controle social do governo, e participação cidadã através da Internet no país. Além dos oito princípios de dados abertos supracitados, Ricardo Matheus utiliza uma abordagem baseada no método *Five Stars of Open Linked Data* de Tim Berners-Lee et al. [6]. Segundo o método, a classificação pode se dar da seguinte forma:

- Uma estrela: disponível na internet (em qualquer formato) e com licença aberta.
- Duas estrelas: disponível em uma estrutura de dados interpretável por máquina.
- Três estrelas: assim como duas estrelas e em um formato não proprietário.
- Quatro estrelas: assim como três estrelas e seguindo padrões do W3C para identificação.
- Cinco estrelas: assim como quatro estrelas e conectando seus dados com o de outras pessoas para prover um contexto.

Portanto, não basta que as informações do governo estejam disponibilizadas para o cidadão, é necessário que ela esteja interligada pelos princípios de Dados Ligados, para assim gerar aplicações mais eficientes e poderosas.

2.2 Conceitos Básicos de Web Semântica

Para compreender a Web Semântica, é preciso primeiro entender o significado de semântica: "em linguística, semântica estuda o significado e a interpretação do significado de uma palavra, de um signo, de uma frase ou de uma expressão em um determinado contexto. Nesse campo de estudo se analisa, também, as mudanças de sentido que ocorrem nas formas linguísticas devido a alguns fatores, tais como tempo e espaço geográfico"[66]. Então, a semântica se refere ao significado das palavras e do contexto no qual elas estão inseridas.

Tim Berners Lee et al. introduziu o conceito de Web Semântica em seu artigo *The Semantic Web - A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities* [7]. A grande ideia por trás da Web Semântica é que as máquinas possam entender o que estão exibindo e assim apresentar melhores resultados para seres humanos. A Figura 2.3 ilustra a estrutura da Web Semântica proposta.

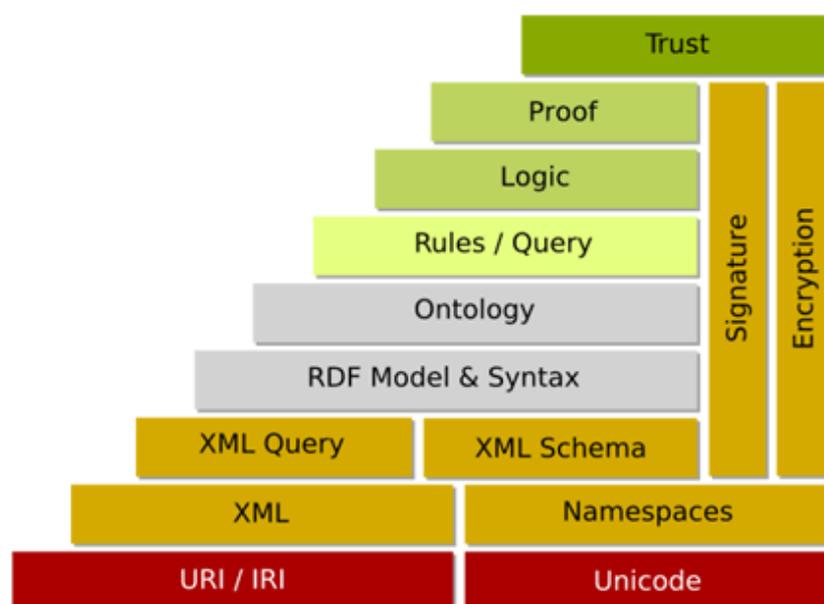


Figura 2.3: Estrutura da Web Semântica. (Fonte Tim Berners Lee et al. [7])

2.2.1 Dados Ligados

A ideia básica do Linked Data é aplicar a estrutura presente na *World Wide Web* no compartilhamento de informação. Um documento Web é uma união de padrões:

- *Uniform Resource Identifiers* (URIs) como um identificar único e global.
- O *Hypertext Transfer Protocol* (HTTP) como meio de acesso universal.
- O *Hypertext Markup Language* (HTML) como formato de conteúdo.

Proposto em 2006 pelo W3C, o termo Dados Ligados se refere a um conjunto de práticas para publicar dados estruturados na Web e assim estruturar esses dados de forma

interligada para que eles se tornem mais úteis. Tim Berners-Lee et al. apresentou essas práticas e elas se tornaram os princípios de Dados Ligados [33], são elas:

1. Usar *Uniform Resource Identifiers* (URIs) como nomes para coisas.
2. Usar URIs HTTP, para que pessoas possam buscar por esses nomes.
3. Quando alguém procura por um URI, forneça informação útil, usando os padrões (RDF, SPARQL).
4. Inclua outros links para URIs, para que mais coisas possam ser descobertas.

A Web se baseia na ideia de que documentos em diferentes servidores se referenciam, criando uma rede de informações. Então utilizando essas ideias e usando o RDF e URIs, podemos aplicar o Linked Data. O RDF é uma forma de armazenar a informação estruturada e interligada e os URIs são identificadores únicos para objetos ou conceitos. Assim podemos permitir a aplicações navegar entre os dados e recolher informação útil.

2.2.2 RDF

O RDF é um dos padrões desenvolvidos pelo W3C, que é uma linguagem de representação de objeto ou alguma coisa que pode ser identificado na Web. A questão da estruturação da informação é importante no que tange o consumo dessa informação por máquinas e não por pessoas. A estrutura do RDF é da forma sujeito->predicado->objeto, onde o sujeito possui algumas propriedades, o predicado, que por sua vez tem conjuntos de valores, objetos, atribuídos às propriedades. A Figura 2.4 exemplifica uma tripla RDF.



Figura 2.4: Estrutura da Tripla RDF

O modelo de dados RDF representa as informações como grafos direcionados, é projetado para a representação integrada de informações originadas por múltiplas fontes. O RDF tem como objetivo ser utilizado como uma linguagem de conexão, capaz de moderar outros modelos de dados utilizados na Web [67].

Em RDF, uma informação é representada em triplas. As três partes da tripla são o sujeito, predicado e objeto. Uma tripla tenta imitar a estrutura básica de uma sentença, como por exemplo: José (Sujeito) tem o apelido (Predicado) Zé (Objeto).

O sujeito de uma tripla é o URI de um recurso ou informação. O objeto pode ser um simples valor literal como um texto, número ou data; ou o URI de outro recurso que é relacionado de alguma forma com o sujeito. O predicado, elemento do meio, identifica qual a relação que existe entre o sujeito e o objeto, por exemplo, este é o nome ou a data de nascimento, ou até alguém que a pessoa conhece no caso de outro recurso. O predicado ainda é identificado por um URI. Os URIs desses predicados vêm de vocabulários, coleções de URIs que podem ser usadas para representar informação sobre um certo domínio.

2.2.3 Ontologias

Para Gruber, uma ontologia é uma especificação formal de uma conceituação compartilhada [30]. A questão filosófica por de trás da ontologia é o estudo das entidades e suas relações. A questão levantada pela ontologia é: quais tipos de coisas existem ou podem existir no mundo, e quais relações essas coisas podem possuir entre si? A ontologia está menos preocupada com o que realmente é do que com o que é possível shirky.

As ontologias são consideradas como um dos pilares da Web Semântica [69]. Apesar de seu conceito se confundir com o de vocabulário, a tendência é a de nomear como ontologia a coleção de itens que é mais complexa e mais formal. Tanto as ontologias quanto os vocabulários representam e formalizam os conceitos e relacionamentos (termos) utilizados para descrever um campo ou área de conhecimento. Eles classificam os termos que podem ser utilizados em uma aplicação particular.

O papel fundamental das ontologias e vocabulários é o de auxiliar a integração dos dados e da informação, evitando ambiguidades entre diferentes conjuntos de dados que referenciam ao mesmo tema. Por exemplo, uma aplicação para as ontologias no campo da saúde seria possível através de uma combinação de dados médicos como sintomas, doenças e tratamentos juntamente com os dados da área farmacêutica, com contribuições de remédios, dosagens e alergias. A combinação desses dois conhecimentos com os dados de um paciente possibilitaria a criação de aplicações que ajudariam na tomada de decisão de médicos e farmacêuticos e no acompanhamento do tratamento de pacientes.

Um exemplo de aplicação desses princípios é o FOAF - *Friend of a Friend*, que nasceu por volta dos anos 2000 pelas mãos de D. Brickley e L. Miller [9]. Esse projeto busca identificar unicamente uma pessoa na Web, detalhando pessoas, ligações com outras pessoas, atividades, interesses, área de atuação, etc. Qualquer pessoa pode usar o FOAF para se descrever, é como criar uma rede social sem um banco de dados centralizado. A Figura 2.5 mostra um exemplo de perfil utilizando o FOAF.

```
@PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@PREFIX foaf: <http://xmlns.com/foaf/0.1/> .
@PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#> .

<#JW>
  a foaf:Person ;
  foaf:name "Jimmy Wales" ;
  foaf:mbox <mailto:jwales@bomis.com> ;
  foaf:homepage <http://www.jimmywales.com/> ;
  foaf:nick "Jimbo" ;
  foaf:depiction <http://www.jimmywales.com/aus_img_small.jpg> ;
  foaf:interest <http://www.wikimedia.org> ;
  foaf:knows [
    a foaf:Person ;
    foaf:name "Angela Beesley"
  ] .

<http://www.wikimedia.org>
  rdfs:label "Wikimedia" .
```

Figura 2.5: Exemplo de Perfil FOAF. (Fonte: Wikipedia [70])

A Figura 2.6 mostra uma nuvem de relacionamentos entre conceitos publicados no projeto Linked Data. Podemos notar a separação por áreas de conhecimento com cores diferentes. Cada círculo representa um repositório diferente, e seu tamanho indica a

dimensão do repositório em relação ao volume de dados armazenados. As arestas ligando os círculos se referem aos relacionamentos entre esses objetos.

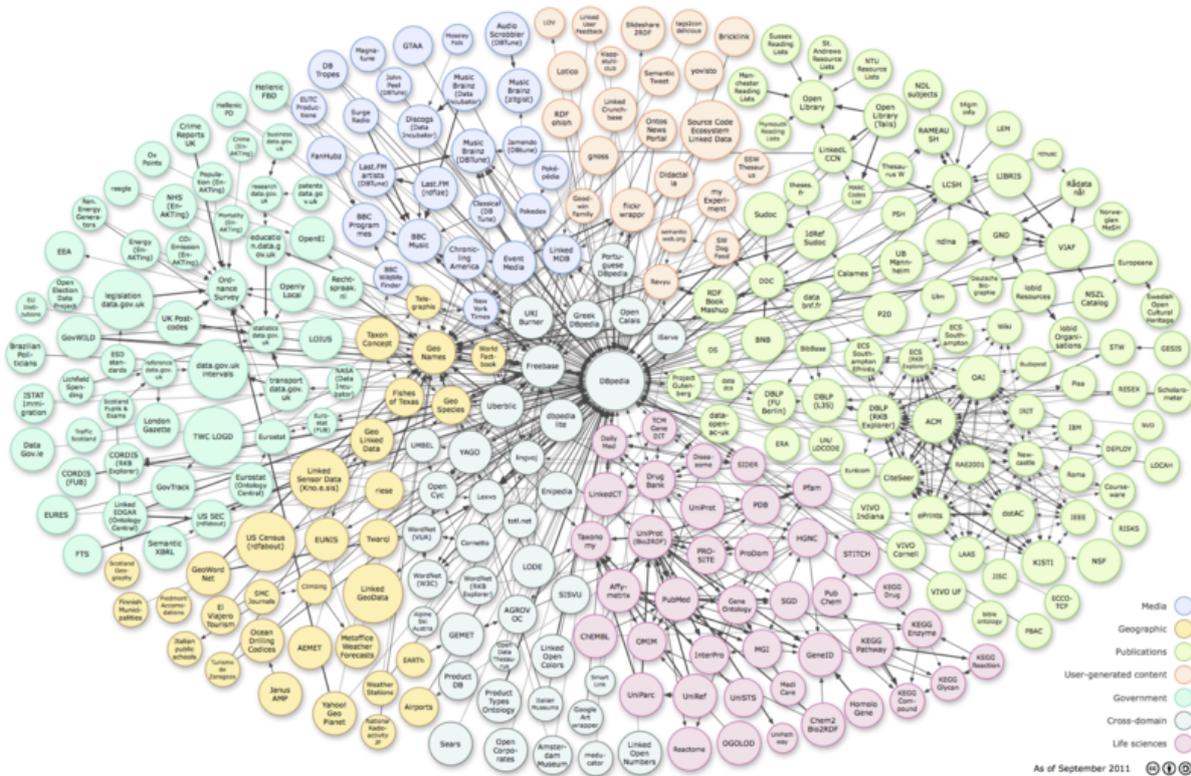


Figura 2.6: Nuvem de Relacionamento DBpedia. (Fonte: DBpedia [13])

2.2.4 SPARQL

SPARQL significa: *Simple Protocol and RDF Query Language*, ela é uma linguagem de consulta e protocolo para RDFs. O SPARQL é descrito de duas maneiras [68]:

1. Uma interface abstrata independente de qualquer realização concreta, implementação ou ligação com outro protocolo.
2. Como ligações HTTP e SOAP para essa interface.

Assim como outras linguagens de consulta, o SPARQL nos permite:

- Recuperar valores de dados estruturados e semiestruturados.
- Explorar dados ao consultar relações desconhecidas.
- Realizar uniões complexas de conjuntos de dados diferentes em uma única e simples consulta.

A estrutura do SPARQL é organizada da seguinte forma, e também apresentada na Figura 2.7:

- Declarações de prefixos, para abreviar URIs.
- Definição do conjunto de dados, informando quais grafos RDF estão sendo consultados.
- A cláusula de resultado, identificando que informação deve ser retornada a partir da consulta.
- O padrão de consulta, especificando o que consultar dentro do conjunto de dados.
- Os modificadores de consulta, limites, ordenação, e outros que podem modificar o resultado final.

```

# prefix declarations
PREFIX foo: <http://example.com/resources/>
...
# dataset definition
FROM ...
# result clause
SELECT ...
# query pattern
WHERE {
    ...
}
# query modifiers
ORDER BY ...

```

Figura 2.7: Estrutura de SPARQL. (Fonte: W3C [68])

Uma consulta SPARQL é executada sobre um conjunto de dados RDF, que consistem em grafos RDF. Essa consulta é executada a partir de SPARQL *endpoints*, que é um serviço que implementa o protocolo SPARQL. Esses *endpoints* podem ser genéricos ou específicos:

- Genérico: consultam qualquer dado RDF acessível pela web.
- Específico: são ligados somente a um conjunto específico de datasets.

Os resultados das consultas SPARQL podem ser retornados em uma variedade de formatos:

- XML: O SPARQL especifica um vocabulário XML para retornar a tabela de resultados.
- JSON: Uma transformação JSON do vocabulário XML, útil em aplicações web.
- CSV/TSV: Representação em texto, ideal para planilhas.
- RDF. Certas consultas SPARQL ativam um gatilho de respostas RDF, que podem ser organizadas em uma variedade de formas (RDF/XML, N-Triples, Turtle, etc.)
- HTML. Quando utilizamos um formulário interativo para consultas.

2.3 Conceitos Básicos de Engenharia Web

A *World Wide Web*, comumente conhecida como Web, está cada vez mais presente no cotidiano humano [57]. O crescimento da Web tem sido exponencial e as organizações continuam a melhorar o escopo de suas aplicações para satisfazer suas necessidades.

Os usuários exigem cada vez a usabilidade, segurança, personalização nas aplicações de hoje em dia. Apesar disso, muitos desenvolvedores e organizações que trabalham com o desenvolvimento de aplicações Web não levam este fator em consideração e acabam por considerar uma página Web sendo simplesmente um documento HTML com hiperlinks para outros documentos. Os requerimentos chave da aplicação e o estudo sobre o design da mesma são desconsiderados, não fazendo o uso de uma metodologia e processos de desenvolvimento de Web Design.

A fim de construir uma aplicação Web mais complexa, independente de sua escala de tamanho, os desenvolvedores devem adotar um processo de desenvolvimento, com metodologias de design e a utilização de melhores ferramentas.

A Engenharia Web se baseia em heurísticas para considerar os aspectos únicos de sistemas Web, utilizando assim uma abordagem sistemática para desenvolver, disponibilizar e manter com sucesso sistemas e aplicações Web com uma elevada qualidade.

2.3.1 OOHDM

O *Object Oriented Hypermedia Design Method* (OODHM) [57] é um método para desenvolvimento de aplicações Web através de instâncias de modelos hipermídia e meta-modelos especializados. Cada um dos modelos foca em diferentes aspectos da aplicação e a compreensão destes modelos auxilia o desenvolvimento de aplicações Web.

O OODHM, através de seus mecanismos e artefatos, permite uma descrição concisa de uma informação mais complexa e também uma especificação de uma navegação complexa seguindo padrões. No OOHDM, as aplicações são construídas seguindo cinco etapas que dão suporte a um modelo incremental de desenvolvimento. As cinco etapas, bem como seus principais artefatos são descritos a seguir [57]:

1. Levantamento de Requisitos: o primeiro passo é coletar requisitos com os interessados da aplicação. Devem-se identificar os atores (quem irá interagir com a aplicação) e quais as tarefas eles deverão realizar na mesma. Após isso, os cenários são criados para auxiliar na construção de diagramas de casos de uso e posteriormente UIDs (Diagramas de Interação do Usuário), que através de elipses representam os estados da aplicação, mostram a informação que é trocada entre o usuário e a aplicação para se realizar uma tarefa.
2. Modelagem Conceitual: nesta etapa, um modelo conceitual do domínio da aplicação é construído utilizando princípios de modelagem orientada a objetos, como classes, conceitos de herança e relações entre as classes.
3. Modelagem Navegacional: no OOHDM, a aplicação é mais vista do ponto de vista navegacional do que conceitual. Os objetos pelos quais o usuário navega não são objetos conceituais, mas sim objetos que são construídos através de um ou mais objetos conceituais a fim de satisfazer as necessidades do usuário de uma melhor forma. Para cada ator é possível definir uma estrutura navegacional diferente que

reflete seus objetos e relações do esquema conceitual baseado nas tarefas. As classes navegacionais são utilizadas para modelar essa etapa e fazem o uso de tipos como nós, âncoras, links e estruturas de acesso para facilitar a modelagem. Os nós no OOHDHDM representam "janelas" lógicas com as quais os usuários irão interagir, são as classes navegacionais constituídas de atributos de uma ou mais classe conceitual. As âncoras são estruturas que ao serem acionadas pelo usuário permitem navegar de uma classe navegacional para a outra. Links são realizações das relações conceituais hipermedia. As estruturas de acesso, como índices (conjunto de âncoras), representam as possíveis formas de navegar dentro da aplicação. A estrutura navegacional de uma aplicação Web é descrita em forma de contexto, que são conjuntos de nós que possuem alternativas similares de navegação e que possuem um significado para a realização de alguma tarefa.

4. Design de Interface Abstrata: o modelo de interface abstrata é construído através da definição dos objetos perceptíveis que contém informação em termos de interface de classes.
5. Implementação: a última etapa mapeia a interface e os objetos navegacionais e podem-se utilizar as mais variadas arquiteturas de programação para conseguir reproduzir os modelos navegacionais desenvolvidos.

A tabela 2.2 reflete os principais artefatos das cinco etapas do OOHDHDM de acordo com Rossi[57]:

Tabela 2.2: Artefatos do OOHDHDM

Etapa OOHDHDM	Artefatos
Levantamento de requisito	Atores Tarefas e UIDS
Modelagem conceitual	Diagrama de Classes Conceituais
Modelagem navegacional	Diagrama de Classes Navegacionais, Diagrama de Contexto Navegacional e Cartões de especificação
Design de Interface Abstrata	ADV, Diagrama de Configuração e ADVcharts
Implementação	Aplicação

2.3.2 SHDM

Para modelar aplicações web mais complexas, como aquelas focadas em Web Semântica, foi desenvolvido um método chamado SDHM (*Semantic Hypermedia Design Method*) [40]. Tendo como base o OOHDHDM, busca-se através do método usar o conhecimento representado em ontologias e processar este conhecimento através da infraestrutura da Web Semântica[41].

Seguindo as mesmas cinco etapas do OOHDHDM, o SHDM se diferencia principalmente na etapa de projeto navegacional, produzindo modelos mais expressivos que são capazes

de representar aplicações Web e até famílias dessas aplicações. A utilização de primitivas para especificar navegações facetadas, que permitem ao usuário formas mais flexíveis de alcançar conjuntos de objetos que são revelantes para executar a tarefa, como categorizando os objetos em classes que fazem sentido e possam exprimir algum significado para o usuário.

O termo faceta foi utilizado primeiramente por Ranganathan apud [40] dentro da área de informação como sendo aspectos, propriedades ou características de uma classe definidas claramente e que são mutuamente exclusivas e exaustivas. Para Maple [43], a análise de facetas é definida como sendo a organização de termos de um domínio de conhecimento em facetas homogêneas, continuando sendo exclusivas. Cada categoria lógica é derivada de um "ancestral" por uma única característica de divisão. Cada uma dessas categorias lógicas deve ser isolada e uma nova característica da divisão deve ser indicada claramente. Em suma, uma faceta pode ser considerada um grupo de termos que representa uma única característica de divisão de um tópico referente a um domínio.

O objetivo da utilização de facetas é a divisão de um tópico referente ao domínio em facetas de alto nível e também dividi-las em componentes menores (subfacetadas) e assim identificar os conjuntos e subconjuntos impossíveis de serem atingidos pela aplicação, melhorando o seu desempenho e a sua transmissão de informação para o usuário[40].

2.4 Conceitos Básicos de Interação Humano-Computador (IHC)

Dentre as formas de lidar com avaliações e desenvolvimento de sistemas que envolvam a interação de seres humanos com dispositivos e sistemas computacionais, destacam-se a Engenharia de Software (ES) e a Interação Humano-Computador (IHC).

No campo da Engenharia de Software, a avaliação é dividida em três fases: verificação, validação e teste [62]. Durante a verificação, é avaliada a compatibilidade do software com as especificações documentadas nos documentos de requisitos. A análise sobre o quanto o sistema atende e irá atender as necessidades do cliente é feita durante a validação. A fase de testes se concentra em avaliações diretamente sobre o código e sobre a execução do mesmo.

Já a visão de uma avaliação sobre a abordagem da IHC tem como foco a interação do usuário com o sistema de computador e os fenômenos relacionados com esse uso. Ela tem como objetivo orientar o avaliador na busca por um sistema interativo de alta qualidade, através de um julgamento de valor sobre a qualidade de uso da solução de IHC proposta e da identificação de problemas na interação e na interface da aplicação que prejudiquem a experiência do usuário durante o uso do sistema [4].

2.4.1 Interação Humano Computador

De acordo com Hewett [35], os objetos de estudos de IHC podem ser agrupados em cinco tópicos relacionados: a natureza da interação humano-computador; o uso de sistemas interativos situados em contexto; características humanas; arquitetura de sistemas computacionais e da interface com usuários; e processos de desenvolvimento preocupados com o uso:

- O estudo sobre a natureza da interação é feito através da investigação do que ocorre enquanto as pessoas utilizam os sistemas interativos em suas atividades. A descrição, a explicação e a previsão desse fenômeno e suas consequências na vida das pessoas envolvidas devem ser possíveis através desse estudo.
- A forma pela qual o contexto de uso influencia a interação de pessoas com os sistemas interativos é evidenciada pelo fato de que uma pessoa está inserida em uma determinada cultura, e possui um modo próprio de realizar atividades e tem crenças, conhecimentos pessoais e linguagens de interação que podem ser diferentes das demais pessoas. O contexto de uso dos desenvolvedores pode ser diferente do contexto de uso de usuário, a pessoa que realmente irá usar o sistema, para isso, esse estudo se torna muito importante e pertinente.
- O conhecimento sobre as características humanas dos usuários permite ao desenvolvedor se aproveitar de suas capacidades e respeitar suas limitações durante a interação. A forma como as pessoas se comunicam e interagem também influencia na interação humano-computador, pois elas tendem a continuar usando esses mesmos mecanismos quando estão lidando com um sistema computacional [55]. Os sentidos, como a visão, audição e o tato e também a capacidade de movimentação e reflexo dos seres humanos são responsáveis pela percepção deles sobre o mundo ao seu redor e sobre a sua forma de agir, e também devem ser considerados nesse estudo.
- O estudo sobre a arquitetura de sistemas computacionais e interfaces com usuário busca o favorecimento da experiência de uso do sistema de computador. Os dispositivos de entrada e saída, por exemplo, são os meios físicos que mediam o contato entre um usuário e um sistema de computador. O projeto da interação deve então levar em conta os modelos conceituais já existentes e conhecidos pelo usuário para facilitar a adoção e o aprendizado do sistema proposto. O conhecimento sobre as tecnologias e dispositivos de interação é fundamental para o processo de desenvolvimento de um software do ponto de vista das formas de interações possíveis.
- O processo de desenvolvimento de um sistema computacional interativo influencia na qualidade do produto final. O conhecimento sobre abordagens de design de IHC, métodos, técnicas, ferramentas e casos de sucesso e insucesso de interfaces ajuda o desenvolvedor do sistema a projetar uma interface de maior qualidade e valor para o usuário.

A Figura 2.8 ilustra a relação desses objetos de estudo.

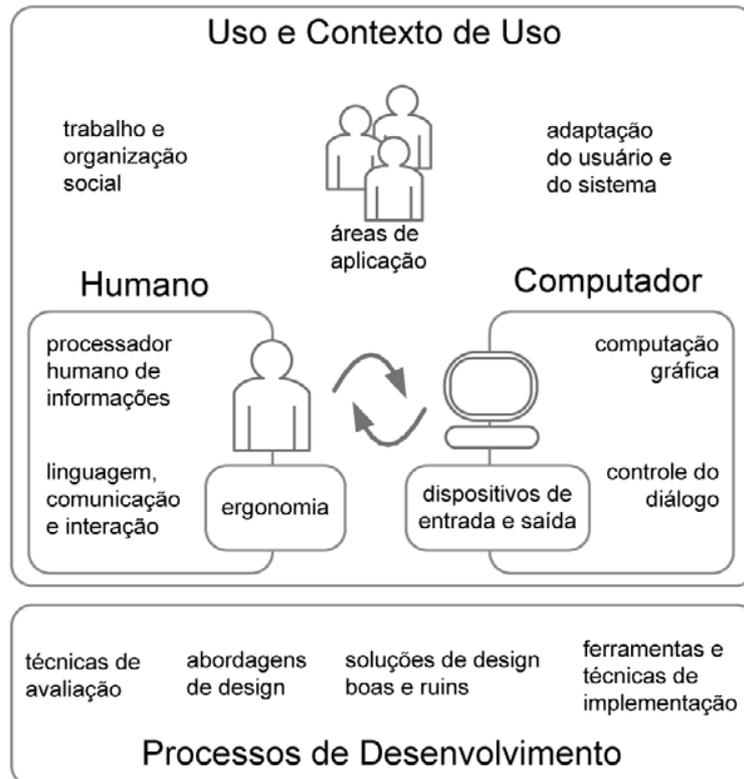


Figura 2.8: Objetivos de estudo em IHC (Fonte: Barbosa e Silva [4])

A forma pela qual um usuário interage com um sistema computacional é através de uma interface. Para Moran [45], interface é a parte de um sistema computacional com a qual a pessoa entra em contato, seja de forma física, perceptiva ou conceitual. A parte física da interface está relacionada com aquilo que o usuário percebe e manipula enquanto a parte conceitual está relacionada com o que o usuário, interpreta, processa e raciocina. Para Souza et al [16], a interface possui componentes de hardware e de software. Os componentes de hardware seriam, por exemplo, teclado, mouse, monitor, câmera, microfone, entre outros. Já os componentes de software seriam aqueles responsáveis pelas seguintes funções:

- Controlar os componentes de hardware.
- Construir dispositivos virtuais para o usuário interagir.
- Gerar símbolos e mensagens que representem informações do sistema.
- Interpretar comandos do usuário.

De acordo com Preece [54], interação é o nome dado ao tipo de comunicação em que o usuário, por meio da interface, gera ações e recebe respostas do sistema computacional. A Figura 2.9 ilustra a interação.

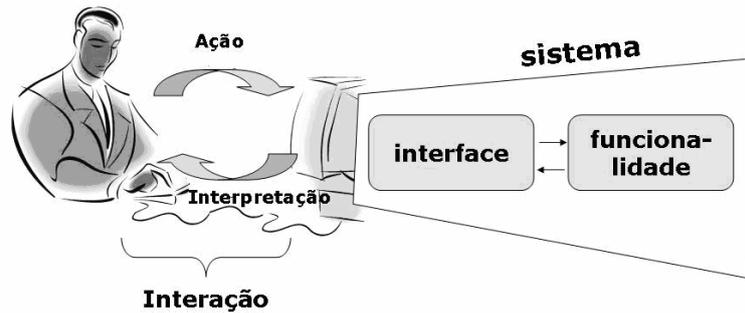


Figura 2.9: Interação entre usuário e sistema computacional (Fonte: Prates e Barbosa[4])

O principal objetivo da Interação Humano-Computador se resume então ao fornecimento de previsões e explicações sobre a interação entre usuário e o sistema e resultados práticos para o design da interface. Com base nesses conhecimentos e resultados, os pesquisadores e desenvolvedores podem aprimorar a interface de suas aplicações e consequentemente a experiência do usuário para com elas [16].

Dentro de IHC, um termo bastante comum e central é a usabilidade. Para Rogers [56], usabilidade é um conceito que tem como foco a busca por uma melhor experiência do usuário, garantindo que produtos interativos sejam realmente efetivos em seu uso, de fácil aprendizagem e prazerosos quanto ao uso do ponto de vista do usuário. Ainda de acordo com Rogers, a usabilidade possui seis objetivos:

- Efetividade no uso: se o produto é realmente bom em fazer o que se ele se propõe a fazer.
- Eficiência no uso: se refere a produtividade que o usuário ganha conforme aprende os conceitos básicos do produto interativo.
- Segurança no uso: se o produto busca evitar que os usuários cometam erros dentro da aplicação, como através da separação de ações críticas no sistema e da recuperação de erros sempre que os mesmos ocorram.
- Ter boa utilidade: se refere ao conjunto de funções providas pelo software, possibilitando os usuários a fazerem o que querem ou o que precisam.
- Facilidade de aprendizado: se o sistema é fácil de usar, ou seja, se o usuário consegue facilmente e em pouco tempo aprender as funcionalidades básicas do sistema.
- Facilidade de lembrar como se usa: se o usuário consegue facilmente se lembrar de como utilizar as funcionalidades básicas do sistema.

Para Souza [16], muitas vezes é difícil atingir esses seis objetivos de forma completa. Sendo assim, cabe aos desenvolvedores identificarem quais desses objetivos são prioritários para um certo sistema computacional em questão, pois atingir o nível de excelência em todos eles pode ser muito dispendioso no desenvolvimento do software.

Ainda de acordo com Souza, além da usabilidade, novos conceitos tornaram-se necessários no cenário de IHC, como a aplicabilidade e a comunicabilidade. Esses três termos norteariam um sistema fácil de usar, aplicar e comunicar. Em seu livro, Souza define [16]:

- Aplicabilidade: tem como base o fato de que todo usuário é especialista em um domínio e o sistema deve servir à essa especialidade e não ao contrário. O sistema computacional deve funcionar como uma ferramenta para o usuário e não exigir que o usuário atenda às exigências e necessidades tecnológicas para usar o software.
- Comunicabilidade: se refere a capacidade do sistema computacional de comunicar os princípios e intenções que guiaram o seu design para o usuário que o usa. Juntamente com a usabilidade, a comunicabilidade tem como pretensão aumentar a aplicabilidade do software.

2.4.2 Projeto de IHC - DECIDE

Para avaliar um sistema computacional do ponto de vista da Interação Humano-Computador, Sharp, Rogers e Preece [60] propuseram um *framework* chamado DECIDE que tem como objetivo orientar o planejamento, a execução e a análise de uma avaliação IHC. Todas as atividades e etapas do *framework* são interligadas e executadas iterativamente.

As atividades e etapas do *framework* DECIDE são descritos por Barbosa e Silva [4] da seguinte maneira:

- D: Determinação dos objetivos da avaliação. O avaliador deve determinar os objetivos gerais da avaliação que irão orientar o restante do planejamento da avaliação, sua execução e a apresentação dos resultados. Além disso, deve-se determinar o porquê e para quem os objetivos determinados são importantes.
- E: Exploração das perguntas a serem respondidas na avaliação. Para cada um dos objetivos definidos na etapa anterior, o avaliador deve formular perguntas específicas a serem respondidas durante a avaliação. Elas devem considerar o perfil e as atividades do público-alvo e são responsáveis por operacionalizar a investigação e o julgamento de valor que serão realizados.
- C: Escolha (*Choice*) dos métodos de avaliação que serão utilizados na avaliação. Tendo como fatores limitantes o prazo, o orçamento, os equipamentos disponíveis e o grau de conhecimento e experiência dos avaliadores, o avaliador deve escolher os métodos mais apropriados para responder as perguntas e atingir os objetivos previamente definidos e esperados.
- I: Identificação e administração das questões práticas da avaliação. Atividades que fazem parte da avaliação, como o recrutamento dos usuários, a preparação dos equipamentos, os prazos e orçamentos disponíveis e a mão-de-obra necessária para conduzir a avaliação são identificadas nessa etapa.
- D: Decisão de como lidar com as questões éticas. Ao se submeter à uma avaliação, o usuário deve ser respeitado e não pode ser prejudicado. Cabe ao avaliador tomar os cuidados éticos necessários para respeitar o usuário e não prejudicá-lo direta ou indiretamente durante toda a avaliação, inclusive na divulgação dos resultados da mesma.
- E: Avaliação (*Evaluation*), interpretação e apresentação dos dados colhidos. Antes de tirar conclusões e divulgar os resultados da avaliação, o avaliador precisa se

atentar a alguns aspectos como: o grau de confiabilidade dos dados; a validade interna, externa e ecológica do estudo.

A definição de cada uma das etapas deixa evidente a dependência de todas as fases do processo, e mudanças e adaptações realizadas em uma única etapa devem ser expandidas e consideradas por todas as demais.

2.4.3 Avaliações em IHC: Método de Avaliação de Comunicabilidade e outros

Segundo Rogers [56], os métodos de avaliação existentes do ponto de vista da Interação Humano-Computador podem ser divididos em três grupos:

- Com a utilização de usuários em ambientes controlados.
- Com a utilização de usuários no próprio local de trabalho.
- Sem a utilização de usuários.

Para Barbosa e Silva [4], os métodos de avaliação ainda podem ser divididos em três outros grupos:

- Avaliação por meio de investigação: o avaliador coleta a impressão dos usuários por meio de questionários, entrevistas e grupos focais. Através disso, o avaliador pode averiguar o design e encontrar problemas na interface que poderão acontecer quando o sistema estiver em funcionamento.
- Avaliação por meio da inspeção: o avaliador busca se colocar no lugar dos usuários enquanto examina e utiliza o sistema computacional. Esse tipo de avaliação não leva em consideração a participação de usuários reais. Alguns exemplos desse método de avaliação são:
 - Avaliação heurística: um especialista utiliza métricas definidas para verificar os requisitos de usabilidade do sistema em análise.
 - Percorso cognitivo: o principal objetivo é verificar a facilidade de aprendizado de um sistema através da tentativa de executar alguma tarefa nele;
 - Método de inspeção semiótica: o foco é a comunicabilidade do sistema e a qualidade da emissão da metacomunicação da interface.
- Avaliação por observação: o avaliador realiza a coleta de dados com usuários em situações reais de uso com o objetivo de identificar os problemas na interface. São exemplos desse método de avaliação:
 - Teste de usabilidade: o objetivo é analisar se os requisitos de usabilidade são atendidos pelo sistema computacional.
 - Método de Avaliação de Comunicabilidade (MAC): o foco é na avaliação sobre a qualidade da comunicação da metamensagem do projetista para os usuários do sistema.

- Prototipação: o protótipo pode ser de baixa ou alta fidelidade, dependendo da tecnologia empregada para sua confecção. Os usuários interagem com o protótipo enquanto o avaliador observa os problemas na interface que o usuário pode encontrar. A Figura 2.10 ilustra os protótipos de baixa e alta fidelidade.

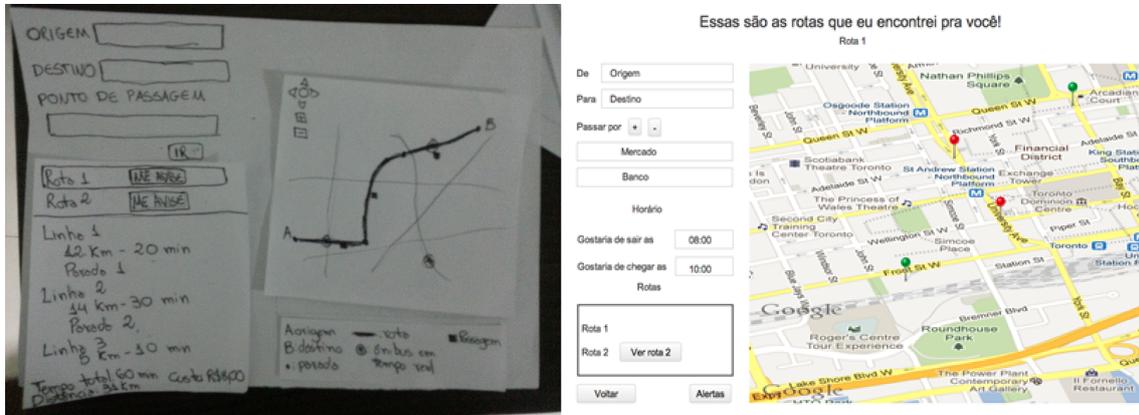


Figura 2.10: Protótipos de baixa fidelidade desenvolvido em papel e alta fidelidade desenvolvido com a ferramenta Justinmind [38].

De acordo com Barbosa e Silva [4], os métodos de avaliação podem ser comparados de diversas formas e cada um deles possui seus pontos fortes e pontos fracos.

A Figura 2.11 classifica os três métodos de inadequado (-) a muito adequado (++++) para cada um dos aspectos relevantes.

		apropriação de tecnologia	alternativas de design	conformidade com padrão	problemas de IHC
investigação	entrevistas	+++	+	-	++
	questionários	++	+	-	++
	grupos de foco	++	+++	-	+++
inspeção	avaliação heurística	-	+++	+++	+++
	percurso cognitivo	+	++	-	+++
	inspeção semiótica	-	++	+	+++
observação	estudo de campo	+++	+	-	+++
	teste de usabilidade	+++	++	-	+++
	aval. de comunicabilidade	+++	++	-	+++
	prototipação em papel	+	+++	-	+++

Figura 2.11: Aspectos avaliados por meio de cada método. (Fonte: Barbosa e Silva [4])

A avaliação, para Barbosa e Silva [4] pode ser feita em dois momentos do desenvolvimento de software. São elas:

- Durante o desenvolvimento do sistema, onde os dados são coletados para reorientar o desenvolvedor do produto antes do mesmo estar pronto. Esse tipo de avaliação recebe o nome de avaliação formativa.
- Após a finalização do desenvolvimento do produto, onde é avaliada a distância que o produto final ficou das metas específicas estabelecidas. Esse tipo de avaliação recebe o nome de avaliação somativa.

A Figura 2.12 ilustra a frequência com que cada um dos métodos são utilizados em cada um dos tipos de avaliação: somativa ou formativa. O símbolo (++) indica o uso mais frequente do método para a avaliação.

		avaliação formativa	avaliação somativa
investigação	entrevistas	++	++
	questionários	++	++
	grupos de foco	++	++
inspeção	avaliação heurística	++	+
	percurso cognitivo	++	+
	inspeção semiótica	+	++
observação	estudo de campo	+	++
	teste de usabilidade	+	++
	aval. de comunicabilidade	+	++
	prototipação em papel	++	+

Figura 2.12: Aspectos avaliados por meio de cada método. (Fonte: Barbosa e Silva [4])

Com relação à análise dos tipos de dados produzidos de acordo com o método de avaliação empregado, sendo que (+++) indica um uso mais adequado do método, Barbosa e Silva [4] comparam eles da forma que pode ser vista na Figura 2.13.

		quantitativos	qualitativos
investigação	entrevistas	++	+++
	questionários	+++	++
	grupos de foco	++	+++
inspeção	avaliação heurística	+	+++
	percurso cognitivo	+	+++
	inspeção semiótica	+	+++
observação	estudo de campo	++	+++
	teste de usabilidade	+++	++
	aval. de comunicabilidade	+	+++
	prototipação em papel	+	+++

Figura 2.13: Frequência com que cada método de avaliação é utilizado nos tipos de avaliação. (Fonte: Barbosa e Silva [4])

O método de avaliação que será utilizado nesse trabalho é o Método de Avaliação de Comunicabilidade (MAC), que tem como origem a Engenharia Semiótica.

Proposta por Clarisse Sieckenius de Souza [15], a Engenharia Semiótica tem como origem a abordagem semiótica para a criação de linguagens de interface do usuário. Após pesquisa realizada pelo Departamento de Informática da Pontifícia Universidade Católica do Rio de Janeiro, a área evoluiu para uma teoria da Interação Humano-Computador.

Uma grande vantagem da visão semiótica, segundo de Souza [15], é a centralização da atenção dos pesquisadores nos signos, o que permite a compreensão do processo de desenvolvimento, uso e avaliação de sistemas computacionais [52]. A avaliação da Engenharia Semiótica se faz sobre a comunicação que o designer objetivou como mensagem e aquilo que o usuário compreendeu, com o objetivo de identificar as falhas de comunicação, mas não de listar as possíveis soluções para os problemas encontrados.

A Engenharia Semiótica tem como base a Semiótica, que de acordo com Santaella [58], é a ciência dos signos e tem como um dos seus idealizadores Peirce. O principal conceito na Semiótica, para Peirce [50] é o de signo. Signo é tudo aquilo que significa algo para alguém, toda e qualquer representação, verdadeira ou falsa, de um conceito ou objeto que será decifrado em um interpretante.

Um signo pode possuir diferentes significados, dependendo de quem o interpreta. Cada uma dessas significações diferentes recebe o nome de interpretante. Um objeto também pode ser representado por diferentes signos, como uma flor, que pode ser representada por uma foto, por um cheiro, ou por se falar ou escrever a própria palavra.

De acordo com Prates e Barbosa [52], os dois conceitos mais importantes dentro da Semiótica são os de significação e o de comunicação.

- Significação: é o processo de criar conteúdo e expressão dos signos baseado em convenções e costumes sociais e culturais, gerando signos conectados ao ambiente que os envolve.
- Comunicação: é processo de criar mensagens compostas por signos de um ou mais sistemas de significação entre os interlocutores. Essas mensagens trafegam dentro de um canal e estão sujeitas a ruídos e obstáculos que interferem na comunicação sobre o canal. A Figura 2.14 ilustra o processo de comunicação.

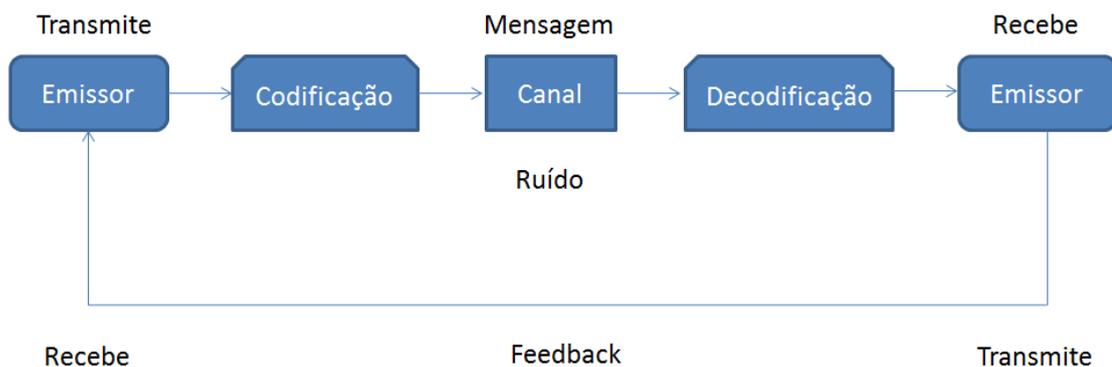


Figura 2.14: Processo de comunicação

Na visão da Engenharia Semiótica, outro conceito que é utilizado é o de tratar um software como sendo um artefato intelectual. Para de Souza [15], os seguintes princípios devem ser satisfeitos para que um objeto de estudo possa ser considerado um artefato intelectual:

- Codificar uma interpretação de uma situação.
- Codificar uma ou mais soluções da situação identificada.
- Ser codificado em uma linguagem específica.
- Os usuários devem ser capazes de utilizar o sistema e compreender a linguagem utilizada para codificar a solução.

Ao entender o software como uma codificação de uma ou mais possíveis soluções através da visão do projetista, de uma situação específica, desenvolvida em uma linguagem de programação que permite que os usuários utilizem essas soluções por meio da interpretação realizada pelo computador daquilo que o projetista codificou, ele pode ser considerado como sendo um artefato intelectual.

Dessa forma, a Engenharia Semiótica entende o software desenhado pelo projetista como uma mensagem para o usuário, que a capta de forma assíncrona dando origem a metacomunicação. De acordo com de Souza [15], esta comunicação pode ser ilustrada com o seguinte trecho:

"Esta é a minha interpretação sobre quem você é, o que eu entendi que você quer ou precisa fazer, de que formas prefere fazê-lo e por quê. Eis, portanto, o sistema que consequentemente concebi para você, o qual você pode ou deve usar assim, a fim de realizar uma série de objetivos associados com esta (minha) visão."

Para Prates [53], um outro conceito importante sobre a área em questão é o de comunicabilidade, definido pela Engenharia Semiótica como a busca por evidenciar a capacidade do projetista de transmitir aos usuários, através da interface, o design tal como foi concebido por ele.

Através de gravações das interações de usuários reais com o sistema a ser avaliado para uma minuciosa inspeção posterior, o Método de Avaliação de Comunicabilidade (MAC) propõe a avaliação da comunicabilidade de um sistema computacional. O MAC tem como principal objetivo avaliar a qualidade da recepção da mensagem pelo usuário.

Segundo Barbosa e Silva [4], a análise realizada através do MAC abrange os prováveis caminhos de interpretação dos usuários, suas intenções de comunicação e as rupturas que aconteceram durante a interação com o próprio sistema ou com algum protótipo. Como resultado da avaliação, tem-se a identificação dos problemas na comunicação da metamensagem do projetista e na comunicação do usuário com o sistema.

As atividades do MAC, de acordo com Barbosa e Silva [4], podem ser descritas como:

- Preparação: Etapa onde os objetivos, métodos que serão utilizados, perfil e número de participantes são definidos. O avaliador ainda pode realizar um teste-piloto para verificar se a avaliação proposta está de acordo com o que foi planejado. A preparação do material relacionado à tarefa que o usuário deverá realizar e a garantia do bom funcionamento das ferramentas de gravação também acontecem nessa fase.
- Coleta de dados: Etapa onde são coletados os dados demográficos e as informações pós-teste do usuário avaliado. O avaliador deve passar tranquilidade ao avaliado e coletar o máximo de informações possíveis da avaliação.
- Interpretação: Etapa onde o avaliador analisa o material obtido a fim de atribuir significados ao que foi colhido. As falhas ou rupturas de comunicação que forem identificadas recebem etiquetas.

- **Consolidação dos resultados:** Etapa onde os resultados e dados coletados são coletados levando-se em consideração a frequência e o contexto em que cada etiqueta ocorreu, a sequência de etiquetas, o nível dos problemas dos usuários dependendo de seus objetivos e outras abordagens e técnicas de IHC que poderiam melhorar a interpretação do pesquisador e avaliador.
- **Relato dos resultados:** Etapa onde a descrição geral da aplicação, métodos utilizados, informações sobre os usuários e a tarefa realizada são documentados.

Segunda Prates e Barbosa [52], o MAC possui treze etiquetas para identificar as rupturas de comunicação:

- **Cadê?:** O usuário tem a intenção de realizar alguma ação na interface, mas não sabe onde encontrá-la.
- **Ué, o que houve?:** O usuário não percebe o retorno do sistema sobre alguma ação realizada ou o sistema não dá esse retorno.
- **E agora?:** O usuário não sabe o que fazer naquele momento para prosseguir na realização da tarefa.
- **Epa!:** O usuário percebe que realizou uma ação de forma errônea ou indesejada e imediatamente a tenta desfazer.
- **Assim não dá:** O usuário tenta seguir por um caminho para realizar a tarefa, e após algum tempo e algumas interações, ele percebe que através desse caminho não conseguirá o resultado estipulado. Então o usuário retorna a um ponto anterior de seu progresso ou ao início para seguir por outro caminho.
- **Onde estou?:** O usuário tenta se comunicar com o sistema através de coisas que seriam apropriadas em outro contexto que não o que ele se encontra.
- **O que é isto?:** O usuário não compreende o significado de um determinado signo no sistema.
- **Por que não funciona?:** O usuário insiste em repetir uma ação que não produz o efeito esperado, mesmo percebendo que o resultado não é o esperado e sem entender o motivo disso.
- **Socorro!:** O usuário não consegue realizar a tarefa e busca auxílio em sistemas de ajuda para concluir a tarefa em questão.
- **Vai de outro jeito:** O usuário não toma conhecimento do melhor caminho projetado para realizar uma tarefa e segue por outro caminho que normalmente é mais longo.
- **Não, obrigado!:** O usuário compreende o caminho preferencial determinado pelo projetista, mas mesmo assim prefere seguir por um caminho alternativo.
- **Para mim está bom...:** O usuário acredita que completou a tarefa estipulada, mas na verdade ele cometeu algum erro e ela não foi concluída completamente.
- **Desisto:** O usuário se conforma em não conseguir completar a tarefa e interrompe a interação.

Após a etiquetagem e a interpretação dos dados, o pesquisador e avaliador devem criar o perfil semiótico do sistema, que é a construção da metamensagem recebida pelo usuário sobre o sistema. De acordo com de Souza [15], o perfil semiótico pode ser descrito como:

"Este é o meu entendimento como designer, de quem você, usuário é, do que aprendi que você quer ou precisa fazer, de que maneiras prefere fazer, e por quê. Este portanto, é o sistema que projetei para você, e esta é a forma como você pode ou deve utilizá-lo para alcançar uma gama de objetivos que se encaixam nesta visão."

2.5 Conceitos Básicos sobre o Domínio: Corrupção

Corrupção vem do latim *corruptus*, que significa quebrado em pedaços. O verbo *corromper* significa "tornar pútrido"[14]. Ela determina a instabilidade e a ineficiência dos serviços prestados, mancha a imagem dos órgãos públicos e cria a imagem de um estado injusto. Não existe uma definição exata de corrupção, pode ser popularmente definida como "o uso incorreto de cargo público para ganhos pessoais"[72], ou ainda um ato suborno. A corrupção pode se manifestar de várias formas, como o suborno, favoritismo, fraude e apropriação.

- Suborno é o pagamento em dinheiro ou serviço e pode partir tanto do servidor público quanto do beneficiário. O suborno pode acontecer na forma de extorsão, conspiratória ou antecipatória [48].
 - Extorsão acontece quando o agente público exige um suborno do beneficiário sob a ameaça de interromper o serviço.
 - Conspiratória é quando o agente público e o beneficiário conspiram para dar e receber suborno em seus respectivos interesses.
 - Antecipatório é quando o beneficiário deseja pagar o suborno, mesmo quando não pedido, para assegurar um serviço ou similar.
- Favoritismo é definido como um mecanismo de abuso de poder na distribuição de recurso público. O nepotismo é uma forma especial de favoritismo, onde são indicadas pessoas próximas e familiares para cargos públicos [24].
- Apropriação é o roubo de propriedade pública e recursos por pessoas que foram confiadas a administrar esses bens.
- Fraude é o ato de enganar o governo por meio de trapagens, sejam elas por meio de manipulação de informação ou pela distorção dos fatos.

Existem diversos fatores que favorecem um ambiente propício a corrupção. Deficiências sistemáticas como falta de informação, falta de supervisão e monitoramento, políticas obsoletas, são algumas delas. Robert Klitgaard, define em seu livro *Controlling Corruption* uma equação para a corrupção [39].

$$\text{Corrupção} = \text{Monopólio} + \text{Discrição} - \text{Prestação de Contas} \quad (2.1)$$

Klitgaard acredita que a corrupção sempre segue essa fórmula. Sua teoria diz que a corrupção surge em um sistema político onde existe um alto grau de monopólio e discrição, com uma baixa demanda por prestação de contas.

O monopólio assegura que os beneficiários não tenham uma alternativa e leva a corrupção por meio de abuso de poder dos órgãos que proveem o serviço. A discricção pode ser má utilizada para manipular votos para alterar uma decisão para favorável ou não favorável. A falta de prestação de contas acentua o problema ajudando os envolvidos a saírem sem serem penalizados.

Klitgaard ainda afirma que a corrupção prospera onde cidadãos não encaram os serviços públicos como um direito e sim como favores.

A prestação de contas pode ser melhorada por um diagnóstico periódico ou auditoria social, que pode ser facilmente aplicada quando se tem informação disponível. A disposição de informação ajudaria a reduzir a corrupção aumentando a transparência.

2.5.1 A Relação Entre Transparência e Corrupção

A *Transparency International* é uma sociedade civil que tem foco no combate à corrupção no mundo, e anualmente é emitido um índice de percepção da corrupção no setor público. O índice é baseado em treze pesquisas abrangendo especialistas em dados abertos e a visão do setor empresarial.

O mapa global ilustrado na Figura 2.15 mostra o panorama mundial da corrupção referente ao ano de 2012, as áreas mais claras (em tons de amarelo), como Estados Unidos e Canadá, são os melhores pontuados no índice, apresentando baixas taxas de percepção da corrupção. As regiões mais escuras (em tons de vermelho) são os países com piores índices, o Brasil se coloca em posição abaixo de países como Omã e Gana, aparecendo apenas na sexagésima nona posição. As Figuras 2.16 e 2.17 mostram respectivamente os melhores classificados e a classificação do Brasil de acordo com o índice emitido pela Transparency International.

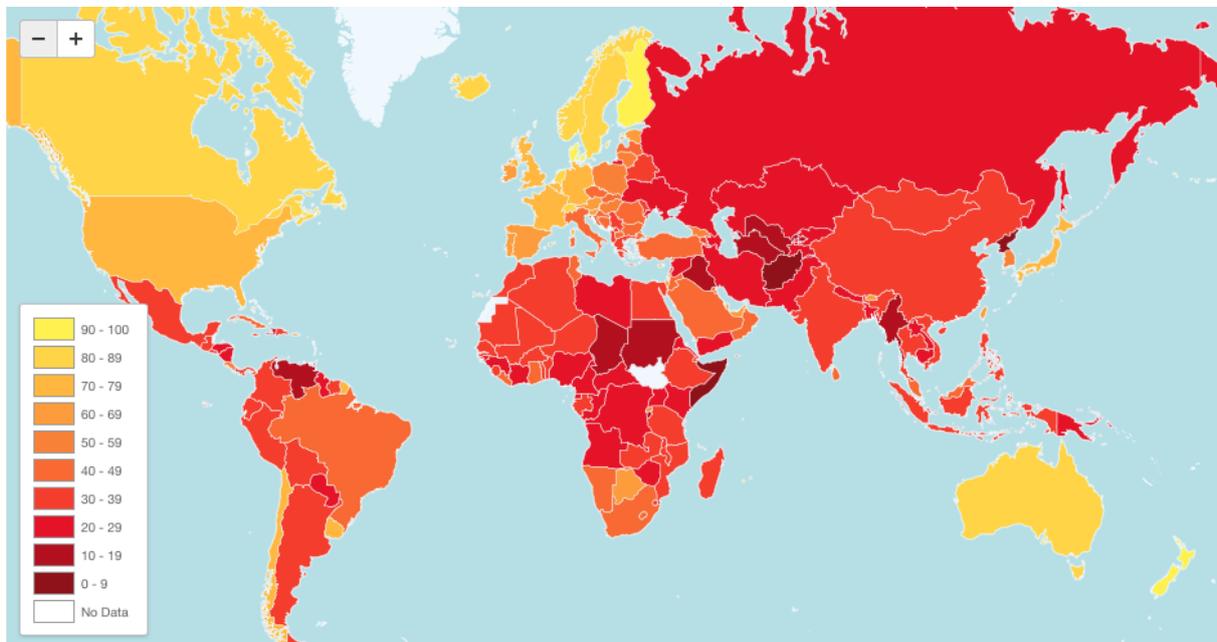


Figura 2.15: Mapa de Percepção de Corrupção - Transparency International [36]

RANK	COUNTRY	SCORE	SURVEYS USED	CI: LOWER	CI: UPPER
1	Denmark	90	7	87	93
1	Finland	90	7	85	95
1	New Zealand	90	7	87	94
4	Sweden	88	7	85	91
5	Singapore	87	9	83	90
6	Switzerland	86	6	81	90
7	Australia	85	8	83	86
7	Norway	85	7	82	87
9	Canada	84	7	80	87
9	Netherlands	84	7	81	88
11	Iceland	82	6	75	89
12	Luxembourg	80	6	75	85
13	Germany	79	8	75	83
14	Hong Kong	77	8	74	80
15	Barbados	76	3	65	87
16	Belgium	75	7	71	78
17	Japan	74	9	70	78
17	United Kingdom	74	8	72	77

Figura 2.16: Tabela de Percepção da Corrupção - Países Melhores Colocados - Transparency International [36]

RANK	COUNTRY	SCORE	SURVEYS USED	CI: LOWER	CI: UPPER
58	Namíbia	48	6	42	54
61	Oman	47	5	35	60
62	Croácia	46	8	41	51
62	Slováquia	46	8	39	53
64	Ghana	45	9	39	51
64	Lesoto	45	5	38	53
66	Kuwait	44	5	37	50
66	România	44	8	38	50
66	Arábia Saudita	44	5	34	55
69	Brasil	43	8	38	49
69	FYR Macedônia	43	6	35	51
69	África do Sul	43	9	39	48
72	Bósnia and Herzegovina	42	7	38	46
72	Itália	42	7	38	46
72	São Tomé and Príncipe	42	3	34	50
75	Bulgária	41	8	35	47
75	Libéria	41	7	35	47
75	Montenegro	41	4	34	49

Figura 2.17: Tabela de Percepção da Corrupção - Brasil - Transparency International [36]

O *Open Data Index* é um projeto da *Open Knowledge Foundation* (OKF) que se baseia na contribuição de especialistas em dados abertos. Os resultados do índice têm como

base os dados coletados no *Open Data Census*. Essa compilação utiliza contribuições da sociedade e praticantes de dados abertos por todo o mundo, e para garantir a qualidade da informação, os dados são checados periodicamente por uma equipe de mais de sessenta contribuidores especialistas. O índice é uma referência mundial para o estado da arte de dados abertos nos governos pelo mundo. Ele provê informações de diversas áreas chave, que refletem a realidade da informação pública em mais de sessenta países. A classificação é liderada por países com alguma tradição na questão dos dados abertos, sendo muitas vezes utilizados como referências para países que buscam a transparência. O Brasil é um país relativamente novo no campo de dados abertos, aparecendo na trigésima sétima posição da tabela. As figuras 2.18 e 2.19 ilustram o ranking de Transparências dos Países de acordo com o *Open Data Index* mais bem colocados e a classificação do Brasil respectivamente.



Figura 2.18: Ranking de Transparência dos Países - Open Data Index [26]

Spain		460
Ireland		460
Isle Of Man		445
Croatia		445
Japan		440
Serbia		440
Ecuador		425
Russian Federation		425
South Korea		425
Poland		420
Taiwan R.O.C.		420
China		415
Indonesia		415
Hungary		415
Brazil		410
Germany		410
Mexico		405

Figura 2.19: Ranking de Transparência Brasil - Open Data Index [26]

É fácil notar que os países que aparecem nas primeiras posições da tabela de percepção da corrupção também aparecem liderando na questão de transparência.

De acordo com a *Transparency International*, transparência significa colocar uma luz nas leis, planos, processos e ações. Entender porque, como, para que e quanto. A transparência assegura que governantes, servidores públicos, gerentes e empresários sejam acompanhados em suas ações e que eles prestem contas de suas atividades, aumentando a confiança entre pessoas e instituições.

Um dos casos mais famosos de transparência combatendo à corrupção ocorreu no Canadá [18] envolvendo organizações de caridade em Toronto. A informação foi disponibilizada pelo próprio governo a pedido de um cidadão. Os dados foram tratados e organizados para listar os maiores receptadores de fundos de caridade, e foi identificado que existiam duas entidades que não eram as maiores e ainda sim somavam uma quantia considerável de dinheiro. Dentro das quinze maiores receptoras estavam quatro fraudulentas, que sozinhas faturaram ilegalmente aproximadamente meio bilhão de dólares apenas no ano de 2005, isso foi facilmente identificável com informação disponível publicamente. É importante observar que a transparência sozinha não pode identificar atos ilícitos, é necessária a participação de um terceiro interessado.

No artigo *Transparency is not Enough: Making Transparency Effective in Reducing Corruption* [42], Lindstedt e Naurin chegam a uma conclusão importante: muito embora a transparência afete a corrupção de várias formas, sozinha ela é insuficiente e ainda afirma que em certos casos a transparência pode aumentar a corrupção. Em linhas gerais Lindstedt e Naurin dizem que:

- A falta de transparência torna a corrupção menos arriscada e mais atraente;
- A falta de transparência torna mais difícil o incentivo de ações limpas por parte de oficiais do governo;
- A falta de transparência torna mais difícil separar os mais honestos e mais eficientes;
- A falta de transparência torna mais difícil a cooperação e instiga o oportunismo;
- A falta de transparência pode prejudicar normas sociais e reduzir a confiança.

O artigo complementa que, o efeito da transparência na corrupção não é incondicional. Em outras palavras, transparência é uma necessidade, mas não é sozinha uma condição suficiente para reduzir a corrupção. É necessário que além de prover o acesso à informação, existam pessoas capazes de processar aqueles dados e ainda formas de incentivar os usos dessa informação.

É preciso ver o outro lado dessas ações também. Uma transparência máxima não necessariamente é ótima. Um setor público que está sempre detalhando todas as suas atividades não será efetivo como deveria em concluir suas atividades. Se você passa mais tempo escrevendo sobre o que você faz, pouco fará. Em seu artigo *Corruption, Connections and Transparency: Does a Better Screen Imply a Better Scene?* [3] Bac traz outro problema da transparência: embora seja mais fácil detectar um agente corrupto com a transparência, também é mais fácil identificar aqueles que devem ser subornados. Seria por exemplo mais fácil identificar aqueles que devem receber dinheiro para facilitar alguma ação de empresa.

2.6 Aplicações de Dados Abertos

Essa seção tem como objetivo descrever e apresentar algumas aplicações desenvolvidas utilizando dados abertos, tanto internacionais quanto nacionais.

2.6.1 Aplicações Internacionais

Data Gov Wiki

A *Data Gov Wiki*, não é meramente uma Wiki, mas sim um site que fornece resultados de busca de dados do governo ligados. Em outubro de 2013 as estatísticas da *Data Gov Wiki* informavam que haviam 417 conjuntos de dados convertidos para RDF que cobriam o conteúdo de 703 dos 5762 conjuntos de dados disponíveis em <http://data.gov>. Esses conjuntos de dados geraram 6.46 bilhões de triplas RDF. Incluem nesses resultados RDFs de conjunto de tabelas fornecidas por data.gov, documentação descrevendo o processo pelo qual passam essas informações [34]. A Figura 2.20 apresenta a estrutura que a *Data Gov Wiki* possui.

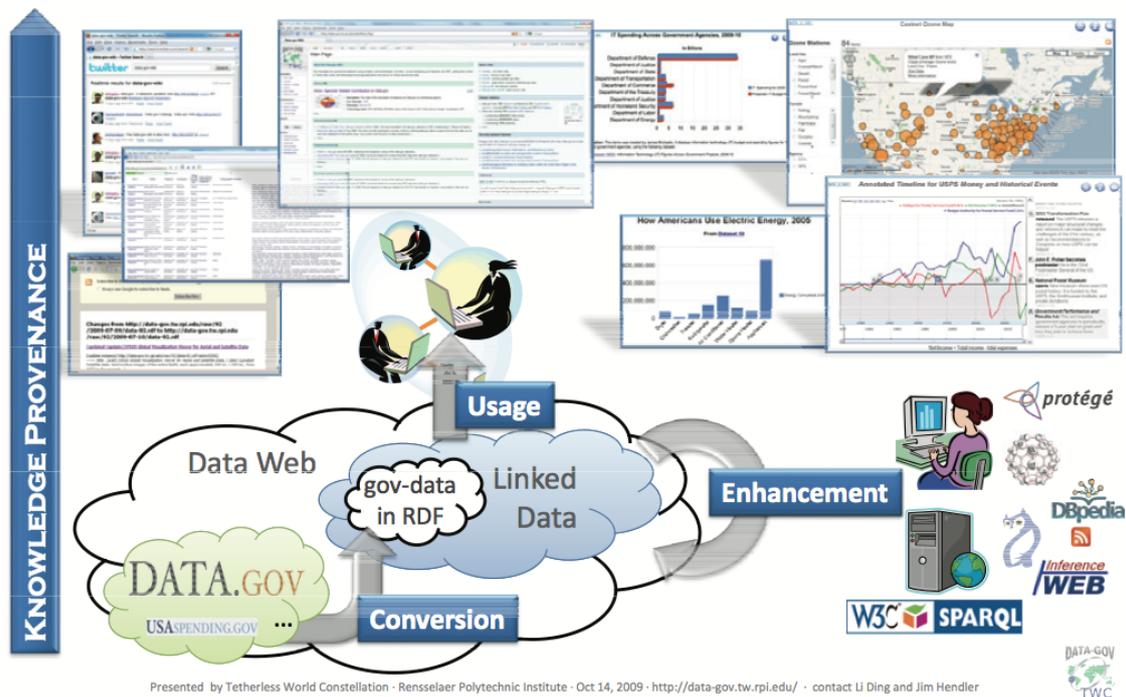


Figura 2.20: Estrutura da Data-Gov Wiki. (Fonte: Data-Gov Wiki [34])

Os passos pelos quais os dados passam na Wiki são:

- Conversão: essa parte do processo é bem direta, primeiramente o dado bruto é recolhido e tratado para ser representado por meio do RDF. Depois os dados convertidos adquirem um URI para que possam ser referenciados unicamente e assim seu conjunto de dados e suas ontologias podem ser utilizadas por terceiros.
- Melhoramento: um dos passos mais complicados do projeto, visa extrair semântica de valores literais dos dados, transformando-os em URIs significativos e ligando os diferentes conjuntos de dados.
- Uso: representa o valor prático de ter Dados Abertos Governamentais ligados. Projetos de ferramentas combinando semântica e padrões de tecnologias web são apresentados.
- Proveniência de Conhecimento: é dada uma atenção especial para a representação e manutenção dos dados tratados pelos 3 passos anteriores. Prover informação é fundamental.

Where does my money go?

A aplicação *Where does my money go* [27] tem como objetivo principal a promoção da transparência e engajamento político popular através de análise e formas criativas de visualização sobre os gastos públicos do Reino Unido. A aplicação é gratuita e não possui nenhum envolvimento partidário e também não tem como objetivo determinar onde e como o governo deveria investir o seu dinheiro.

Os dados utilizados pela aplicação são provenientes do *Country Regional Analysis* publicado anualmente pelo *HM Treasury*.

Dividida em duas principais funcionalidades, a aplicação permite que o usuário informe seu salário e visualize quantas libras ele paga de imposto sobre o que recebe. Além disso, o valor pago em impostos é dividido em áreas de investimento, como educação, saúde e cultura, e para cada uma dessas áreas é mostrado o valor investido através do salário informado. A segunda funcionalidade diz respeito aos investimentos totais do governo em cada área, e ao selecionar uma determinada área, como educação, a aplicação mostra quanto foi investido em cada uma das subcategoria dessa área, como universidades, escola primária e escola secundária.

Similarmente ao *Where does my money go*, no Brasil existe a aplicação chamada Para onde foi o meu dinheiro, que possui o mesmo objetivo da aplicação inglesa, porém não possui a funcionalidade de calcular o investimento público a partir de um salário informado.

A Figura 2.21 ilustra a visualização da aplicação *Where does my money go* sobre como o investimento anual do Reino Unido foi realizado, dividido em categorias.

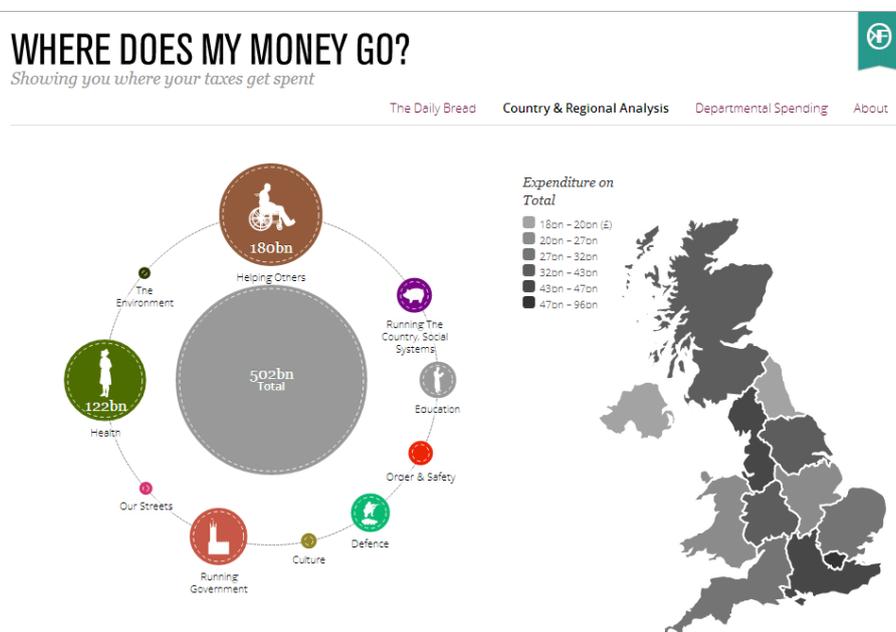


Figura 2.21: Visualização do *Where does my money go*. (Fonte: *Where does my money go does* [27])

Fix my street

A aplicação *Fix my street* [46] não tem como foco principal a denúncia sobre corrupção ou sobre o mau investimento público. Seu diferencial está na participação popular na alimentação de dados da mesma.

A aplicação funciona com a denúncia de problemas em ruas e locais do Reino Unido. O usuário informa o código postal ou o nome da rua, localiza o local do problema no mapa, o detalha, e a própria aplicação encaminha o problema para o órgão responsável por resolver o problema.

Os demais usuários da aplicação podem visualizar todos os problemas, comentar sobre eles e realizar buscas sobre quais problemas estão em determinadas áreas do Reino Unido.

Quando algum problema que foi denunciado é corrigido, o mesmo é atualizado na aplicação. Dessa forma, a população inglesa consegue acompanhar as ações governamentais que são tomadas a partir de suas denúncias, promovendo uma transparência para todos os envolvidos.

A Figura 2.22 ilustra a visualização de um problema que foi reportado na aplicação *Fix my street*.



Figura 2.22: Visualização de um problema reportado na aplicação *Fix my street*. (Fonte: *Fix my street* [46])

2.6.2 Aplicações Brasileiras

Atualmente o portal brasileiro de dados abertos <http://dados.gov.br>, disponibiliza em sua página inicial algumas das aplicações desenvolvidas pela população, utilizando dados abertos.

Queremos Saber

A aplicação brasileira Queremos Saber [32] foi criada pela Comunidade Transparência Hacker e pela *Open Knowledge Foundation Brasil* para facilitar o acesso aos dados abertos governamentais pela população brasileira e visualizar o quanto as instituições estão cumprindo seu dever de disponibilizar essas informações.

O usuário acessa a aplicação, informa o órgão público que deseja receber a informação e escreve um pedido explicando aquilo que deseja saber. A aplicação envia a solicitação para o órgão responsável e publica no site a resposta para que não somente a pessoa que requisitou a informação a tenha, mas também, todas as outras envolvidas ou que por ventura venham a desejá-la, não fazendo necessária uma nova requisição ao órgão público emissor do dado. Todos os dados que foram requisitados no site ficam disponíveis como ilustrado na Figura 2.23.

Queremos Saber

Início Criar uma solicitação Ver pedidos Ver órgãos públicos Ler blog Ajuda

Visualizar ou buscar pedidos

Palavras-chave

Feitos entre e

Exibindo [todos os pedidos](#) | [pedidos atendidos](#) | [pedidos mal sucedidos](#) | [pedidos pendentes](#)

Buscar

76 pedidos de acesso a Informação encontrados

Dados Abertos

Concluída por [Câmara dos Deputados](#) para [Paola Lima](#) em 25 de Setembro de 2013.

Concluída.

[1]https://www.camara.leg.br/internet/in/Servico de Informacao ao Cidadão Solicitante: Paola Lima As...

Iluminação Pública em Goiânia

Concluída parcialmente por [Prefeitura de Goiânia - GO](#) para [Lidiane](#) em 19 de Setembro de 2013.

Concluída parcialmente.

Boa tarde Sra. Lidiane, Favor entrar em contato na Central de Atendimento ao Cidadão - 156, na opção 9, com endereço completo do local (bairro, rua...

Figura 2.23: Visualização das solicitações de dados no Queremos Saber. (Fonte: Queremos Saber [32])

Para Onde Foi o Meu Dinheiro

A Aplicação Para Onde Foi o Meu Dinheiro [49], tem foco no gasto público. Desenvolvida pela Rede Nossa São Paulo e com o apoio do W3C Brasil o aplicativo ajuda o cidadão a monitorar os orçamentos municipal, estadual e municipal. A ferramenta permite também ao internauta saber quais empresas ou pessoas físicas receberam os valores pagos pelo governo estadual. A Figura 2.24 mostra a interface da aplicação.



Figura 2.24: Tela da Aplicação Para Onde Foi o Meu Dinheiro. (Fonte: Para Onde Foi o Meu Dinheiro [49])

Obras do PAC

A Aplicação Obras do PAC [37], desenvolvida por uma empresa privada do estado de São Paulo, fornece uma apresentação visual do dinheiro investido nas obras do Programa de Aceleração do Crescimento (PAC). Os gastos podem ser vistos por períodos, estados. A aplicação ainda fornece um painel para que o usuário possa cruzar informações de investimentos do PAC. A interface da aplicação é retratada na Figura 2.25.

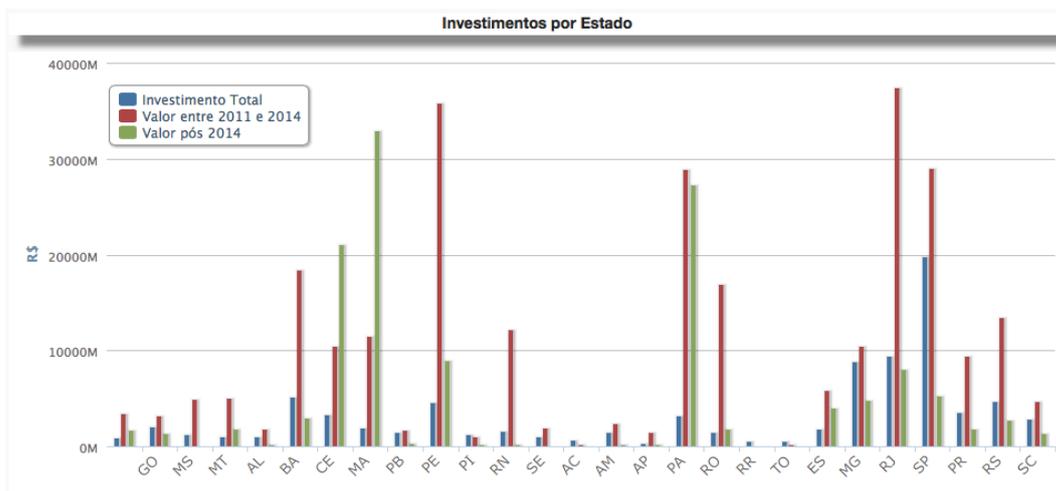


Figura 2.25: Visualização de Dados do PAC. (Fonte: Obras do PAC [37])

Reclamações Procon

Desenvolvido por um grupo de estudantes da Universidade Tecnológica Federal do Paraná, campus Londrina, a aplicação Reclamações Procon [17] utiliza dados disponibilizados pelo Procon sobre as reclamações relativas ao ano de 2011. A Figura 2.26 mostra a interface do Reclamações Procon.

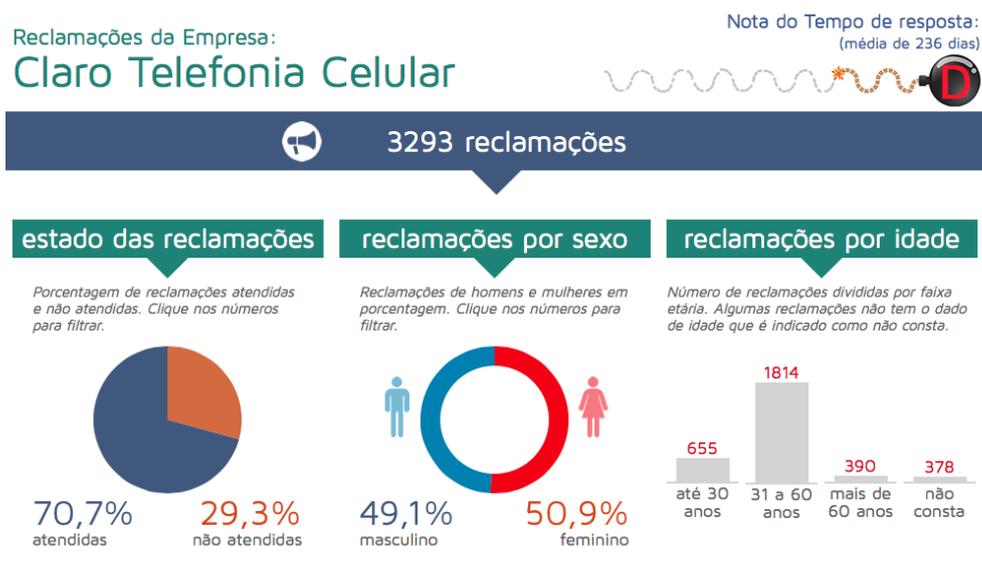


Figura 2.26: Visualização de Dados da Empresa Claro em 2011. (Fonte: Reclamações Procon [17])

Aeroportos Brasil

A Aeroportos Brasil [63] foi desenvolvida por uma empresa multinacional com sede na Espanha, em parceria com uma empresa brasileira (que também desenvolveu o Obras do PAC). A aplicação mostra o movimento de aeronaves, passageiros, carga e mala postal nos sessenta e seis aeroportos administrados pela Infraero. Em aeronaves, estão computados pousos e decolagens. Em passageiros, embarques e desembarques. A Figura 2.27 ilustra os dados colhidos na aplicação.



Figura 2.27: Dados Referentes a Aeronaves em 2011. (Fonte: Aeroportos Brasil [63])

Radar Parlamentar

O PoliGNU, um grupo de estudantes da Escola Politécnica da USP (EPUSP), desenvolveu a aplicação Radar Parlamentar [51] para analisar a semelhança de comportamento das votações dos partidos em projetos de lei em casa legislativas. Os dados são apresentados em gráficos bidimensionais, onde círculos representam os partidos e a distância entre eles representa a sua semelhança. A Figura 2.28 mostra a relação de semelhança dos partidos de acordo com a aplicação.

Basômetro

O Basômetro [20] é uma ferramenta interativa que permite medir o apoio dos parlamentares ao governo e acompanhar como eles se posicionaram nas votações legislativas. Cada um é representado por uma bolinha com a cor do partido. Quanto mais próxima ela está do governo (no alto), maior é a taxa de governismo.

Os votos dos parlamentares foram obtidos nos sites da Câmara dos Deputados e do Senado. Foram consideradas todas as votações nominais, as únicas em que o voto individual é computado, que ocorreram desde o início de 2011. A Figura 2.29 mostra a interface da aplicação.

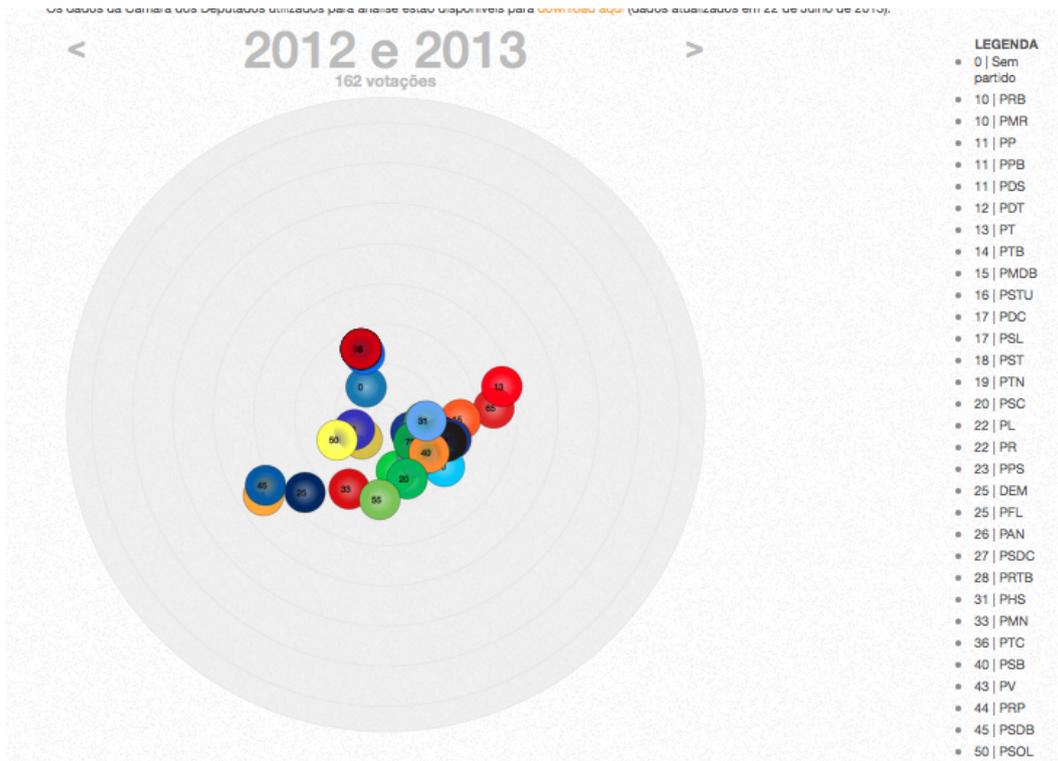


Figura 2.28: Semelhança entre partidos na Câmara dos Deputados. (Fonte: Radar Parlamentar [51])

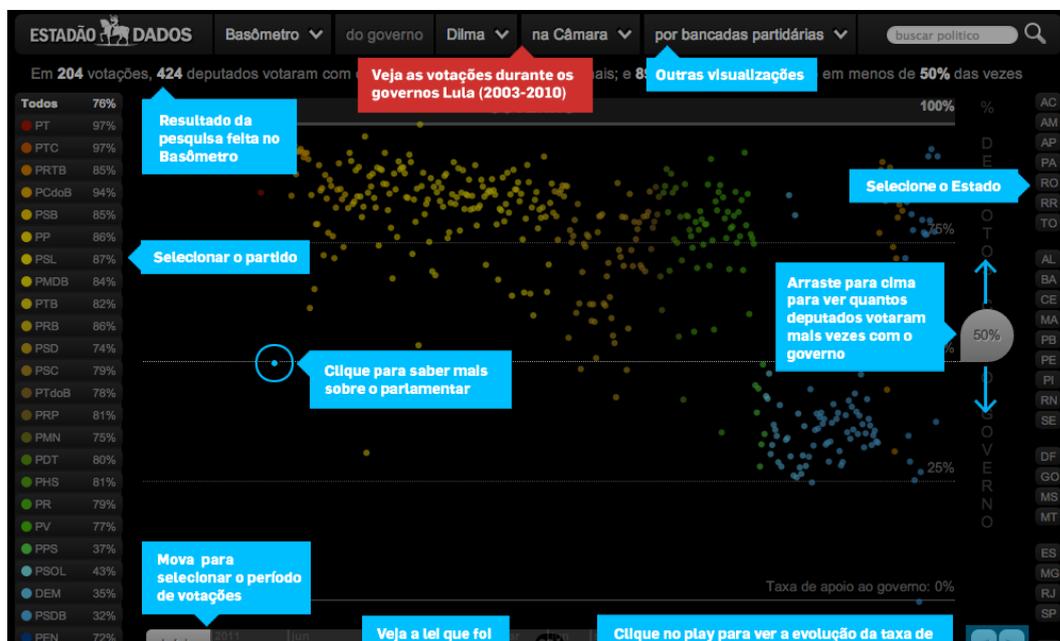


Figura 2.29: Basômetro, apoio dos parlamentares ao governo. (Fonte: Basômetro [20])

Capítulo 3

Proposta de Trabalho

A utilização das diferentes abordagens propostas nesse trabalho tem como principal objetivo auxiliar o projetista de uma aplicação que envolva dados abertos. O método OOHDm da Engenharia Web tem o propósito de guiar o projetista no desenvolvimento de artefatos que retratem e modelem navegações mais complexas para servirem de base na criação de um protótipo. Inicialmente um protótipo em papel foi desenvolvido com base nessas modelagens. Os diagramas navegacionais auxiliam o projetista na apresentação da informação e na navegação entre as telas de uma forma lógica. Um protótipo digital foi construído a partir do primeiro protótipo de papel para ser avaliado com o MAC da Engenharia Semiótica. Dessa forma é possível avaliar a qualidade com que a mensagem prevista pelo projetista chega ao receptor. O capítulo 3 tem como objeto descrever a proposta em termos técnicos, apresentar a fonte de dados para o protótipo e os artefatos gerados através da abordagem de desenvolvimento proposta.

3.1 Utilização de um Conjunto dos Dados Abertos Governamentais

Os dados abertos governamentais que serão utilizados no protótipo são:

- **Gastos Diretos:** Os gastos diretos dizem respeito às licitações realizadas no período e trazem informações como o valor, descrição e quem foi o contratado. Os dados foram retirados do Portal da Transparência [8].
- **Prestação de contas:** Os dados da prestação de conta dizem respeito às despesas e às receitas que os candidatos, partidos e comitês tiveram na época de eleição. Os dados foram retirados do site do TSE [65].
- **Resultados das eleições:** Os resultados das eleições consistem em dados sobre os vencedores das eleições por urna e por zona eleitoral. Os dados foram retirados do site do TSE [65].
- **Cadastro nacional de empresas inidôneas e suspensas:** Os dados de empresas inidôneas dizem respeito às empresas que tiveram algum problema no cumprimento de seus contratos e portanto foram classificadas dessa forma. Os dados foram retirados do Portal da Transparência [8].

3.2 Modelagem da Aplicação

A aplicação foi modelada através do método OOHDH apresentado no capítulo 2. A tabela 3.1 destaca os artefatos de cada etapa e os que serão apresentados nesta seção.

Etapa OOHDH	Artefatos
Levantamento de requisitos	Tarefas UIDS
Modelagem conceitual	Diagrama de Classes Conceituais
Modelagem navegacional	Diagrama de Classes Navegacionais, Diagrama de Contexto Navegacional

3.2.1 1ª Etapa do OOHDH: Levantamento de requisitos

As principais tarefas que o usuário com o perfil de eleitor poderá realizar na aplicação são:

1. O usuário deseja ver informações sobre os políticos que mais receberam e mais gastaram durante a campanha eleitoral de um período.
2. O usuário deseja ver informações sobre as empresas que mais ganharam licitações e quanto patrocinaram na campanha eleitoral de um período.
3. O usuário deseja poder avaliar uma despesa realizada por um político.

Os UIDs desenvolvidos para as tarefas apresentadas foram, respectivamente:

1. Tarefa 1: Políticos.

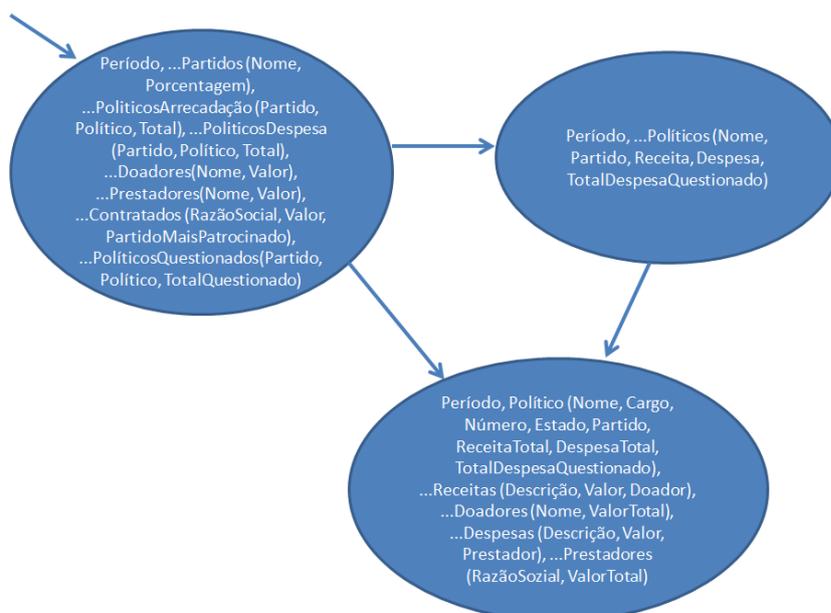


Figura 3.1: Tarefa 1

2. Tarefa 2: Empresas.

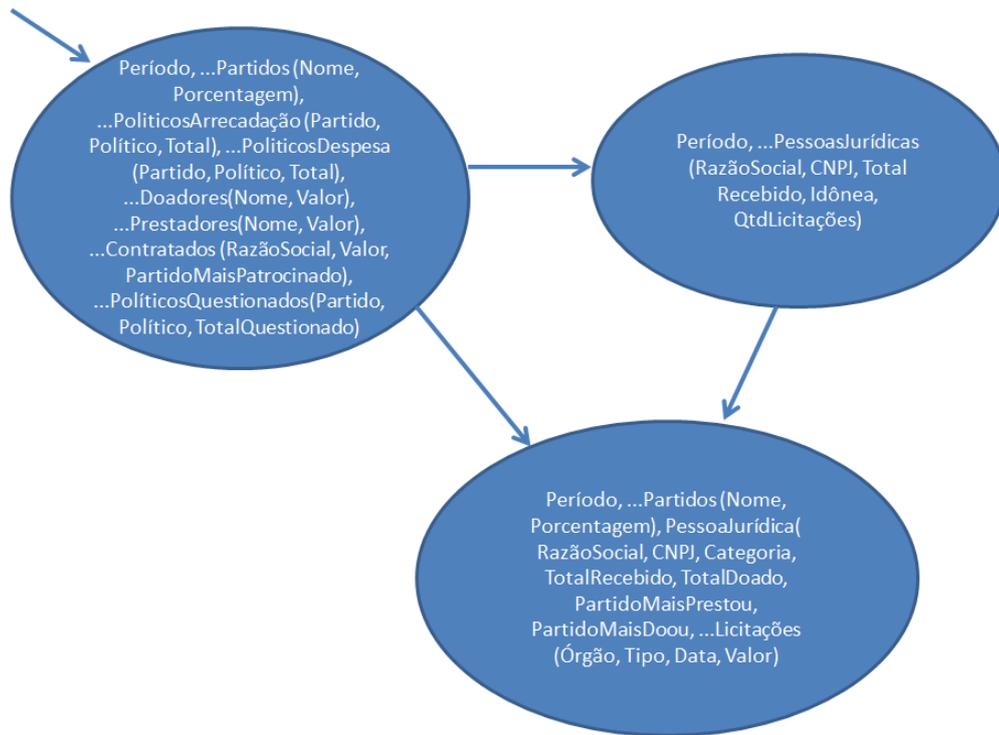


Figura 3.2: Tarefa 2

3. Tarefa 3: Despesas.

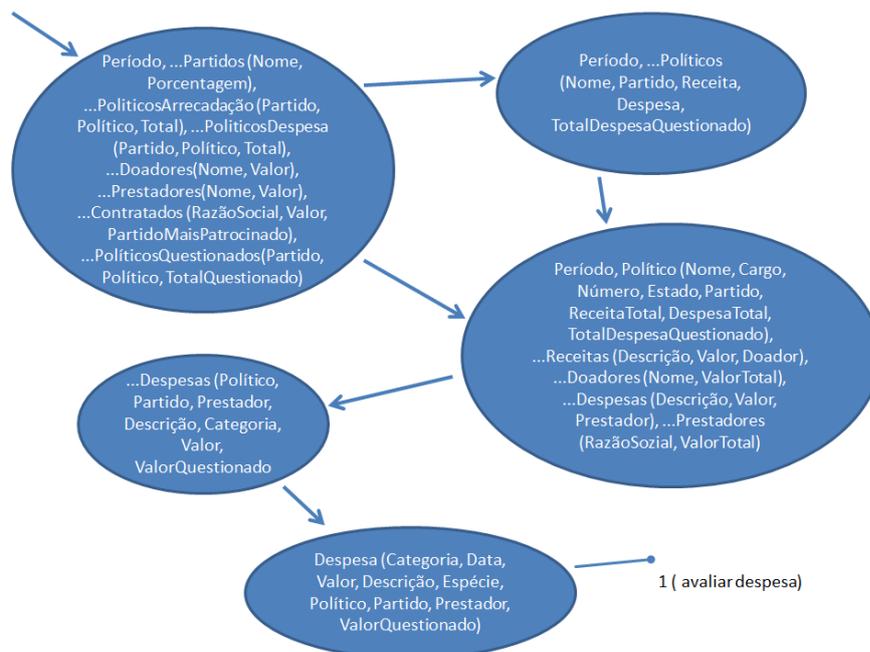


Figura 3.3: Tarefa 3

3.2.2 2ª Etapa do OOHD: Modelagem Conceitual

A partir dos UIDs desenvolvidos para todo o protótipo e seguindo os princípios e diretrizes apresentados por Rossi [59], foi construído o Diagrama de Classes Conceituais apresentado a seguir na Figura 3.4:

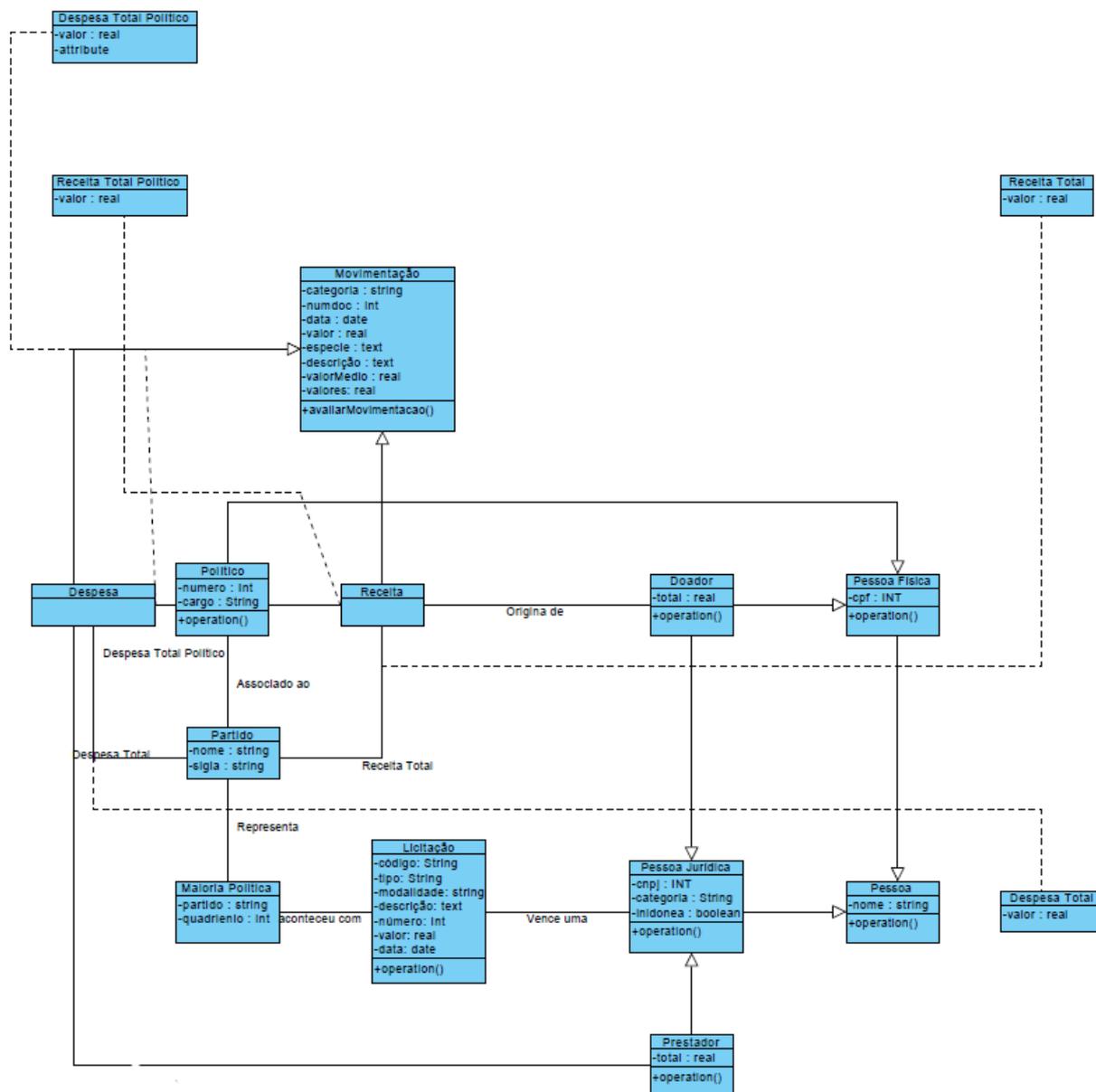


Figura 3.4: Diagrama de Classes Conceituais

A classe conceitual Movimentação foi utilizada como sendo uma superclasse de Despesas e Receitas e as classes Pessoa Física e Jurídica foram utilizadas para criar uma separação entre Doador e Prestador, sendo que o prestador pode ser somente uma pessoa jurídica, enquanto o doador também pode ser uma pessoa física. Os vencedores de licitações (contratados), também só podem ser pessoas jurídicas.

3.3 Protótipos de Baixa Fidelidade para a Avaliação

Com base nos modelos desenvolvidos na etapa anterior, as telas de navegação foram elaboradas. O Diagrama de Classes Conceituais permite ao projetista uma visão holística do domínio trabalhado e como ele será aplicado ao projeto. As modelagens navegacionais se apresentam como um guia para o desenvolvimento das telas, as classes navegacionais mostram quais informações devem ser apresentadas e o contexto navegacional explicita as navegações entre as telas de forma clara para o projetista. Esse tipo de abordagem é muito importante em situações onde a equipe que implementará as telas não é a mesma da etapa de modelagem. Primeiramente, foi desenvolvido um protótipo em papel que abrangeria toda a aplicação com o intuito de se utilizar dele para a realização da avaliação de comunicabilidade através do Método de Avaliação de Comunicabilidade. A vantagem do protótipo de baixa fidelidade em papel é que possui uma construção mais rápida e deixa claro para o usuário que irá interagir com ele que aquilo se trata de um protótipo, sem fins estéticos, deixando a atenção do usuário voltada para a informação e para a interação. A seguir são apresentadas algumas imagens do protótipo construído. A primeira imagem (Figura 3.7) retrata a tela de licitações vencidas por uma pessoa jurídica em um período e a página inicial do protótipo, enquanto a segunda imagem (Figura 3.8) revela a listagem de políticos e a página de perfil de um único político.

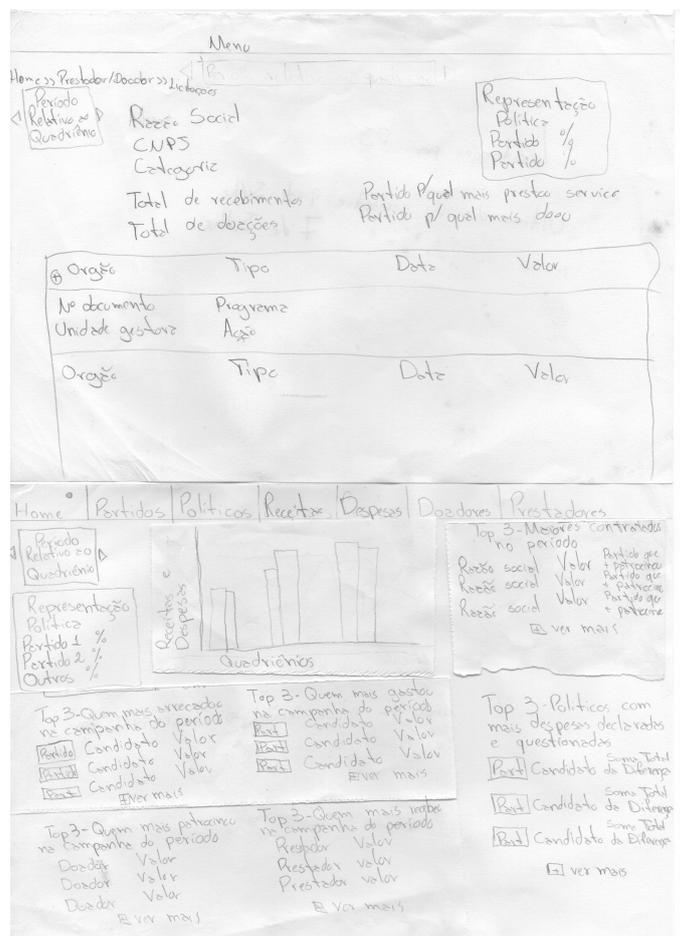


Figura 3.7: Protótipo de Papel - Tela Inicial

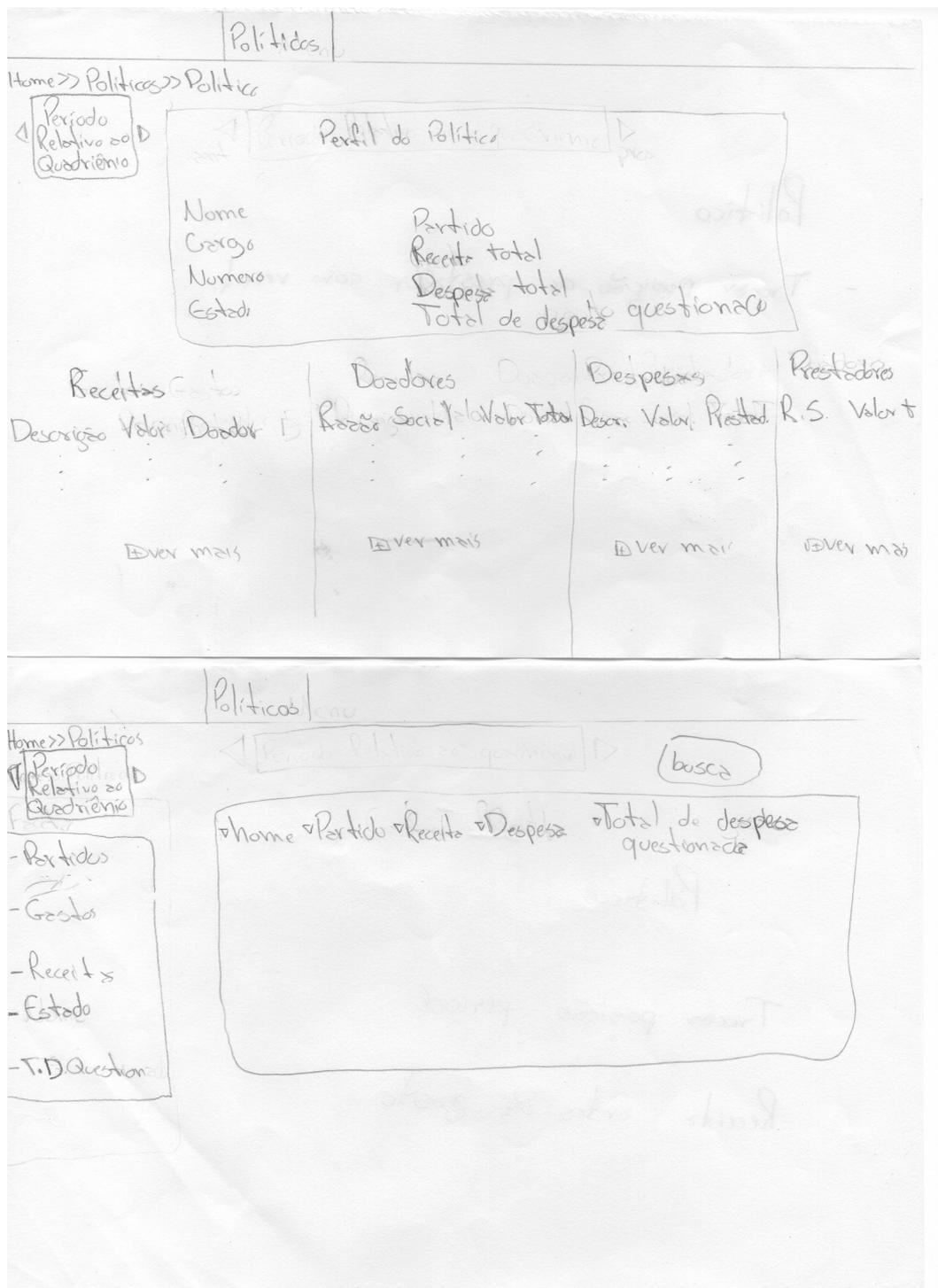


Figura 3.8: Protótipo de Papel - Perfil do Político

Com o objetivo de facilitar a interação do usuário com o protótipo e permitir uma navegação mais fácil, foi construído um segundo protótipo, ainda de baixa fidelidade, digital. Algumas modificações no design de algumas páginas foram realizadas para melhorar a informação. As Figuras 3.9, 3.10 e 3.11 se referem ao protótipo construído através da ferramenta *WireframeSketcher* [71].

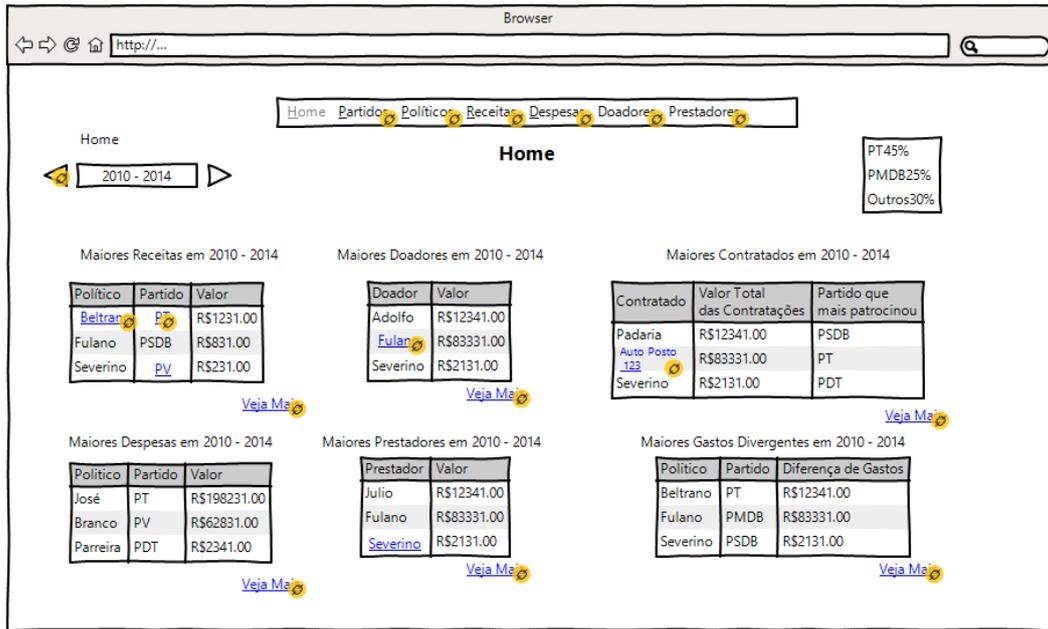


Figura 3.9: Protótipo Digital - Tela Inicial

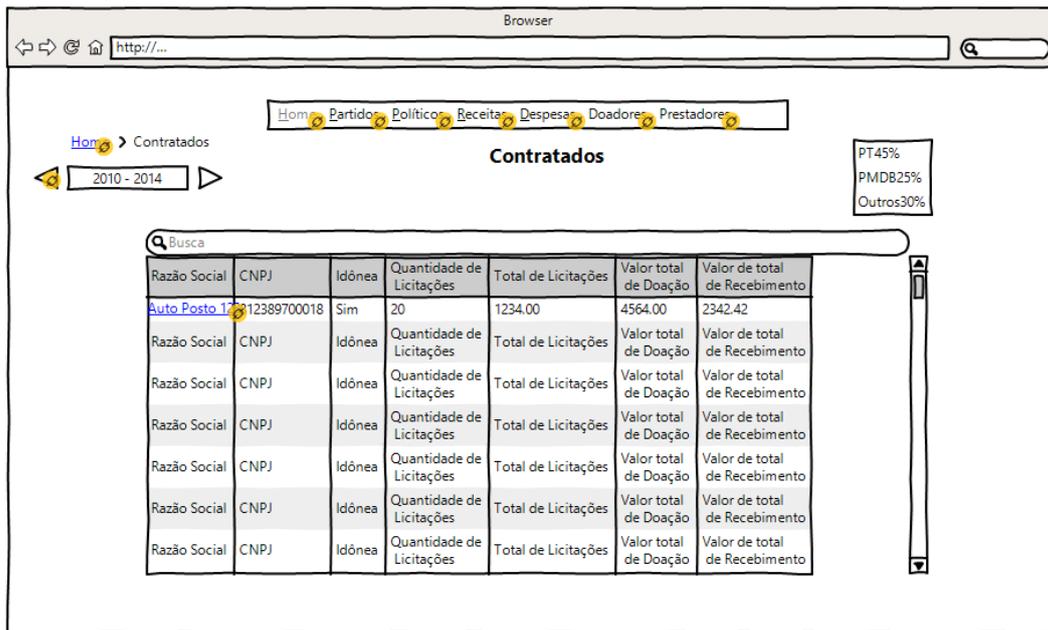


Figura 3.10: Protótipo Digital - Licitações de uma Pessoa Jurídica

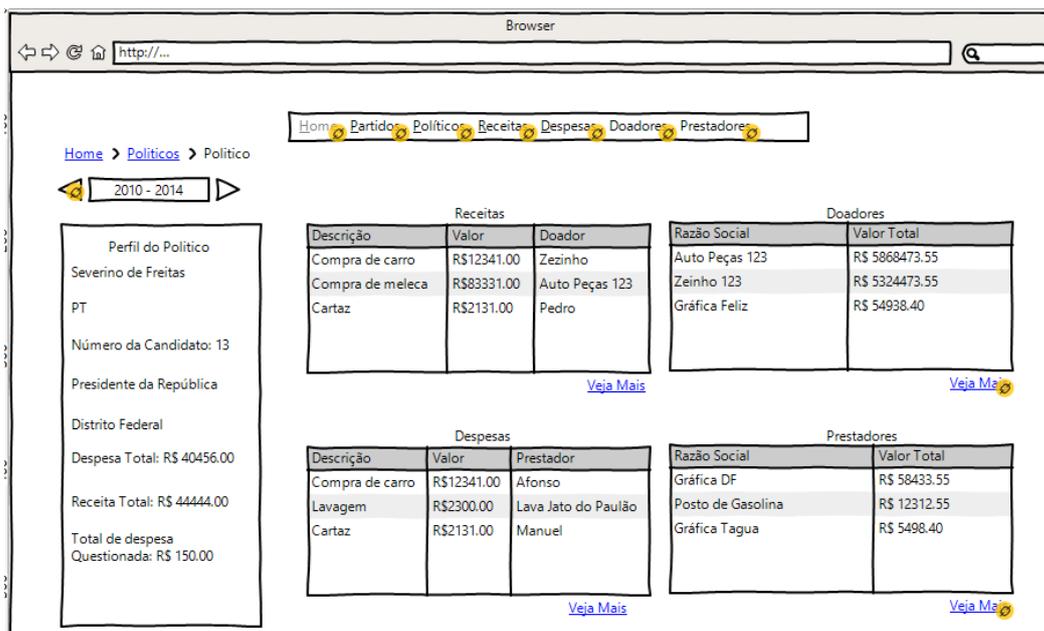


Figura 3.11: Protótipo Digital - Perfil do Político

3.4 Utilização de Método para a Avaliação da Aplicação

Para melhorar a usabilidade e a comunicabilidade do protótipo, o Método de Avaliação da Comunicabilidade será utilizado com o objetivo de identificar rupturas de comunicação nos protótipos construídos.

Os documentos utilizados para a realização do MAC se encontram na área de apêndices deste trabalho. O roteiro de avaliação auxilia o avaliador na realização do experimento com os passos que deve seguir. O termo de consentimento é o documento que será entregue ao usuário para que seja lido e assinado, permitindo o uso do experimento realizado como fonte de dados e concordando com os termos apresentados. Antes de realizar o MAC, o usuário também responderá um pré-questionário demográfico para ajudar na classificação do perfil do usuário em questão. O documento Informações ao Participante é um documento que será lido pelo avaliador, explicando como funcionará a avaliação e como o usuário pode e deve proceder nesta. O documento de Tarefas será entregue ao usuário e contém todas as tarefas que ele irá realizar na avaliação através do protótipo. São elas:

- Verificar quais são os partidos de maior receita no período.
- Listar os doadores dos partidos PT e PV.
- Listar as maiores licitações da empresa Auto Posto 123.
- Informar o valor corrigido para a despesa de valor mais alto no período.

O questionário pós-teste é um documento que será entregue ao usuário após a realização das tarefas para colher um *feedback* sobre o protótipo da aplicação.

3.5 Conversão dos Dados para o Formato RDF

Para possibilitar a utilização de consultas SPARQL nos dados coletados do Portal Brasileiro de Dados Abertos, estes devem ser convertidos para o formato RDF. Este formato permite que propriedades sejam criadas ou aproveitadas para servirem de predicado nas triplas e conectarem objetos, recursos e tipos mais simples. Dessa forma, algumas propriedades foram definidas e serão utilizadas no vocabulário do trabalho proposto. A Tabela 3.2 ilustra as propriedades, ou predicados, definidas e a quais colunas dos dados elas se referem.

Tabela 3.2: Propriedades do RDF projetado

Sujeito	Predicado	Objeto/Predicado	Predicado	Objeto
URI da receita	hasDate	Data da Receita		
URI da receita	hasPrice	Valor da Receita		
URI da receita	hasDesc	Descrição da Receita		
URI da receita	hasType	Tipo da Receita		
URI da receita	hasSource	Fonte do Recurso		
URI da receita	hasKind	Espécie do Recurso		
URI da receita	hasDonator	URI do Doador	hasCpfCnpj	CPF/CNPJ do Doador
URI da receita	HasAdressee	URI do Candidato	hasCpf	CPF do Candidato
URI da receita	HasAdressee	URI do Candidato	hasPosition	Cargo do Candidato
URI da receita	HasAdressee	URI do Candidato	hasParty	Partido do Candidato

Através dessa estrutura, é possível gerar o arquivo com os dados no formato RDF, o qual permite consultas SPARQL que serão realizadas por conhecedores do domínio e da tecnologia de consulta.

Capítulo 4

Implementação da Proposta

Este capítulo tem por objetivo relatar a aplicação do Método de Avaliação de Comunicabilidade e a implementação da solução proposta no capítulo 3 deste trabalho.

4.1 Aplicação do Método de Avaliação de Comunicabilidade (MAC)

A aplicação do MAC consiste na identificação das etiquetas e rupturas de comunicação da interface proposta através de filmagens da interação dos usuários com o protótipo realizando as atividades definidas. Através da análise dos resultados encontrados é possível propor soluções para melhorar a interação dos usuários com o protótipo.

A fim de diversificar a amostra e conseguir interações de pessoas de diversas áreas de conhecimento e atuação no mercado de trabalho, o perfil dos usuários que realizaram a avaliação está descrito na Tabela 4.1. A avaliação relacionada à execução das tarefas descritas no capítulo 3 pode ser analisada através da Figura 4.1 onde os resultados estão divididos por usuário. A quantidade de usuários foi determinada seguindo a sugestão dos especialistas do método apud Barbosa [4]

Tabela 4.1: Perfil dos Usuários da Avaliação

Nome	Faixa Etária	Escolaridade	Área de Estudo	Área de Atuação
Usuário 1	20-30	Superior Incompleto	Exatas	Política
Usuário 2	20-30	Pós-Graduação	Humanas	Direito Penal
Usuário 3	31-60	Superior Completo	Saúde	Aposentado
Usuário 4	20-30	Superior Incompleto	Exatas	Engenharia Mecânica
Usuário 5	15-19	Médio Incompleto	Nenhuma	Estudante
Usuário 6	20-30	Médio	Nenhuma	Nenhuma
Usuário 7	20-30	Superior Incompleto	Humanas	Política

Usuário	Tarefa	Cadê	Ué, o que houve?	E agora ?	Epa!	Assim não dá	Onde estou ?	O que é isto?	Por que não funciona	Socorro!	Vai de outro jeito	Não, obrigado	Para mim está	Desisto	Total
1	1					1					1				2
1	2	1				1					1		1		4
1	3					1					1		1		3
1	4														0
2	1												1		1
2	2	2				1		1						1	5
2	3														0
2	4												1		1
3	1														0
3	2												1	1	2
3	3												1		1
3	4												1		1
4	1	1		1		2					1		1		6
4	2	1				2					1			1	5
4	3														0
4	4														0
5	1												1		1
5	2					1							1		2
5	3										1		1		2
5	4												1		1
6	1	1				3							1		5
6	2					5					1			1	7
6	3												1		1
6	4	1											1		2
7	1	1											1		2
7	2	2			1	2								1	6
7	3														0
7	4												1		1
Total		10	0	1	1	19	0	1	0	0	7	0	17	5	61

Figura 4.1: Resultado do MAC

A tabela torna visível a grande quantidade de etiquetas do tipo "Cadê", "Assim não dá", "Vai de outro jeito" e "Para mim está bom...". Enquanto as três primeiras podem indicar que o usuário estava um tanto quanto perdido no protótipo e tentava realizar as tarefas por meio de tentativa e erro clicando em todas as opções possíveis do protótipo e analisando o resultado, a última indica que os usuários não foram até o final, principalmente não explorando as classes navegacionais propostas e acessando o "perfil" de políticos, partidos e empresas. Através da análise das etiquetas encontradas, foi possível enumerar as seguintes deficiências de comunicabilidade no protótipo proposto:

1. A quantidade excessiva de informação na página inicial do protótipo, faz com que muitos dos usuários a ignorem e tentem realizar as tarefas utilizando o menu e não as informações resumidas e categorizadas dessa tela.
2. A falta de uma descrição sobre o que é a dominância política e a porcentagem dos partidos faz com que essa informação seja inútil para os usuários e as vezes até ignorada.
3. Não ficou claro a possível navegação pela aplicação através dos itens dentro das tabelas. Dessa forma, os usuários navegavam somente pelas abas do menu e não pelos perfis de políticos e partidos para conseguir mais informações.
4. A falta de explicação dos termos utilizados, bem como a falta de intimidade dos usuários com os termos de política, fez com que algumas tarefas não ficassem claras

e eles navegassem aleatoriamente pela aplicação para entender o significado dos termos.

5. A quantidade de informação exibida em qualquer tabela do protótipo é relativamente grande, o que dificulta a visualização dos números e valores por parte do usuário.

As deficiências encontradas foram corrigidas da seguinte forma no novo protótipo:

- Criação de uma página FAQ com perguntas pertinentes ao protótipo e explicação sobre o significado dos termos utilizados.
- Inserção do texto "Dominância Política" sobre a caixa com a porcentagem dos partidos no poder.
- Criação dos botões de "prox" e "ant" na tela de partido e políticos para permitir a navegação em contexto.
- Criação da "obs" nas telas de receitas e despesas para indicar que elas não estão agrupadas para partido e como o usuário deve proceder para encontrar esse tipo de informação.

As Figuras 4.2, 4.3, 4.4 e 4.5 refletem respectivamente as principais mudanças no novo protótipo nas telas de FAQ, Tela Inicial, Partido e Receitas:

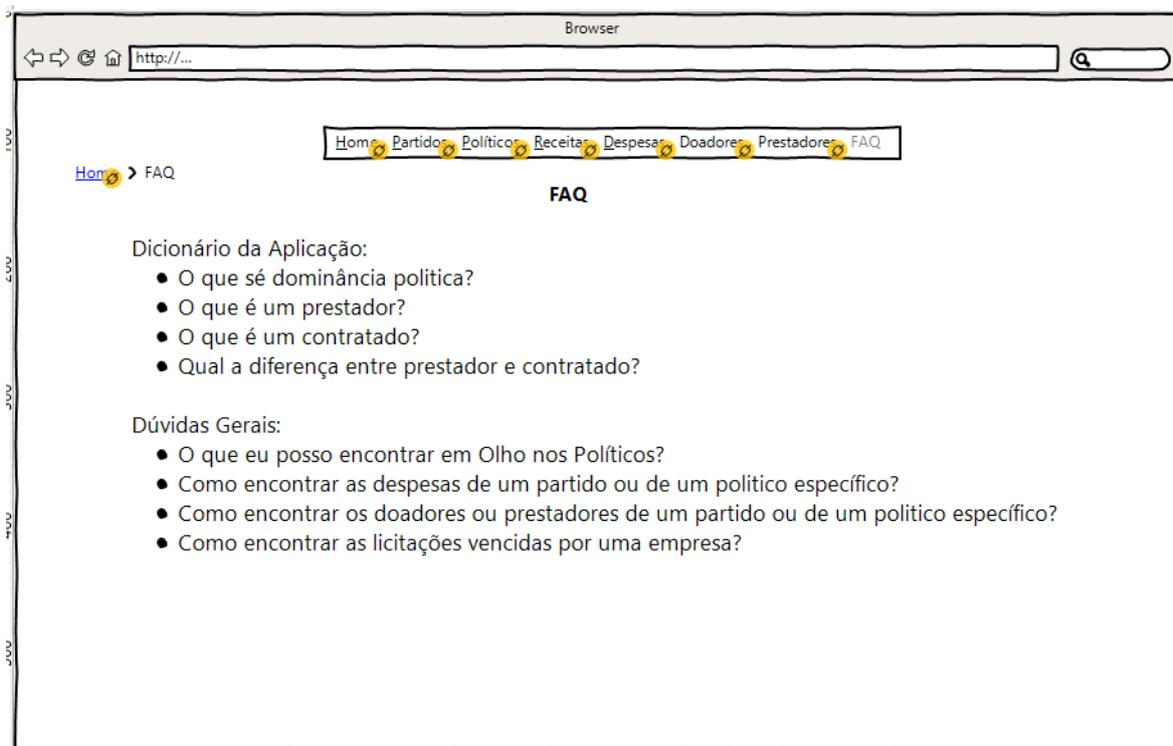


Figura 4.2: Nova Tela - FAQ

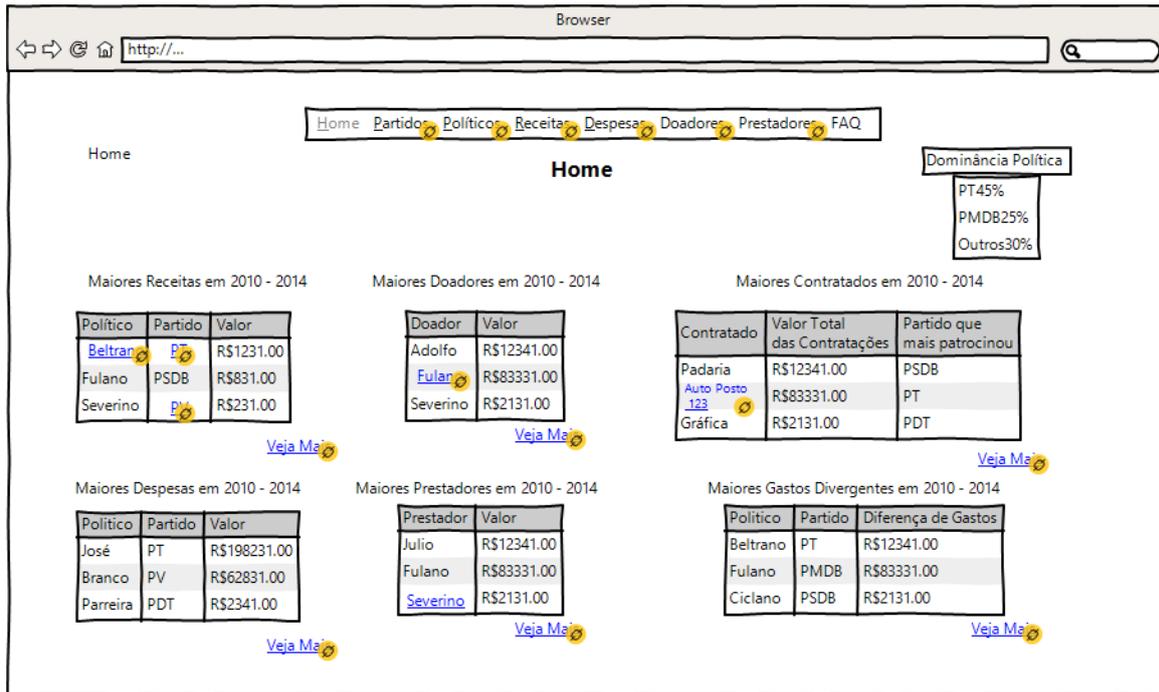


Figura 4.3: Nova Tela - Tela Inicial

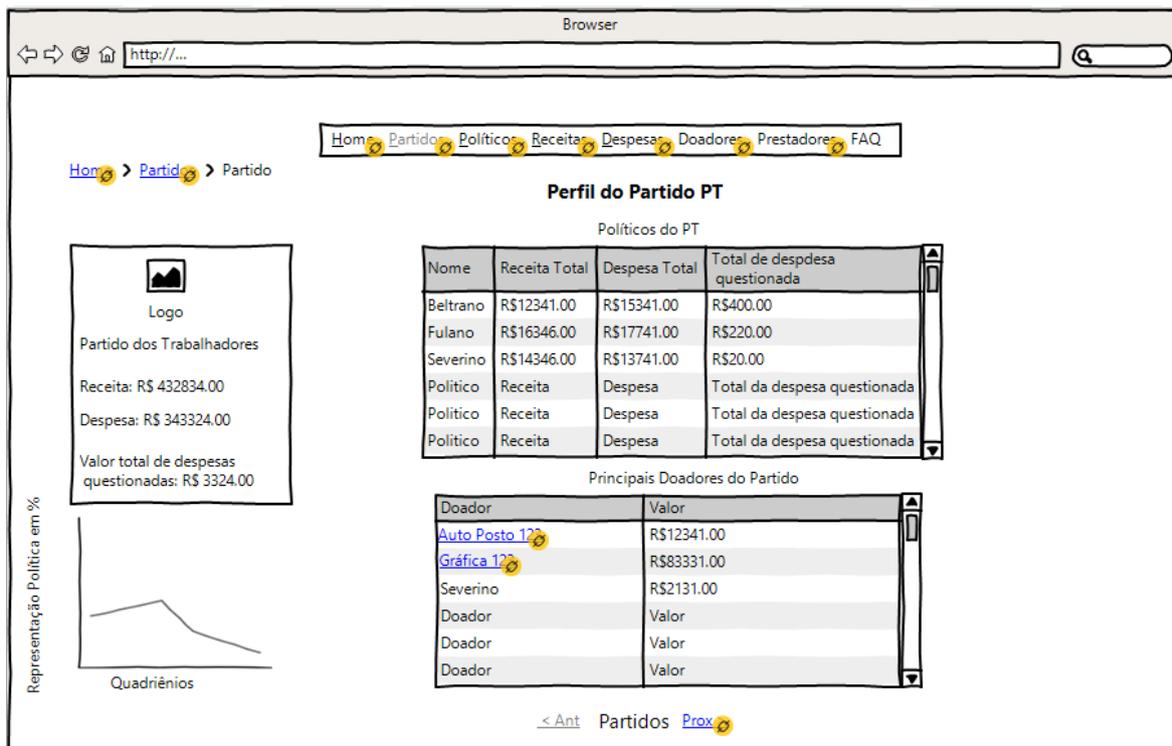


Figura 4.4: Nova Tela - Partido

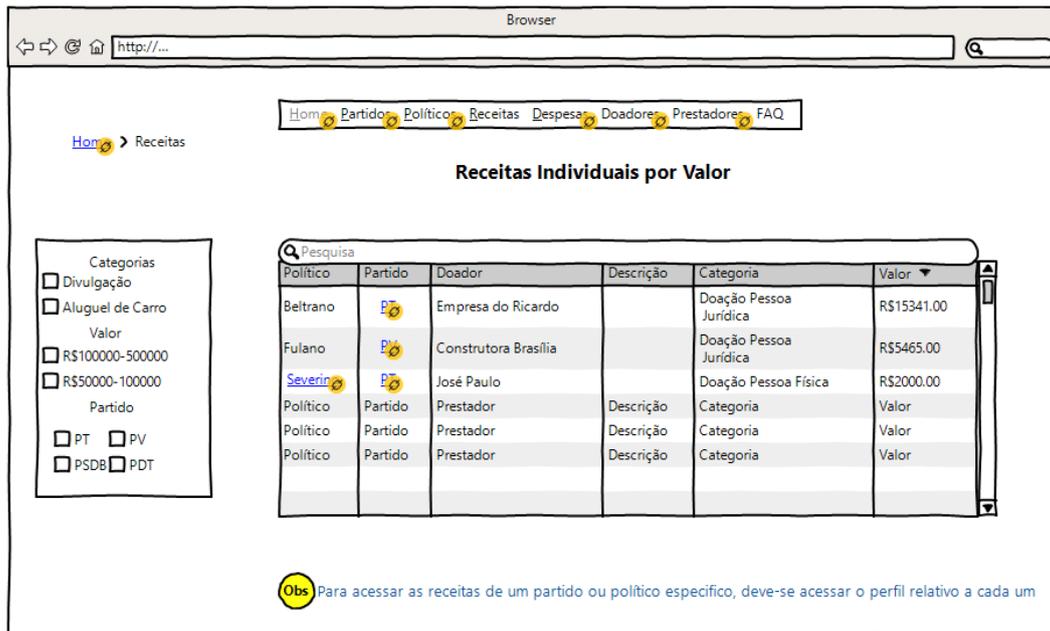


Figura 4.5: Nova Tela - Receitas

4.2 Ambiente Tecnológico Utilizado

Essa seção descreve o ferramental tecnológico utilizado para o início dos trabalhos que foram projetados no capítulo 3. Será descrito nesta seção desde o tratamento dos dados até as questões de linguagem de programação.

4.2.1 Google Refine

Um dos grandes desafios de trabalhar com dados abertos é a qualidade do que é fornecido. Apesar do portal brasileiro de dados abertos [22] fornecer uma boa variedade de dados em formato aberto, os dados apresentados precisam passar por um processo de limpeza para que efetivamente possam ser utilizados por uma aplicação. Abordando especificamente o conjunto de dados utilizados neste trabalho, prestações de contas fornecidas pelo site do TSE [65] não possui, por exemplo, um padrão de nomes de doadores. Uma mesma pessoa jurídica é apresentada de diversas formas, como é possível visualizar na Figura 4.6.

Uma ferramenta de grande ajuda no tratamento de dados é o *Google Refine* ou *Open Refine* [29]. Ela auxilia na limpeza desses dados permitindo que as instâncias semelhantes, como mostradas na Figura 4.6, possam ser agrupadas em uma só, dentre outras opções. Outro ponto muito importante na escolha da ferramenta foi a possibilidade de exportar os dados em um formato RDF a partir da extensão *RDF Refine* [47] da Universidade Nacional da Irlanda, Galway. A ferramenta possibilita ainda que os dados sejam ligados, fazendo com que partido, UF e cargo político remetam ao seu respectivo dado no DBpedia (Figura 4.7). Utilizando o *Google Refine* ainda foi possível relacionar as colunas com propriedades de outras ontologias e a criação de novas, como pode ser visto na Figura 4.8.

Cluster & Edit column "Nome do doador"

This feature helps you find groups of different cell values that might be alternative representations of the same thing. For example, the two strings "New York" and "new york" are very likely to refer to the same concept and just have capitalization differences, and "Gödel" and "Godel" probably refer to the same person. [Find out more ...](#)

Method Keying Function 973 clusters found

Cluster Size	Row Count	Values in Cluster	Merge?	New Cell Value
11	44	<ul style="list-style-type: none"> • CONSTRUÇÕES E COMÉRCIO CAMARGO CORREA S/A (14 rows) • CONSTRUÇÕES E COMERCIO CAMARGO CORREA S/A (13 rows) • CONSTRUÇÕES E COMERCIO CAMARGO CORREA S/A (8 rows) • CONSTRUÇÕES E COMÉRCIO CAMARGO CORREA S.A. (3 rows) • CONSTRUÇÕES E COMERCIO CAMARGO CORREA S/A (2 rows) • CONSTRUÇÕES E COMERCIO CAMARGO CORREA S-A (1 rows) • CONSTRUÇÕES E COMERCIO CAMARGO CORREA SA (1 rows) • CONSTRUÇÕES E COMERCIO CAMARGO CORREA S.A. (1 rows) • CONSTRUÇÕES E COMERCIO CAMARGO CORREA S/A (1 rows) • CONSTRUÇÕES E COMERCIO CAMARGO CORRÉA SA (1 rows) • CONSTRUÇÕES E COMÉRCIO CAMARGO CORRÉA S/A (1 rows) 	<input type="checkbox"/>	CONSTRUÇÕES E COMÉRCIO CA
8	21	<ul style="list-style-type: none"> • ACUCAR GUARANI S/A (7 rows) • ACUCAR GUARANI SA (6 rows) 	<input type="checkbox"/>	ACUCAR GUARANI S/A

Choices in Cluster

2 — 11

Rows in Cluster

0 — 390

Average Length of Choices

8 — 95

Length Variance of Choices

0 — 4.5

Select All Deselect All

Merge Selected & Re-Cluster Merge Selected & Close Close

Figura 4.6: Pessoa Jurídica Repetida

UF	uf_uri	Sigla Partido	partido_uri	Número candida	Cargo	cargoURI
São Paulo (Bundesstaat) Choose new match	http://dbpedia.org/resource/S%C3%A3o_Paulo_(state)	PMDB Choose new match	http://dbpedia.org/resource/PMDB	151	Senator Choose new match	http://dbpedia.org/resource/Senator
São Paulo (Bundesstaat) Choose new match	http://dbpedia.org/resource/S%C3%A3o_Paulo_(state)	PMDB Choose new match	http://dbpedia.org/resource/PMDB	151	Senator Choose new match	http://dbpedia.org/resource/Senator
São Paulo (Bundesstaat) Choose new match	http://dbpedia.org/resource/S%C3%A3o_Paulo_(state)	PMDB Choose new match	http://dbpedia.org/resource/PMDB	151	Senator Choose new match	http://dbpedia.org/resource/Senator
São Paulo (Bundesstaat) Choose new match	http://dbpedia.org/resource/S%C3%A3o_Paulo_(state)	PMDB Choose new match	http://dbpedia.org/resource/PMDB	151	Senator Choose new match	http://dbpedia.org/resource/Senator
São Paulo (Bundesstaat) Choose new match	http://dbpedia.org/resource/S%C3%A3o_Paulo_(state)	PMDB Choose new match	http://dbpedia.org/resource/PMDB	151	Senator Choose new match	http://dbpedia.org/resource/Senator
São Paulo (Bundesstaat) Choose new match	http://dbpedia.org/resource/S%C3%A3o_Paulo_(state)	PMDB Choose new match	http://dbpedia.org/resource/PMDB	151	Senator Choose new match	http://dbpedia.org/resource/Senator
São Paulo (Bundesstaat) Choose new match	http://dbpedia.org/resource/S%C3%A3o_Paulo_(state)	PMDB Choose new match	http://dbpedia.org/resource/PMDB	151	Senator Choose new match	http://dbpedia.org/resource/Senator
São Paulo (Bundesstaat) Choose new match	http://dbpedia.org/resource/S%C3%A3o_Paulo_(state)	PMDB Choose new match	http://dbpedia.org/resource/PMDB	151	Senator Choose new match	http://dbpedia.org/resource/Senator
São Paulo (Bundesstaat) Choose new match	http://dbpedia.org/resource/S%C3%A3o_Paulo_(state)	PMDB Choose new match	http://dbpedia.org/resource/PMDB	151	Senator Choose new match	http://dbpedia.org/resource/Senator

Figura 4.7: Dados Ligados Google Refine

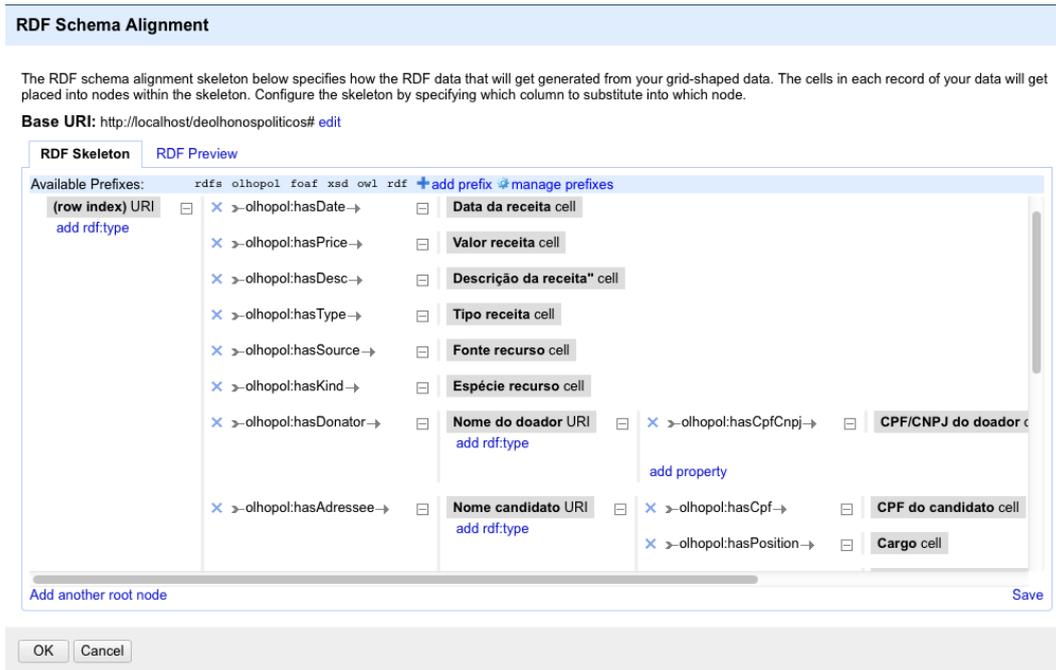


Figura 4.8: Alinhamento de RDF

4.2.2 Virtuoso

Um dos objetivos deste trabalho é utilizar as novas tecnologias semânticas que permitam o entendimento dos dados pelos computadores. Tendo em vista esse objetivo, foi utilizado um servidor de triplas, diferente dos relacionais que são mais comuns em aplicações Web na atualidade. Um dos requisitos mais importantes é que, além de armazenar a informação, o servidor seja capaz de realizar inferências, ou seja, ele tem de ter um sistema de inferências incluído em seu código. Três grandes aplicações merecem destaque nesse cenário: Sesame [1], Jena [28] e Virtuoso [61].

O Virtuoso é o servidor adotada pelo DBpedia e muitas outras grandes aplicações semânticas por ser um servidor universal. Ele compreende o armazenando em triplas, relacional, inferência, e ambiente de execução. Um dos fatores de destaque é que o Virtuoso possui uma versão *open source* e outra comercial. A arquitetura do Virtuoso pode ser observada na Figura 4.9.

O Virtuoso fornece um SPARQL *Endpoint*, que é um serviço que permite que um humano ou máquina consultem os dados através da linguagem SPARQL. Esse *Endpoint* possibilita que linguagens como o Python, PHP e outras possam realizar consultas semânticas através da biblioteca apropriada. Nesse trabalho a biblioteca para PHP SparqlLib [31] foi a escolhida para trabalhar junto com o Virtuoso.

4.2.3 Endpoint Olho nos Políticos

O Virtuoso fornece uma plataforma sólida de armazenamento e inferência, o que permite que um usuário que tenha conhecimento do vocabulário possa realizar consultas que não foram previstas durante o projeto da aplicação. Para fins de estudo e prototipação, foram utilizados três conjunto de dados: prestação de contas de políticos, gastos diretos e

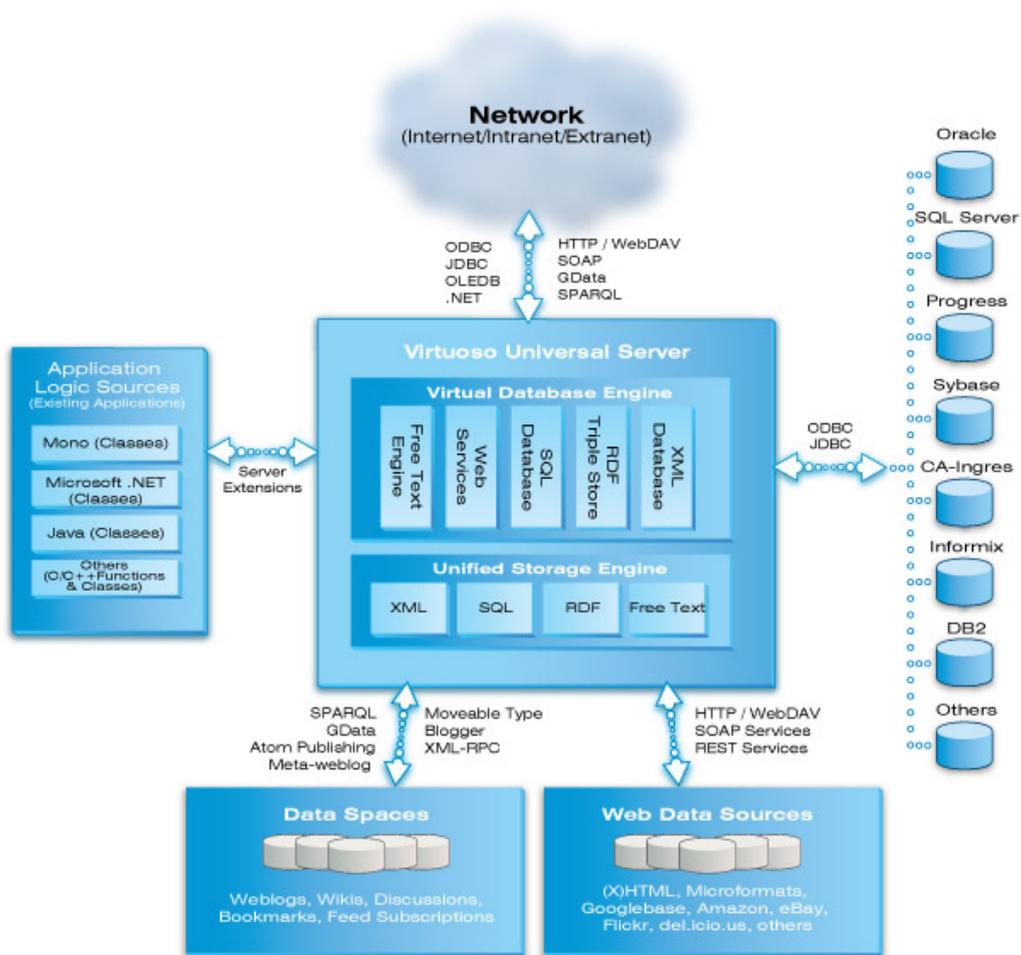


Figura 4.9: Arquitetura do Virtuoso. (Fonte: Virtuoso [61])

lista de empresas suspensas e inidôneas. Todos os conjuntos são abertos para uso público. Existe uma observação quanto ao conjunto de dados de gastos diretos que é parcialmente aberto, como informações sobre pessoas físicas que são omitidas, pois não são de propriedade do governo e assim não podem ser abertas para uso público. Foi desenvolvido um protótipo do que seria o *Endpoint* final da aplicação, a intenção é permitir que usuários possam procurar por conta própria informações que a aplicação não fornece diretamente, tanto por questões de projeto quanto por demanda dessa informação. A Figura 4.10 mostra a interface do *Endpoint* criado. A Figura 4.11 retrata o resultado de uma consulta que teve como objetivo identificar alguma empresa que já patrocinou campanha eleitoral de algum político, ganhou alguma licitação, e já foi classificada como empresa inidônea.

Endpoint Olho nos Políticos

Esse endpoint é um protótipo para consultas relativas ao conjunto de dados do Olho nos políticos.

Esse dataset contém informações triplicadas do [Portal de Dados Abertos Brasileiro](#) e do portal do [TSE](#).

```
SELECT Distinct ?Empresalni ?CNPJ
WHERE{
?Licitacao <http://localhost/deolhonospoliticos#hasWinner> ?Prestador.
?Prestador <http://localhost/deolhonospoliticos#hasCpfCnpj> ?CNPJ.
?Politico <http://localhost/deolhonospoliticos#hasDonator> ?Doador.
?Doador <http://localhost/deolhonospoliticos#hasCpfCnpj> ?CNPJ.
?Inidonea <http://localhost/olhonospoliticos/empresas#hasCnpjIni> ?CNPJ.
?Inidonea <http://localhost/olhonospoliticos/empresas#hasName> ?Empresalni.
}
```

Consultar

Figura 4.10: Consulta no *Endpoint*

Número de resultados: 1 resultados.

EmpresaIni	CNPJ
AUTOPEL AUTOMACAO COMERCIAL E INFORMATICA LTDA.	6698091000167

Figura 4.11: Resultado da Consulta no *Endpoint*

Capítulo 5

Considerações Finais

A utilização de dados abertos governamentais é um fenômeno que possibilita que a população tenha conhecimento das atividades governamentais, entretanto é necessário que existam conhecedores de domínio e de computação para que esse conhecimento tenha alcance adequado. A simples disponibilização dos dados não é suficiente para que a informação seja extraída facilmente daquele conjunto de dados.

Ao prototipar uma aplicação baseada nos dados disponibilizados, é importante planejá-la e desenvolvê-la seguindo princípios já existentes da computação, visando um bom desempenho, usabilidade e comunicabilidade. O uso de um método da Engenharia Web, como o *Object Oriented Hypermedia Design Method* (OOHDM), permite a construção de artefatos que auxiliam ao projetista criar aplicações mais robustas e com navegações mais complexas. A inclusão das facetas provenientes do método SHDM possibilita uma classificação do domínio que faça sentido e se mostrou positiva durante a avaliação do protótipo.

Uma avaliação de comunicabilidade como o Método de Avaliação de Comunicabilidade (MAC), da Interação Humano-Computador e da Engenharia Semiótica, permite que o projetista da aplicação avalie a informação que está sendo transmitida. Assim, pode-se evitar o retrabalho e propor uma solução mais intuitiva e com maior comunicabilidade. A partir dos resultados obtidos com o MAC, foi possível identificar algumas rupturas que mostraram que os termos utilizados não foram compreendidos adequadamente pelo usuário. Embora a navegação tenha sido projetada antes do desenvolvimento, o primeiro protótipo ainda apresentou falhas. Este resultado corrobora com a tese de que o MAC é um complemento importante para as abordagens de Engenharia Web no desenvolvimento de um protótipo de um sistema de informação.

Através do Virtuoso e a utilização do *Google Refine* [29] para limpar os dados e construir o esqueleto RDF para a posterior conversão dos dados para esse formato, foi possível inserir um significado semântico aos dados que antes eram difíceis de serem trabalhados. A aplicação de semântica nos dados do portal possibilitou uma inferência complexa fosse feita de forma bem simples. A partir de uma única consulta, foi possível encontrar um indicador de ações fraudulentas no conjunto de dados escolhidos. A grande dificuldade da aplicação dessa área envolve o tratamento dos dados e a escolha do ambiente de execução.

Devido a limitações de tempo para a realização do trabalho, outras rodadas do MAC que deveriam ter sido realizadas não o foram, pois cada uma das rodadas e suas respectivas análises demandariam muito tempo. O *Endpoint* construído se encontra somente

em servidor local, devido ao custo envolvido para contratar um servidor que oferecesse suporte para qualquer uma das ferramentas para a utilização de arquivos RDF. O cenário ideal seria a aplicação final implementada a partir das telas utilizadas nos protótipos e o *Endpoint* na mesma aplicação.

5.1 Resultados deste trabalho

Os resultados obtidos neste trabalho foram:

- Em Dados Abertos: foram utilizados dados de prestações de contas políticas eleitorais, gastos diretos e relação de empresas inidôneas para identificar um possível doador de campanha, que posteriormente foi contratado pelo governo e veio a ser considerado inidôneo.
- Em Engenharia Web: foram utilizados artefatos do OOHDM, os diagramas de classes conceituais, classes navegacionais e de contexto navegacional apresentados no capítulo 3 para dar suporte para o desenvolvimento do protótipo.
- Em Web Semântica: os conjuntos de dados utilizados não possuíam qualquer significado semântico. A semântica foi introduzida utilizando o *Google Refine* [29] e os dados foram ligados as páginas do DBpedia relativas a partido, cargo, estado. Os dados foram exportados em um formato de triplas RDF que permitiu o desenvolvimento de um *Endpoint* para consultas SPARQL.
- Em Interação Humano-Computador: o MAC foi aplicado no protótipo digital para avaliar a comunicabilidade do protótipo apresentado no capítulo 3.

5.2 Limitações

As limitações encontradas para a implementação deste trabalho foram:

- Falta de dados no formato RDF e atual formato dos dados disponibilizados através do portal Dados.gov.br [22]. O trabalho identificou que o conjunto de dados brasileiros se encontra no nível de duas estrelas, seguindo a classificação proposta por Berners-Lee [6].
- Alguns conjuntos de dados utilizados estão apenas parcialmente abertos, pois contêm informações que não são de domínio público.
- A limpeza de um grande conjunto de dados para que pudesse ser utilizado na prática.
- A transformação do formato dos dados para o formato RDF para que seja possível incluir um valor semântico.

5.3 Trabalhos Futuros

Os possíveis assuntos e trabalhos que podem, a partir deste trabalho, serem aprofundados e desenvolvidos são:

- O desenvolvimento de uma aplicação completa para utilização em larga escala da população que siga a abordagem utilizada nesse trabalho.
- Conversão total dos dados escolhidos para o formato RDF.
- Realizar novas rodadas de avaliação de comunicabilidade no protótipo após as correções dos problemas observados na primeira avaliação.

Apêndice A

Roteiro da Avaliação

O documento "Roteiro da Avaliação" tem como objetivo fornecer diretrizes para que o avaliador conduza a avaliação.

- Verificar se todos os equipamentos estão funcionando;
- Agradecer a participação da pessoa e ter um a conversa de aquecimento antes da avaliação em si;
- Entregar o Termo de Consentimento ao participante, junto com o Questionário Demográfico;
- Verificar se o participante está confortável com o computador utilizado, fornecer tanto o Mac quanto o Windows como opções;
- Ler o documento Informações ao Participante para o avaliado;
- O usuário deve começar quando o avaliador iniciar o cronômetro;
- Identificar os problemas de comunicabilidade e ambiguidade. Fazer as anotações;
- Quando o usuário terminar, o avaliador deve parar o tempo e anotar na Ficha do Observador;
- Após a avaliação, fornecer o questionário pós-avaliação;
- Agradecer a participação;
- Verificar se as informações estão identificadas corretamente e armazená-las corretamente.

Apêndice B

Termo de Consentimento

O documento "Termo de Consentimento" tem como objetivo explicar para o avaliado uma ideia sobre a avaliação, como ele pode proceder e também servir como autorização para que os avaliadores utilizem os dados coletados dele.

Você foi convidado(a) para participar de um teste de avaliação da Interação Humano Computador de um projeto de pesquisa que está utilizando o Método de Avaliação de Comunicabilidade (MAC) para o avaliar o projeto de uma aplicação Web. Neste teste, buscamos verificar a usabilidade e a comunicabilidade da interação do usuário com a aplicação, durante a realização de uma tarefa específica. Por essa razão, solicitamos seu consentimento para a realização desse teste e, em seguida, uma breve entrevista. Para tanto, é importante que você tenha algumas informações:

- Os dados coletados durante o teste destinam-se estritamente a atividades de pesquisa e desenvolvimento;
- A equipe dessa pesquisa tem o compromisso de divulgar os resultados obtidos para fins acadêmicos. A divulgação desses resultados pauta-se no respeito a sua privacidade e o anonimato é preservado em quaisquer documentos elaborados;
- O consentimento para o teste é uma escolha livre, feita mediante a prestação de todos os esclarecimentos necessários sobre a pesquisa;
- A realização do teste pode ser interrompida a qualquer momento, segundo a disponibilidade do participante. Nesse caso, a equipe se compromete a descartar o teste para fins da avaliação a que se destinaria;
- Nossa equipe encontra-se disponível para contato através do telefone (61) 9645-5962, ou pelo email lshimabuko@gmail.com. Caso queira tirar alguma dúvida, procure por Lucas.

De posse das informações acima, gostaríamos que você se pronunciasse acerca do teste.

Dou meu consentimento para sua realização.

Não autorizo sua realização.

Brasília, ____ de _____ de 2014

Nome do Participante

Nome do avaliador

Apêndice C

Questionário Demográfico

O documento "Questionário Demográfico" tem como objetivo coletar informações pessoais sobre o avaliado e nivelar seu conhecimento sobre o domínio política.

1. Gênero: () Masculino () Feminino

2. Idade: _____ anos

3. Escolaridade:

() Ensino Fundamental Incompleto

() Ensino Fundamental

() Ensino médio Incompleto

() Ensino médio

() Ensino superior incompleto

() Ensino superior completo. Qual? _____

() Pós-graduação (Lato-sensu). Qual? _____

() Pós-graduação (Stricto-sensu) - Mestre ou Doutor. Qual? _____

Para as questões 4 e 5 utilize as seguintes faixas:

1 - Menos de 1 ano;

2 - 1 a 2 anos;

3 - 3 a 4 anos;

4 - 5 a 6 anos;

5 - Mais de 6 anos;

4. Tempo de experiência com computadores: _____

5. Tempo de experiência com navegação em sites usando:

a) Desktops/Notebooks: _____

b) Tablets: _____

c) Smartphones: _____

Para a questão 6 utilize as seguintes faixas:

- 1 - Menos de 1 hora;
- 2 - 1 a 2 horas;
- 3 - 3 a 4 horas;
- 4 - 5 a 6 horas;
- 5 - Mais de 6 horas;
6. Tempo médio de uso diário de navegação em sites usando:
- a) Desktops/Notebooks: _____
- b) Tablets: _____
- c) Smartphones: _____
7. Qual grau de habilidade você considera que possui para navegar em sites na Internet?
- a) Em Desktops/Notebooks: () Pouca () Moderada () Muita
- b) Em Tablets: () Pouca () Moderada () Muita
- c) Em Smartphones: () Pouca () Moderada () Muita
8. O quanto você se sente familiarizado com o assunto política? Marque uma opção.
- () - Muito pouco
- () - Pouco
- () - Razoável
- () - Familiarizado
- () - Muito Familiarizado
9. Com que frequência você lê notícias sobre política? Marque uma opção.
- () - Nunca
- () - Ocasionalmente. Quando? _____
- () - Regularmente
10. Quais os principais veículo de informação que você utiliza para adquirir informações sobre política?
- TV - () Globo () Record () Band () Globo News () Record News () Band News
() Outro
- Se marcou outro, qual? _____
- Publicações Impressas - () Veja () Época () Correio Braziliense () Folha ()
Estadão () Outro
- Se marcou outro, qual? _____
- Internet - Qual Site? _____
- Outro - Qual? _____

11. Coloque os meios de comunicação da questão anterior em ordem de importância para você, sendo 1 a de maior importante e 4 a de menor importância:

- TV

- Publicações Impressas

- Internet

- Outros

12. Nos diga quais foram as últimas notícias sobre política que mais te chamou a atenção.

13. Quais atividades você realiza durante o período de eleições?

- Assistir debates

- Pesquisar sobre o passado do seu candidato

- Acompanhar notícias

- Outro. Qual? _____

14. Qual sua área de formação?

- Exatas. Qual? _____

- Humanas. Qual? _____

- Saúde. Qual? _____

- Nenhuma.

15. Qual a sua área de trabalho?

Apêndice D

Informações ao Participante

O documento "Informações ao Participante" é lido totalmente para o avaliado explicando detalhadamente sobre a avaliação.

1. Olá, meu nome é Pedro e este é Lucas, estamos fazendo uma monografia visando a conclusão de curso na área de Computação da Universidade de Brasília, e uma das etapas consiste em avaliação de comunicabilidade;
2. Isto significa que estamos avaliando um sistema computacional e as mensagens que os programadores desejam passar para as pessoas que utilizarão o sistema;
3. É importante ressaltar que não estamos avaliando você e sim o sistema computacional (definir comunicabilidade); A capacidade do sistema transmitir a mensagem que o projetista do sistema esperava.
4. Devemos informá-lo(a) que a tela do computador será gravada por um software e uma câmera será utilizada para registrar a tela, o teclado, o mouse e seus movimentos.
5. Essas imagens serão utilizadas apenas para fins acadêmicos, principalmente para que nós pesquisadores possamos avaliar a comunicabilidade do sistema. Suas imagens não serão divulgadas sem a sua permissão expressa;
6. Estamos procurando falhas na comunicação do sistema que serão encontradas no decorrer da tarefa. Para tanto solicitamos que você utilize uma técnica chamada think aloud (pense alto). Essa técnica consiste em que o usuário fale o que quer que esteja pensando durante a execução tarefa. Principalmente as dificuldades encontradas ao O propósito desse método é tornar explícita toda a informação que será útil para a avaliação. Ex.: (Na página do Google) Onde está o link para acessar o meu email? (clica em link errado) Não era aqui.
7. Não é interessante que durante a avaliação exista interação entre nós pesquisadores e você, assim não poderemos dialogar com você até o final do teste;
8. Para a avaliação você deverá realizar algumas tarefas pré definidas. Pedimos que você realize as tarefas na ordem proposta. Por favor, leia a folha de tarefas neste momento e nos explique o que você deverá fazer (PAUSA, tirar dúvidas apenas relativas a tarefa nunca sobre o sistema);

9. Por fim, informamos que a avaliação só terminará quando você disser que acabou as tarefas, ou que não gostaria de prosseguir realizando as tarefas.
10. As perguntas que surgirem durante a interação serão anotadas e respondidas ao final da avaliação do sistema. Isto se deve ao fato de que todos os avaliados precisam ter o mesmo nível de conhecimento do sistema ao iniciar a avaliação;
11. Por favor lembre-se de falar durante a avaliação, não se esqueça, pense alto! (think aloud);
12. A partir desse momento você pode começar a primeira tarefa.

Apêndice E

Tarefas

O documento "Tarefas" tem como objetivo listar todas as tarefas que o usuário deverá realizar na avaliação e ficará de posse do usuário até que este termine a avaliação.

Imagine que você é um eleitor que busca informações sobre os políticos que participarão da eleição da sua cidade. Você deve utilizar a aplicação para encontrar informações sobre as campanhas eleitorais de candidatos e partidos, e vencedores de licitações do período. Para isso você deve realizar as seguintes tarefas:

Obs.: O sistema permite que você insira um valor para um gasto de campanha caso discorde daquele que foi declarado pelo político, a média desses valores informados é chamada de valor corrigido. Exemplo: o político informa que gastou R\$ 50,00 em 10 litros de combustível. Se você discorda desse valor basta efetuar a ação de corrigir despesa e colocar o valor que acredita ser o correto.

- Tarefa 1:
Verificar quais são os partidos de maior receita no período.
- Tarefa 2:
Listar os doadores dos partidos PT e PV.
- Tarefa 3:
Listar as maiores licitações da empresa Auto Posto 123 no período.
- Tarefa 4:
Informar o valor corrigido para a despesa de valor mais alto no período.

Apêndice F

Questionário Pós-Teste

O documento "Questionário Pós-Teste" é aplicado ao usuário após a realização das tarefas e tem o objetivo de verificar o entendimento do protótipo e da realização de suas tarefas por parte do usuário.

1. Você considera que concluiu com sucesso as tarefas? Por favor marque abaixo para cada uma delas:

- a) Tarefa 1: () Não () Sim () Parcialmente
- b) Tarefa 2: () Não () Sim () Parcialmente
- c) Tarefa 3: () Não () Sim () Parcialmente
- d) Tarefa 4: () Não () Sim () Parcialmente

2. Por favor preencha o questionário de feedback a seguir. Gostaríamos de saber da sua opinião a respeito da interação. Avalie com notas de de 1 a 5, sendo 1 Muito Ruim e 5 Muito Bom:

1. Facilidade de aprendizado: avalie se as funcionalidades foram fáceis de serem aprendidas pela primeira vez, ou seja, se o protótipo do sistema é sugestivo e intuitivo;

2. Efetividade no uso: avalie o quão bom o protótipo do sistema é em fazer o que se propõe;

- Facilidade de Aprendizado: ()
- Efetividade no Uso: ()

3. Através do protótipo do sistema você percebeu alguma relação entre patrocinadores de campanha e vencedores de licitação? Se sim, qual?

4. Se você notasse que uma empresa que patrocinou uma campanha de um partido que está no poder no período ganhasse licitações, você consideraria ser um ato recriminável? Justifique.

Espaço livre para sugestões.

Referências

- [1] Aduna. Open RDF - Sesame. Disponível em: <http://www.openrdf.org>. [Acesso em Junho de 2014]. 60
- [2] S. Auer, C. Bizer, G. Kobilarov, J. Lehmann, R. Cyganiak, and Z. Ives. Dbpedia: A nucleus for a web of open data. In *The Semantic Web. Lecture Notes in Computer Science*, volume 4825, page 722, 2007. 6
- [3] M. Bac. Corruption, connections and transparency: Does a better screen imply a better scene? *Public Choice*, 107(1-2):87–96, 2001. [Acesso em Outubro de 2013]. 36
- [4] S.D.J Barbosa and B.S. Silva. *Interação Humano-Computador*. Editora Elsevier, 1 edition, 2010. v, 2, 20, 22, 23, 24, 25, 26, 27, 29, 54
- [5] C. Bellamy and J. Taylor. Informatisation and new public management: An alternative agenda for public administration. *Public Policy and Administration.*, 7(29):29–41, 1992. v, 8
- [6] T. Berners-Lee. Linked data. Disponível em: <http://www.w3.org/DesignIssues/LinkedData.html>. [Acesso em Outubro de 2013]. 12, 64
- [7] T. Berners-Lee, J. Hendler, and O. Lassila. The semantic web. *Scientific American*, 284(5):34–43, May 2001. [Acesso em Outubro de 2013]. v, 2, 13
- [8] Governo Brasileiro. Dados governamentais - gastos diretos. Disponível em: <http://www.portaltransparencia.gov.br/downloads/>. [Acesso em Outubro de 2013]. 44
- [9] D. Brickley and L. Miller. FOAF. Disponível em: <http://xmlns.com/foaf/spec/>. [Acesso em Outubro de 2013]. 15
- [10] Presidência da República. Lei complementar 131. Disponível em: http://www.planalto.gov.br/ccivil_03/leis/lcp/lcp131.htm. [Acesso em Outubro de 2013]. 12
- [11] Presidência da República. Lei nº 12.527. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/l12527.htm. [Acesso em Outubro de 2013]. 3, 12
- [12] Open Gov Data. 8 principles of open government data. Disponível em: <http://www.opengovdata.org/home/8principles>. [Acesso em Outubro de 2013]. 9

- [13] DBpedia. DBpedia. Disponível em: <http://wiki.dbpedia.org/>. [Acesso em Outubro de 2013]. v, 16
- [14] E. de Freitas. O que é corrupção? Disponível em: <http://www.mundoeducacao.com/geografia/o-que-corrupcao.htm>. [Acesso em Outubro de 2013]. 31
- [15] C. de Souza. *The Semiotic Engineering of Human-Computer Interaction*. The MIT Press, 2005. 27, 28, 29, 31
- [16] C. de Souza, J. Leite, R. Prates, and S. Barbosa. Projeto de interfaces de usuário: Perspectivas cognitivas e semióticas. *Anais das Jornadas de Atualização em Informática*, 4,5,6,7,9,24, 1999. [Acesso em Outubro de 2013]. 22, 23
- [17] Universidade Federal do Paraná. Reclamações Procon. Disponível em: <http://www.reclamacoesprocon.com.br>. [Acesso em Outubro de 2013]. vi, 41
- [18] D. Eaves. Case study: How open data saved canada 3.2 billion. Disponível em: <http://eaves.ca/2010/04/14/case-study-open-data-and-the-public-purse/>. [Acesso em Outubro de 2013]. 35
- [19] D. Eaves. The three laws of open government data. Disponível em: <http://eaves.ca/2009/09/30/three-law-of-open-government-data/>. [Acesso em Outubro de 2013]. 8
- [20] Grupo Estado. Basômetro. Disponível em: <http://estadaodados.com/basometro/>. [Acesso em Outubro de 2013]. vi, 42, 43
- [21] M. Ackerman et al. Corruption - how, and why, to avoid it. Technical report, Centre for Economic Research and Graduate Education, Charles University - Economic Institute, Academy Sciences, Prague - Zurich, June 1999. v, 6, 7
- [22] Governo Federal. Portal brasileiro de dados abertos. Disponível em: <http://dados.gov.br>. [Acesso em Junho de 2014]. 1, 58, 64
- [23] Governo Federal. Portal da transparência. Disponível em: <http://www.portaltransparencia.gov.br>. [Acesso em Junho de 2014]. 1
- [24] O. Fjeldstad and J. Andvig. Corruption: A review of contemporary research, 2001. Acesso em Outubro de 2013. 31
- [25] Open Knowledge Foundation. Open data handbook. Disponível em: http://opendatahandbook.org/pt_BR/what-is-open-data/index.html. [Acesso em Outubro de 2013]. 6
- [26] Open Knowledge Foundation. Open data index. Disponível em: <https://index.okfn.org/country/>. [Acesso em Outubro de 2013]. v, 34, 35
- [27] Open Knowledge Foundation. Where does my money go. Disponível em: <http://wheredoesmymoneygo.org/sources.html>. [Acesso em Outubro de 2013]. v, 37, 38

- [28] The Apache Software Foundation. Apache - Jena. Disponível em: <https://jena.apache.org>. [Acesso em Junho de 2014]. 60
- [29] Google. Open Refine/Google Refine. Disponível em: <https://github.com/OpenRefine>. [Acesso em Junho de 2014]. 58, 63, 64
- [30] T. Gruber. A translation approach to portable ontology specifications. *Knowledge Acquisition*, 5(2):199–220, June 1993. 15
- [31] C. Gutteridge. Sparql RDF library for PHP. Disponível em: <http://graphite.ecs.soton.ac.uk/sparqllib/>. [Acesso em Junho de 2014]. 60
- [32] Comunidade Transparência Hacker and OKFN. Queremos saber. Disponível em: <http://www.queremossaber.org.br>. [Acesso em Outubro de 2013]. v, 39, 40
- [33] T. Heath and C. Bizer. *Linked Data, Evolving the Web into a Global Data Space*. Books for Change, 2011. 14
- [34] J. Hendler, D. McGuinness, and L. Ding. Data-gov wiki. Disponível em: <http://data-gov.tw.rpi.edu/wiki>. [Acesso em Outubro de 2013]. v, 36, 37
- [35] T. Hewett and R. et al. Baecker. Curricula for human-computer interaction. Disponível em: <http://old.sigchi.org/cdg/>, 1992. Acesso em Outubro de 2013. 20
- [36] Transparency International. Corruption perception index. Disponível em: <http://www.transparencya.org/cpi2012/results>. [Acesso em Outubro de 2013]. v, 32, 33
- [37] IT4biz. Painel do PAC. Disponível em: <http://www.it4biz.com.br/apps/dados.gov.br/obrasdopac/>. [Acesso em Outubro de 2013]. v, 41
- [38] Justinmind. Justinmind prototyper. Disponível em: <http://www.justinmind.com/>. [Acesso em Outubro de 2013]. v, 26
- [39] R. E. Klitgaard. *Controlling corruption*. Berkeley: U. of California Press, 1998. 31
- [40] F. Lima. *Modelagem Semântica de Aplicações na WWW*. Doutorado em ciência da computação, Pontifícia Universidade Católica do Rio de Janeiro, 2003. 19, 20
- [41] F. Lima and D. Schwabe. Application modeling for the semantic Web. In *Web Congress, 2003. Proceedings. First Latin American*, pages 93–102. IEEE, 2003. [Acesso em Outubro de 2013]. 19
- [42] C. Lindstedt and D. Naurin. Transparency is not Enough: Making Transparency Effective in Reducing Corruption. *International Political Science Review*, 31(3):301–322, June 2010. [Acesso em Outubro de 2013]. 35
- [43] A. Maple. Faceted access: A review of the literature. Disponível em: <http://garamond.stanford.edu/depts/music/mlatest/BCC/BCC-Historical/Bcc95/95WGFAM2.html>. [Acesso em Outubro de 2013]. 20

- [44] R. Mateus, M. Ribeiro, and J. Vaz. New perspectives for electronic government in Brazil: The adoption of open government data in national and subnational governments of Brazil. In J. Gil-Garcia, N. Helbig, and A. Ojo, editors, *Proceedings of the 6th International Conference on Theory and Practice of Electronic Governance*, pages 22–29. ACM Press, October 2012. [Acesso em Outubro de 2013]. 12
- [45] T. Moran. The Command Language Grammar: a representation for the user interface of interactive computer systems. *International Journal of Man-Machine Studies*, 15(1):3–50, July 1981. Acesso em Outubro de 2013. 22
- [46] mySociety and Young Foundation. Fix my street. Disponível em: <http://www.fixmystreet.com/>. [Acesso em Outubro de 2013]. v, 38, 39
- [47] Galway National University of Ireland. RDF Refine. Disponível em: <http://refine.deri.ie>. [Acesso em Junho de 2014]. 58
- [48] S. Paul. *Holding the State to Account: Lessons of Bangalore's Citizen Report Cards*. Books for Change, 2002. 31
- [49] Rede Nossa São Paulo and W3C Brasil. Para onde foi o meu dinheiro? Disponível em: <http://www.paraondefoiomeudinheiro.org.br>. [Acesso em Outubro de 2013]. v, 40
- [50] C. Peirce. *The Essential Peirce: Selected Philosophical Writings*. Indiana University Press, 1998. 28
- [51] PoliGNU. Radar parlamentar. Disponível em: <http://radarparlamentar.polignu.org>. [Acesso em Outubro de 2013]. vi, 42, 43
- [52] R. Prates and S. Barbosa. *Introdução à teoria e prática da Interação Humano Computador fundamentada na Engenharia Semiótica*. Sociedade Brasileira de Computação e Editora PUC-Rio, 2007. 28, 30
- [53] Raquel O. Prates, Clarisse S. de Souza, and Simone D. J. Barbosa. Methods and tools: A method for evaluating the communicability of user interfaces. *interactions*, 7(1):31–38, January 2000. 29
- [54] J. Preece, Y. Sharp, D. Benyon, S. Holland, and T. Carey. *Human-Computer Interaction*. Assison-Wesley, 1994. 22
- [55] B. Reeves and C. Nass. *The Media Equation : How People Treat Computers, Television, and New Media Like Real People and Places*. Center for the Study of Language and Inf, 2003. Acesso em Outubro de 2013. 21
- [56] Y. Rogers, H. Sharp, and J. Preece. *Interaction Design*. Wiley, Reino Unido, 2011. 23, 25
- [57] G. Rossi, O. Pastor, D. Schwabe, and L. Olsina. *Web Engineering - Modelling and Implementing Web Applications*. Springer, 2007. 2, 18, 19
- [58] L. Santaella. *O que é Semiótica*. Editora Brasiliense, 1983. 28

- [59] D. Schwabe and G. Rossi. An object oriented approach to web-based applications design. *Theory and Practice of Object Systems*, 4(4):207–225, October 1998. [Acesso em Outubro de 2013]. 47, 48
- [60] H. Sharp, Y. Rogers, and J. Preece. *Interaction design: beyond human-computer interaction*. NY: John Wiley and Sons, 2007. 24
- [61] Open Link Software. Virtuoso open link. Disponível em: <http://virtuoso.openlinksw.com>. [Acesso em Junho de 2014]. vi, 60, 61
- [62] I. Sommerville. *Engenharia de Software*. Editora Pearson, 9 edition, 2010. 20
- [63] Stratebi and IT4biz. Aeroportos Brasil. Disponível em: <http://ison.stratebi.es/aerobrasil/>. [Acesso em Outubro de 2013]. vi, 42
- [64] J. Tauberer. *Open Government Data*. Joshua Tauberer, 1.1b edition, 2012. 10
- [65] TSE. Dados TSE. Disponível em: <http://www.tse.jus.br/eleicoes/estatisticas/repositorio-de-dados-eleitorais>. [Acesso em Outubro de 2013]. 44, 58
- [66] Grupo Virtuoso. Semântica. Disponível em: <http://www.soportugues.com.br/secoes/seman/>. [Acesso em Outubro de 2013]. 13
- [67] W3C. RDF. Disponível em: <http://www.w3.org/RDF/>. [Acesso em Outubro de 2013]. 1, 14
- [68] W3C. SPARQL protocol for RDF. Disponível em: <http://www.w3.org/TR/rdf-sparql-protocol/>. [Acesso em Outubro de 2013]. v, 16, 17
- [69] W3C. W3C - Vocabulary. Disponível em: <http://www.w3.org/standards/semanticweb/ontology>. [Acesso em Outubro de 2013]. 15
- [70] Wikipedia. Ontologia FOAF. Disponível em: [http://en.wikipedia.org/wiki/FOAF_\(ontology\)](http://en.wikipedia.org/wiki/FOAF_(ontology)). [Acesso em Outubro de 2013]. v, 15
- [71] WireFrameSketcher. Wireframesketcher download. Disponível em: <http://wireframesketcher.com/download.html>. [Acesso em Outubro de 2013]. 50
- [72] Distance Learning Program Youth for Governance. Introduction to corruption. Disponível em: <http://info.worldbank.org/etools/docs/library/35970/mod03.pdf>. [Acesso em Outubro de 2013]. 31