

UNIVERSIDADE DE PASSO FUNDO
INSTITUTO DE CIÊNCIAS EXATAS E GEOCIÊNCIAS
PROGRAMA DE PÓS-GRADUAÇÃO EM COMPUTAÇÃO APLICADA

OPEN SMART CITY VIEW

**Uma Solução Computacional para Manipulação
e Apresentação de Dados Governamentais**

Abertos

Diego Antonio Lusa

Dissertação apresentada como requisito parcial
à obtenção do grau de Mestre em Computação
Aplicada na Universidade de Passo Fundo.

Orientador: Prof. Dr. Roberto dos Santos Rabello

Coorientador: Prof. Dr. Cristiano Roberto Cervi

Passo Fundo

2016

CIP – Catalogação na Publicação

L968o Lusa, Diego Antonio
Open smart city view : uma solução computacional para
manipulação e apresentação de dados governamentais /
Diego Antonio Lusa. – 2016.
82 f. : il. color. ; 30 cm.

Orientador: Prof. Dr. Roberto dos Santos Rabello.
Coorientador: Prof. Dr. Cristiano Roberto Cervi.
Dissertação (Mestrado em Computação) – Universidade
de Passo Fundo, 2016.

1. Programas de computador. 2. Proteção de dados. 3.
Governo. I. Rabello, Roberto dos Santos, orientador. II.
Cervi, Cristiano Roberto, coorientador. III. Título.


CDU: 004.42

Catalogação: Bibliotecária Marciéli de Oliveira - CRB 10/2113

**ATA DE DEFESA DO
TRABALHO DE CONCLUSÃO DE CURSO DO ACADÊMICO**

DIEGO ANTONIO LUSA

Aos seis dias do mês de janeiro do ano dois mil e dezesseis, às 9 horas, realizou-se, no Instituto de Ciências Exatas e Geociências, prédio B5, da Universidade de Passo Fundo, a sessão pública de defesa do Trabalho de Conclusão de Curso **“OPEN SMART CITY VIEW: Uma Solução Computacional para Manipulação e Apresentação de Dados Governamentais Abertos”**, de autoria de Diego Antonio Lusa, acadêmico do Curso de Mestrado em Computação Aplicada do Programa de Pós-Graduação em Computação Aplicada – PPGCA/UPF. Segundo as informações prestadas pelo Conselho de Pós-Graduação e constantes nos arquivos da Secretaria do PPGCA, o aluno preencheu os requisitos necessários para submeter seu trabalho à avaliação. A banca examinadora foi composta pelos doutores Roberto dos Santos Rabello, Cristiano Roberto Cervi, Ana Carolina Bertoletti De Marchi e Carina Friedrich Dorneles. Concluídos os trabalhos de apresentação e arguição, a banca examinadora considerou o candidato APROVADO. Foi concedido o prazo de até quarenta e cinco (45) dias, conforme Regimento do PPGCA, para o acadêmico apresentar ao Conselho de Pós-Graduação o trabalho em sua redação definitiva, a fim de que sejam feitos os encaminhamentos necessários à emissão do Diploma de Mestre em Computação Aplicada. Para constar, foi lavrada a presente ata, que vai assinada pelos membros da banca examinadora e pela Coordenação do PPGCA.



Prof. Dr. Roberto dos Santos Rabello
Presidente da Banca Examinadora
(Orientador)



Prof. Dr. Cristiano Roberto Cervi
(Coorientador)



Prof. Dra. Carina Friedrich Dorneles – UFSC
(Avaliador Externo)



Prof. Dra. Ana Carolina Bertoletti De Marchi
(Avaliador Interno)



Prof. Dr. Carlos Amarel Hölbig
Coordenador do PPGCA

Dedico este trabalho, em primeiro lugar, a Deus, a quem devo a oportunidade e sabedoria tão essenciais para dar cabo à tarefa. A Ele devo também a força e a esperança que dia após dia me alimentam com a certeza das coisas que estão por vir, pela excelência da fé. De mesmo modo, dedico a minha esposa amada, que com grande compreensão soube entender meus períodos de ausência, sempre me incentivando com zelo e amor. A você, Sandra, meu amor, meu respeito e a minha certeza de que muitas outras vitórias ainda estão por vir. Dedico também este trabalho a meus pais, Ivanir e Marines, base daquilo que sou. Agradeço pelo amparo, amor e carinho e pelo incentivo que sempre abundou em todos os projetos que realizei em minha vida. A minha conquista certamente é fruto da força e do amor que sempre nos uniu.

AGRADECIMENTOS

Meu especial agradecimento ao Prof. Dr. Roberto dos Santos Rabello e Prof. Dr. Cristiano Roberto Cervi pelo inestimável auxílio conferido durante o desenvolvimento deste trabalho. Também estendo meu agradecimento à Prefeitura Municipal de Passo Fundo, na pessoa da senhora Eliana Fátima De Zorzi, Coordenadora de Tecnologia da Informação, por disponibilizar seu tempo para nos atender pessoalmente e pela cordialidade e disposição em nos responder a todas as dúvidas levantadas durante a execução do trabalho. Por fim, agradeço a todos os colegas pelo companheirismo demonstrado no decorrer destes dois anos. Certamente a etapa se finda mas a amizade permanece.

OPEN SMART CITY VIEW

RESUMO

O cidadão inteligente, no exercício das suas atribuições frente a cidade inteligente, empodera-se por meio das informações que tem a seu dispor. Dentre as fontes de informação disponíveis, encontram-se os dados governamentais abertos. Todavia, tais dados são disponibilizados à sociedade em seu estado bruto e em diferentes sítios, o que impõem certas dificuldades para seu consumo. Em vista disso, neste trabalho apresenta-se a arquitetura Open Smart City View, um arcabouço arquitetural que agrega técnicas de integração de dados e data warehousing para dar condições a visualização integrada e simplificada de informações oriundas de conjuntos de dados governamentais abertos coletados de diversas fontes. Dividida em quatro camadas funcionais, a arquitetura apresenta-se como uma ferramenta de empoderamento que abstrai os aspectos técnicos associados à coleta e consumo dos conjuntos de dados ao passo que mantém o foco do cidadão no entendimento das informações extraídas. Desta forma, a arquitetura vai ao encontro do maior empoderamento do cidadão por meio de informações de qualidade, corroborando diretamente com as iniciativas *open data* e de transparência pública, bem como com a consolidação do perfil do cidadão inteligente.

Palavras-Chave: cidades inteligentes, cidadão inteligente, dados governamentais abertos, empoderamento.

OPEN SMART CITY VIEW

ABSTRACT

The smart citizen, in the exercise of its duties front of smart city, empowers itself through the information it has at its disposal. Among the available informations sources are the open government data. However, such data is made available to society in its raw state and in different places, which impose certain difficulties in its consumption. In view of this, the present work introduces the Open Smart City View architecture, an architectural framework that aggregates data integration techniques and data warehousing to give conditions to integrated and simplified view of information coming from open government data sets collected from various sources. Divided into four functional layers, the architecture presents itself as an empowerment tool that abstracts the technical aspects associated with the collection and use of data sets while maintaining the focus of the citizens' understanding of the information extracted. Thus, the architecture is consistent with the greater empowerment of citizens through quality information, corroborating directly with the open data and public transparency initiatives, as well as with the consolidation of the smart citizen profile.

Keywords: smart cities, smart citizens, open government data, empowerment.

LISTA DE FIGURAS

Figura 1.	Atividades fundamentais de uma cidade inteligente. Fonte primária.	21
Figura 2.	Modelo de responsabilidades × aspectos habilitadores do <i>Smart City Council</i> . Adaptado de [11].	23
Figura 3.	Relacionamento entre os eixos de atuação e sistemas de uma cidade inteligente [18].	24
Figura 4.	Diagrama esquemático de um cubo OLAP. Fonte primária.	30
Figura 5.	Representação esquemática de registro fictício de tabela fato. Fonte primária.	31
Figura 6.	Apresentação gráfica do relacionamento entre tabelas fato e tabelas dimensão em um esquema do tipo estrela. Fonte primária.	32
Figura 7.	Modelo simplificado das camadas funcionais da arquitetura OSCV. Fonte primária.	35
Figura 8.	Camada Fontes de Dados. Fonte primária.	38
Figura 9.	Camada Coleta e Processamento. Fonte primária.	40
Figura 10.	PDI - Exemplo de <i>transformation</i> e seus componentes de construção. Fonte primária.	43
Figura 11.	PDI - Exemplo de <i>job</i> e seus componentes de construção. Fonte primária.	43
Figura 12.	Protocolo de desenvolvimento de integrações ETL na arquitetura OSCV. Fonte primária.	46
Figura 13.	Diagrama conceitual da Camada Web. Fonte primária.	48
Figura 14.	Inter-relacionamento entre as metáforas de interação da Camada Web. Fonte primária.	50
Figura 15.	Diagrama de classes - conceitos de apresentação de fato, perspectivas e perspectivas relacionadas. Fonte primária.	53
Figura 16.	Seção de critérios de filtragem - Tela de apresentação de informações. Fonte primária.	58
Figura 17.	Diagrama de sequência simplificado do processo de renderização de visualizações. Fonte primária.	59
Figura 18.	Seção de renderização - Tela de apresentação de informações. Fonte primária.	60
Figura 19.	Seção de renderização contendo duas visualizações - Tela de apresentação de informações. Fonte primária.	61
Figura 20.	Camada Dispositivos Interação. Fonte primária.	62

Figura 21.	Tela de listagem das apresentações de fato cadastradas na plataforma.	
	Fonte primária.	73
Figura 22.	Tela de manutenção de apresentações de fato. Fonte primária.	74
Figura 23.	Tela de manutenção de perspectivas. Fonte primária.	75

LISTA DE TABELAS

Tabela 1.	Lista de características desejáveis para identificação de fontes de dados na arquitetura OSCV. Fonte primária.	37
Tabela 2.	Conjunto de dados das Fontes de Interesse. Fonte primária.	39
Tabela 3.	Listagem de atributos do conceito "Apresentação de Fato". Fonte primária.	54
Tabela 4.	Exemplo de comando SQL de extração de dados associado a uma perspectiva. Fonte primária.	55
Tabela 5.	Listagem de atributos do conceito "Perspectiva". Fonte primária.	56
Tabela 6.	Trecho do arquivo contendo dados de crimes em municípios - SSP/RS	76
Tabela 7.	Trecho do arquivo contendo dados da legislação de municípios brasileiros - IBGE.	77
Tabela 8.	Trecho do arquivo contendo dados de violência contra a mulher - DADOS.RS	78
Tabela 9.	Trecho do arquivo contendo dados de crimes de trânsito - DADOS.RS	79
Tabela 10.	Trecho do arquivo contendo dados de Produto Interno Bruto de municípios - FEE	80
Tabela 11.	Trecho do arquivo contendo dados demográficos - FEE	81
Tabela 12.	Trecho do arquivo contendo dados de frota de veículos - DENATRAM	82

LISTA DE SIGLAS

TIC – Tecnologias de Informação e Comunicação
DGA – Dados Governamentais Abertos
OSCV – Open Smart City View
IBM – International Business Machines
W3C – World Wide Web Consortium
ISO – International Standardization Organization
URI – Uniform Resource Identifier
ROLAP – Relational Online Analytical Processing
SGBD – Sistema Gerenciador de Banco de Dados
ONG – Organização Não Governamental
API – Application Programming Interface
JSF – Java Server Faces
MVC – Modelo-Visão-Controlador
PNG – Portable Network Graphics
SEMISH – Seminário Integrado de Software e Hardware
ICWI – International Conference WWW/Internet
CAPES – Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
SSP – Secretaria de Segurança Pública
IBGE – Instituto Brasileiro de Geografia e Estatística
FEE – Fundação de Economia e Estatística Siegfried Emanuel Heuser
DENATRAM – Departamento Nacional de Trânsito

SUMÁRIO

1	INTRODUÇÃO	14
2	FUNDAMENTAÇÃO TEÓRICA	17
2.1	CIDADES INTELIGENTES E SEUS ASPECTOS PROMOTORES	18
2.2	CIDADES INTELIGENTES E SEUS OBJETIVOS	20
2.3	CIDADES INTELIGENTES E SUAS ÁREAS-CHAVE	22
2.4	CIDADES INTELIGENTES, GOVERNO PARTICIPATIVO E O PAPEL DOS DADOS GOVERNAMENTAIS ABERTOS	25
2.5	DATA WAREHOUSE, FATOS E DIMENSÕES	29
2.6	TRABALHOS RELACIONADOS	32
3	ARQUITETURA OPEN SMART CITY VIEW	35
3.1	CAMADA FONTES DE DADOS	36
3.1.1	Fontes de Interesse Iniciais da Arquitetura OSCV	38
3.2	CAMADA COLETA E PROCESSAMENTO	40
3.2.1	Tabelas de Interface	45
3.3	CAMADA WEB	47
3.3.1	Apresentação de Fatos, Perspectivas e Perspectivas Relacionadas	49
3.3.2	Visões Funcionais	51
3.3.2.1	Visão de Backend	52
3.3.2.2	Visão de Frontend	57
3.4	CAMADA DISPOSITIVOS DE INTERAÇÃO	61
3.5	CONSTATAÇÕES, PERCEPÇÕES E EXPECTATIVAS	63
4	CONSIDERAÇÕES FINAIS	65
4.1	DISSEMINAÇÃO DO CONHECIMENTO	66
4.2	TRABALHOS FUTUROS	67
	REFERÊNCIAS	69
	APÊNDICE A – Backend - Gestão de apresentação de fatos	73
	APÊNDICE B – Backend - Manter apresentação de fatos	74

APÊNDICE C – Backend - Manter perspectivas	75
ANEXO A – SSP: Conjunto de dados sobre crimes	76
ANEXO B – IBGE: Conjunto de dados sobre legislação de municípios	77
ANEXO C – DADOS.RS: Conjunto de dados sobre violência contra a mulher . . .	78
ANEXO D – DADOS.RS: Conjunto de dados sobre crimes de transito	79
ANEXO E – FEE: Conjunto de dados sobre Produto Interno Bruto	80
ANEXO F – FEE: Conjunto de dados sobre demografia	81
ANEXO G – DENATRAM: Conjunto de dados sobre frota de veículos	82

1. INTRODUÇÃO

Nos primórdios da civilização humana, os indivíduos que residiam nas cidades representavam uma pequena parcela da população total. Grande parte das pessoas da época encontrava-se no campo, trabalhando a terra e subsistindo de atividades primárias, como a agricultura. No decorrer do tempo, este cenário foi sendo alterado e as cidades gradativamente passaram a galgar lugar de destaque, até efetivamente tornarem-se a expressão máxima da sociedade contemporânea.

Atualmente, para mais de 50% da população mundial [1, 2], as cidades representam uma real visão de liberdade, criatividade, oportunidade e prosperidade [3]. Isso equivale a cerca de 3.5 bilhões de pessoas vivendo em apenas 2% da superfície total do planeta, consumindo 75% dos recursos energéticos e emitindo cerca de 80% de todo o carbono produzido no mundo [4].

Ao agregar tamanhas populações em espaços deveras reduzidos, as dificuldades naturalmente enfrentadas pelas cidades crescem exponencialmente, tornando a garantia da qualidade de vida e de trabalho dos cidadãos um objetivo difícil de ser alcançado. Tal cenário em muito decorre do desequilíbrio entre o crescimento populacional e a oferta de infraestrutura básica e prestação de serviços. A crescente demanda imposta pela quantidade cada vez maior de usuários, aliado às expectativas crescentes por serviços mais eficientes fazem com que as cidades operem constantemente no limite de seus recursos.

Em decorrência da sobrecarga infligida as cidades, surgem diversos problemas no meio urbano, como aqueles relacionados ao transporte público, saúde, educação e segurança, por exemplo. Como resultado, cria-se um ambiente caótico e estressante que reduz a qualidade de vida dos cidadãos, tornando as cidades, em especial, as de maior população, um local pouco aprazível para se viver.

Para fazer frente aos desafios impostos pelas superpopulações, as cidades se veem obrigadas a gerir com mais inteligência seus recursos e infraestruturas. Neste sentido, as Tecnologias de Informação e Comunicação (TIC) destacam-se como uma importante ferramenta de apoio, ao passo que seu uso busca capacitar os diversos sistemas urbanos por meio da coleta, distribuição e processamento dos dados de forma automática e em tempo real.

Desta forma, o uso de modernas tecnologias no meio urbano, associado a investimentos no elemento social e humano, visam prover às cidades meios para um crescimento mais sustentável e aos cidadãos, maior qualidade de vida e de trabalho. Essa nova concepção dada às cidades, fortemente baseada no uso das TICs, deu origem ao que hoje denomina-se **cidade inteligente** [5].

As cidades inteligentes, além de eficientes, precisam preservar aspectos fundamentais da vida em sociedade, como a espontaneidade, a serendipidade e a sociabilidade de seus cidadãos [1]. Adicionalmente, o cidadão é visto como o elemento diferencial da cidade inteligente [6] e sua interação com o meio urbano caracteriza-se principalmente pelo constante consumo das informações e proximidade com o poder público [7], características que lhe conferem a designação de “**cidadão inteligente**”.

Para desempenhar seu papel, o cidadão inteligente pode consumir informações de diversas fontes. A informação, portanto, representa um de seus principais insumos de empoderamento e a disponibilidade de fontes de informações de qualidade define um importante instrumento de potencialização e qualificação de seu papel na cidade inteligente. Dentre tais fontes de informação, merecem especial atenção àquelas relacionadas aos dados produzidos pelos órgãos públicos durante a execução de suas atividades. Estes dados, em muitos países do globo, tem acesso público garantido por lei, de modo que qualquer cidadão tem direito a obtê-los. Ao tornarem-se públicos, os conjuntos de dados produzidos pelas entidades governamentais recebem o nome de Dados Governamentais Abertos (DGA), representando um importante insumo de empoderamento a disposição do cidadão.

Contudo, embora disponibilizados aos cidadãos como subsídio para um maior empoderamento, os dados governamentais abertos muitas vezes têm sua eficácia limitada pelas dificuldades associadas a sua coleta, tratamento e agregação. Essas dificuldades, de modo geral, relacionam-se com a natureza bruta das informações disponibilizadas, bem como com a carência de ferramentas de software aptas a consumir tais dados e disponibilizar ao cidadão as informações resultantes de modo mais natural e simplificado.

Assim sendo, se observado que as cidades inteligentes pressupõem uma mudança cultural, impulsionada por iniciativas de maior empoderamento e participação dos cidadãos na comunidade [8, 7], a construção de ferramentas que facilitam o acesso, consumo e o entendimento dos dados públicos abertos não é apenas desejável mas também necessária. A importância da criação e uso de tais ferramentas vai além dos aspectos técnicos e fundamenta-se na contribuição que seu uso traria à consolidação do papel do cidadão em sua comunidade, servindo ao processo de inclusão social e política por meio do empoderamento pela informação.

A grande questão, contudo, está em coletar e transformar os conjuntos de dados abertos existentes em informação de valor para o cidadão. A multiplicidade de fontes, formatos e representações conceituais dos dados impõem desafios a qualquer ferramenta de software que tenha por objetivo sua manipulação. De modo complementar, a forma de apresentação das informações também representa um grande desafio, visto que o conteúdo precisa servir a construção de novos conhecimentos, a curiosidade e ao desejo de participação ativa nas questões da comunidade por parte do cidadão.

Imbuído neste desafio, o presente trabalho visa apresentar uma arquitetura de software voltada especificamente à coleta, tratamento e apresentação de informações oriundas de conjuntos de dados governamentais abertos coletados de diversas fontes. Desta forma, constituem-se como objetivos específicos do trabalho a identificação de fontes de dados governamentais abertos de interesse, a captura e processamento dos conjuntos de dados elencados de tais fontes por meio de técnicas de ETL, a averiguação da viabilidade de uso de técnicas de Inteligência Artificial e Big Data na concepção da arquitetura e, por fim, a aplicação dos princípios de dados governamentais abertos durante o processo de concepção do sistema de software baseado na arquitetura proposta.

Denominada por Open Smart City View (OSCV), a arquitetura proposta busca dar condições ao cidadão de conhecer e avaliar a realidade do seu município por meio da visualização

integrada e simplificada das informações coletadas dos conjuntos de dados abertos, corroborando assim com o processo de empoderamento do cidadão. Constituída por diferentes camadas funcionais, a arquitetura OSCV agrega modernas tecnologias de integração de dados, *data warehousing* e desenvolvimento Web para fornecer ao cidadão uma eficaz ferramenta de empoderamento. Desta forma, por meio da apresentação simplificada das informações municipais, a arquitetura busca colaborar com o aprimoramento da atuação do cidadão inteligente frente as diversas questões sociais, políticas e econômicas que o cercam no âmbito de seu município.

A fim de melhor descrever a arquitetura OSCV e os conceitos que a envolvem, o presente trabalho encontra-se organizado conforme segue: no Capítulo 2 apresenta-se o referencial teórico que contextualiza os conceitos de cidade inteligente, cidadão inteligente e dados governamentais abertos. Também apresenta-se nesta mesma seção trabalhos relacionados com a temática abordada. Na sequência, o Capítulo 3 descreve em detalhes a arquitetura OSCV, suas camadas funcionais e os aspectos técnicos relevantes para sua concepção e funcionamento. Por fim, os meios de disseminação do conhecimento, trabalhos futuros e considerações finais são abordados por meio do Capítulo 4.

2. FUNDAMENTAÇÃO TEÓRICA

Em 2008, pela primeira vez na história da humanidade, 50% da população mundial encontrava-se nas áreas urbanas [9, 1, 2]. Em termos de cenário futuro, segundo estimativas, 80% dos indivíduos estarão vivendo nas cidades até o final do presente século [10], situação praticamente oposta há de três séculos atrás, quando a representatividade da população urbana frente à população mundial era de apenas 5% [10]. Este intenso fluxo migratório observado, além de confirmar uma tendência consolidada há algumas décadas, traz consigo grandes desafios para os centros urbanos, impondo aos sistemas essenciais e de apoio das cidades acentuada sobrecarga.

O processo migratório observado está intimamente ligado as novas oportunidades que as cidades têm oferecido aos cidadãos. Bons empregos, acesso a saúde e educação de qualidade, entretenimento e cultura são alguns dos benefícios almejados pelos indivíduos migrantes [11]. Neste movimento, as grandes cidades tornam-se o destino final escolhido pela grande maioria, fato comprovado pelas estimativas de crescimento urbano que atribuem a apenas 600 cidades a responsabilidade por $\frac{2}{3}$ do crescimento mundial até 2025 [2].

As estimativas apenas reforçam um movimento social facilmente perceptível, no qual as grandes cidades atuam como verdadeiros imãs, atraindo grande número de pessoas todos os anos, para integrar os já populosos centros urbanos. Parte disso se deve a capacidade que as grandes cidades apresentam de viabilizar uma expressiva escala de produção, ostentar um estilo de vida diferenciado e armazenar grande quantidade de riqueza na medida que mantêm considerável capacidade de resistência a fenômenos adversos [12]. Todos estes ingredientes apenas evidenciam o papel de destaque conquistado pelas grandes cidades e acabam por fomentar o ingresso constante de novos indivíduos aos grandes centros urbanos.

Como resultado deste crescimento urbano desenfreado, as grandes cidades se veem diante de desafios de proporções gigantescas, tanto em termos de espaço físico quanto de gestão, segurança, serviços e uso de recursos naturais. É necessário garantir que serviços essenciais como os de distribuição de água, energia elétrica e coleta de esgoto funcionem a contento e não entrem em colapso. Também é preciso que os serviços médicos de urgência e emergência estejam capacitados a atender toda a população, na mesma proporção que a demanda exige. O mesmo se espera para os serviços de segurança pública, transporte e controle de trânsito. Contudo, na medida que a demanda pelos serviços passa a superar substancialmente a capacidade de oferta, torna-se extremamente oneroso, senão impossível, atender as necessidades básicas da população dentro dos parâmetros de qualidade desejados.

Em vista disso, as preocupações com os efeitos negativos oriundos de um crescimento descontrolado das cidades começaram a pautar as discussões de pensadores desde a década de 1990. Nesta época, discutiam-se novas políticas de planejamento urbano baseados na teoria do crescimento inteligente (*Smart Growth*). Segundo esta teoria, o crescimento é uma grande oportunidade para o progresso, desde que respeitados alguns princípios básicos, tais como preservação dos espaços

naturais, a oferta diversificada de meios de transporte, a tomada de decisões de forma previsível, justa e rentável e uma maior colaboração da comunidade nas decisões de desenvolvimento locais [13].

As ideias contidas na teoria do crescimento inteligente representaram as preocupações iniciais de se prover meios às cidades para consolidação de um processo de crescimento mais sustentável. A partir de 2005, os conceitos manifestos pelo crescimento inteligente serviram de base a grandes players da indústria tecnológica mundial, como IBM, Cisco e Siemens para dar significado a expressão **cidade inteligente**¹, termo então aplicado para fazer referência ao aglomerado de tecnologias e sistemas de informação destinados à interoperabilidade da infraestrutura e dos serviços urbanos [10].

Nascia, desta forma, a ideia inicial de utilizar Tecnologias de Informação e Comunicação na gestão das cidades, a fim de agregar maior eficiência e inteligência aos processos existentes. Com isso, a cidade dita inteligente passa a ser vista como um “sistema de sistemas que integra e otimiza um conjunto de diferentes sistemas independentes, públicos e privados, para alcançar um novo nível de eficácia e eficiência” [14].

Da mesma forma, conjuntamente à tecnologia, a cidade inteligente passa a investir em capital humano e social para dar suporte ao crescimento econômico sustentável e melhores condições de vida, prezando neste processo pela consolidação de um governo participativo [15].

Assim, no decorrer do desenvolvimento da cidade inteligente, diferentes fatores tecnológicos e humanos operam em conjunto para promover a evolução do ambiente urbano, mas por não existir uma única receita a seguir, a transformação das cidades inteligentes pode se dar de diferentes modos e sob diferentes visões, conforme descrito na sequência.

2.1 CIDADES INTELIGENTES E SEUS ASPECTOS PROMOTORES

Atualmente, as cidades inteligentes são tema recorrente nas agendas políticas internacionais, o que atesta a grande importância que as mesmas possuem para o futuro das sociedades urbanas no mundo. Contudo, a despeito dos benefícios decorrentes da sua construção, observa-se que o processo de conversão de uma cidade convencional em uma cidade inteligente é resultado majoritário da magnitude dos problemas vivenciados, os quais impõem às cidades a constante necessidade por transformação urgentes.

O crescimento populacional é um destes problemas de caráter transformador [11]. Em vista das dificuldades advindas do grande crescimento populacional, as cidades têm investido maior poder econômico, político e tecnológico para otimizar suas operações fundamentais, na tentativa de mitigar os efeitos negativos da sobrecarga dos sistemas básicos [2].

Outro problema observado é o crescimento do estresse da população. Na maioria dos casos, a infraestrutura urbana não se desenvolve na mesma proporção que a população cresce e, em razão disso, o sistema como um todo fica saturado e deixa de atender às necessidades dos habitantes. Como resultado, há um crescimento natural do descontentamento da população com

¹Do inglês, *smart city* ou *smarter city*

a ineficácia dos serviços prestados, culminando com exigências recorrentes por educação, saúde e transporte público de qualidade e pelo uso responsável dos recursos existentes.

Além das exigências por melhor qualidade na prestação dos serviços, há também a necessidade de maior individualização e adaptação dos mesmos. A ideia de que um único modelo de prestação de serviços atende a todos tornou-se ultrapassado e não satisfaz mais as expectativas dos cidadãos. Devido a esta insatisfação, as autoridades governamentais precisam conviver com as constantes pressões da população por melhores abordagens, as quais por vezes são difíceis de conceber, considerando-se que o orçamento disponível é limitado [2].

No âmbito econômico, o processo de globalização impõe às cidades a necessidade de estar em constante movimento, absorvendo e criando tecnologias. Quanto melhor uma cidade se apresentar frente ao mercado, maiores são as chances de captar investimentos, gerar empregos e de manter e atrair capital humano.

Segundo Dirks e Keeling [2], nas economias desenvolvidas, a prestação de serviços tornou-se a principal atividade econômica, abarcando $\frac{3}{4}$ (três quartos) de todo o comércio e, consequentemente, superando a cadeia produtiva primária. Em uma economia de abrangência global, fundamentada primordialmente no setor de prestação de serviços, é natural que as cidades tornem-se um local propício para o estabelecimento de novos negócios, visto que nelas estão concentrados boa parte dos recursos necessários, como capital humano e infraestrutura de comunicação.

Com relação às questões ambientais, observa-se que os ambientes urbanos são os maiores consumidores de energia e também os maiores emissores de CO₂ na atmosfera [11]. Em vista disso, qualquer ação que vise mitigar os impactos ambientais deve considerar, a priori, ações dentro das cidades. Além do papel de grandes vilãs ambientais, as cidades mostram-se também como grandes vítimas de eventos climáticos extremos, que ceifam milhares vidas e causam grandes prejuízos financeiros todos os anos.

Como aspectos transformadores positivos, intimamente relacionados à popularização das cidades inteligentes, pode-se citar a evolução tecnológica acelerada, acompanhada por um intenso declínio nos custos de aquisição. Recentes avanços tecnológicos têm proporcionado novas possibilidades às cidades e por meio da instrumentação, comumente feita com uso de sensores, torna-se possível obter dados de serviços fundamentais outrora impossíveis de serem coletados. Além do sensoriamento, diferentes sistemas presentes nas cidades podem ser interconectados, permitindo que as informações fluam entre eles de modo automatizado. Conjuntamente, instrumentação e interconexão trazem a tona informações de grande relevância para processos de tomada de decisão [2].

A evolução acelerada das tecnologias de hardware e software, associada a custos mais acessíveis de aquisição dos produtos e serviços, têm possibilitado às cidades fazerem uso de novas tecnologias para resolver antigos problemas de modo mais eficiente, criativo e menos dispendioso. Segundo Harrison e Donnelly [10], a aplicação de sistemas de informação nas cidades gera os seguintes benefícios:

- Redução do consumo de recursos, como energia e água.

- Otimização do uso da infraestrutura existe.
- Disponibilizar novos serviços aos cidadãos, atualizados em tempo real.
- Otimizar as atividades empresariais por meio da publicação de informações de valor sobre a cidade em tempo real.
- Identificar e revelar picos de consumo de recursos para proativamente redimensionar a infraestrutura e assim melhorar a resiliência dos serviços.

Em uma cidade inteligente, as tecnologias de informação e o conhecimento são alavancas para a inovação. O conhecimento, quando adequadamente gerido, permite tornar a cidade mais produtiva e competitiva. Por sua vez, as tecnologias de informação permitem criar espaços ricos de interação entre os cidadãos, setor privado e governo. Quando associadas, TICs e gestão de conhecimento permitem juntas criar um modelo de comunidade que promove a construção de uma economia mais sustentável e competitiva, mais preocupada com os interesses comuns e com o desenvolvimento dos cidadãos [6].

Os principais avanços tecnológicos que permitiram inserir novas formas de inteligência na cidade surgiram a partir da segunda metade da década 1990 e podem ser classificados em três “ondas” principais [3]. A primeira onda refere-se ao surgimento da World Wide Web, por meio da qual muitas cidades ingressaram na era digital, disponibilizando acesso a sistemas de informação e aplicações aos cidadãos. A segunda onda tecnológica coincide com o surgimento da Web 2.0, que introduziu novos conceitos a recém chegada Web, destacando-se as ideias de compartilhamento, colaboração e comunidades virtuais. Por sua vez, a terceira e última onda iniciou-se por volta de 2009, com a popularização das redes sem fio e sistemas embarcados, tornando disponível às cidades um novo conjunto de ferramentas para construção de espaços digitais [3].

Dentro do contexto das tecnologias de informação e comunicação aplicadas às cidades inteligentes, a Internet desempenha papel fundamental ao atuar como veículo de comunicação, compartilhamento e processamento de informações, transferência e análise de dados e computação distribuída. O surgimento da Internet das Coisas (IoC) ², conjuntamente com a maior adoção das tecnologias Web, tornaram possível tratar diversos desafios sociais por meio do emprego da tecnologia [17]. Desta forma, as TICs têm gerado grandes benefícios às cidades e aos cidadãos, em especial, para àquelas cidades que fazem uso massivo das ferramentas tecnológicas disponíveis.

2.2 CIDADES INTELIGENTES E SEUS OBJETIVOS

Conforme descrito, os desafios de ordem social, ambiental e econômico, associados à rápida evolução tecnológica, atuam como molas propulsoras para a concepção de cidades mais inteligentes. Todavia, esse processo de aperfeiçoamento das cidades não pode ser atribuído unicamente à

²Segundo glossário de termos do Gartner Inc. [16], Internet das Coisas é “uma rede de objetos físicos que contém tecnologia embebida para comunicar e sensoriar ou interagir com seus estados internos ou com o ambiente externo.”

mitigação dos problemas enfrentados. Os reais objetivos da construção de cidades inteligentes vão além do simples trato às dificuldades locais, estando de fato diretamente ligados com melhorias na qualidade de vida e de trabalho da população e com a implantação de um modelo sustentável de uso dos recursos disponíveis.

Para alcançar seus objetivos fundamentais, a cidade inteligente emprega um grande aparato tecnológico de hardware e software, visando basicamente prover recursos para a execução de três atividades fundamentais, a saber: **coleta**, **comunicação** e **análise de dados** [11]. Juntas, tais atividades definem um ciclo virtuoso de melhoria contínua, conforme pode-se observar pela Figura 1.

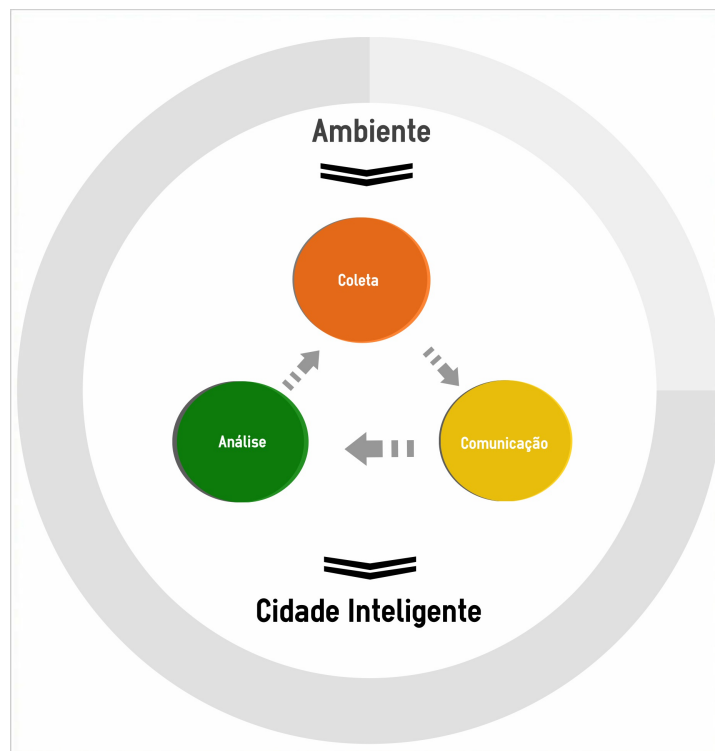


Figura 1. Atividades fundamentais de uma cidade inteligente. Fonte primária.

A coleta é a primeira das etapas. Nela, um grande conjunto de sensores, monitores e medidores inteligentes espalhados pela cidade recolhem dados em tempo real das mais variadas fontes (consumo de gás, água, eletricidade, pontos de engarrafamento, lotação do sistema público de transporte, entre outros). Além de precisos, os dados coletados por tais instrumentos comumente permitem identificar a localização no espaço, uma informação indispensável para avaliação e resposta eficaz aos problemas.

Os dados capturados na etapa de coleta são transmitidos por meio dos canais de comunicação disponíveis, sejam eles com ou sem fio. É importante que todos os interessados nestes dados tenham condições de acessá-los em tempo real. Para isso, é imprescindível que o ambiente ofereça plenas condições de conectividade em todo o lugar, para todo e qualquer dispositivo existente [11].

Na sequência, os dados anteriormente coletados e comunicados são analisados por ferramentas analíticas, com o objetivo de convertê-los em informações de valor. Tais informações visam,

principalmente, aperfeiçoar as decisões tanto humanas quanto de máquinas dentro do contexto da cidade inteligente [11]. Objetivamente, a etapa de análise de dados serve a três propósitos básicos, que são: **conhecimento do presente, aperfeiçoamento e predição**.

O conhecimento da situação presente torna a cidade inteligente ciente da sua situação em tempo real. Os serviços de monitoramento fazem uso dos dados coletados pelos sensores espalhados no ambiente para gerar uma visão sumarizada dos serviços e estados da cidade, passíveis de compreensão e interpretação pelos operadores humanos. A partir disso, torna-se possível conhecer, por exemplo, quais ruas encontram-se congestionadas, qual é a posição atual dos veículos de transporte público ou quais locais encontram-se com falta de energia elétrica ou água. Tecnicamente, qualquer aspecto de uma cidade inteligente pode ser monitorado, desde que seja possível coletar dados do mesmo.

O aperfeiçoamento refere-se à conversão do conhecimento extraído, tanto da situação atual quanto dos fatos históricos, em ações de intervenção que geram algum tipo de melhoria no ambiente [11]. Por meio do aperfeiçoamento, espera-se que os serviços e infraestrutura da cidade estejam em constante aprimoramento, de modo que a experiência do cidadão ao interagir com o ambiente seja sempre a mais aprazível possível.

Já a predição visa antever situações futuras com base nos dados históricos e presentes. Ter a capacidade de prever com grau de certeza aceitável é, sem sombra de dúvidas, uma das habilidades analíticas mais importantes. Isso porque, em conhecendo a probabilidade de ocorrência de um evento futuro, pode-se tomar ações imediatas que permitam mitigar seus efeitos, quando negativos, ou potencializá-los, quando positivos [11].

2.3 CIDADES INTELIGENTES E SUAS ÁREAS-CHAVE

Em uma cidade inteligente, as atividades de coleta, comunicação e análise de dados, auxiliadas e potencializadas pelas TICs, direcionam esforços para o aprimoramento de uma ou mais áreas-chave. Cada área-chave (ou domínio) corresponde a um macro sistema de fundamental importância para a cidade inteligente e normalmente envolve sistemas como os de distribuição de energia elétrica, água, governo, segurança e mobilidade.

Embora havendo áreas-chave comuns entre as diferentes concepções que buscam caracterizar uma cidade inteligente, não há um modelo único que estabeleça o que faz parte ou não do seu escopo. O que se observa é a existência de diversas visões, concebidas e apresentadas em sua maioria por grandes *players* mundiais da área de tecnologia. Dentre elas, merece especial atenção a visão de responsabilidades × aspectos habilitadores do *Smart City Council* [11], o modelo de áreas fundamentais descrito pelo *The Committee of Digital and Knowledge-based Cities of UCLG* [6] e a visão da International Business Machines (IBM) para cidades inteligentes [18].

No modelo apresentado pelo *Smart City Council*, uma cidade inteligente é vista como uma relação entre responsabilidades e aspectos habilitadores. As responsabilidades referem-se aos serviços fundamentais providos pela cidade a seus cidadãos e, conforme o modelo (Figura 2), resumem-se a

oito: água, energia, ambiente, telecomunicações, saúde e serviços humanos, pagamentos, transporte e infraestrutura física. Entrecortando os serviços essenciais, encontram-se os aspectos habilitadores que são responsáveis por tornar os serviços essenciais da cidade mais inteligentes por meio da aplicação das tecnologias de informação e comunicação disponíveis. Conforme explica Berst [11], são os aspectos habilitadores que tornam uma cidade inteligente de fato inteligente.

		Responsabilidades								
		Aspectos Universais	Pagamentos	Segurança Pública	Saúde e Serviços Humanos	Transporte	Telecomunicações	Energia	Ambiente Construído	Água e Esgoto
Aspectos Tecnológicos Habilitadores	Instrumentação e Controle									
	Conectividade									
	Interoperabilidade									
	Segurança e Privacidade									
	Gestão de Dados									
	Recursos Computacionais									
	Analytics									

Figura 2. Modelo de responsabilidades × aspectos habilitadores do *Smart City Council*. Adaptado de [11].

Na concepção do *The Committee of Digital and Knowledge-based Cities of UCLG* [6], uma cidade inteligente consiste em um modelo que considera a presença e convergência de seis áreas-chave fundamentais, identificadas como economia, mobilidade, cidadania, qualidade de vida, ambiente e governo. Cada uma das áreas é avaliada com base em um conjunto específico de indicadores. A presença de valores positivos para os indicadores considerados, conjuntamente com a adequada combinação de elementos, como infraestrutura, comunicação, desenvolvimento econômico e engajamento social na gestão participativa dos recursos naturais, compõem o conjunto de características necessárias a uma cidade inteligente.

O modelo apresentado por [6] chama atenção especialmente por sua definição da área-chave de governo. Nele, a cidade inteligente busca aproximar o governo dos cidadãos, a fim de tornar transparente suas ações à sociedade. Para tal, faz uso intensivo das TICs para criar novos canais de comunicação, possibilitando ao cidadão maior participação nos processos de tomada de decisão. De modo complementar, as TICs são aplicadas também à criação de novos serviços públicos, mais eficientes e menos dispendiosos, com grande influência na economia e vida dos cidadãos.

Na visão IBM — um dos grandes *players* mundiais em tecnologias voltadas às cidades inteligentes — uma cidade, para ser considerada inteligente, deve fazer uso otimizado de todas as

informações interconectadas existentes para melhor entender, controlar e otimizar o uso de recursos limitados. Além disso, deve atuar como uma entidade única, composta de inúmeros sistemas interconectados que suportam os diferentes domínios da cidade, provendo uma visão holística da mesma em tempo real [19].

As tecnologias de informação e comunicação pervasivas são utilizadas para tornar os sistemas da cidade instrumentados, interconectados e inteligentes. A instrumentação ou digitalização dos sistemas indica que as tarefas por eles exercidas transformam-se em pontos de dados e que o sistema como um todo torna-se mensurável. Por sua vez, a interconexão significa que diferentes partes de um sistema central podem ser acopladas de forma a trocar dados e gerar informação. De posse das informações, mecanismos de inteligência tornam possível modelar padrões de comportamento ou resultados prováveis, convertendo-os em conhecimento que serve de suporte a tomada de decisões [2].

Em sua maioria, as cidades são compostas por diferentes sistemas, os quais ditam em razão da sua eficiência e eficácia como a cidade trabalha e o quão competente ela é em atingir seus objetivos. Estes sistemas caracterizam-se por estarem intimamente interconectados, de modo que devem ser considerados tanto no seu aspecto individual quanto global. Por meio da otimização destes sistemas, as cidades têm maiores condições de orientar um crescimento econômico sustentável, e por conseguinte, prover maior prosperidade aos cidadãos. De modo complementar, os líderes municipais, de posse de um maior volume de informações de qualidade, tornam-se capazes de tomar melhores decisões e de antecipar e resolver problemas de modo proativo [18, 2].



Figura 3. Relacionamento entre os eixos de atuação e sistemas de uma cidade inteligente [18].

Os diferentes sistemas que compõem uma cidade inteligente estão associados a três eixos básicos (Figura 3): **planejamento e gerenciamento**, **infraestrutura** e **humano**. O eixo de planejamento e gerenciamento inclui sistemas como os de segurança pública, planejamento urbano e governo e órgãos administrativos. Com relação a infraestrutura, considera-se sistemas de abas-

tecimento de água e energia, transporte e elementos ambientais. Por fim, no eixo humano estão presentes os sistemas de assistência social, assistência médica e educação [18].

Deste modo, ao analisar as três visões apresentadas, percebe-se que elas descrevem as cidades inteligentes sob diferentes perspectivas. Contudo, em todas elas fica evidente a presença das tecnologias de informação e comunicação no processo de interconexão dos sistemas fundamentais, na coleta e processamento dos dados produzidos pela cidade e no aperfeiçoamento dos serviços prestados à sociedade. Também é notória a importância do eixo humano e da consolidação de um governo participativo, atento as demandas sociais e em constante interação com os cidadãos da cidade inteligente.

2.4 CIDADES INTELIGENTES, GOVERNO PARTICIPATIVO E O PAPEL DOS DADOS GOVERNAMENTAIS ABERTOS

Analisando-se as áreas-chave definidas pelas três visões de cidade inteligente descritas, observa-se a comum presença de tópicos relacionados com o governo local. Conforme descreve-se no modelo proposto pelo *Committee of Digital and Knowledge-based Cities of UCLG* [6], o governo de uma cidade inteligente deve aproximar-se dos cidadãos, criando meios para estreitar os laços de comunicação existentes. Assim sendo, além de propriamente aperfeiçoar o funcionamento da máquina pública, a aplicação das tecnologias de informação e comunicação deve fortalecer a construção de um governo amplamente participativo e colaborativo.

Um governo participativo representa a oportunidade de repensar os tradicionais modelos de gestão que comumente afastam os cidadãos das decisões. Um modelo de governo mais aberto, transparente, democrático e responsivo gera um campo fértil para a inovação, tanto social quanto econômica, dentro da cidade inteligente [1]. Todavia, para construção de um governo participativo é preciso que os cidadãos estejam engajados e comprometidos com as causas locais. Desta forma, incutir na comunidade estes valores representa um importante passo na construção das bases de um modelo de governo diferenciado.

Um dos meios encontrados por diversos países do mundo para aproximar a população das questões públicas é a de disponibilização de dados governamentais em formatos abertos na Internet. Além de propriamente visar a construção de uma relação de confiança entre governo e cidadão, tais iniciativas estão atreladas fortemente a novos e importantes conceitos aplicados à governança pública, como o movimento Open Data e o Government 2.0 [20].

De modo superficial, Government 2.0 diz respeito ao uso de tecnologias da Web 2.0 em favor das práticas de gestão pública. Contudo, seu significado é mais amplo, abarcando uma nova forma de governo, redescoberta e reimaginada a partir do princípio. As bases do movimento Government 2.0 estão relacionadas à colaboração, à construção conjunta de soluções em prol da sociedade e ao desenvolvimento de um ambiente favorável para a concepção de práticas inovadoras [21]. Dentro do contexto de Government 2.0, os dados abertos apresentam-se como uma importante ferramenta estratégica, prestando-se a quatro importantes funções: transparência e prestação de contas, gera-

ção de valor social e econômico, aprimoramento de serviços públicos e criação de novos negócios e empregos [20].

As iniciativas de abertura dos dados governamentais têm contribuído para a geração de valor em diversas áreas. No que se refere às questões de transparência pública e prestação de contas, por exemplo, os dados governamentais abertos representam uma resposta às exigências dos cidadãos, que permitem à sociedade analisar, identificar e responsabilizar possíveis casos de corrupção, contribuindo para o fortalecimento do processo democrático [20].

No âmbito econômico, iniciativas de abertura de dados governamentais têm grande potencial para agregação de riquezas em diversos setores. Conforme explica McKinsey [22], a abrangência das ações de abertura de dados, mesmo não exploradas ou percebidas em sua totalidade, têm capacidade de injetar anualmente na economia global cerca de 3 trilhões de dólares.

Contudo, os benefícios dos dados abertos vão além dos aspectos econômicos e de transparência pública, estendendo-se, conforme explicam Gray et. al [23], à geração de valor para as seguintes áreas:

- Participação popular.
- Empoderamento dos cidadãos.
- Desenvolvimento ou melhora de produtos e serviços privados.
- Inovação.
- Melhora na eficiência de serviços governamentais.
- Melhora na efetividade de serviços governamentais.
- Medição do impacto das políticas.
- Geração de conhecimentos novos a partir da combinação de fontes de dados e padrões.

O empoderamento do cidadão, um dos benefícios oriundos da abertura dos dados governamentais, é de fundamental importância para a construção de cidades mais inteligentes. Ao empoderar o cidadão com novas fontes de conhecimento, o processo de construção e inovação bottom-up da cidade é fortemente impulsionado. Neste modelo, as muitas partes que compõem a cidade contribuem de forma ativa, colaborativa e participativa na melhoria do ambiente urbano [24].

Quando os dados governamentais tornam-se disponíveis à multidão, um processo de inteligência coletiva é iniciado. Novas ideias, métodos e práticas são arquitetados, processos são refinados e o próprio cidadão é aperfeiçoado pelos resultados gerados. Além disso, a adoção de uma abordagem mais centrada no cidadão auxilia as autoridades municipais a permanecerem alinhadas com os interesses da população, algo indispensável à qualquer cidade inteligente [24].

A abertura dos dados governamentais abertos vai também ao encontro da consolidação do papel do cidadão inteligente (*smart citizen*), um cidadão diferenciado, que produz e faz uso

de informações para conhecer mais do espaço onde vive e com isso propor soluções criativas e inovados para sua cidade [8]. Adicionalmente, o cidadão inteligente busca estar sempre próximo do poder público, monitorando a situação da cidade e reivindicando seu espaço de direito nos processos decisórios [7].

Uma das principais iniciativas de abertura de dados mundialmente reconhecidas é o portal **Data.gov**³, criado e mantido pelo governo dos Estados Unidos. Publicado em maio de 2009 como parte da iniciativa de abertura do governo instaurada no país, o portal conta atualmente com mais de 150.000 conjuntos de dados disponíveis para acesso público. Além de servir como repositório de dados, o Data.gov tem foco em três elementos fundamentais, que são transparência, participação e colaboração [25].

No quesito transparência, o Data.gov visa aperfeiçoar a descoberta, o entendimento e o uso dos dados públicos abertos para prestação de contas e extração de valores econômicos e sociais ainda desconhecidos. Quanto à participação, o portal busca prover mecanismos de incentivo aos cidadãos que promovam a criação de novas formas de uso dos dados disponíveis e de algum modo, gerem impacto positivo na vida dos cidadãos. Somado a isso, diversos mecanismos de colaboração foram introduzidos ao portal para que os cidadãos pudessem participar ativamente da sua construção, consumindo dados existentes e/ou produzindo novos [25].

A partir da abertura dos dados pelo governo americano, um vasto conjunto de *cases* surgiram, muitos deles protagonizados por iniciativas desenvolvidas pelo governo das cidades, em colaboração direta com a iniciativa privada e cidadãos em geral. Nova York é uma destas cidades que desenvolveram ações próprias de abertura dos dados. O portal NYC Open Data⁴ oferece a população uma grande quantidade de conjuntos de dados, que podem ser acessados diretamente pelo portal ou via APIs específicas de consulta.

Ainda como incentivo ao desenvolvimento de novos softwares, a cidade de Nova York promove a competição *BigApps*⁵, que propõe aos participantes o desenvolvimento de novas soluções para os desafios enfrentados pela cidade. Um dos aplicativos vencedores da competição de 2014 foi o “*Water Status*”, criado com a finalidade de engajar o cidadão no controle dos reservatórios de água da cidade [26]. Outras iniciativas voltadas à mobilidade urbana, cultura e saúde também estiveram entre as ideias premiadas no concurso.

No Brasil, a abertura dos dados governamentais ocorreu pela sanção da lei 12.527, também conhecida como Lei de Acesso a Informação Pública ou Lei da Transparência, em 18 de novembro de 2011. A partir dela, cidadãos comuns puderam ter livre acesso à dados dos três poderes da União, estados, Distrito Federal e municípios [27, 28].

De modo semelhante ao ocorrido em outros países, a partir da abertura dos dados governamentais no Brasil, novos portais e aplicativos foram criados e disponibilizados aos cidadãos, utilizando fundamentalmente os conjuntos de dados abertos como matéria-prima. Dentre eles, destaca-se o

³<https://www.data.gov/>

⁴<https://nycopendata.socrata.com/>

⁵<http://www.nycbigapps.com>

Portal Brasileiro de Dados Abertos ⁶, uma ferramenta mantida pelo governo federal para facilitar a busca e obtenção de dados públicos. Outras iniciativas semelhantes são encontradas também nas esferas estadual e municipal, todas regimentadas pelas diretrizes presentes na Lei 12.527.

A cidade do Rio de Janeiro foi uma das cidades que desenvolveram um portal para disponibilizar seus dados à população. Sob o nome de **Dados Rio**⁷, o portal conta atualmente com mais de 1200 conjuntos de dados e também com alguns aplicativos para dispositivos móveis, como o 1746⁸, utilizado como canal de comunicação entre cidadãos e governo municipal.

Contudo, para que os dados governamentais abertos tenham pleno uso em diferentes iniciativas e sirvam ao propósito de agregar valor à sociedade, eles devem apresentar um conjunto básico de propriedades ou características. Tais características visam garantir livre acesso, reutilização e redistribuição dos conjuntos existentes por qualquer interessado [23].

Conforme explica Diniz [29], o conjunto de características esperadas para os dados governamentais abertos são:

- Independência de plataforma tecnológica.
- Uso de formatos padronizados, recomendados por organizações como a *World Wide Web Consortium* (W3C) e *International Standardization Organization* (ISO).
- Inexistência de vínculos ou dependências com seus repositórios de origem. O conjunto de dados deve ser autocontido.
- Uso de formatos manipuláveis por máquinas.
- Presença de descrição dos conjuntos de dados (metadados).
- Uso de *Uniform Resource Identifiers* (URIs) para identificar os objetos sempre que possível.
- Uso de informações semânticas no código das páginas web que hospedam os conjuntos de dados.

Além de facilmente reproduzíveis e reutilizáveis, os dados governamentais abertos não podem fazer discriminação de público alvo. Também devem ser completos, atuais, de cunho primário e não sujeitos à patentes, direitos de uso ou qualquer outra licença senão livre. Todas estas características visam possibilitar a interoperabilidade dos dados entre diferentes sistemas e entre diferentes conjuntos de dados [23].

Para serem disponibilizados abertamente, os dados governamentais precisam atender aos princípios técnicos e também às diretrizes legais. Albano (2014) [28] explica que nem todos os dados podem ser publicados abertamente e deve-se adequar qualquer iniciativa à respectiva legislação vigente. Respeitando-se as premissas legais envolvidas, todo dado torna-se passível de publicação.

⁶<http://dados.gov.br/>

⁷<http://data.rio.rj.gov.br>

⁸<http://www.1746.rio.gov.br>

Considerando-se os aspectos filosóficos e técnicos que envolvem a questão dos dados governamentais abertos é possível perceber a grande aplicabilidade que os mesmos têm no âmbito das cidades inteligentes. No cumprimento do seu dever, o governo registra grandes volumes de dados, que se adequadamente minerados, podem ser utilizados ativamente no processo de aperfeiçoamento dos serviços e da infraestrutura das cidades. Mais importante que isso é a possibilidade que as informações extraídas têm de capacitar o cidadão enquanto ente social, na construção de valores importantes, como participação e engajamento com as questões relacionadas a comunidade.

2.5 DATA WAREHOUSE, FATOS E DIMENSÕES

Segundo o Instituto Gartner, um data warehouse "é uma arquitetura de armazenamento designada para armazenar dados extraídos de sistemas transacionais, sistemas operacionais e de fontes externas" [16]. Já na concepção de Elmasri et. al [30, apud W. H. Inmon 1992], data warehouse caracteriza-se por ser uma coleção de dados orientados a assuntos de forma integrada, não volátil, variando no tempo a fim de favorecer os processos de tomada de decisão.

Diferentemente dos sistemas ditos transacionais, um data warehouse armazena os dados em um formato multidimensional e não normalizado. Esse tipo de modelagem visa relacionar os dados de modo a favorecer os processos de agregação e sumarização, em detrimento dos processos de atualização e inserção. Portanto, um data warehouse é modelado especificamente para favorecer a recuperação de informações.

O processo de modelagem multidimensional compreende a definição de estruturas de dados denominadas **fatos** e **dimensões**. Considerando-se um data warehouse desenvolvido sobre um banco de dados relacional (ROLAP⁹), as estruturas de dados utilizadas são denominadas **tabelas fato** e **tabelas dimensão**. Por meio da relação entre fatos e dimensões torna-se possível manipular os dados através de estruturas abstratas chamadas **cubos** (até três dimensões) ou **hipercubos** (acima de três dimensões) [30].

O uso de cubos e hipercubos na manipulação dos dados facilita a observação das informações sob diferentes perspectivas, conforme demonstrado na Figura 4. Nela, os dados de um fato fictício referente a frota de veículos podem ser manipulados nas perspectivas de cidade, ano e tipo de veículo.

⁹Relational Online Analytical Processing

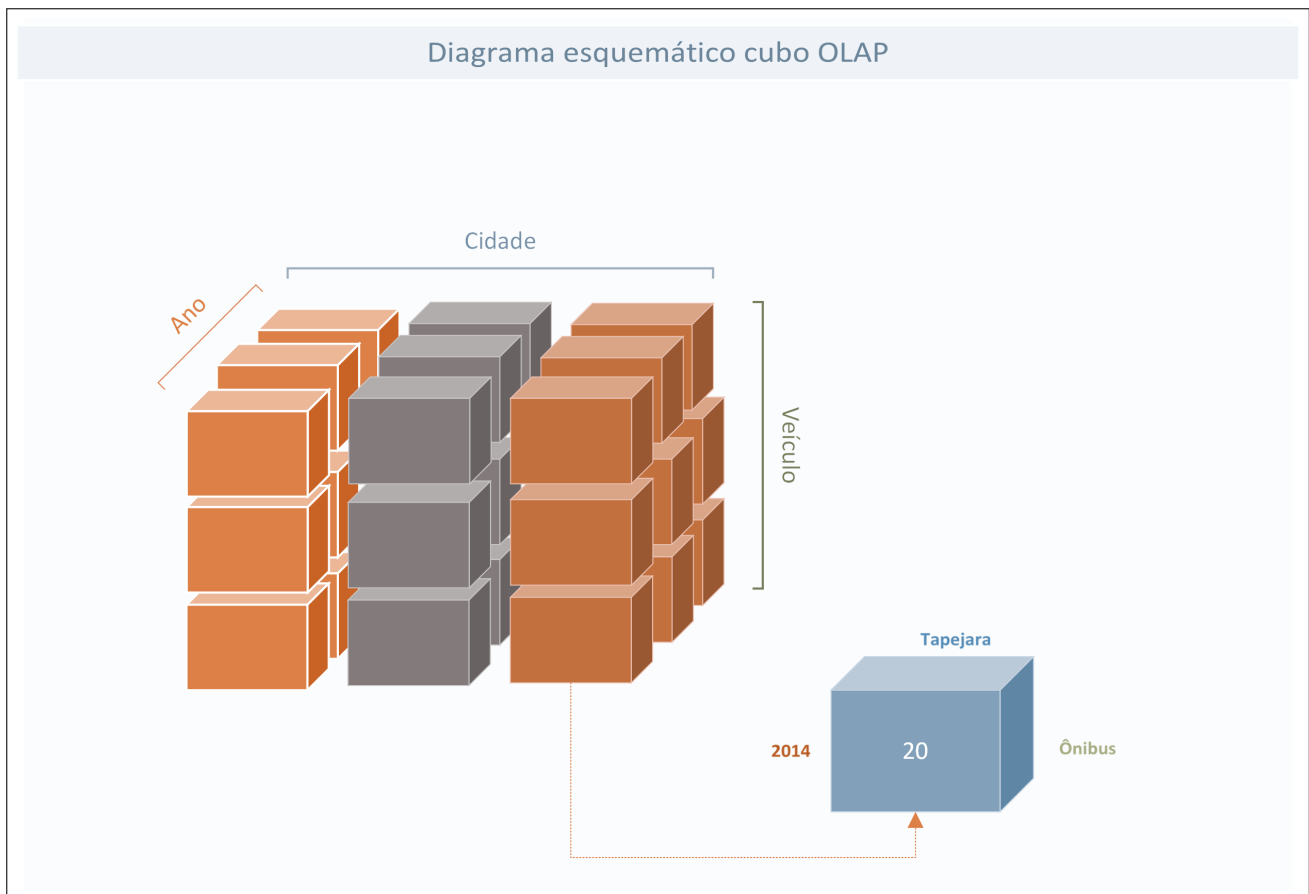


Figura 4. Diagrama esquemático de um cubo OLAP. Fonte primária.

Conforme descrito anteriormente, um data warehouse compõe-se de tabelas fato e tabelas dimensão. Tecnicamente, tanto fatos quanto dimensões são idênticos, visto que sua implementação é feita por meio do mesmo objeto em um SGBD relacional: a relação. O que diferencia ambas as estruturas ocorre em nível conceitual. Uma tabela fato é responsável por armazenar uma ou mais variáveis observadas acerca de um evento. Cada linha em uma tabela fato representa a medida de um determinado evento, registrado em um certo nível de detalhe conhecido como **granularidade**¹⁰. Quanto mais detalhes um registro fato apresentar, menor será sua granularidade [31]. Desta forma, pode-se empiricamente afirmar que a riqueza de detalhes é inversamente proporcional a granularidade dos dados em um data warehouse.

Uma tabela fato é composta basicamente por dois elementos: medidas e referências às dimensões. As medidas representam a quantificação do evento. Comumente são valores numéricos associados a quantidades ou valores. Estas medidas são sumarizadas, agregadas ou contadas nos processos de consulta face a disposição das dimensões escolhidas. Portanto, as referências às dimensões presentes em um registro fato qualificam o evento no seu contexto [31], permitindo com isso a visualização das informações sob diferentes perspectivas.

A Figura 5 representa esquematicamente um registro fictício de tabela fato. Nela, observam-se referências as dimensões ano, cidade e veículo e uma medida, de valor 428. Embora sendo valores

¹⁰Do inglês, *grain*

hipotéticos, uma interpretação possível para o exemplo é a de que, para a cidade de código 1, no ano de código 5 haviam 428 veículos do tipo 2.

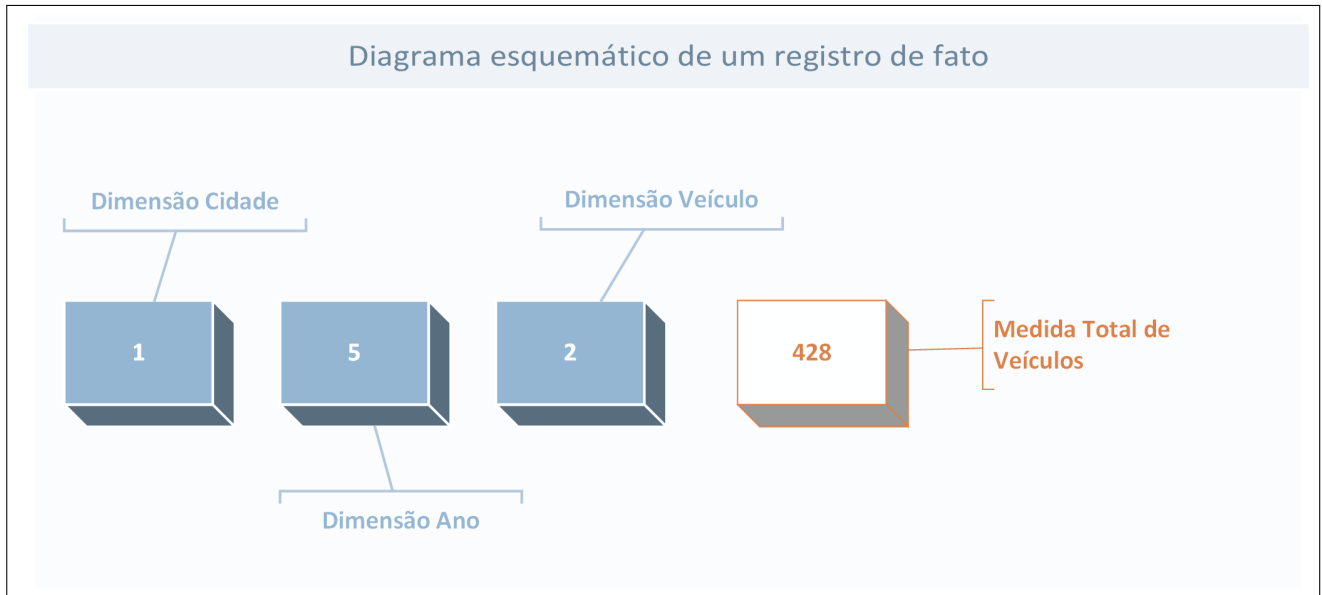


Figura 5. Representação esquemática de registro fictício de tabela fato. Fonte primária.

Em um data warehouse, fatos e dimensões compõem as estruturas fundamentais de armazenamento de dados. De acordo com as ligações lógicas criadas entre estas estruturas, o esquema do data warehouse recebe uma qualificação específica. Conforme explica Kimball et. al [31], em um data warehouse do tipo ROLAP, o esquema apresenta-se como sendo do tipo estrela (*star schema*). Em termos práticos, um esquema em estrela significa que uma ou mais tabelas dimensão encontram-se associadas a uma tabela fato por meio de chaves estrangeiras. Assim, se disposto graficamente em um plano, o modelo gerado assemelha-se a uma estrela, conforme demonstra a Figura 6.

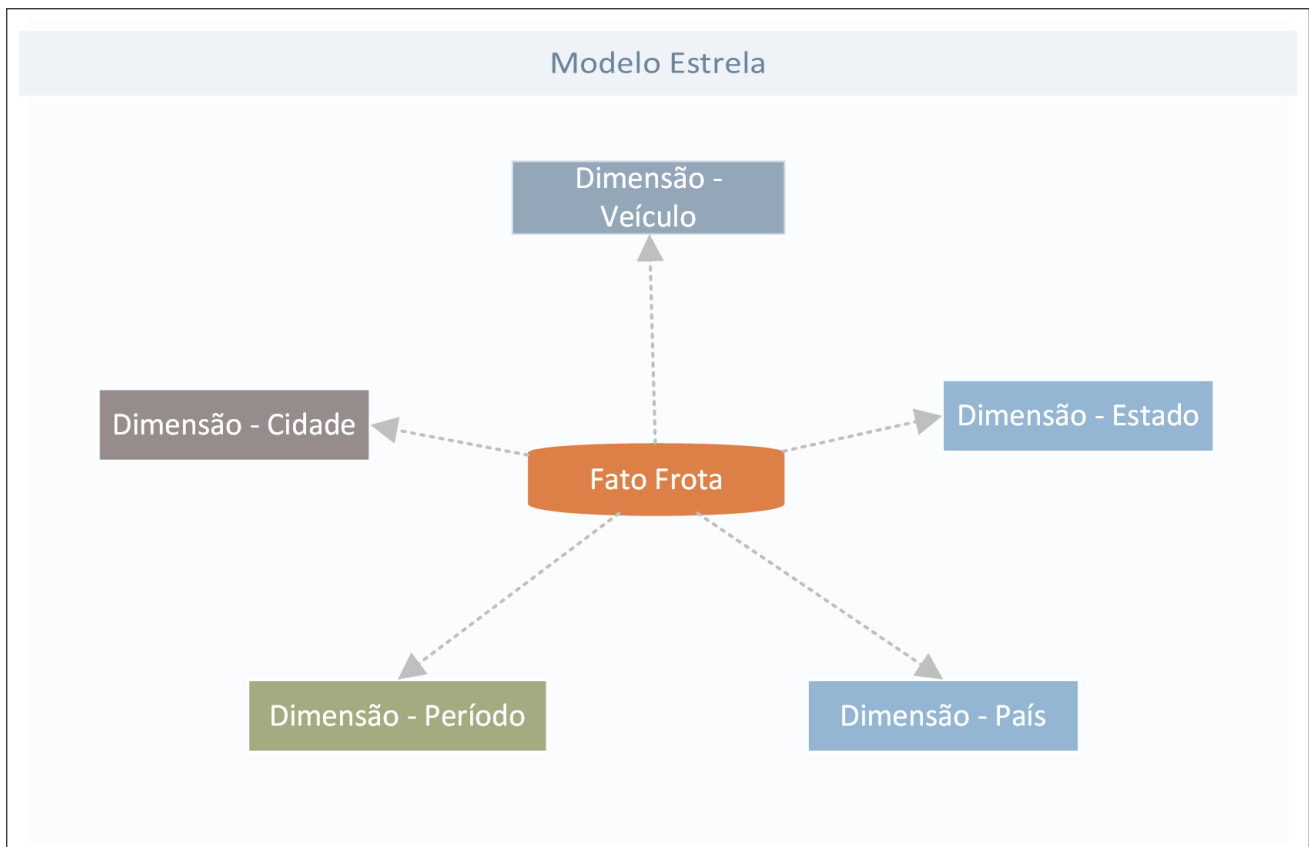


Figura 6. Apresentação gráfica do relacionamento entre tabelas fato e tabelas dimensão em um esquema do tipo estrela. Fonte primária.

2.6 TRABALHOS RELACIONADOS

Na literatura encontram-se diversos estudos referentes a dados governamentais abertos, muitos deles relacionados com os processos de transparência pública. Em Pinho [32], por exemplo, é feito um estudo sobre a real contribuição dos portais de transparência para a democratização e a melhoria de vida dos cidadãos brasileiros. No estudo de caráter exploratório foram tomados 10 portais governamentais, sendo nove de estados da União (Bahia, Rio Grande do Sul, São Paulo, Minas Gerais, Rio de Janeiro, Paraná, Pernambuco, Santa Catarina e Goiás) mais o portal de transparência do Distrito Federal. Por meio de um instrumento de análise próprio, os autores chegaram a conclusão de que os portais analisados dispunham de recursos tecnológicos adequados, contudo apresentavam deficiência de mecanismos de interatividade com o cidadão, tornando, de certa forma, passivo o papel do cidadão frente ao processo de transparência pública.

Já no trabalho de Sales e Martins [33], buscou-se analisar os impactos da Lei de Responsabilidade Fiscal (LRF) na Administração Pública brasileira. Com um viés voltado ao Direito, o estudo faz uma importante discussão sobre os aspectos de transparência pública. Como conclusões, os autores afirmam que o desenvolvimento de instrumentos que aproximem os governantes da população é de suma importância. Adicionalmente, afirmam que se faz necessário uma mudança cultural da população brasileira, no sentido de que haja maior participação ativa da mesma nos processos de transparência pública.

No trabalho de Böhm et al. [34], por sua vez, é proposto um protótipo de ferramenta Web, denominado *GovWILD*, voltado à integração e saneamento de dados governamentais abertos em larga escala. Todos os conjuntos de dados utilizados para o trabalho foram selecionados de portais específicos do governo norte-americano e da União Europeia e a temática abordada pelos mesmos condicionou-se a financiamentos governamentais e de fundos públicos.

No processo de integração dos conjuntos de dados, os autores utilizaram um ambiente clusterizado Hadoop¹¹, no qual os scripts de integração, construídos por meio da linguagem Jaql, foram executados. Durante o processo de integração, submeteu-se os conjuntos de dados a cinco etapas distintas de processamento, definidas pelos autores como preparação, saneamento e mapeamento, detecção de entidades, fusão de dados e pós-processamento. Ao final, os dados resultantes do processo de integração foram disponibilizados para uso na ferramenta Web, desenvolvida especificamente para dar condições ao usuário final de responder questões envolvendo políticos, companhias e financiamentos públicos [34].

Além de aplicar tecnologias de Big Data no processo de integração de dados, o trabalho de Böhm et al. [34] cita diversas dificuldades enfrentadas pelos autores diretamente relacionados ao trato dos dados governamentais abertos. Segundo eles, "a exploração de dados governamentais abertos não é uma atividade fácil" e os desafios se tornaram evidentes logo no início, na atividade de descoberta das fontes de dados mais adequadas aos objetivos do trabalho.

Adicionalmente, Böhm et al. [34] citam dificuldades relacionadas com a complexidade de acesso a portais governamentais e com a natureza heterogênea dos dados abertos, que por vezes encontram-se espalhados em diferentes silos pela internet, disponibilizados em diferentes formatos, semânticas e estruturas. Além disso, os autores ressaltam a existência de discrepâncias consideráveis na qualidade dos dados disponibilizados que corroboraram para os desafios enfrentados no processamento dos mesmos durante a etapa de integração dos dados.

Graves e Hendler [35], por outro lado, ao discorrer sobre dados governamentais abertos, afirmam existir uma importante parcela da população que pode beneficiar-se diretamente dos mesmos, mas que são impedidos de fazê-lo pelo desconhecimento em realizar operações necessárias relacionadas com a coleta, processamento, combinação e uso adequado dos dados disponíveis. Segundo os autores, as causas de tal situação são várias, mas a falta de perícia e de conhecimentos técnicos é o mais crítico de todos.

Como alternativa a tal situação, Graves e Hendler [35] propõem o uso de visualizações como meio de comunicação entre os dados governamentais abertos e a população. Adicionalmente, os autores descrevem um conjunto de características desejáveis às ferramentas de software responsáveis pela criação destas visualizações. Tais características foram definidas após análise de um questionário próprio aplicado a 257 pessoas, em sua maioria, trabalhadores do setor público e privado, acadêmicos e participantes de ONGs¹².

¹¹<https://hadoop.apache.org>

¹²Organização Não Governamental

O conjunto de características (ou recomendações) desejáveis às ferramentas de visualização, segundo os autores, seriam: facilidade de criação das visualizações, presença de mecanismos de exploração de dados e metadados, facilidade de comunicação com os provedores de dados, viralização e compartilhamento e, por fim, o reaproveitamento de visualizações. Na questão da facilidade de criação, o objetivo é tornar o processo de geração das visualizações o mais simples possível, fazendo uso de poucos cliques no processo. Com relação aos mecanismos de exploração, conforme descrevem os autores, é importante descrever o processo de geração das visualizações, bem com os conjuntos de dados utilizados. Da mesma forma, a possibilidade de compartilhar visualizações via redes sociais e reaproveitá-las para geração de outras novas mostrou-se importante [35].

Com base nas características elencadas, Graves e Hendler [35] desenvolveram um protótipo de software a fim de validar suas proposições. Contudo, o processo de validação com os usuários finais foi assinalado pelos autores como um dos trabalhos futuros intencionados. Já em suas conclusões, os autores ressaltam que as limitações de uso dos dados governamentais abertos por parte de quem eles definem como “cidadãos comuns” podem ser mitigadas pelo uso de técnicas de visualização, desde que hajam ferramentas de software aptas a disponibilizá-las à sociedade.

Conforme apontam os estudos acima apresentados, existem dificuldades relacionadas ao consumo dos dados governamentais abertos, tanto no cenário nacional quanto internacional. A ausência de mecanismo de interatividade, que aproxime os cidadãos das questões públicas, contribui para um cenário de passividade social, no qual o cidadão abdica do seu direito de participar ativamente em sua comunidade. De modo complementar, observa-se que as dificuldades relacionadas ao uso de dados governamentais abertos pela sociedade pode ser mitigado com uso de tecnologia da informação, por meio da criação de softwares aptos a integrar dados e apresentar as informações resultantes a sociedade de modo trivial e simplificado.

3. ARQUITETURA OPEN SMART CITY VIEW

A arquitetura Open Smart City View (OSCV), proposta neste trabalho, caracteriza-se por ser um modelo tecnológico funcional de coleta, processamento e apresentação de informações relevantes acerca de um determinado município. Os insumos de entrada da OSCV são única e exclusivamente conjunto de dados governamentais abertos, coletados de diversas fontes segundo diretrizes que atendam as necessidades de informação elencadas.

Para alcançar seus objetivos, a arquitetura OSCV divide-se em quatro camadas funcionais de operação (Figura 7). A primeira camada, identificada como *Fontes de Dados*, refere-se ao agregado de todas as fontes das quais deseja-se extrair conjuntos de dados governamentais abertos. A segunda camada, nomeada de *Coleta e Processamento*, consolida os processos de extração, saneamento, carga e transformação dos diferentes conjuntos de dados governamentais abertos elencados.

Os dados extraídos pela camada de *Coleta e Processamento* são armazenados em um data warehouse multidimensional, constituído de tabelas fato e dimensão. Por sua vez, as tabelas fato e dimensão alimentam a camada imediatamente seguinte, denominada *Camada Web*, incumbida pelo atendimento às requisições dos diversos dispositivos que interagem com a arquitetura.

A última camada, definida como *Dispositivos de Interação*, descreve os diversos consumidores de informação que podem interagir diretamente com a *Camada Web*. Estes dispositivos representam as diferentes interfaces que o cidadão da cidade inteligente poderá utilizar para extrair informações do data warehouse central da arquitetura OSCV.

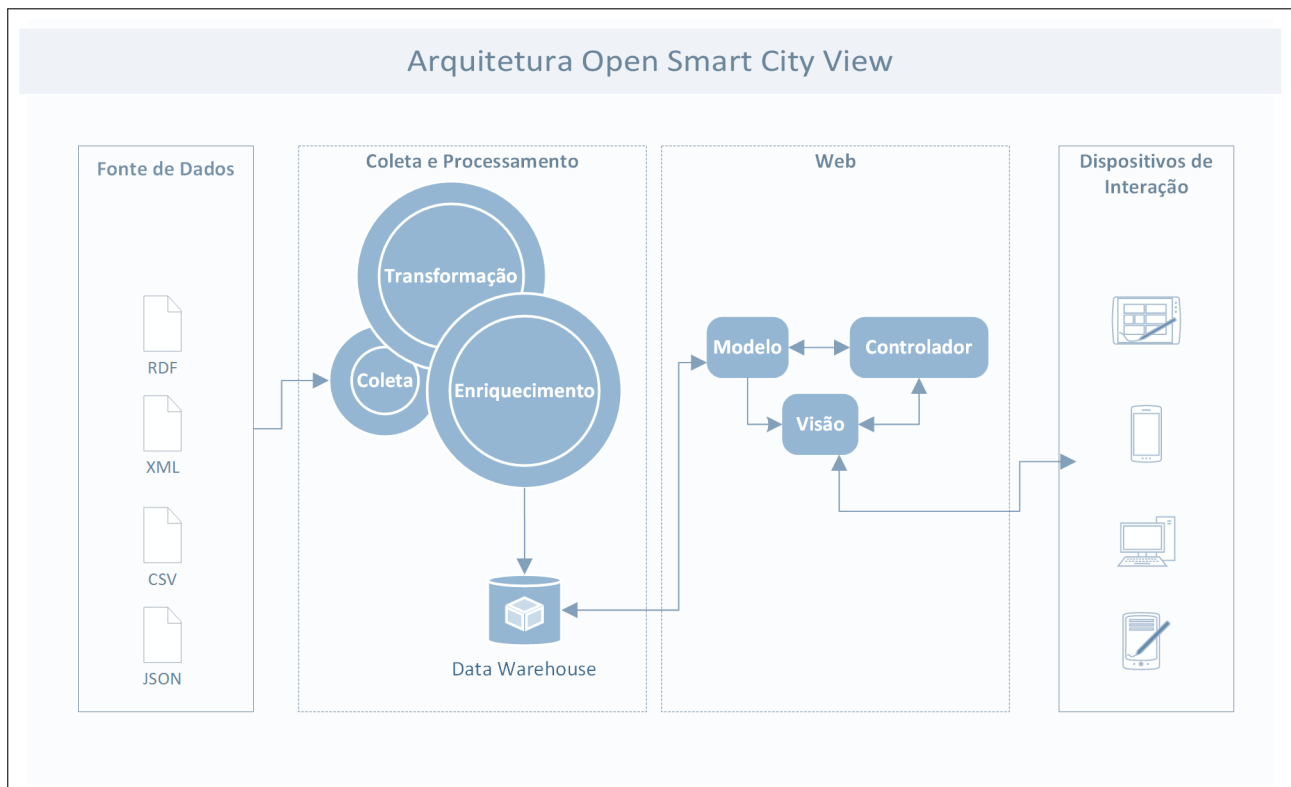


Figura 7. Modelo simplificado das camadas funcionais da arquitetura OSCV. Fonte primária.

A Figura 7 apresenta o diagrama funcional de alto nível da arquitetura proposta. Detalhes sobre cada uma das camadas funcionais são descritos na sequência.

3.1 CAMADA FONTES DE DADOS

Fontes de Dados representa a primeira camada da arquitetura OSCV, responsável pela identificação e mapeamento das fontes de dados governamentais abertos de interesse. Entende-se por fontes de interesse qualquer sítio hospedado na Internet, criado e mantido por uma entidade pública, que disponibilize dados governamentais abertos relevantes para a caracterização de um ou mais municípios.

Com relação ao formato dos dados, considera-se unicamente arquivos em formatos legíveis por máquina, disponíveis em sítios governamentais para captura manual ou automatizada. Quando a captura dos arquivos se der de forma manual, ela torna-se de responsabilidade do(s) administradores da plataforma. Havendo, contudo, possibilidades técnicas de automatizar a coleta dos conjuntos de dados, a obtenção dos arquivos é delegada à camada de Coleta e Processamento.

De qualquer modo, para coletar dados de uma determinada fonte é preciso realizar um processo de análise prévia. Neste processo deve-se identificar as fontes e os conjuntos de dados que melhor satisfazem as necessidades de informação elencadas no contexto que se apresenta. Também é preciso analisar detalhadamente a estrutura dos conjuntos de dados, identificando as variáveis e observações contidas, o nível de granularidade dos registros e a periodicidade da atualização dos mesmos.

Arquivos com granularidade baixa¹ e alta frequência de atualização, por exemplo, são mais adequados do que arquivos com baixa frequência de atualização e granularidade alta. Todos os arquivos que atendem aos critérios de granularidade, informação e periodicidade elencados tornam-se aptos a tornarem-se insumos da camada imediatamente posterior da arquitetura.

Com relação à granularidade dos conjuntos de dados, ou seja, ao nível de sumarização das observações disponibilizadas, é preciso levar em consideração que, no mínimo, hajam meios de se obter observações em nível municipal. Níveis mais altos de granularidade não são úteis porque conflitam com a premissa básica da arquitetura OSCV, que é coletar dados e extrair informações de conjuntos de dados **em nível municipal**.

A Tabela 1 apresenta, de forma resumida, as características desejáveis e/ou imperativas para as fontes de dados de interesse no âmbito da arquitetura OSCV. Tais características tem por objetivo orientar a escolha da fonte mais adequada, levando em consideração as características conceituais e técnicas associadas à OSCV.

¹Granularidade refere-se ao nível de segregação dos registros de dados. Quanto mais baixo, mais detalhes encontram-se disponíveis.

Tabela 1. Lista de características desejáveis para identificação de fontes de dados na arquitetura OSCV. Fonte primária.

Característica	Recomendável	Necessário
Fonte de dados é um sítio governamental		Sim
Fonte de dados é a entidade origem da informação	Sim	
Fonte de dados disponibiliza DGAs para download publicamente		Sim
Conjuntos de DGAs encontram-se disponíveis em formatos legíveis por máquina		Sim
A granularidade dos dados encontra-se em nível municipal ou inferior		Sim
Alta frequência de atualização dos conjuntos de dados	Sim	
Disponibilidade de APIs para consumo automático dos conjuntos de dados	Sim	

Do ponto de vista técnico, cada conjunto de DGAs elencados requer um processo específico de integração na camada de *Coleta e Processamento*. Isso ocorre porque cada conjunto é composto por arquivos com estrutura e metadados particulares, que precisam ser processados de modo específico. Desta forma, para garantir que os conjuntos de dados sejam adequadamente consumidos pelos scripts de integração, torna-se imprescindível identificar e documentar detalhadamente suas características estruturais.

Outra importante consideração acerca das fontes de interesse refere-se à coleta dos conjuntos de dados na origem. Embora hajam sítios Web que disponibilizam APIs para consumo de conjuntos de dados, estes representam as exceções. O que se observa na maioria dos casos são páginas Web onde pode-se somente efetuar download dos arquivos. Este formato de publicação, embora atendendo aos critérios de disponibilização de DGAs impostos pela lei, dificulta a construção de processos de coleta automatizada dos dados.

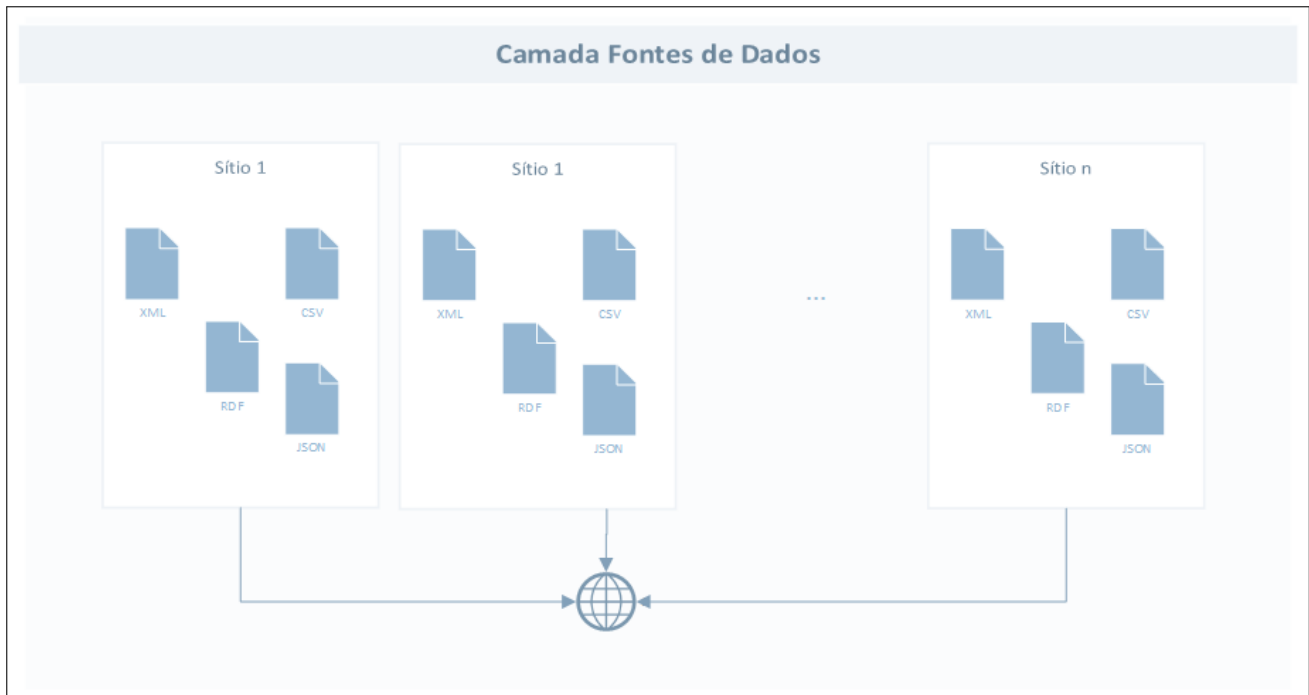


Figura 8. Camada Fontes de Dados. Fonte primária.

Na Figura 8, a camada Fonte de Dados é ilustrada conceitualmente. Por meio dela percebe-se que as fontes de interesse podem ser várias, e em cada uma podem haver um ou mais conjuntos de dados utilizados. Além disso, o acesso as fontes ocorre sempre por meio da Internet, desconsiderando qualquer acesso a Intranets governamentais restritas no processo. Somente aquilo que de fato é disponibilizado ao acesso público pode tornar-se insumo de entrada da arquitetura OSCV.

3.1.1 Fontes de Interesse Iniciais da Arquitetura OSCV

Em decorrência do sancionamento da Lei de Transparência, a quantidade de fontes de dados governamentais abertos no Brasil cresceu consideravelmente nos últimos anos, visto que, por força da lei, inúmeras entidades públicas viram-se obrigadas a disponibilizar em seus portais institucionais dados referentes as atividades realizadas. Entretanto, este processo de publicação parece não seguir um protocolo bem definido e cada entidade governamental tende a disponibilizar seus dados conforme suas diretrizes particulares. Por conseguinte é comum encontrar, por exemplo, dados acerca de um mesmo tópico em formatos e granularidades diferentes publicados em vários portais governamentais ao mesmo tempo.

Face a esta multiplicidade de opções de fontes e conjuntos de dados, combinado à complexidade inerente de processamento dos mesmos, optou-se por escolher um subconjunto inicial de fontes de dados para homologar a estrutura funcional da arquitetura, bem como seus aspectos técnicos internos. A escolha destas fontes iniciais não seguiu qualquer protocolo técnico e sua relevância no âmbito da arquitetura OSCV limita-se a testificar o alcance dos objetivos firmados.

Desta forma, ao delimitar o subconjunto inicial de fontes de interesse, chegou-se a quantia de sete diferentes tópicos (ou indicadores) de informação, os quais foram subdivididos em quatro categorias distintas para melhor orientar as atividades. As categorias elencadas, a saber, demografia, segurança, mobilidade e economia, encontram-se descritas na Tabela 2. Quanto a categoria "geral", sua inclusão foi necessária para fornecer dados de estados e municípios, fundamentais ao funcionamento da arquitetura. Exemplos de conteúdo dos arquivos escolhidos para processamento encontram-se na seção de anexos, conforme indicado da coluna Exemplo de Conteúdo.

Tabela 2. Conjunto de dados das Fontes de Interesse. Fonte primária.

Contexto	Conjunto de Dados	Fonte	Formato dos Arquivos	Granularidade	Exemplo de Conteúdo
Economia	PIB	Fundação de Economia e Estatística	CSV	A nível municipal, por ano	Anexo E
Mobilidade	Frota de Veículos	Departamento Nacional de Transito (Denatram)	XLS	A nível municipal, por mês	Anexo G
Mobilidade	Crimes de Transito	Dados RS	CSV	A nível municipal, por data e hora	Anexo D
Segurança	Registro de Crimes	Secretaria de Segurança Pública (RS)	XLS	A nível municipal, por ano e mês	Anexo A
Segurança	Violência contra a mulher	Dados RS	TXT	A nível municipal, por ano e mês	Anexo C
Geral	Dados de Municípios	Instituto Brasileiro de Geografia e Estatística	XSL	A nível municipal	Anexo B
Demografia	Dados Demográficos	Fundação de Economia e Estatística	XLS	A nível municipal, por ano	Anexo F

O processamento de cada um dos conjuntos de dados elencados das fontes de interesse iniciais foi realizado por meio de scripts ETL especificamente desenvolvidos para seu consumo. Conforme protocolo de integração definido para a arquitetura OSCV, cada conjunto de dados originou um novo script ETL. Estes scripts, bem como a definição das interfaces de carga, são de responsabilidade da camada Coleta e Processamento, a qual será descrita em detalhes na sequência.

3.2 CAMADA COLETA E PROCESSAMENTO

Os conjuntos de dados disponíveis nas fontes de interesse representam os insumos fundamentais para o sistema de software e servem de entrada para a camada imediatamente posterior, definida como **Coleta e Processamento**. Nesta etapa, os conjuntos de dados serão coletados, preferencialmente por via automática, dos seus locais de origem. Após coletados, cada conjunto de dados passa por um pipeline de operações para extração, transformação e carga dos registros no data warehouse da arquitetura, perfazendo um processo conhecido como ETL (**E**xtract, **T**ransform and **L**oad).

A Figura 9 apresenta esquematicamente o diagrama da camada **Coleta e Processamento**. Nela observa-se a interrelação entre os diferentes componentes, em específico, dos scripts ETL com as interfaces de carga do data warehouse. Esta interrelação, bem como seus componentes participantes, são detalhados no decorrer desta seção.

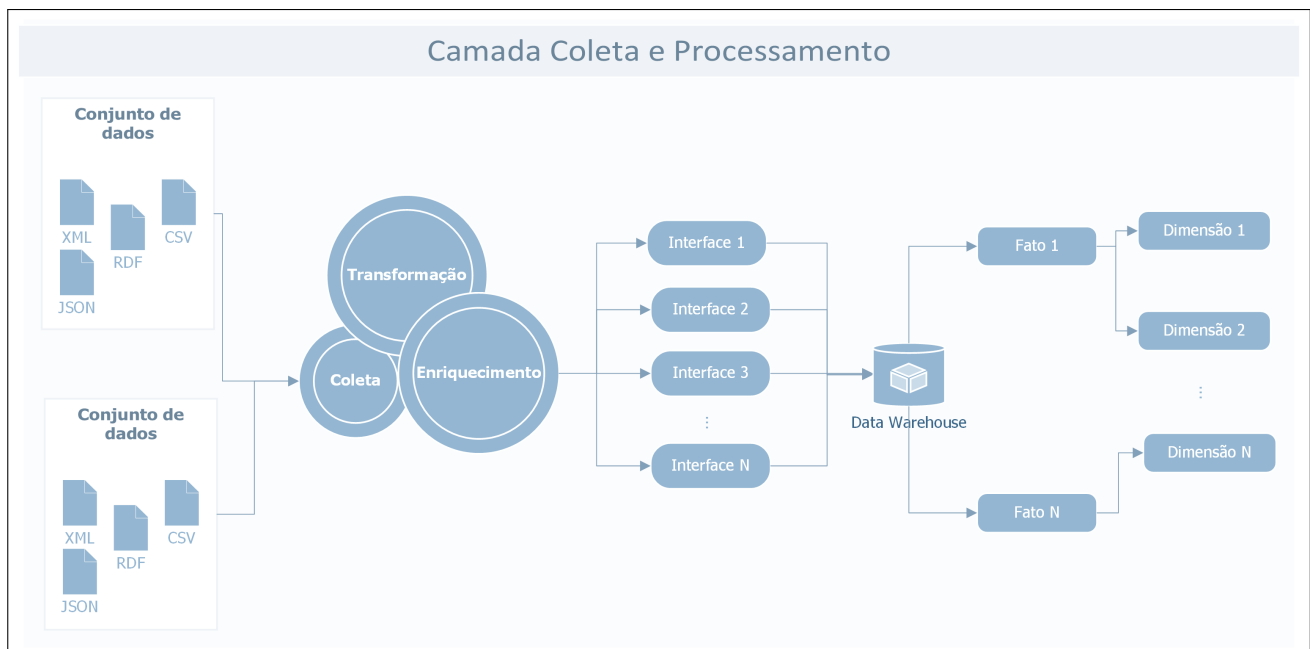


Figura 9. Camada Coleta e Processamento. Fonte primária.

Considerando-se que os dados governamentais abertos são disponibilizados ao público seguindo diretrizes específicas quanto à tipificação dos arquivos, espera-se consumir arquivos em formatos não-proprietários e compreensíveis por máquinas, como HTML, TXT, XML, JSON, CSV, RDF e TXT. Todavia, outros formatos de arquivo, inclusive proprietários, são comumente encontrados nos portais de entidades públicas brasileiras. Um destes formatos é o XLS, atrelado a patentes da Microsoft Corporation para uso na ferramenta Microsoft Excel.

De qualquer modo, cada formato de arquivo exige, por parte dos scripts ETL, tratamento específico, de acordo com suas características estruturais. Desta forma, para conceber qualquer script de integração de dados, é preciso conhecer as principais características destes formatos, apresentadas na sequência.

- **XML**: O formato XML (**EX**tensible **M**arkup **L**anguage) é amplamente utilizado no intercâmbio de informações na Web. Caracteriza-se por ser uma linguagem de marcação extensível, na qual gera-se arquivos contendo conjuntamente os dados e a descrição associada. Uma importante característica do formato XML é ser legível tanto por máquinas quanto para seres humanos [23].
- **HTML**: Semelhante ao XML, o formato HTML (**H**yper**T**ext **M**arkup **L**anguage) é uma linguagem de marcação não-extensível voltada a apresentação de dados, utilizada especialmente na construção de páginas na Web. Embora sua manipulação possa ser feita de modo análogo ao formato XML, conjunto de dados disponibilizados em HTML exigem maior esforço computacional para extração dos dados de interesse, principalmente em processos de saneamento.
- **JSON**: O formato JSON é o acrônimo de **J**ava**S**cript **O**bject **N**otation e também é muito utilizado na troca de informações na Web. Um dos principais pontos positivos do formato JSON é a simplicidade de processamento, visto que boa parte das linguagens de programação de alto nível tem suporte nativo a ele. Além disso, dados serializados em JSON transportam uma quantidade menor de meta-informações se comparados ao XML e HTML, o que comumente lhe rende o adjetivo de “formato leve” [23].
- **CSV**: Dados disponibilizados em formato CSV (**C**omma **S**eparated **V**alues) apresentam uma estrutura tabular simplificada. Cada registro é representado por uma linha e a divisão das colunas é feita por um caractere separador (normalmente ponto-e-vírgula). Contudo, embora simples de manipular, dados em formato CSV tem maior probabilidade de inconsistência e podem impor dificuldades de processamento em determinados casos [23].
- **RDF**: O formato RDF (**R**esource **D**escription **F**ramework) é um modelo de troca de informações na Web recomendado pelo W3C, considerado também um dos pilares da Web Semântica. Dados em formato RDF têm a unidade de informação estruturada em sujeito, predicado e objeto, formando uma tripla. Em vista da sua estrutura, o RDF permite que dados estruturados e semi-estruturados sejam misturados entre diferentes aplicações de forma relativamente simples [36].
- **XLS**: Formato binário de dados criado pela Microsoft Corporation para armazenar dados gerados pela ferramenta Microsoft Excel. Neste tipo de arquivo, os dados ficam armazenados em uma ou mais planilhas constituídas de linhas e colunas. Cada unidade de informação, denominada célula, é referenciadas através da coluna e linha correspondentes [37]. Diferentemente dos demais formatos, o padrão XLS envolve direitos de propriedade e patentes, o que está em desacordo com os princípios de publicação de dados governamentais abertos. Mesmo assim, é muito comum encontrar conjuntos de dados disponíveis neste formato em diversos sítios governamentais.
- **TXT**: O formato TXT é um dos formatos mais simplificados de armazenamento de dados. Nele, as informações encontram-se estruturadas numa sequência de linhas, também conhecidas

como registros. Além disso, neste tipo de formato, os dados são armazenados textualmente, tornando-o inadequado para representar tipos de informação não textuais. Afora isso, o formato TXT está em conformidade com os princípios de publicação de dados governamentais abertos, sendo comumente usado em conjunto de dados disponibilizados por entidades governamentais.

No contexto da arquitetura OSCV, a camada **Coleta e Processamento** tem a importante incumbência de alimentar o data warehouse, inserindo novos registros nas tabelas fato e dimensão. Neste processo, os scripts ETL leem os conjuntos de dados abertos das fontes de interesse em seu estado bruto, executam atividades de saneamento e transformação, e por fim, armazenam os dados resultantes nas tabelas de interface. Neste processo, os scripts ETL são de fundamental importância para o adequado funcionamento da arquitetura OSCV como um todo.

Embora a construção de scripts ETL possa ser feita por meio de qualquer linguagem de programação, o uso de ferramentas especializadas apresenta-se como uma boa prática, em especial, pela maior confiabilidade do código final gerado. Outros importantes benefícios advindos do uso de ferramentas especializadas são o aprimorado controle de execução dos scripts e o menor tempo de desenvolvimento necessário para tornar funcional as integrações.

Para desenvolvimento dos scripts de processamento dos arquivos da OSCV foi utilizado a ferramenta Pentaho Data Integrator (PDI) ², em sua versão *community*. Também conhecido como Kettle, o PDI apresenta-se com um dos softwares Open Source para ETL bem aceitos no mercado. Nele, o desenvolvimento se dá de modo intuitivo, utilizando representações gráficas de componentes, os quais, quando postos em conjunto e sob determinada ordem, executam as atividades de ETL definidas pelo desenvolvedor. Desta forma, por meio da adequada configuração dos componentes e de sua organização em processos sequenciais e/ou paralelos, torna-se possível transportar e transformar dados entre diferentes tecnologias, como arquivos, filas JMS ³, XML, bancos de dados, entre outros.

Outra importante característica dos componentes do PDI refere-se a sua natureza funcional. Dentro do contexto de um script ETL, os componentes classificam-se em quatro grupos fundamentais e mutuamente excludentes, que são: *transformations*, *jobs*, *hops* e *steps*. As *transformation* são os elementos que atuam diretamente no processamento dos dados. É por meio delas que o processo de ETL é efetivamente realizado. Tecnicamente, uma *transformation* é composta por um ou mais *steps*, cada um executando uma parte específica da tarefa de ETL. Ligando os diferentes *steps* que compõem uma *transformation* estão os *hops*, os quais atuam como canais por onde os dados trafegam de um *step* para outro no fluxo de ETL [38]. É o que demonstra a Figura 10, ao apresentar um exemplo de *transformation* construída no PDI para processamento de um dos conjuntos de dados elencado das fontes de interesse iniciais da arquitetura OSCV.

²<http://community.pentaho.com/projects/data-integration/>

³Java Message Service

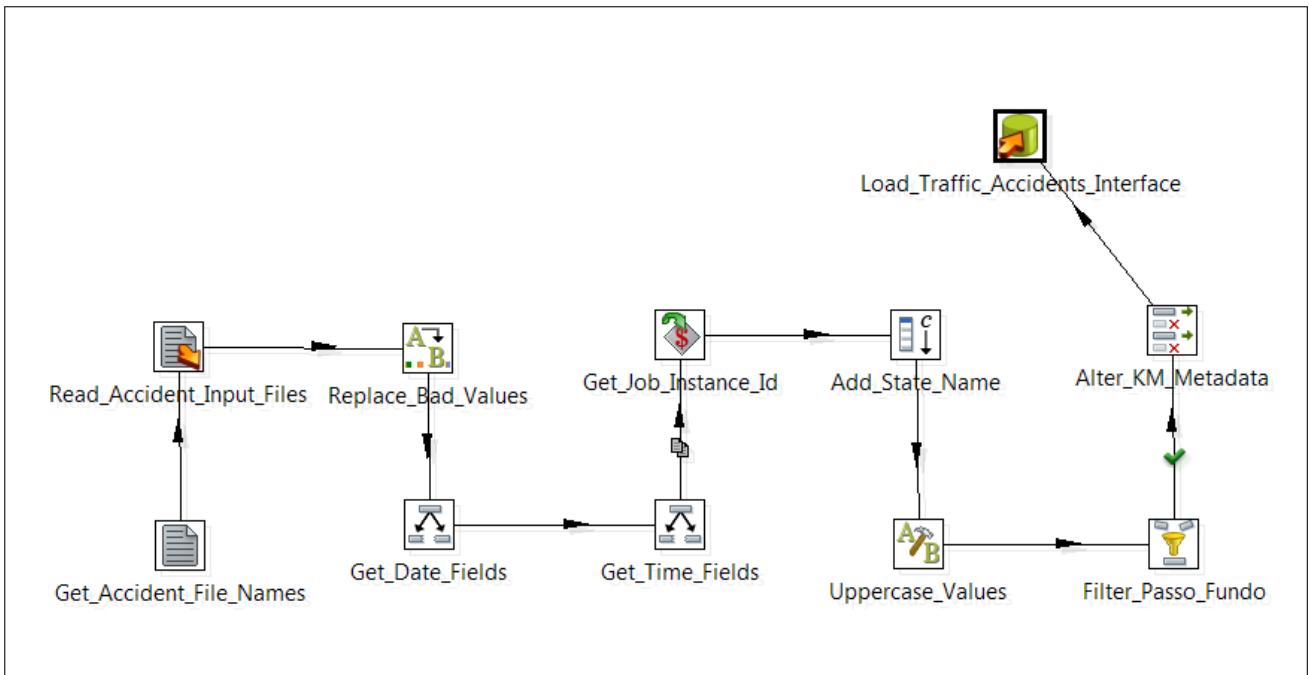


Figura 10. PDI - Exemplo de *transformation* e seus componentes de construção. Fonte primária.

Os *jobs*, por sua vez, são componentes utilizados para realizar tarefas de controle. Diferentemente das *transformations* que executam seus *hops* em paralelo, os *jobs* garantem a sequência de execução. Sua construção, contudo, se dá de modo semelhante às *transformations*, visto que utiliza-se um ou mais *job entries* ligados entre si por meio de *job hops*. Os *job entries* são os componentes que executam tarefas específicas dentro de um *job*, enquanto que os *job hops* representam a sequência de execução, podendo ou não realizar validações condicionais durante a execução [38]. Conforme apresentado pela Figura 11, cada componente do *job* recebe um ícone que o identifica em relação aos demais. Fluxos de sucesso e erro também são diferenciados por cores e símbolos nas setas que determinam o fluxo.

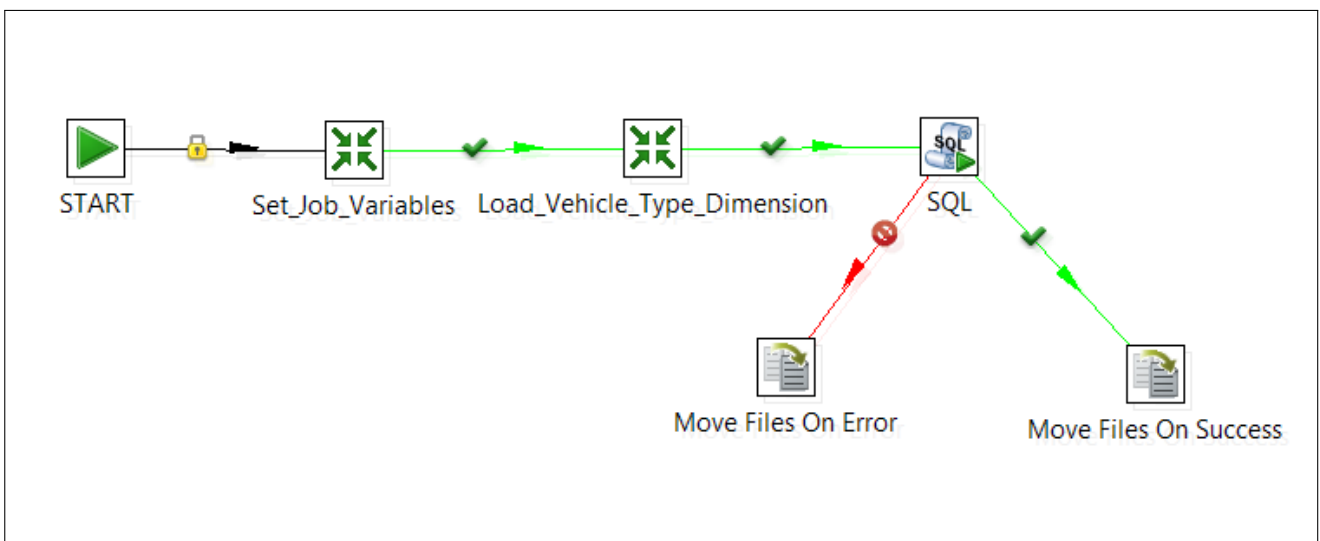


Figura 11. PDI - Exemplo de *job* e seus componentes de construção. Fonte primária.

Comumente, um *job* atua na orquestração de uma ou mais *transformations*, preparando as condições necessárias para execução e controlando os fluxos das exceções que porventura ocorreram durante a execução do processo de ETL. Em vista disso, os *jobs* podem ser agendados para executar em períodos de tempo pré-determinados de forma automática, conforme as necessidades de integração identificadas.

Outra importante questão relacionada à construção dos scripts ETL refere-se ao padrão de desenvolvimento adotado. Embora não hajam receitas prontas para solucionar todo e qualquer problema, a adoção de boas práticas é enfaticamente recomendável. Padronização de nomenclatura, protocolo de tratamento de erros, uso de técnicas performáticas e um arcabouço estrutural bem definido são bons exemplos de práticas benéficas aos desenvolvimento ETL.

Para os scripts da arquitetura OSCV fora definido um arcabouço estrutural baseado em tabelas de interface e *stored procedures*. Esta estruturação visa mitigar erros de carga no data warehouse da arquitetura, de modo a garantir a qualidade dos dados armazenados. Outra importante responsabilidade da estruturação escolhida refere-se a simplificação do processo de construção dos scripts ETL, de modo a reduzir o tempo necessário para desenvolver uma carga específica de dados.

Estruturalmente, o data warehouse da arquitetura OSCV caracteriza-se por ser do tipo ROLAP, ou seja, está implementado sobre um banco de dados relacional, mantido pelo Sistema Gerenciador de Banco de Dados PostgreSQL⁴, versão 9.4. A escolha do SGBD PostgreSQL sustenta-se principalmente pelo seu amplo uso no mercado e na academia, estabilidade funcional e disponibilidade de recursos avançados para manipulação de dados.

Em vista das características específicas da arquitetura OSCV, bem como da natureza sensível das operações de manipulação das tabelas fato e dimensões, a concessão de acesso direto às tabelas fato e dimensão não se mostra uma boa prática. Além disso, na OSCV as demandas analíticas envolvidas podem variar no decorrer do tempo. Isso significa que as tabelas fato e dimensão não são fixas em estrutura e número. A medida que novas demandas de informação forem surgindo, ou mesmo, novas fontes de interesse forem sendo adicionadas, existe a probabilidade de acréscimo de tabelas fato ou mesmo de tabelas dimensão no data warehouse.

Considerando-se ainda que o número de scripts ETL tenderá a crescer em quantidade, haja vista a provável agregação de novas fontes de dados a OSCV com o passar do tempo, o acesso direto as tabelas fato e dimensão poderia tornar-se um ponto crítico de inconsistências, dado que cada integração que manipulasse tais estruturas deveria preocupar-se em mantê-las consistentes durante as execuções.

Não obstante, o acesso direto as tabelas fato e dimensão obrigaria o desenvolvedor das integrações a conhecer integralmente a implementação do data warehouse, o que tornaria o processo de desenvolvimento mais oneroso e predisposto a erros. Em decorrência disso, uma condição de alto acoplamento entre a estrutura do data warehouse e os scripts de ETL seria estabelecida, algo indesejável em qualquer arquitetura de software.

⁴<http://www.postgresql.org>

Idealmente, boas arquiteturas de software devem prezar pelo baixo acoplamento de código a fim de mitigar o impacto das mudanças entre os diferentes módulos constituintes. O baixo acoplamento, portanto, é uma característica almejada para arquitetura OSCV, e para obtê-lo, definiu-se uma camada intermediária para os processos de integração ETL. Essa camada é baseada no conceito de tabelas de interface, as quais são descritas na sequência.

3.2.1 Tabelas de Interface

As tabelas de interface desenvolvidas para a arquitetura OSCV assemelham-se a uma área de armazenamento intermediário, na qual as integrações ETL efetuam a publicação dos dados a serem carregados para as tabelas fato e dimensão. Em sua estrutura, contemplam dois conjuntos de atributos específicos: atributos de controle e atributos de dados.

Os atributos de controle têm por função dar subsídios ao processamento dos registros da tabela de interface. Este processamento, para qualquer tabela de interface, deve ser executado por uma e somente uma *stored procedure* criada especificamente para este fim. Outra importante característica relacionada aos atributos de controle refere-se a sua repetição em todas as tabelas de interface. Em vista disso, é possível discriminar cada um deles, bem como sua função de controle.

- **Identificador de linha:** Identifica de forma única cada linha armazenada na tabela de interface, representando a chave primária da mesma.
- **Identificador de lote:** Identifica o conjunto de um ou mais registros agrupados logicamente. Por convenção, a cada execução de um script ETL gera-se um novo identificador de lote, específico para a execução em questão.
- **Data da última atualização:** Estampa de tempo atribuída ao registro no momento de sua criação e atualizado quando alterações nos valores de atributos são efetuadas. Sua presença na tabela de interface justifica-se para fins de monitoria e rastreabilidade.
- **Situação de processamento:** Mantém a codificação de situações possíveis que um registro pode assumir no decorrer do seu ciclo de vida na tabela de interface. Os valores aceitos são **P**, **I**, **S** e **E**, onde.
 - **P:** Indica que o registro encontra-se *pendente* de processamento. É a situação inicial de todo registro recém criado.
 - **I:** Sinaliza que o registro encontra-se *em processamento* naquele instante. A situação é temporária e permanece enquanto operações estão sendo executadas sobre o registro.
 - **S:** Situação final, indicando que houve *sucesso* no processamento do registro.
 - **E:** Situação final, indicando que houve *erro* no processamento do registro.
- **Descrição da situação de processamento:** Armazena, em caráter suplementar, a descrição do erro capturado durante o processamento.

Quanto aos atributos de dados, estes são específicos para cada tabela de interface e dependem exclusivamente de quais tabelas fato e dimensão serão alimentadas no processo. Isso porque o protocolo ETL desenvolvido para a arquitetura OSCV prevê que uma tabela de interface pode alimentar diferentes tabelas fato e dimensão, ou seja, que a mesma tabela de interface pode ser compartilhada para várias integrações envolvendo o mesmo assunto.

Outra importante característica do protocolo ETL é que tanto tabelas fato quanto dimensão são alimentadas por meio de *stored procedures* especializadas. Assim, o processamento do lote de dados da interface é desacoplando do processo de carga do data warehouse em termos funcionais. Neste cenário, mantendo-se as assinaturas das *stored procedures* especializadas, qualquer alteração interna feita na forma de carga dos dados para o data warehouse seria imperceptível aos mecanismos de integração em execução.

Nestes termos, ao script ETL cabe tão somente ler os arquivos de origem, obter um identificador de lote, armazenar os dados na respectiva tabela de interface e executar a *stored procedure* de carga conforme sequencia descrita por meio do diagrama funcional exibido pela Figura 12.

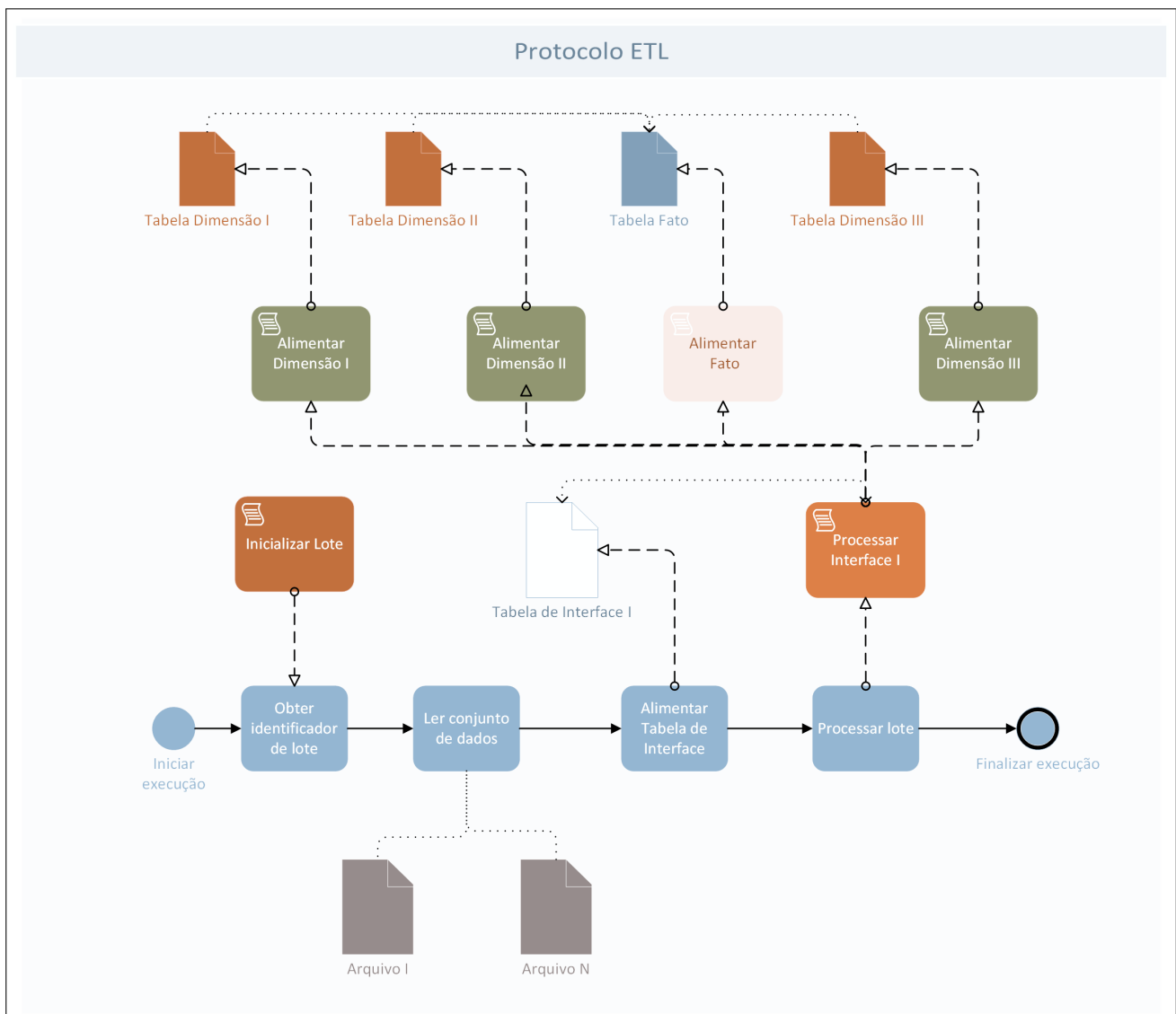


Figura 12. Protocolo de desenvolvimento de integrações ETL na arquitetura OSCV. Fonte primária.

Desta forma, ao prezar pelo baixo acoplamento entre as etapas que compõem o processo de ETL, entende-se que a expansão da arquitetura, por meio da agregação de novas fontes de dados governamentais abertos, pode ser feita de modo prático e seguro. Além disso, a abstração dos processos internos de carga do data warehouse reforça consideravelmente a qualidade dos dados adicionados, mitigando possíveis inconsistências que ultrapassaram os critérios de saneamento definidos pelos próprios scripts ETL.

Toda as preocupações envolvidas com a camada de *Coleta e Processamento* são plenamente justificáveis pela sua importância no contexto da arquitetura OSCV. Sem o adequado funcionamento dos processos que alimentam a arquitetura como um todo, os objetivos da OSCV não seriam alcançados. Tampouco a camada Web, que segue no fluxo estrutural da arquitetura, teria condições de fornecer ao cidadão qualquer informação de valor.

3.3 CAMADA WEB

Uma vez transferidos para o data warehouse, os dados tornam-se disponíveis para uso na *Camada Web*. A principal função da *Camada Web* é prover serviços de acesso e consumo para os dados armazenados, o que lhe confere, de certa forma, características de um software OLAP (**O**nline **A**nalYTical **P**rocessing).

Dentro da arquitetura OSCV, a Camada Web situa-se como uma interface analítica entre os dados mantidos no data warehouse e os diversos dispositivos de interação que podem vir a consumi-los (Figura 13), abstraindo qualquer aspecto relacionada a origem ou natureza dos dados, bem como qualquer regra de negócio vinculada ao seu consumo. Nenhum aspecto técnico ou necessidade de conhecimento prévio, portanto, deve se fazer presente na interação do cidadão com a *Camada Web*.

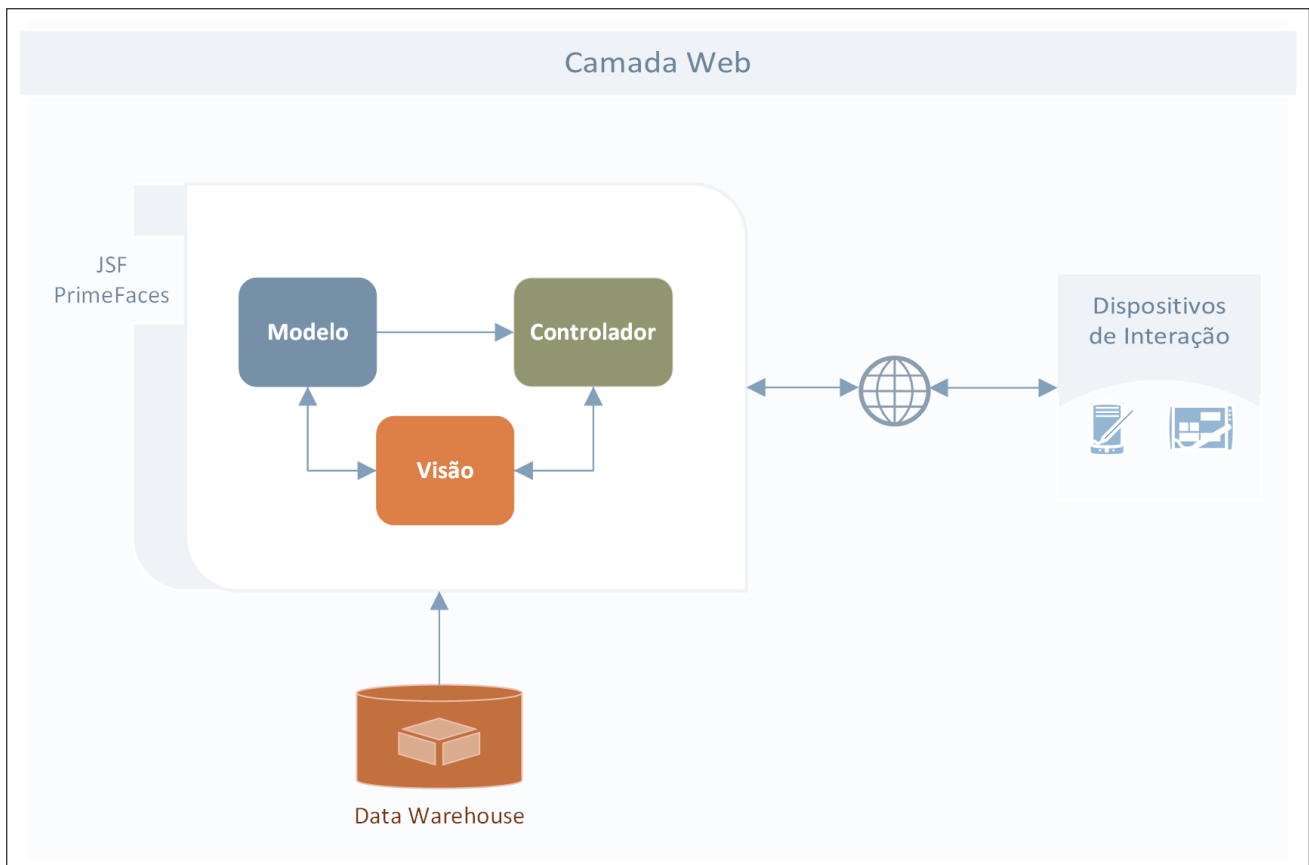


Figura 13. Diagrama conceitual da Camada Web. Fonte primária.

Em vista dos requisitos não funcionais associados à Camada Web, em especial, no que se refere a grande quantidade de usuários que podem vir a fazer uso da mesma e a necessidade de resposta no menor tempo possível, torna-se imprescindível o uso de tecnologias e *design patterns* condizentes com as demandas identificadas. Neste sentido, optou-se pelo uso de tecnologias consagradas no mercado para desenvolvimento de aplicações Web de médio e grande porte, como é o caso do Java Server Faces (JSF).

Conforme descreve Deitel & Deitel [39], Java Server Faces é um "framework de aplicativo Web que simplifica o design da interface com o usuário de um aplicativo e separa ainda mais a apresentação de um aplicativo Web da sua lógica de negócio". Por sua vez, na visão da Oracle Corporation [40], JSF refere-se a um conjunto de APIs para apresentação de componentes visuais e controle dos seus respectivos estados, eventos, validação de entrada do usuário, definição de navegação, internacionalização e acessibilidade. Juntas, estas APIs constituem um arcabouço arquitetural que permite uma clara separação entre a lógica da aplicação e a sua apresentação.

Em termos de *design patterns*, o JSF acomoda a implementação do padrão Modelo-Visão-Controlador, comumente conhecido pelo acrônimo MVC. Neste padrão arquitetural há forte apelo pela separação entre as camadas de uma aplicação, em especial, na separação dos modelos de dados criados para representar a percepção do mundo real da sua apresentação final na interface com o usuário por meio de componentes gráficos [41]. Assim, para atingir o adequado desacoplamento entre modelo e apresentação, o padrão MVC orienta o desenvolvimento da aplicação em três camadas distintas, a saber, camada de modelo, de visão e de controle.

Quanto ao alinhamento com as necessidades tecnológicas da arquitetura OSCV, o conjunto de características arquiteturais do JSF mostrou-se adequado aos requisitos elencados para o desenvolvimento da **Camada Web**. O primeiro ponto positivo refere-se à incorporação do JSF ao Java EE. Essa característica permite que o código desenvolvido possa ser executado em qualquer *container* Java EE do mercado, como Glassfish, JBoss Application Server, Weblogic Server, IBM WebSphere Application Server, entre outros. Além disso, ao executar em um *container* Java EE, a aplicação automaticamente herda todos os benefícios associados, como escalabilidade, redundância e alta disponibilidade, por exemplo.

Outra característica do JSF adequada ao desenvolvimento da *Camada Web* relaciona-se à alta propensão para modificações que a camada apresenta. É de conhecimento prévio, portanto, que novas funcionalidades poderão ser agregadas à OSCV no decorrer do tempo, o que torna a preocupação com aspectos de qualidade uma constante no projeto. Neste sentido, o uso do JSF e, por conseguinte, de um modelo arquitetural MVC, reduz consideravelmente o acoplamento de código, visto que os modelos de dados ficam separados dos aspectos relacionados a sua apresentação. Além disso, o desenvolvimento baseado em componentes, aliado às robustas APIs de programação providas pelo JSF tornam o processo de desenvolvimento menos oneroso em termos de esforço.

Definidos os aspectos tecnológicos envolvidos, o passo seguinte refere-se a concepção da metáfora de interação com o usuário. Conforme outrora exposto, uma das grandes preocupações da arquitetura OSCV é tornar o consumo de informações trivial, minimamente sujeito a conhecimentos prévios do funcionamento da plataforma.

Para alcançar este importante objetivo, foram desenvolvidos três conceitos básicos para orientar a interação do usuário com os dados armazenados no data warehouse. Estes conceitos foram batizados de **apresentações de fato**, **perspectivas** e **perspectivas relacionadas**, conforme descreve-se na sequência.

3.3.1 Apresentação de Fatos, Perspectivas e Perspectivas Relacionadas

Tendo em vista que a arquitetura OSCV visa oferecer ao cidadão informações de valor de modo trivial, entende-se que a forma mais adequada de fazê-lo seja prezando pela simplicidade na apresentação. Interfaces simplificadas, objetivas e funcionais têm maior probabilidade de atingir seus objetivos se comparadas a interfaces poluídas e superdimensionadas. Além disso, considerando-se a variabilidade dos usuários da OSCV, tanto em aspectos culturais quanto instrucionais, optar pelo simples mostra-se ser a melhor escolha.

Adicionalmente, observa-se que a *Camada Web* tem papel fundamental na OSCV pela sua proximidade com o usuário final. O próprio sucesso da arquitetura depende diretamente da boa relação de interação entre usuários (cidadãos) e a *Camada Web*. Além disso, problemas relacionados ao uso do software podem impactar negativamente no interesse do cidadão pela busca de informações acerca do seu município e, conseqüentemente, prejudicar seu processo de empoderamento.

Deste modo, atendendo a premissa básica de prezar pela simplicidade, optou-se por organizar as informações apresentadas ao usuário em três diferentes conceitos ou metáforas, que são: **apresentação de fatos**, **perspectivas** e **perspectivas relacionadas**. Estes conceitos relacionam-se entre si e, em conjunto, objetivam tornar a experiência de consumo de informações na OSCV agradável e simplificada. A natureza do relacionamento entre os conceitos é apresentado por meio da Figura 14.

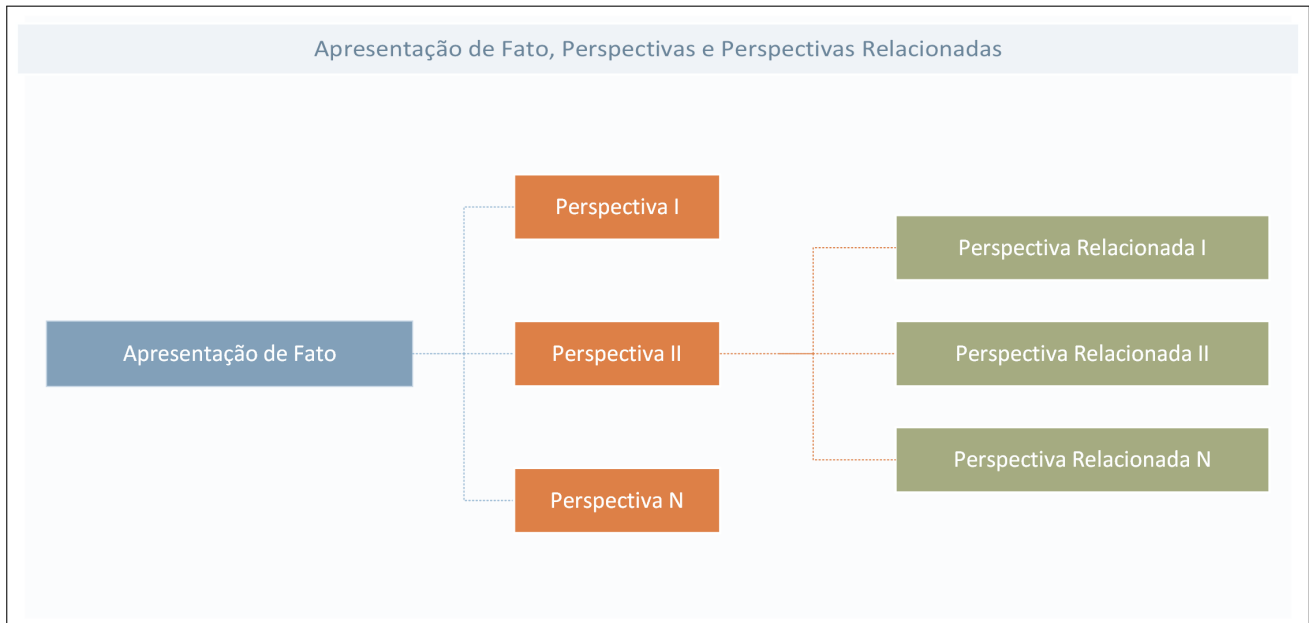


Figura 14. Inter-relacionamento entre as metáforas de interação da Camada Web. Fonte primária.

Cada um dos conceitos apresentados pela Figura 14 foram agregados à *Camada Web* seguindo as convenções e restrições apresentadas a seguir:

- **Apresentação de Fatos:** O data warehouse da arquitetura OSCV poderá conter diversas tabelas fato. Uma apresentação de fato refere-se ao grupo de perspectivas que fazem uso de uma tabela fato em particular para apresentar ao cidadão informações relevantes.

O administrador do sistema poderá manter as informações de apresentação de fatos, adicionando novos, alterando ou mesmo removendo os desnecessários. Desta forma, apresentações de fato são informações de natureza dinâmica na arquitetura OSCV.

- **Perspectivas:** Uma perspectiva compreende uma observação particular feita sobre dados de uma tabela fato. Diz respeito a questão a ser respondida por meio de consulta ao data warehouse e relaciona-se como as apresentações de fato na proporção de 1 para n , ou seja, uma apresentação de fatos deve ter uma ou muitas perspectivas associadas.

Um exemplo de perspectiva, considerando-se o contexto de uma apresentação de fato "Violência Urbana", poderia ser o total de homicídios ocorridos em uma determinada cidade nos últimos 10 anos. Da mesma forma que as apresentações de fato, as perspectivas são mantidas pelo administrador do sistema e, portanto, são informações de natureza dinâmica na arquitetura OSCV.

- **Perspectivas relacionadas:** As perspectivas relacionadas visam dar condições para o correlacionamento de informações na arquitetura OSCV. Diferentes fatos podem ter uma relação implícita estabelecida, a qual somente torna-se visível quando as observações de ambos os fatos são analisadas de forma conjunta.

Uma perspectiva relacionada associa-se a uma perspectiva sob os seguintes critérios:

- Perspectivas relacionadas não são obrigatórias;
- Perspectivas relacionadas **podem** expor uma correlação real entre as informações ou uma simples casualidade. A interpretação fica sempre a critério do usuário.
- Uma perspectiva pode possuir no mínimo 0 (zero) e no máximo n perspectivas relacionadas.

A aplicação prática dos conceitos de apresentação de fato, perspectivas e perspectivas relacionadas visa facilitar a interação do usuário da arquitetura com os dados armazenados na mesma. Considerando-se, por exemplo, uma apresentação de fato chamada “Frota de Veículos”, uma perspectiva possível seria “Evolução do número de automóveis”. Já uma perspectiva relacionada à “Evolução do número de automóveis” poderia ser “Evolução dos acidentes fatais de trânsito”, por exemplo. A perspectiva relacionada, neste caso, contrapõe um assunto distinto do tratado pela perspectiva, de modo a expor uma possível relação entre elas. Neste exemplo, uma correlação plausível seria o crescimento proporcional do número de automóveis e de acidentes fatais de trânsito.

Naturalmente, o conhecimento dos detalhes relacionados às metáforas de interação são de exclusividade do administrador da plataforma. O usuário final, enquanto consumidor de informações, não terá necessidade qualquer de conhecê-los para interagir com a camada Web. Isso é possível devido a separação da Camada Web em duas visões funcionais distintas, especificamente projetadas para cada tipo de usuário da plataforma.

3.3.2 Visões Funcionais

A camada Web é responsável tanto pela apresentação de informações ao cidadão quanto pela administração e configuração da plataforma Open Smart City View. Desta forma, as interfaces de acesso público, desenvolvidas especificamente para interação com o usuário final, integram a visão funcional denominada *frontend*. Já o conjunto de interfaces projetadas especificamente para interação com os administradores da plataforma compõem a visão funcional de *backend*.

Mesmo operando dentro de um mesmo contexto e de forma complementar, as visões de *frontend* e *backend* apresentam características próprias. Certamente uma das características que as diferencia profundamente é a abstração de conceitos técnicos. Enquanto na visão de *frontend* eles são encapsulados em sua totalidade, na visão de *backend* procura-se torná-los presentes, visto que deles depende o bom funcionamento da plataforma como um todo.

Na sequência, as visões funcionais de *backend* e *frontend* são descritas em maiores detalhes.

3.3.2.1 Visão de Backend

Embora sendo imperceptíveis ao usuário final (cidadão), os conceitos de apresentação de fato, perspectivas e perspectivas relacionadas precisam ser de conhecimento do administrador da arquitetura. A adequada configuração dos metadados destes conceitos no *backend* permite ao cidadão interagir com os dados armazenados no data warehouse. Configurações incorretas ou inexistentes impedem o cidadão de obter as visualizações de interesse, comprometendo assim o processo de empoderamento e os próprios objetivos da OSCV.

Segundo a hierarquia definida para os conceitos ou metáforas de interação da OSCV (Figura 14), o ponto de partida para configurar os metadados inicia pelas apresentações de fato. Na sequência, configuram-se as perspectivas associadas à apresentação e, por fim e de forma opcional, as perspectivas relacionadas à perspectiva principal.

Internamente, cada um dos conceitos é representado por classes. Estas classes relacionam-se entre si conforme o Diagrama de Classes apresentado pela Figura 15. Pelo diagrama percebe-se que o conceito “Perspectiva Relacionada” representa apenas um autorelacionamento de agregação da classe Perspectiva. Por sua vez, o relacionamento entre apresentação de fato e perspectivas é representado por uma composição, cuja multiplicidade mínima é 1. Isso significa que uma perspectiva não pode existir sem estar relacionada a uma apresentação de fato e, no mínimo, uma apresentação de fato relaciona-se com uma perspectiva.

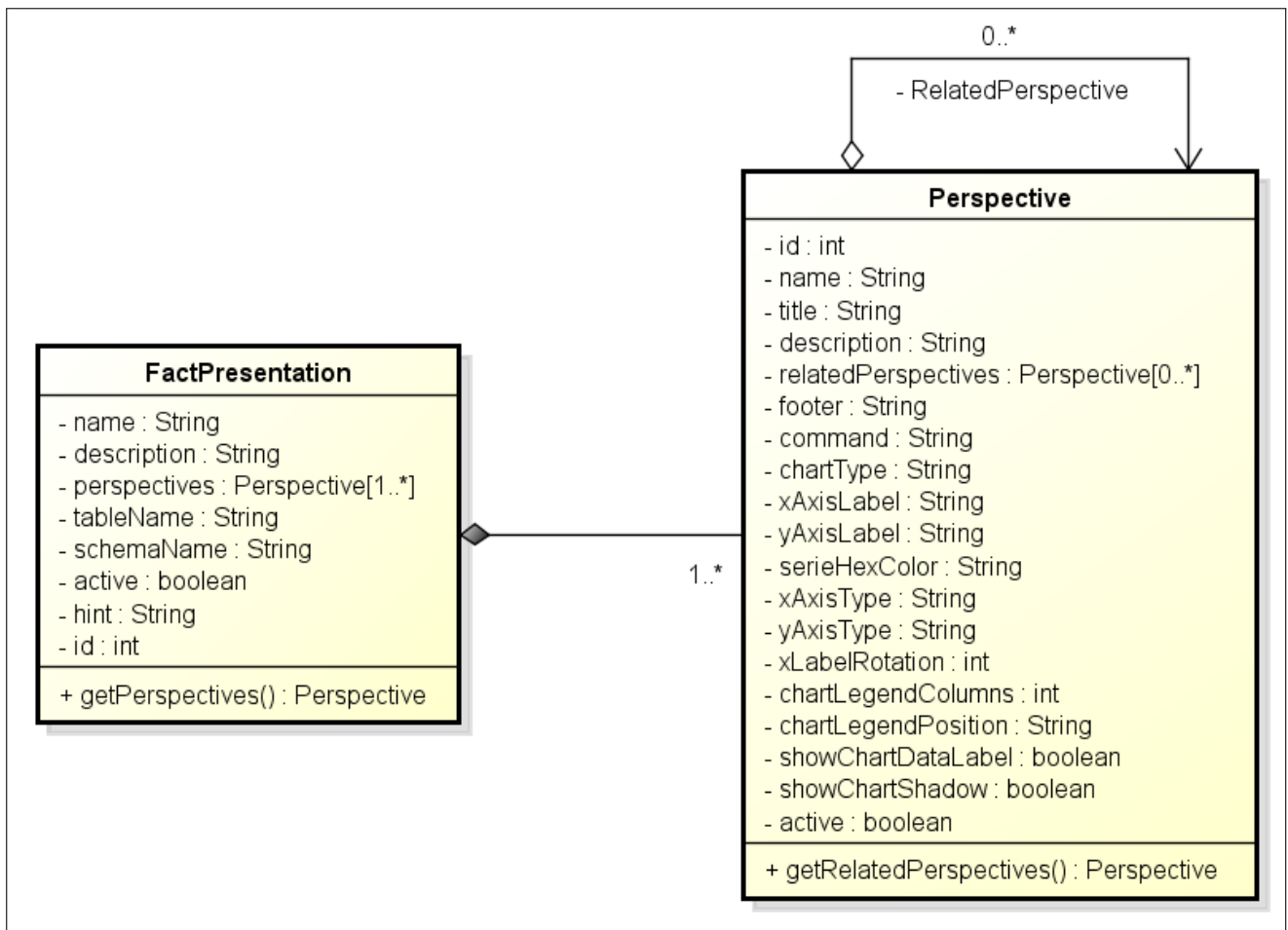


Figura 15. Diagrama de classes - conceitos de apresentação de fato, perspectivas e perspectivas relacionadas. Fonte primária.

Com relação ao conceito "apresentação de fato", optou-se por definir apenas oito atributos de informação. Estes atributos armazenam informações que são voltadas à interação com o usuário na visão de *frontend* e informações de controle, utilizadas pelos mecanismos internos. A lista de atributos referentes a apresentações de fato consta na Tabela 3, acompanhadas de uma breve descrição da função exercida por cada um. Detalhes das interfaces com o usuário relacionadas à manutenção de apresentações de fato encontram-se descritas nos Apêndices A e B.

Tabela 3. Listagem de atributos do conceito "Apresentação de Fato". Fonte primária.

Atributo	Descrição
Identificador	Código interno único gerado para cada apresentação de fato no momento de seu cadastro. Uso interno.
Lista de perspectivas	Relação de perspectivas associadas à apresentação de fato. A presença de ao menos uma perspectiva é obrigatória para tornar a apresentação de fato visível ao usuário.
Nome	Identifica a apresentação de fato para o usuário final. Seu valor deve ser único em relação as demais apresentações cadastradas.
Descrição	Texto descritivo relacionado a apresentação de fato. O valor armazenado é apresentado ao usuário na visão de frontend para fins de informação.
Nome de Tabela	Para uso futuro. Identifica qual é o nome da tabela fato utilizada do data warehouse. Uso interno.
Esquema	Para uso futuro. Indica o esquema ao qual pertence a tabela fato no data warehouse. Uso interno.
Ativo	Controla o ciclo de vida da apresentação. Somente apresentações de fato ativas são apresentadas ao usuário. Uso interno.
Dica	Para uso futuro. Seu objetivo é fornecer ao usuário dicas no momento da escolha da apresentação de fato na visão de frontend.

No caso das perspectivas e perspectivas relacionadas, o conjunto de atributos de informações elencados é consideravelmente maior se comparado às apresentações de fato. Isso porque as perspectivas são diretamente responsáveis por definir a forma como os dados são extraídos do data warehouse e exibidos ao usuário na visão de *frontend*.

Dentre os atributos de informações mais importantes de uma perspectiva (Vide Tabela 5) está o comando de extração de dados. Neste, deve-se obrigatoriamente informar uma instrução SQL válida com um comando DML ⁵ de SELECT que atenda as seguintes restrições:

- Aceita-se apenas o retorno de duas colunas, onde a primeira refere-se à categoria da informação e a segunda, sua respectiva medida numérica (agregação).
- A instrução de SELECT deve fazer uso somente das tabelas fato e tabelas dimensão do data warehouse.
- Aceitam-se somente como parâmetros nomeados os seguintes: `min__year`, `max__year`, `state__id` e `city__id`. Durante a renderização da perspectiva na visão de frontend, os parâmetros recebem os valores correspondentes às escolhas do usuário.

Um exemplo de comando SQL que atende todas as restrições acima descritas encontra-se transposto na Tabela 4.

⁵Data Manipulation Language

Tabela 4. Exemplo de comando SQL de extração de dados associado a uma perspectiva. Fonte primária.

```
1.  select
2.    dd.year,
3.    cast( sum(f.quantity) as integer)
4.  from fact.fact_crime f
5.    inner join dimension.dim_city dc on f.city_id = dc.id
6.    inner join dimension.dim_state ds on f.state_id = ds.id
7.    inner join dimension.dim_crime dct on f.crime_type_id = dct.id
8.    inner join dimension.dim_date dd on f.date_id = dd.id
9.  where
10.   dc.id = :city_id
11.   and ds.id = :state_id
12.   and dd.year between :min_year and :max_year
13.   and dct.crime_name = 'LATROCÍNIO'
14.  group by
15.   dd.year
16.  order by
17.   dd.year desc
```

Detalhes acerca dos demais atributos de informação que compõem o conceito de perspectiva encontram-se descritos na Tabela 5. Vale lembrar que os atributos são idênticos tanto para perspectivas quanto perspectivas relacionadas, visto que ambos os conceitos são implementados pela mesma classe. Outra importante característica dos atributos de informação de perspectivas refere-se as parametrizações de geração de gráficos de plano cartesiano. Portanto, quando menciona-se os eixos X e Y, está sendo feito referência ao eixo de abscissas e eixo de ordenadas, respectivamente.

Tabela 5. Listagem de atributos do conceito "Perspectiva". Fonte primária.

Atributo	Descrição
Identificador	Código interno único gerado para cada perspectiva no momento de seu cadastro. Uso interno.
Título	Título da perspectiva apresentado ao usuário na visão de <i>frontend</i> .
Texto de rodapé	Informação textual, de caráter opcional, posta no rodapé durante a renderização da perspectiva ao usuário na visão de <i>frontend</i> .
Nome	Identificação textual da perspectiva apresentada ao usuário na visão de <i>frontend</i> .
Descrição	Informação textual que serve de apoio ao usuário no momento da análise da perspectiva
Comando	Comando responsável pela extração dos dados do data warehouse. Uso interno.
Tipo de gráfico	Define o tipo de gráfico que será exibido ao usuário. São aceitos gráficos em linha, em barras e em formato de pizza. Uso interno.
Título do eixo X	Título do eixo X no plano cartesiano do gráfico apresentado ao usuário na visão de <i>frontend</i> .
Título do eixo Y	Título do eixo Y no plano cartesiano do gráfico apresentado ao usuário na visão de <i>frontend</i> .
Cor da série	Para uso futuro. Visa definir a cor da série de dados. Uso interno.
Tipo do eixo X	Define o tipo de dado do eixo X. São aceitos dados categóricos, temporais e lineares. Uso interno.
Tipo do eixo Y	Define o tipo de dado do eixo Y. São aceitos dados categóricos, temporais e lineares. Uso interno.
Rotação do título de dados em X	Ângulo de rotação do título dos dados do eixo X. Uso interno.
Número de colunas para legenda	Número de colunas para apresentação da legenda do gráfico. Uso interno.
Posição da legenda	Posição da legenda no gráfico. Aceitam-se valores para as posições superior, inferior, esquerda e direita. Uso interno.
Apresentar título dos dados em X	Define se serão exibidos os identificadores dos dados referentes a uma determinada posição do gráfico. Uso interno.
Apresentar sombra	Ativa/desativa o recurso de sombreado no gráfico. Uso interno.
Perspectivas relacionadas	Conjunto de zero ou mais perspectivas relacionadas disponibilizadas ao usuário na visão de <i>frontend</i> .
Ativo	Controla o ciclo de vida da perspectiva. Somente perspectivas de fato ativas são apresentadas ao usuário. Uso interno.

A tela de *backend* disponibilizada para manter dados de perspectivas é apresentada no Apêndice C. Quanto a tela de manutenção de perspectivas relacionadas, a semelhança é tamanha com a tela de manutenção de perspectivas que optou-se por não apresentá-la. Conjuntamente com a tela de manutenção de apresentação de fatos, as telas de manutenção de perspectivas (relacionadas ou principais) compõem as principais interfaces administrativas da arquitetura OSCV, as quais, pela sua importância, têm acesso restringido por meio de autenticação via usuário e senha.

3.3.2.2 Visão de Frontend

A visão de *frontend* é responsável pela interação com o usuário final da plataforma Open Smart City View. Basicamente, as interfaces que compõem a visão de frontend visam prover informações ao cidadão por meio da apresentação gráfica dos dados armazenados no data warehouse. Nesta camada, os conceitos de apresentação de fato, perspectiva e perspectiva relacionada são completamente abstraídos, tornando-se simplesmente critérios de filtragem a disposição do usuário final.

Dentre as interfaces que compõem a visão de *frontend*, a de apresentação de informações certamente é a mais importante. Sua função é dar condições para o cidadão interagir com os dados armazenados na arquitetura, por meio de buscas e análise das visualizações apresentadas. Assim, ao atender o principal objetivo da OSCV, a interface de apresentação de fatos mostra-se também como a mais importante de toda a arquitetura.

Estruturalmente, a interface de apresentação de informações encontra-se dividida em duas seções distintas. A seção superior contém os critérios de filtragem, sendo utilizada especificamente pelo usuário para determinar as características de sua busca. Já a seção central corresponde a área de renderização, na qual os resultados da pesquisa do usuário são exibidos.

Na seção de critérios de filtragem estão disponíveis à escolha do usuário os seguintes filtros:

- **Estado:** Lista de Unidades da Federação ativas na plataforma. A escolha é unitária e obrigatória.
- **Cidade:** Lista de cidades ativas e pertencentes a Unidade da Federação escolhida. A escolha é unitária e obrigatória.
- **Período:** Intervalo de anos para filtragem. Informação obrigatória.
- **Assunto:** Lista de assuntos disponíveis para visualização. Internamente, esta informação é oriunda das apresentações de fato configuradas pelo administrador da plataforma. A escolha é unitária e obrigatória.
- **Lista de visualizações:** Conjunto de visualizações disponíveis para escolha, relacionadas ao fato selecionado. Representam as perspectivas e perspectivas relacionadas à apresentação de fato escolhida. É obrigatório a escolha de ao menos uma perspectiva.

Na Figura 16, apresenta-se a seção de critérios de filtragem da tela de apresentação de informações. As perspectivas são exibidas ao usuário em um menu em formato de árvore (ao centro da imagem), sendo o primeiro nível reservado para as perspectivas principais e o segundo, para as relacionadas. Neste menu, a escolha de uma perspectiva relacionada impõe, de modo automático, a escolha da perspectiva principal correspondente. Já o oposto não é verdadeiro, ou seja, a escolha de uma perspectiva principal não obriga a escolha de nenhuma relacionada.

Figura 16. Seção de critérios de filtragem - Tela de apresentação de informações. Fonte primária.

No momento que o usuário solicita o processamento de sua pesquisa por meio do botão “Apresentar” (Figura 16), uma sequência de operações é executada internamente para construir a visualização gráfica dos dados. Este processo inicia-se com a validação da entrada do usuário, a fim de garantir que todas as informações obrigatórias tenham sido recebidas. Na sequência, cada uma das perspectivas escolhidas pelo usuário é convertida numa representação gráfica correspondente conforme parametrização informada pelo administrador da arquitetura na visão de *backend*.

Os dados necessários à representação gráfica das perspectivas, sejam elas principais ou relacionadas, são extraídos diretamente do data warehouse por meio da execução da instrução armazenada no atributo **comando** do objeto (Tabela 5). Quando a perspectiva for do tipo principal, os dados obtidos do data warehouse irão compor a **série de dados principal de um novo gráfico** na tela de apresentação. Já os dados resultantes da execução da instrução de uma perspectiva relacionada originam uma **nova série de dados no gráfico correspondente da perspectiva principal**.

Todo o processo de renderização gráfica requer a interação de diferentes componentes internos para alcançar os resultados almejados. O primeiro destes componentes (classe), responsável pela interceptação das solicitações enviadas pelo usuário, chama-se *ContentProviderController*. No modelo arquitetural MVC, o componente *ContentProviderController* atua na camada de controle, recebendo e enviando informações à camada de visão, que, neste caso, refere-se a tela de apresentação de informações.

Também participa do processo o componente *PresentationViewBuilder*, o qual atua diretamente na construção das visualizações por meio do uso de objetos específicos da biblioteca Primefaces ⁶. Este componente é de suma importância, visto que interpreta os metadados contidos nas perspectivas e orquestra todo o processo de construção dos modelos gráficos com base neles. É de sua responsabilidade, por exemplo, processar cada perspectiva escolhida pelo usuário,

⁶<http://www.primefaces.org/>

coletar os dados do data warehouse conforme comando de extração e construir as séries de dados correspondentes no modelo gráfico final.

Além dos componentes *ContentProviderController* e *PresentationViewBuilder*, outras interações se fazem necessárias no processo de renderização. Estas interações, devido a sua complexidade, encontram-se descritas no Diagrama de Sequência apresentado pela Figura 17. Salienta-se que o diagrama representa uma simplificação das interações existentes, contemplando apenas as principais trocas de mensagem envolvidas no processo.

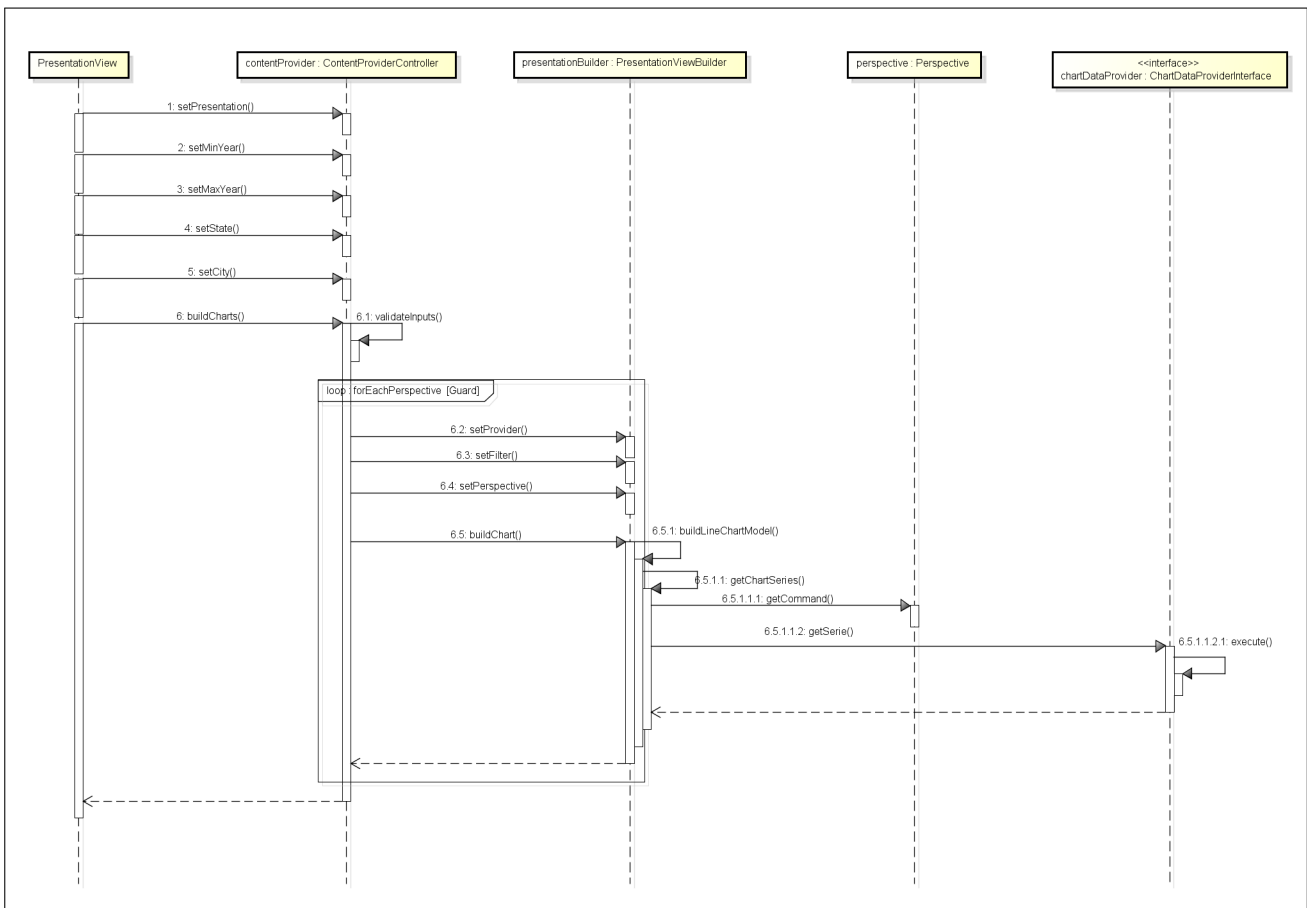


Figura 17. Diagrama de sequência simplificado do processo de renderização de visualizações. Fonte primária.

Supondo-se que o usuário escolha uma perspectiva principal e duas relacionadas, o modelo gráfico resultante será composto de três séries de dados, plotadas conjuntamente no mesmo plano cartesiano. Todas as três séries irão obrigatoriamente compartilhar o mesmo eixo de abcissas (eixo X), mas terão seu próprio eixo de coordenadas (eixo Y). Desta forma, a despeito das diferenças de grandeza de coordenadas entre as séries componentes do gráfico, a disposição conjunta das mesmas no formato de múltiplos eixos Y tende a facilitar a observação imediata de possíveis correlações entre as séries.

Quando apresentado em tela, todo gráfico acompanha um conjunto de informações auxiliares para facilitar sua interpretação. Na parte superior, em formato de título, consta o nome da perspectiva principal. Já na parte inferior, logo abaixo do gráfico, encontra-se a descrição de

cada uma das perspectivas (principal e relacionadas) utilizadas para compor a visualização. Estas informações estão dispostas em um painel do tipo *accordion*, que permite ao usuário expandir os detalhes da perspectiva desejada, enquanto as demais são automaticamente recolhidas (Figura 18).

Adicionalmente, à esquerda e imediatamente abaixo do gráfico, consta o botão “Exportar”. Por meio dele o usuário tem disponível a opção de exportar o gráfico apresentado em tela em formado PNG. Além disso, é possível salvar em arquivo, nos formatos XML e CSV, os dados que foram utilizados na composição das séries do gráfico.

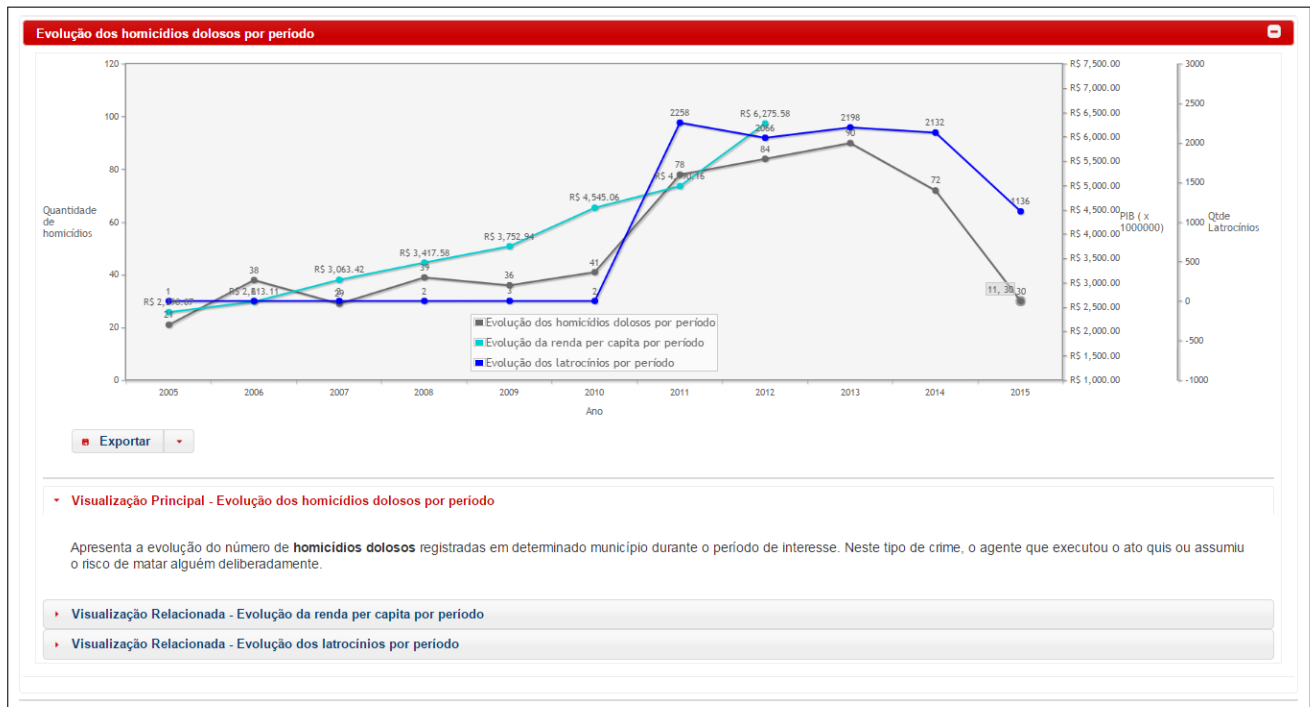


Figura 18. Seção de renderização - Tela de apresentação de informações. Fonte primária.

Visto que o usuário tem a sua disposição a possibilidade de escolha de uma ou mais perspectivas acerca de um determinado assunto (seção de critérios de filtragem), a seção de renderização poderá conter, naturalmente, uma ou mais visualizações, sempre em igual número à quantidade de perspectivas principais selecionadas.

O que se apresenta na Figura 18, portanto, é o resultado da renderização de apenas uma perspectiva principal, associada a duas relacionadas. Havendo-se diversas perspectivas principais, essa mesma estrutura de visualização é repetida para cada uma delas, alterando-se apenas as informações não estáticas. Neste caso, para simplificar a observação dos resultados, o usuário poderá minimizar uma ou mais visualizações, de modo a manter o foco em apenas uma delas, conforme demonstra a Figura 19.

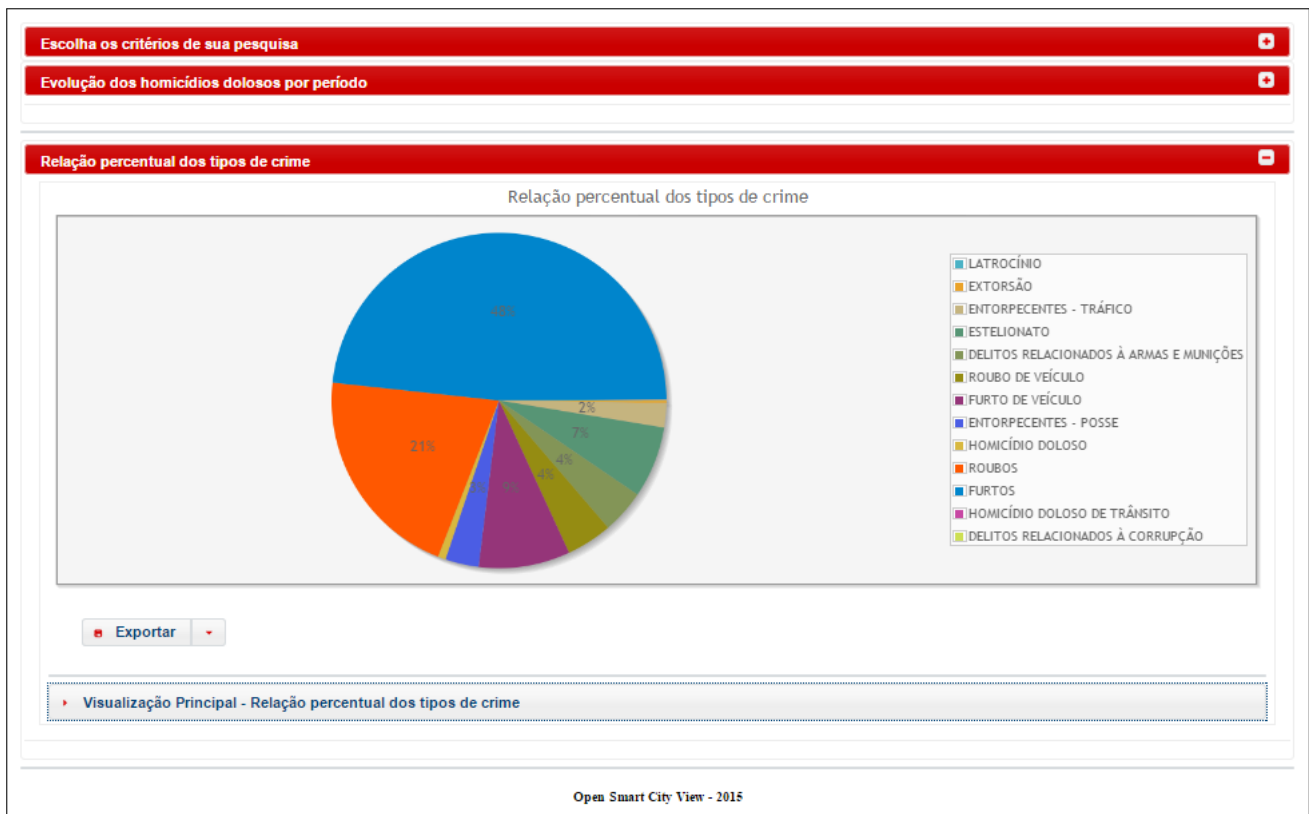


Figura 19. Seção de renderização contendo duas visualizações - Tela de apresentação de informações. Fonte primária.

A visão de *frontend* (e, por conseguinte, de *backend*) foi concebida para estar disponível na Internet, beneficiando-se diretamente da ubiquidade e pervasividade que a rede naturalmente possui. Em sua versão inicial, qualquer dispositivo dispo de um Web Browser adequadamente instalado encontra-se apto a acessá-la. Contudo, no decorrer do processo de evolução da plataforma, espera-se agregar novos e diferentes dispositivos de interação, como smartphones, eletrodomésticos inteligentes, automóveis, entre outros. Todo este universo de dispositivos é mapeado por uma camada específica da arquitetura OSCV, denominada **Dispositivos de Interação**.

3.4 CAMADA DISPOSITIVOS DE INTERAÇÃO

Uma das características que define as cidades como inteligentes é a sua interconectividade, ou seja, o amplo uso de tecnologias de informação e comunicação na gestão e integração de processos e infraestruturas. De modo análogo, o cidadão inteligente é reconhecido por sua capacidade de conectar-se a cidade, interagindo por meio dos diversos canais de comunicação disponibilizados.

Desta forma, sendo um dos objetivos da arquitetura Open Smart City View tornar-se uma importante ferramenta de empoderamento do cidadão inteligente, é de suma importância que a sua amplitude de cobertura estenda-se a todos os canais de comunicação tecnologicamente disponíveis no contexto urbano.

Com vistas a atender tal premissa, é preciso considerar que um amplo universo de dispositivos tecnológicos poderão vir a interagir diretamente com a arquitetura. Embora inicialmente restrita a dispositivos que dispõem de Web Browser instalado, espera-se que, a medida que a arquitetura evoluir, dispositivos móveis, dispositivos vestíveis, eletrodomésticos inteligentes ou mesmo automóveis possam comunicar-se com arquitetura.

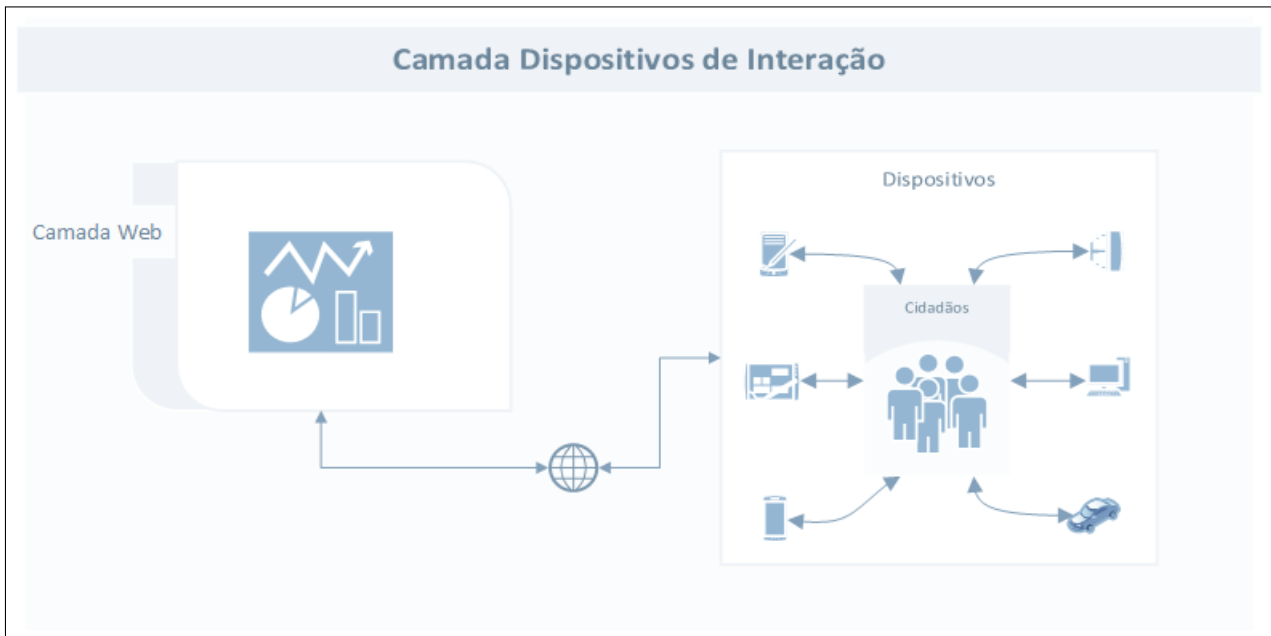


Figura 20. Camada Dispositivos Interação. Fonte primária.

Para representar este universo de diferentes dispositivos aptos a interagir com a arquitetura OSCV definiu-se uma camada específica em seu modelo, denominada **Dispositivos de Interação** (Figura 20). Embora relacionada com aspectos tecnológicos, a camada Dispositivos de Interação tem como pilar fundamental o cidadão. Portanto, sua função concentra-se em mapear os diferentes dispositivos que o cidadão pode – ou poderá num futuro próximo — utilizar para consumir as informações disponíveis na OSCV.

Outra importante função desempenhada pela camada **Dispositivos de Interação** diz respeito as interfaces de integração com outros sistemas. Além dos dispositivos tecnológicos, a arquitetura OSCV tem por meta oferecer mecanismos de integração específicos para permitir que outros softwares possam fazer uso dos dados armazenados no seu data warehouse. Estes mecanismos serão implementados por meio de tecnologias amplamente utilizadas no mercado, como Web Services, API Rest ou mesmo filas JMS. Com isso espera-se alavancar a criação de novas ferramentas de empoderamento ao cidadão, bem como agregar ainda maior valor aos dados armazenados na arquitetura.

3.5 CONSTATAÇÕES, PERCEPÇÕES E EXPECTATIVAS

O estado bruto dos dados governamentais abertos, ao passo que favorece o processamento automatizado, impõe também sérias restrições de uso ao cidadão. Da mesma forma que o minério em estado bruto adquire valor a medida que passa por um processo de refinamento, os dados abertos têm seu real valor evidenciado a medida que convertidos em informações relevantes à sociedade. Contudo, para tal é necessário o uso de ferramentas de software específicas, aptas a converter os dados em informação de valor.

Se não houverem ferramentas de software aptas a entremear o consumo dos dados governamentais, a simples disponibilidade dos mesmos publicamente não é suficiente para o pleno exercício da cidadania. Além disso, ações de transparência pública perdem relevância frente a sociedade devido ao cerceamento (intencional ou não) implicitamente imposto a população por meio das dificuldades de consumo dos dados disponibilizados.

Neste contexto, a concepção e desenvolvimento da arquitetura OSCV apresenta significativa relevância. Ao intermediar o processo de consumo dos dados governamentais abertos, a arquitetura exige o cidadão da necessidade de conhecer a priori qualquer aspecto tecnológico, conceitual ou funcional associado aos conjuntos de dados. Adicionalmente, o tempo necessário ao cidadão para buscar as fontes de dados e selecionar os melhores conjuntos no método tradicional é diretamente investido na interpretação dos resultados gerados pela própria arquitetura.

Por meio da OSCV, portanto, o cidadão tem condições de empoderar-se a despeito de qualquer questão de segunda ordem relacionada à natureza ou estrutura dos dados de origem. Seu foco se mantém no que as informações têm a dizer sobre seu município e, por conseguinte, em qual será seu posicionamento enquanto cidadão frente a elas. Assim, em função de seu papel no processo, a arquitetura OSCV se apresenta como uma importante ferramenta de empoderamento, cuja contribuição social volta-se à consolidação do papel do cidadão inteligente no âmbito de sua comunidade.

No que tange aos aspectos tecnológicos, a OSCV apresenta características desafiadoras. A concepção e desenvolvimento das camadas de *Coleta e Processamento* e *Web*, por exemplo, envolve acentuada complexidade. No caso da camada de Coleta e Processamento, a existência de diferentes scripts de integração e de um data warehouse em constante evolução requer controles apurados e *expertise* por parte dos administradores. Diferentes tecnologias precisam trabalhar em conjunto de forma harmoniosa e consistente. Além disso, a preocupação com requisitos não-funcionais como desempenho devem ser constantes, haja vista que lentidão na recuperação das informações ou indisponibilidades afetam diretamente a percepção de uso da arquitetura por parte do cidadão.

Ainda na *Camada Web*, outro grande desafio refere-se a forma de apresentação das informações ao cidadão. A grande dificuldade, neste ponto específico, está no equilíbrio entre a devida simplicidade e a qualidade da informação oferecida. Não é de interesse da arquitetura explicar o porquê dos fatos apresentados, visto que a dúvida corrobora para o autoquestionamento e de certa forma predispõe o cidadão a novas atitudes no ambiente urbano. Contudo, a arquitetura tem o

dever de explicar o que está sendo apresentado de maneira clara e objetiva, o que acredita-se ter sido alcançado por meio do uso das metáforas de interação definidas como apresentação de fatos, perspectivas e perspectivas relacionadas.

Em relação ao estado da arte, ao avaliar-se a arquitetura OSCV em relação aos trabalhos apresentados por Böhm et al. [34] e Graves e Hendler [35], citados na Seção 2.6, observam-se importantes diferenças. No trabalho de Böhm et al., os conjuntos de dados escolhidos restringem-se àqueles condizentes com questões envolvendo políticos, companhias e financiamentos públicos. Na OSCV, todo conjunto de dados é virtualmente passível de ser consumido, visto que a temática de atuação da arquitetura é determinada pelas demandas de informação existentes. Em termos de apresentação das informações, a OSCV vai ao encontro do trabalho de Graves e Hendler [35], no que tange o uso visualizações como meio de comunicação. Todavia, em termos funcionais, a arquitetura OSCV contém funcionalidades adicionais, destacando-se as etapas de Fontes de Dados e Coleta e Processamento, não consideradas por Graves e Hendler [35].

Por fim, a arquitetura OSCV é concebida com a importante percepção de que precisa estar em constante evolução. Para tal, é imprescindível o uso de práticas, tecnologias e padrões arquiteturais condizentes, que propiciem condições para um processo de evolução natural e eficiente. Deste modo, a concepção e desenvolvimento da OSCV primou pelo uso de *design patterns* consagrados, como o MVC, implementado por meio de tecnologia robusta, expansível e reconhecida no mercado, como é o caso do Java EE. Quanto aos mecanismos de integração, houve a preocupação de se utilizar ferramentas especializadas de ETL e de se definir um protocolo único que serve de base a qualquer novo script criado para alimentar o data warehouse da OSCV.

4. CONSIDERAÇÕES FINAIS

Os desafios vivenciados no dia a dia das cidades demandam respostas rápidas, proativas e eficazes em termos de resultados. Os meios tradicionais de gestão e monitoramento do ambiente urbano muitas vezes não apresentam a capacidade necessária para atender a contento as demandas impostas. Neste momento, as cidades se veem obrigadas a agregar maior inteligência a seus processos, encontrando nas Tecnologias de Informação e Comunicação aliadas indispensáveis no atingimento de suas metas.

Contudo, apesar de fundamentadas no uso intensivo das TICs em seus diferentes sistemas, as cidades inteligentes mantêm seu foco no cidadão. A tecnologia apresenta-se como coadjuvante, sustentando as iniciativas que visam aperfeiçoar a experiência de vida das pessoas, mas sem interferir em demasia na espontaneidade dos indivíduos e nas relações interpessoais. Deste modo, o cidadão criativo e inovador, que se mantém em constante contato com o governo de sua cidade, designa o esteriótipo de cidadão almejado pelas cidades inteligentes.

O cidadão inteligente, mais do que qualquer outro, busca empoderar-se para melhor interagir com sua cidade. Para ser efetivo, este processo de empoderamento precisa ser constantemente suprido com informações de valor acerca dos diferentes tópicos que compõem o contexto social, econômico, cultural e ambiental no qual o cidadão encontra-se inserido. Neste sentido, a presença de fontes de informação de qualidade torna-se fator determinante para que o cidadão inteligente possa desempenhar plenamente seu papel no âmbito de sua comunidade.

Os dados governamentais abertos, enquanto insumo de empoderamento, tem seu real valor diminuído pelas dificuldades que se apresentam na sua obtenção e consumo. Embora de acesso público, uma parcela significativa dos cidadãos abre mão de seus benefícios por não estar apta a fazer o adequado uso dos dados disponíveis. Por conseguinte, estes mesmos cidadão implicitamente afastados do seu direito legítimo de exercício da cidadania, deixam de contribuir de modo mais ativo na comunidade porque não dispõem de informações suficientes para fazê-lo.

Neste contexto, a arquitetura Open Smart City View (OSCV) tem campo fértil para atuação. Os benefícios de seu uso, a priori, vão ao encontro do maior empoderamento do cidadão por meio de informações de qualidade. De modo complementar, ao extrair o verdadeiro valor dos dados governamentais abertos, a OSCV também corrobora significativamente para as iniciativas de transparência pública existentes, dando maior significado às ações e aos dados feito públicos pelas entidades governamentais.

Desafios comumente enfrentados na concepção da arquitetura OSCV, como a falta de documentação dos metadados, inconsistências, duplicidades e a falta de padronização dos portais governamentais brasileiros representam os desafios enfrentados pelos próprios cidadãos ao tentar responder suas questões por meio dos dados governamentais abertos. Além de comprovar o cenário exposto por este trabalho, estes problemas representaram, no contexto de concepção da plataforma OSCV, uma das tarefas mais custosas em termos de desenvolvimento. Todavia, o enfrentamento

de tais desafios reafirmou com ainda maior propriedade a importância que arquitetura OSCV tem enquanto ferramenta de empoderamento.

Outra importante constatação obtida, diretamente associada ao enfrentamento dos problemas relatados, diz respeito a padronização da representação dos dados. A inexistência de um padrão seguido com fidelidade para representar informações de mesma natureza é, em grande parte, a origem de muitas das dificuldades observadas na obtenção e consumo dos dados governamentais abertos. Entende-se, desta forma, que o desenvolvimento de um modelo único de representação, como uma ontologia, por exemplo, traria benefícios ímpares ao processo de transparência pública. Indo além, esse mesmo modelo de representação poderia ser estendido aos portais de transparência, com vistas a facilitar a interação do cidadão com as informações ali fornecidas.

Portanto, diante do exposto, entende-se que a arquitetura OSCV tem plenas condições de potencializar os benefícios dos dados governamentais abertos e, por conseguinte, contribuir com o fortalecimento da iniciativa *open data*. Considera-se ainda que os benefícios da arquitetura OSCV tendem a consolidar-se no decorrer do tempo, por meio de um processo natural de evolução fortemente baseado no feedback dos cidadãos. Além disso, ao intermediar o acesso as informações contidas nos diferentes conjuntos de dados abertos, a OSCV corrobora para a consolidação do perfil do cidadão inteligente, o que denota que a natureza da contribuição deste trabalho e por conseguinte da arquitetura OSCV, extrapola o escopo teórico da Computação e estende-se à aplicação prática dos conceitos em benefício da sociedade de modo geral.

4.1 DISSEMINAÇÃO DO CONHECIMENTO

Compartilhar conhecimento é parte fundamental de qualquer produção científica. A ciência, em sua essência, se constrói e reconstrói por meio das múltiplas expressões que dialogam entre si de modo complementar ou opositivo. Neste sentido, no decorrer do desenvolvimento deste trabalho, desenvolveram-se atividades especificamente voltadas à disseminação do conhecimento gerado, as quais encontram-se elencadas na sequência:

- Redação e submissão de artigo intitulado *Open Smart City View - Uma Arquitetura para Manipulação e Apresentação de Dados Governamentais Abertos em Nível Municipal* ao Seminário Integrado de Software e Hardware - SEMISH ¹. Sendo o principal fórum científico do Congresso da Sociedade Brasileira de Computação (CSBC), o evento possui estrato B4 na classificação Qualis de conferências realizada pela CAPES. Na oportunidade, o artigo submetido fora aceito na categoria *short paper*, razão pela qual optou-se por não dar seguimento à publicação do mesmo neste veículo.
- Redação e submissão de artigo intitulado *Open Smart City View - An Architecture for Open Government Data Manipulation and Presentantion at City Level* à conferencia International Conference WWW/Internet (ICWI) ². Na sua 14^o edição, o evento realizou-se entre os dias

¹<http://csbc2015.cin.ufpe.br>

²<http://internet-conf.org>

24 e 26 de outubro de 2015, na cidade de Dublin, Irlanda. O artigo submetido foi aceito na categoria *full paper*, tendo sido publicado nos anais do evento e nos mecanismos de indexação utilizados pelo mesmo. Quanto ao estrato, a conferência ICWI encontra-se classificada como B2, com índice H igual a 16.

- Apresentação de resumo do presente trabalho na Semana do Conhecimento³. O evento, organizado pela Universidade de Passo Fundo, encontra-se em sua segunda edição, integrando a XXV Mostra de Iniciação Científica e IX Mostra de Extensão. A apresentação do trabalho durante o evento ocorreu no dia 4 de novembro de 2015.

4.2 TRABALHOS FUTUROS

Considerando-se as características da arquitetura OSCV, em especial, o seu potencial enquanto ferramenta analítica, torna-se possível traçar diversos trabalhos futuros associados a mesma. O primeiro deles refere-se a validação da arquitetura em um ambiente real, por meio da disponibilização da ferramenta desenvolvida ao público (cidadãos). O processo avaliativo, neste caso, deve englobar tanto os aspectos técnicos da arquitetura, como tempo de resposta às solicitações feitas pelo usuário, quanto os aspectos sociais, na forma dos efeitos observados no processo de empoderamento dos cidadãos. Para este último, torna-se necessário definir uma metodologia de análise específica, visto que a variável em estudo é de complexa observação.

A automação do processo de coleta dos conjuntos de dados das fontes de interesse representa outra linha de trabalho almejada. Na atual versão da arquitetura, os arquivos são obtidos de forma manual e submetidos para consumo do script ETL correspondente. Pretende-se, contudo, tornar esse processo automático, eliminando qualquer necessidade de intervenção humana. Para tal, uma das abordagens tecnológicas plausíveis são os Web Crawlers, softwares especificamente criados para navegar na Web de forma autônoma e metódica. Assim, a tarefa de rastrear e obter os conjuntos de dados ficaria a cargo dos Web Crawlers, bem como a inicialização da integração ETL responsável pelo consumo dos arquivos, quando estes fossem encontrados nas fontes de interesse.

Outro trabalho futuro diretamente associado aos objetivos da arquitetura OSCV é a implementação de comparadores entre municípios. Atualmente o usuário da plataforma tem a sua disposição visualizações relacionadas a uma cidade em particular. Todavia, entende-se que a possibilidade de comparar um determinado município em relação a outro pode trazer a tona importantes percepções ao cidadão como, por exemplo, evidenciar incoerências no uso do dinheiro público ou mesmo indicar a existência de bons resultados na aplicação das políticas públicas. Diante disso, a comparação entre diferentes municípios pode vir a ter grande relevância como facilitador do processo de empoderamento do cidadão.

Também é de interesse tornar a arquitetura OSCV acessível por meio de aplicativos para dispositivos móveis, especificamente desenvolvidos com uso de tecnologia nativa de cada plataforma operacional. A relevância deste trabalho está fundamentada na grande penetração que as tecnologias

³<http://www.upf.br/semanadoconhecimento>

móveis têm na sociedade de modo geral. Ao torná-la mais próxima do cidadão, a arquitetura OSCV pode beneficiar-se da pervasividade da plataforma móvel para aprimorar sua interação com o usuário final.

Além disso, pretende-se num curto prazo de tempo, aprimorar também a capacidade inferencial e interativa da arquitetura por meio do uso de algoritmos de inteligência artificial voltados aos processos de mineração de dados ⁴ e de recomendação em tempo real. No que diz respeito à mineração de dados, os algoritmos seriam aplicados diretamente no data warehouse da arquitetura a fim de coletar e identificar padrões, anomalias, associações ou qualquer outra relação não explícita nos dados. Os resultados gerados pelo processo de mineração poderia então ser apresentado ao cidadão como uma nova ferramenta de empoderamento.

Já a aplicação dos sistemas de recomendação manteria foco no aperfeiçoamento dos processos de interação com o usuário. Em suma, a ideia assemelha-se ao que ocorre em sistemas de e-commerce, nos quais são lançadas recomendações de compra ao usuário com base no seu perfil. De modo análogo, o uso de técnicas de recomendação traria à arquitetura OSCV a possibilidade de personalização e, se bem sucedida, tornaria a experiência de uso mais agradável e objetiva para o cidadão.

Por fim, outro trabalho de grande valia para a arquitetura OSCV diz respeito a aplicação de ferramentas de Big Data na camada de *Coleta e Processamento*. Contudo, a despeito dos trabalhos anteriormente relacionados, a aplicação de tais ferramentas justifica-se somente a longo prazo. Isso porque, a curto prazo, a arquitetura OSCV não atende aos três pilares fundamentais para definição de um cenário de Big Data, que são **variedade**, **volume** e **velocidade** [42].

Tal fato explica-se pelas características das fontes de dados governamentais abertos atualmente disponíveis. Em um cenário de Big Data, os dados apresentam grande variabilidade quanto a sua estrutura, havendo predominância de informações semi-estruturadas e não estruturadas. Adicionalmente, a velocidade com que os dados são produzidos e consumidos normalmente é de tempo real e o volume gerado extrapola as quantias tidas como normais. O que se observa é que nenhuma das fontes de dados analisadas até então apresenta tais características, o que torna a aplicação de ferramentas de Big Data uma superestimação tecnológica desnecessária no momento.

Todavia, é possível que num futuro não muito próximo, os dados governamentais abertos sejam disponibilizados por novos meios e em diferentes formatos. Neste tempo, então, as ferramentas de Big Data poderão ser acopladas aos processos existentes da arquitetura OSCV para atender demandas realmente ajustadas a um real cenário de Big Data.

⁴Tradução do termo original *data mining*

REFERÊNCIAS BIBLIOGRÁFICAS

- [1] TOWNSEND, A. *Smart Cities: Big Data, Civic Hackers, and the Quest for a New Utopia*. [S.l.]: W. W. Norton, 2013. ISBN 9780393082876.
- [2] DIRKS, S.; KEELING, M. *A vision of smarter cities*. United States of America, 2009, 18 p.
- [3] SCHAFFERS, H. et al. Smart Cities as Innovation Ecosystems sustained by the Future Internet. p. 65, 2012. FIREBALL White paper. Disponível em: <<https://hal.inria.fr/hal-00769635>>.
- [4] AOUN, C. *O alicerce da cidade inteligente: eficiência urbana*. [S.l.], 2013, 16 p. Disponível em: <http://www2.schneider-electric.com/documents/support/white-papers/smart-cities/998-1185469_smart-city-cornerstone-urban-efficiency_BR.pdf>.
- [5] SCHAFFERS, H.; RATTI, C.; KOMNINOS, N. Special Issue on Smart Applications for Smart Cities - New Approaches to Innovation: Guest Editors' Introduction. *Journal of theoretical and applied electronic commerce research*, v. 7, n. 3, p. 9–10, 2012. ISSN 0718-1876. Disponível em: <http://www.scielo.cl/scielo.php?script=sci_arttext&pid=S0718-18762012000300005&lng=en&nrm=iso&tlng=en>.
- [6] AZKUNA, I. naki. *Smart Cities Study*. Bilbao, Biscay, Espanha, 2012.
- [7] GASCÓ-HERNÁNDEZ, M. *Open Government: Opportunities and Challenges for Public Governance*. [S.l.]: Springer Science & Business Media, 2014.
- [8] LEMOS, A. Cidades inteligentes. *GV-Executivo*, v. 12, n. 2, p. 46–49, 2013. Disponível em: <<http://bibliotecadigital.fgv.br/ojs/index.php/gvexecutivo/article/viewFile/20720/19454>>. Acesso em: Ago. 10, 2015.
- [9] DIGITAL, C. of; CITIES, K. based. *Comisión de Ciudades Digitales y del Conocimiento*. 2014. Disponível em: <http://www.cities-localgovernments.org/committees/cdc/Upload/formations/smartcitiesstudy_en.pdf>. Acesso em: Abr. 10, 2014.
- [10] HARRISON, C.; DONNELLY, I. A. A theory of smart cities. *International Society for the Systems Sciences*, n. 55, 2011.
- [11] BERST, J. *Smart Cities Readiness Guide*. Redmound, WA, USA, 2013.
- [12] MACROPLAN. *A Gestão das grandes cidades brasileiras: resultados & práticas*. [S.l.], 2013, 174 p.
- [13] EMERINE, D. et al. This is smart growth. 2006. Disponível em: <<http://trid.trb.org/view.aspx?id=792715>>.

- [14] NAPHADE, M.; BANAVAR, G.; HARRISON, C. Smarter cities and their innovation challenges. *Computer*, p. 32–39, 2011. Disponível em: <http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5875937>.
- [15] CARAGLIU, A.; Del Bo, C.; NIJKAMP, P. Smart Cities in Europe. *Journal of Urban Technology*, v. 18, n. 2, p. 65–82, 2011. ISSN 1063-0732.
- [16] INCORPORATED, G. *IT Glossary*. 2014. Disponível em: <<http://www.gartner.com/it-glossary/>>. Acesso em: Jul. 10, 2014.
- [17] CELINE, I.; KOTOULAS, S. Smart cities. *IEEE Internet Computing*, v. 12, n. 6, p. 8–11, 2013.
- [18] IBM. *Cidades Mais Inteligentes*. 2014. Disponível em: <http://www.ibm.com/smarterplanet/br/pt/smarter_cities/overview/index.html>. Acesso em: Jun. 20, 2014.
- [19] COSGROVE, M. et al. *Front cover Smarter Cities Series : Introducing the IBM City Operations and Management Solution*. United States of America, 2011, 26 p.
- [20] ARRIBAS-BEL, D. Accidental, open and everywhere: Emerging data sources for the understanding of cities. *Applied Geography*, v. 49, p. 45–53, maio 2014. ISSN 01436228. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0143622813002178>>.
- [21] O'REILLY, T. Government as a Platform. *Innovations: Technology, Governance, Globalization*, v. 6, n. 1, p. 13–40, jan. 2011. ISSN 1558-2477. Disponível em: <http://www.mitpressjournals.org/doi/abs/10.1162/INOV_a_00056>.
- [22] MANYIKA, J. et al. *Open data: Unlocking innovation and performance with liquid information*. [S.l.], 2013.
- [23] GRAY, J.; DIETRICH, D.; MCNAMARA, T. *Manual dos dados abertos: governo*. [S.l.], 2010, 58 p. Disponível em: <http://www.w3c.br/pub/Materiais/PublicacoesW3C/Manual_Dados_Abertos_WEB.pdf>.
- [24] SCHUURMAN, D.; BACCARNE, B.; De Marez, L. Smart Ideas for Smart Cities: Investigating Crowdsourcing for Generating and Selecting Ideas for ICT Innovation in a City Context. *Journal of theoretical and applied electronic commerce research*, v. 7, n. 3, p. 11–12, 2012. ISSN 0718-1876. Disponível em: <http://www.scielo.cl/scielo.php?script=sci_arttext&pid=S0718-18762012000300006&lng=en&nrm=iso&tlng=en>.
- [25] LAKHANI, K.; AUSTIN, R.; YI, Y. Data. gov. p. 1–20, 2010. Disponível em: <http://www.data.gov/media/2013/10/attachments/hbs_datagov_case_study_0.pdf>.
- [26] GREENBERG, E.; HICKEY, J. *Water Status - Engage Your Water*. Disponível em: <<http://waterstatus.splashthat.com/>>. Acesso em: Out. 19, 2014.

- [27] GOVERNO FEDERAL. *Conheça seu direito*. Disponível em: <<http://www.acaoainformacao.gov.br/assuntos/conheca-seu-direito>>. Acesso em: Set. 16, 2014.
- [28] ALBANO, C. S. *Dados Governamentais Abertos: Proposta de um modelo de produção e utilização de informações sob a ótica conceitual da cadeia de valor*. 217 p. Tese (Tese de Doutorado em Administração) — Universidade de São Paulo, 2014. Disponível em: <<http://www.teses.usp.br/teses/disponiveis/12/12139/tde-03062014-170642/>>.
- [29] DINIZ, V. Como conseguir dados governamentais abertos. *III Congresso Consad de Gestão Pública*, p. 19, 2010. Disponível em: <http://www.escoladegoverno.pr.gov.br/arquivos/File/Material_CONSAD/paineis_III_congresso_consad/painel_13/como_conseguir_dados_governamentais_abertos.pdf>.
- [30] ELMASRI, R. et al. *Sistemas de banco de dados*. 6. ed. [S.l.]: Pearson Addison Wesley, 2011. 788 p.
- [31] KIMBALL, R.; ROSS, M. *The data warehouse toolkit: The definitive guide to dimensional modeling*. [S.l.]: John Wiley & Sons, 2013.
- [32] PINHO, J. A. G. D. Investigando portais de governo eletrônico de estados no Brasil: muita tecnologia, pouca democracia. *Revista de Administração Pública*, v. 42, n. 3, p. 471–493, 2008. ISSN 0034-7612.
- [33] SALES, T. S.; MARTINS, A. L. P. Planejamento, transparência, controle social e responsabilidade na administração pública após o advento da lei de responsabilidade fiscal. *Nomos*, v. 34, 2014. ISSN 1807-3840.
- [34] BÖHM, C. et al. GovWILD: integrating open government data for transparency. *Proceedings of the 21st international conference companion on World Wide Web - WWW '12 Companion*, p. 321–324, 2012. Disponível em: <<http://dl.acm.org/citation.cfm?id=2188039>\n<http://dl.acm.org/citation.cfm?doid=2187980.2188039>>.
- [35] GRAVES, A.; HENDLER, J. Visualization tools for open government data. In: ACM. *Proceedings of the 14th Annual International Conference on Digital Government Research*. [S.l.], 2013. p. 136–145.
- [36] W3C. *Resource Description Framework (RDF)*. Disponível em: <<http://www.w3.org/RDF/>>. Acesso em: Out. 02, 2014.
- [37] OPENOFFICE.ORG. *OpenOffice.org's Documentation of the Microsoft Excel File Format*. 2015. Disponível em: <<https://www.openoffice.org/sc/excelfileformat.pdf>>. Acesso em: Out. 31, 2015.

- [38] CASTERS, M.; BOUMAN, R.; DONGEN, J. V. *Pentaho Kettle solutions: building open source ETL solutions with Pentaho Data Integration*. [S.l.]: John Wiley & Sons, 2010.
- [39] DEITEL, P.; DEITEL, H. *Java - Como Programar*. 8. ed. [S.l.]: Pearson Prentice Hall. 1144 p. ISBN 9788576055631.
- [40] ORACLE CORPORATION. *JavaServer Faces Technology Overview*. 2015. Disponível em: <<http://www.oracle.com/technetwork/java/javaee/overview-140548.html>>. Acesso em: Out. 10, 2015.
- [41] FOWLER, M. *Model View Controller*. 2015. Disponível em: <<http://martinfowler.com/eaDev/uiArchs.html#ModelViewController>>. Acesso em: Out. 08, 2015.
- [42] GANTZ, J.; RAISEL, D. *The digital universe in 2020: big data, bigger digital shadows, and biggest growth in the far east*.

APÊNDICE A – BACKEND - GESTÃO DE APRESENTAÇÃO DE FATOS

Gestão de apresentações de fatos

Ajuda

O data warehouse da arquitetura OSCV pode conter diversas tabelas fato. Uma **apresentação de fato** refere-se ao grupo de perspectivas que fazem uso de uma tabela fato em particular para fornecer ao cidadão informações relevantes.

	Código	Nome	Tabela Fato	Dica	Situação
<input type="radio"/>	1	Segurança	fact_vehicte_fleet		Ativo
<input checked="" type="radio"/>	2	Frota	Nova tabela	Ajuda	Inativo

Adicionar **Editar** **Remover**

Open Smart City View - 2015

Figura 21. Tela de listagem das apresentações de fato cadastradas na plataforma. Fonte primária.

APÊNDICE B – BACKEND - MANTER APRESENTAÇÃO DE FATOS

Manter apresentação de fato
☰

Ajuda

Uma apresentação de fato refere-se ao grupo de perspectivas que fazem uso de uma tabela fato em particular para apresentar ao cidadão informações relevantes.

Uma perspectiva compreende a uma observação particular feita sobre dados de uma tabela fato. Diz respeito a alguma questão a ser respondida por meio de consulta ao data warehouse e relaciona-se como as apresentações de fato na proporção de 1 para n, ou seja, uma apresentação de fatos pode ter uma ou muitas perspectivas associadas.

Nome *
Segurança

Esquema *
fact

Tabela *
fact_security

Ativo *

Ajuda *
Informações de ajuda

Descrição

O fato **Segurança** descreve informações relevantes ao cidadão referentes a crimes cometidos em determinado município.

Voltar
Salvar

Perspectivas Associadas
☰

Código	Nome	Título	Tipo de Gráfico
<input type="radio"/> 2	Homicídios	Evolução dos homicídios no município	LINE
<input checked="" type="radio"/> 3	Violência contra a mulher	Evolução da violência contra mulher no município	BAR

Adicionar
Editar
Remover

Open Smart City View - 2015

Figura 22. Tela de manutenção de apresentações de fato. Fonte primária.

APÊNDICE C – BACKEND - MANTER PERSPECTIVAS

Opções

Gerenciar perspectivas relacionadas a apresentação de fato [Frota]

* Ajuda

Nome *
Homicídios

Título *
Evolução dos homicídios no município

Rodapé
Informações relevantes de rodapé

Etique eixo X
X label

Rotação etiqueta em X
30

Tipo eixo X
Categorico

Etique eixo Y
TEste

Tipo de Eixo Y
Categorico

Colunas para legenda *
1

Cor da série

Tipo de Gráfico *
Em linha

Posição da Legenda *
Esquerda

- Voltar

Comando *

```
select d.city_name, sum(f.quantity)
from fact.fac_vehicle_fleet f inner join dimension.dim_city d
on d.id = f.city_id
inner join dimension.dim_date dd on dd.id = f.date_id
where dd.year between :min_year and :max_year
group by d.city_name
order by 1
```

Descrição *

Apresenta a evolução do número de homicídios apresentada por um determinado município durante o período de interesse.

Exibir legenda de dados
 Exibir sombra no gráfico
 Ativo

Salvar

Perspectivas relacionadas

	Código	Nome	Título	Tipo de Gráfico
•	8	Evolução da renda per capita	Evolução da renda per capita no município	BAR
Adicionar Editar Remover				

Open Smart City View - 2015

Figura 23. Tela de manutenção de perspectivas. Fonte primária.

ANEXO A – SSP: CONJUNTO DE DADOS SOBRE CRIMES

Tabela 6. Trecho do arquivo contendo dados de crimes em municípios - SSP/RS

Municípios	Homicídio Doloso	Homicídio Doloso de Trânsito	Furtos	Furto de Veículo	Roubos	Latrocínio	Roubo de Veículo	Extorsão	Extorsão Mediante Sequestro	Estelionato
ACEGUA	0	0	25	2	1	0	0	0	0	0
AGUA SANTA	0	0	18	3	0	0	0	0	0	1
AGUDO	0	0	79	6	2	0	0	0	0	2
AJURICABA	0	0	19	0	1	0	0	0	0	4
ALECRIM	0	0	44	0	0	0	1	0	0	1
ALEGRETE	2	0	357	12	50	0	0	0	0	30
ALEGRIA	0	0	14	0	0	0	0	0	0	0
ALMIRANTE TAMANDARE DO SUL	0	0	7	0	1	0	0	0	0	0
ALPESTRE	0	0	46	6	1	0	2	0	0	0
ALTO ALEGRE	0	0	4	0	1	0	1	0	0	1
ALTO FELIZ	0	0	6	0	0	0	0	0	0	0
ALVORADA	62	0	1.123	179	1.394	2	213	1	0	117
AMARAL FERRADOR	0	0	20	1	0	0	0	0	0	0
AMETISTA DO SUL	1	0	28	7	0	0	0	0	0	1
ANDRE DA ROCHA	0	0	10	1	0	0	0	0	0	1
ANTA GORDA	0	0	11	1	1	0	0	0	0	1
ANTONIO PRADO	0	0	61	7	2	0	1	0	0	11
ARAMBARE	0	0	69	1	5	0	0	0	0	2
ARARICA	0	0	14	0	8	0	5	1	0	3
ARATIBA	0	0	13	0	2	0	0	0	0	2
ARROIO DO MEIO	0	0	105	11	13	0	1	1	0	16
ARROIO DO PADRE	0	0	2	0	0	0	0	0	0	0
ARROIO DO SAL	2	0	235	11	23	0	1	0	0	7
ARROIO DO TIGRE	0	0	76	5	6	0	0	0	0	2
ARROIO DOS RATOS	0	0	70	1	12	1	0	0	0	2
ARROIO GRANDE	1	0	121	5	4	0	0	0	0	5
ARVOREZINHA	0	0	54	5	6	0	0	0	0	1
AUGUSTO PESTANA	0	0	32	2	1	0	0	0	0	4
AUREA	0	0	8	1	1	0	0	0	0	0
BAGE	7	0	929	39	160	1	11	3	0	46
BALNEARIO PINHAL	4	0	263	21	40	1	7	2	0	17
BARAO	0	0	11	2	1	0	0	0	0	2
BARAO DE COTEGIPE	1	0	31	1	3	1	1	0	0	1
BARAO DO TRIUNFO	1	0	33	2	6	0	0	0	0	1

ANEXO B – IBGE: CONJUNTO DE DADOS SOBRE LEGISLAÇÃO DE MUNICÍPIOS

Tabela 7. Trecho do arquivo contendo dados da legislação de municípios brasileiros - IBGE

UF	NOME_UF	UF_MUNIC	MUNIC	Nome_Município	LEI	DTCRIACAO	DTINSTALA	PROV1	NOME_PROV1	PROV2
11	Estado de Rondônia	1100015	00015	Alta Floresta D'Oeste	104	20/05/1986	31/12/1986	1100080	Costa Marques	0
11	Estado de Rondônia	1100023	00023	Ariquemes	6448	11/10/1977	21/11/1977	1100205	Porto Velho	0
11	Estado de Rondônia	1100031	00031	Cabixi	201	01/06/1988	01/01/1989	1100064	Colorado do Oeste	0
11	Estado de Rondônia	1100049	00049	Cacoal	6448	11/10/1977	21/11/1977	1100205	Porto Velho	0
11	Estado de Rondônia	1100056	00056	Cerejeiras	71	05/08/1983	30/12/1984	1100064	Colorado do Oeste	0
11	Estado de Rondônia	1100064	00064	Colorado do Oeste	6921	16/06/1981	06/10/1983	1100304	Vilhena	0
11	Estado de Rondônia	1100072	00072	Corumbiara	377	13/02/1992	01/01/1993	1100064	Colorado do Oeste	1100304
11	Estado de Rondônia	1100080	00080	Costa Marques	6921	16/06/1981	05/10/1983	1100106	Guajará-Mirim	0
11	Estado de Rondônia	1100098	00098	Espigão D'Oeste	6921	16/06/1981	01/02/1983	1100189	Pimenta Bueno	0
11	Estado de Rondônia	1100106	00106	Guajará-Mirim	991	12/07/1928	13/09/1943	5199904	Guajará-Mirim	0
11	Estado de Rondônia	1100114	00114	Jaru	6921	16/06/1981	01/02/1983	1100023	Ariquemes	1100122
11	Estado de Rondônia	1100122	00122	Ji-Paraná	6448	11/10/1977	22/11/1977	1100205	Porto Velho	0
11	Estado de Rondônia	1100130	00130	Machadinho D'Oeste	198	11/05/1988	01/01/1989	1100023	Ariquemes	1100114
11	Estado de Rondônia	1100148	00148	Nova Brasilândia D'Oeste	157	19/06/1987	01/01/1989	1100254	Presidente Médici	0
11	Estado de Rondônia	1100155	00155	Ouro Preto do Oeste	6921	16/06/1981	01/02/1983	1100122	Ji-Paraná	0
11	Estado de Rondônia	1100189	00189	Pimenta Bueno	6448	11/10/1977	24/11/1977	1100205	Porto Velho	0
11	Estado de Rondônia	1100205	00205	Porto Velho	757	02/10/1914	13/09/1943	1399903	Porto Velho	0
11	Estado de Rondônia	1100254	00254	Presidente Médici	6921	16/06/1981	01/02/1983	1100122	Ji-Paraná	0
11	Estado de Rondônia	1100262	00262	Rio Crespo	376	13/02/1992	01/01/1993	1100023	Ariquemes	1100130
11	Estado de Rondônia	1100288	00288	Rolim de Moura	71	05/08/1983	28/12/1984	1100049	Cacoal	0
11	Estado de Rondônia	1100296	00296	Santa Luzia D'Oeste	100	11/05/1986	31/12/1986	1100189	Pimenta Bueno	1100288
11	Estado de Rondônia	1100304	00304	Vilhena	6448	11/10/1977	23/11/1977	1100106	Guajará-Mirim	1100205

ANEXO C – DADOS.RS: CONJUNTO DE DADOS SOBRE VIOLÊNCIA CONTRA A MULHER

Tabela 8. Trecho do arquivo contendo dados de violência contra a mulher - DADOS.RS

Município	Ano	Mes	DiadoMes	DiadaSemana	Turno	Hora	Tentado	TipoLocal	Idadevitima	IdadeAutor
ACEGUA RS	2006	6	2	sexta-feira	Madrugada	01:00:00	Consumado		16	45
ACEGUA RS	2007	10	22	segunda-feira	Manh.,	09:30:00	Consumado	Residência	12	32
ACEGUA RS	2011	1	18	terÁa-feira	Madrugada	00:30:00	Consumado	Residência	15	27
AGUASANTA RS	2006	1	1	domingo	Madrugada	00:01:00	Consumado	Residência	9	19
AGUASANTA RS	2008	1	30	quarta-feira	Madrugada	00:01:00	Consumado	Residência	14	34
AGUASCLA RAS(VIAMAO)RS	2010	4	3	s-bado	Tarde	14:00:00	Consumado		10	63
AGUDO RS	2006	5	10	quarta-feira	Tarde	12:00:00	Consumado	Out ros	17	22
AGUDO RS	2006	12	30	s-bado	Manh.,	06:30:00	Tentado	Residência	13	28
AGUDO RS	2007	3	1	quinta-feira	Manh.,	08:00:00	Consumado	Residência	10	23
AGUDO RS	2007	4	12	quinta-feira	Madrugada	00:15:00	Tentado	ViaPublica	15	30
AGUDO RS	2007	12	8	s-bado	Manh.,	10:00:00	Tentado	Residência	10	71
AGUDO RS	2009	1	15	quinta-feira	Noite	18:00:00	Consumado	Residência	9	14
AGUDO RS	2009	8	25	terÁa-feira	Tarde	14:00:00	Tentado	Out ros	20	38
AGUDO RS	2009	8	29	s-bado	Noite	23:40:00	Tentado		17	0
AGUDO RS	2009	9	10	quinta-feira	Noite	22:40:00	Tentado		22	0
AGUDO RS	2010	1	9	s-bado	Tarde	15:30:00	Tentado	Residência	6	30
AGUDO RS	2010	4	3	s-bado	Madrugada	02:00:00	Consumado		49	0
AGUDO RS	2010	12	24	sexta-feira	Tarde	12:30:00	Consumado	Out ros	14	31
AGUDO RS	2011	3	17	quinta-feira	Manh.,	08:00:00	Consumado	Out ros	14	25
AGUDO RS	2011	4	3	domingo	Noite	19:00:00	Tentado	Residência	40	79
AGUDO RS	2011	4	6	quarta-feira	Tarde	14:30:00	Tentado	Residência	16	42
AGUDO RS	2011	5	21	s-bado	Madrugada	05:20:00	Tentado	ViaPublica	49	38
AGUDO RS	2011	5	22	domingo	Noite	20:30:00	Consumado	Residência	42	27
AGUDO RS	2011	5	25	quarta-feira	Noite	19:00:00	Tentado	Residência	14	37
AGUDO RS	2011	6	18	s-bado	Madrugada	04:00:00	Tentado	Out ros	17	35
AGUDO RS	2011	10	16	domingo	Tarde	16:00:00	Consumado	ViaPublica	10	0
AJURICABA RS	2007	11	14	quarta-feira	Tarde	16:00:00	Tentado	Residência	28	49
AJURICABA RS	2008	2	14	quinta-feira	Manh.,	09:00:00	Consumado	Out ros	3	0
AJURICABA RS	2008	3	1	s-bado	Manh.,	09:00:00	Consumado	Residência	16	43
AJURICABA RS	2010	1	1	sexta-feira	Manh.,	08:00:00	Consumado	Residência	7	13
AJURICABA RS	2010	9	29	quarta-feira	Manh.,	10:30:00	Tentado	Residência	38	56
AJURICABA RS	2010	12	11	s-bado	Madrugada	00:01:00	Consumado	Residência	10	42

ANEXO D – DADOS.RS: CONJUNTO DE DADOS SOBRE CRIMES DE TRANSITO

Tabela 9. Trecho do arquivo contendo dados de crimes de trânsito - DADOS.RS

Data	Hora	Município	Nro da Ocorrência	Tipo de Via	Tipo Logradouro	Logradouro	KM	Natureza Ocorrência	Tipo Ocorrência	Historico	Teor Alcoólico
01/05/2011	14:00	Seberi	364	Municipal	Rua	FLORES DA CUNHA	1359	Choque Com Objeto Fixo	Homicídio De Trânsito		
14/08/2011	13:20	Sede Nova	618	Municipal	Estrada	LINHA PELETI		Choque Com Objeto Fixo	Homicídio De Trânsito		
15/04/2011	17:30	Segredo	966	Municipal	Rua	PADRE JOAO PASA		Colisão	Homicídio De Trânsito		
02/08/2011	11:45	Selbach	203	Municipal	Estrada	LINHA FLORESTA		Capotagem	Outros Acidentes		
04/04/2011	10:00	Selbach	113	Estadual	Rodovia	RS 223		Tombamento	Outros Acidentes		
11/08/2011	08:40	Sentinela do Sul	1664	Municipal	Rua	HERMES RODEL		Colisão	Homicídio De Trânsito		
19/12/2011	01:10	Serafina Corrêa	1489	Estadual	Rodovia	RS 129	144	Colisão	Homicídio De Trânsito		
16/06/2011	08:30	Sertão	206	Estadual	Rodovia	RS 135	40	Colisão	Homicídio De Trânsito		
11/04/2011	11:00	Sertão Santana	600	Municipal	Estrada	DA MARAVILHA		Capotagem	Outros Acidentes		
05/04/2011	18:15	Severiano de Almeida	4289	Federal	Rodovia	BR 153	17	Atropelamento	Homicídio De Trânsito		
13/10/2011	15:30	Severiano de Almeida	316	Federal	Rodovia	BR 153	16	Colisão	Homicídio De Trânsito		
07/04/2011	13:00	Sinimbu	144	Estadual	Rodovia	BR 153		Atropelamento	Homicídio De Trânsito		
26/09/2011	23:30	Sobradinho	8059	Estadual	Rodovia	RS 481		Colisão Lateral	Homicídio De Trânsito		
04/03/2011	16:50	Sobradinho	578	Estadual	Rodovia	RS 400	44	Atropelamento	Lesão Corporal Culposa		
11/12/2011	22:10	Sobradinho	3022	Municipal	Rua	EUCLIDES BENTO PEREIRA		Colisão	Homicídio De Trânsito		
27/04/2011	12:30	Sobradinho	1055	Municipal	Rua	ELOI DE OLIVEIRA BRITO	658	Atropelamento	Homicídio De Trânsito		
24/11/2011	17:30	Soledade	5719	Estadual	Rodovia	RS 153	106	Tombamento	Homicídio De Trânsito		
10/09/2011	03:45	Soledade	4546	Federal	Rodovia	BR 386	244	Colisão	Homicídio De Trânsito		
13/01/2011	19:30	Soledade	290	Federal	Rodovia	BR 386	198	Tombamento	Outros Acidentes		
29/12/2011	08:15	Soledade	27 / 6368	Estadual	Rodovia	RS 153	88	Colisão	Homicídio De Trânsito		
21/04/2011	15:00	Soledade	2012	Federal	Rodovia	BR 386	236	Colisão	Homicídio De Trânsito		
13/03/2011	19:00	Soledade	1397	Estadual	Rodovia	RS 332	78	Colisão	Homicídio De Trânsito		
12/06/2011	10:20	Soledade	117730	Estadual	Rodovia	RS 153	100	Colisão Lateral	Lesão De Trânsito		
23/06/2011	18:20	Tapejara	1215	Estadual	Rodovia	RS 463	3	Atropelamento	Homicídio De Trânsito		
30/01/2011	21:20	Tapera	95	Estadual	Rodovia	RS 332	147	Capotagem	Outros Crimes		
29/04/2011	13:55	Tapes	977	Federal	Rodovia	BR 116	350,5	Colisão	Homicídio De Trânsito		
04/03/2011	16:50	Tapes	523	Federal	Rodovia	BR 116		Colisão	Homicídio De Trânsito		
26/02/2011	20:00	Tapes	475	Federal	Rodovia	BR 116	359,8	Colisão	Homicídio De Trânsito		
25/02/2011	20:00	Tapes	473	Federal	Rodovia	BR 116	372,9	Colisão	Homicídio De Trânsito		
22/02/2011	06:50	Tapes	429	Federal	Rodovia	BR 116	372,4	Colisão	Homicídio De Trânsito		
19/02/2011	06:30	Tapes	416	Municipal	Rua	MELVIN JONES		Tombamento	Encontro De Cadáver		
03/12/2011	00:35	Tapes	2573 / 488574	Estadual	Rodovia	RS 717	5	Tombamento	Lesão De Trânsito		
01/01/2011	06:15	Tapes	2	Municipal	Rua	SYDIA ALBUQUERQUE JARDIM	1000	Colisão	Homicídio De Trânsito		
04/09/2011	17:00	Tapes	1845	Municipal	Estrada	DO CAPAO DA MOCA	10	Tombamento	Homicídio De Trânsito		
09/08/2011	03:00	Tapes	1651	Federal	Rodovia	BR 116	371,6	Choque Com Objeto Fixo	Homicídio De Trânsito		
17/07/2011	18:05	Tapes	1503	Federal	Rodovia	BR 116	363	Atropelamento	Homicídio De Trânsito		
18/02/2011	21:30	Taquara	924	Estadual	Rodovia	RS 020	49	Colisão Lateral	Homicídio De Trânsito		
13/02/2011	23:30	Taquara	792	Estadual	Rodovia	RS 115	3	Colisão	Homicídio De Trânsito		
12/10/2011	04:00	Taquara	5575	Estadual	Rodovia	RS 115	3	Colisão Lateral	Homicídio De Trânsito		
30/01/2011	22:00	Taquara	527	Estadual	Rodovia	RS 239	61	Colisão	Homicídio De Trânsito		
21/09/2011	14:00	Taquara	5175	Estadual	Rodovia	RS 115	2	Atropelamento	Homicídio De Trânsito		
17/09/2011	03:00	Taquara	5101	Estadual	Rodovia	RS 115	3	Choque Com Objeto Fixo	Homicídio De Trânsito		
23/07/2011	16:35	Taquara	3978	Municipal	Rua	EMILIO BADERMANN	1140	Colisão	Lesão De Trânsito		
16/07/2011	19:15	Taquara	3866	Estadual	Rodovia	RS 020	41	Colisão	Homicídio De Trânsito		
08/07/2011	20:30	Taquara	3700	Estadual	Rodovia	RS 020		Colisão	Falecimento		
08/07/2011	16:45	Taquara	3691	Estadual	Rodovia	RS 239	59	Colisão	Homicídio De Trânsito		
20/06/2011	17:55	Taquara	3368	Estadual	Rodovia	RS 020	47	Colisão	Lesão Corporal Seguida De Morte		
05/05/2011	07:00	Taquara	2474	Estadual	Rodovia	RS 239	46	Colisão	Homicídio De Trânsito		
24/04/2011	18:30	Taquara	2247	Estadual	Rodovia	RS 242	10	Capotagem	Homicídio De Trânsito		

ANEXO E – FEE: CONJUNTO DE DADOS SOBRE PRODUTO INTERNO BRUTO

Tabela 10. Trecho do arquivo contendo dados de Produto Interno Bruto de municípios - FEE

Municípios	Código	1999 (R\$ mil)	2000 (R\$ mil)	2001 (R\$ mil)	2002 (R\$ mil)	2003 (R\$ mil)	2004 (R\$ mil)	2005 (R\$ mil)	2006 (R\$ mil)	2007 (R\$ mil)	2008 (R\$ mil)	2009 (R\$ mil)	2010 (R\$ mil)
Aceguá	4300034	-	-	38371,76	46840,98	60463,13	79677,52	72228,14	67309,79	76587,51	87540,46	115740,9	133773,7
Água Santa	4300059	23001,94	23956,02	27778,82	33936,97	49663,43	44366,41	32114,28	49437,7	68496,57	83091,1	81634,91	96517,36
Agudo	4300109	87575,75	100570,8	106582,2	143889,9	179453,9	180105,5	168366,8	188511	202383,6	232095,5	267519,4	274205,9
Ajuricaba	4300208	38648,97	39448,91	48930,74	52165,52	84132,2	70240,1	59833,13	78662,94	119040,7	126339,3	151134	151134
Alecrim	4300307	28018,04	30501,2	34487,32	37454,32	45993,96	44878,12	44452,73	50550,1	58678,14	64079,19	68997,46	80604,82
Alegrete	4300406	414547,7	407197,6	482480,9	528119,4	648317,6	763009,4	745086,6	820066,3	1104147	1133620	1247546	1247546
Alegria	4300455	18577,51	20895,93	23246,81	24810,95	40936	33886,52	26155,44	32131,64	44732,37	46073,24	48876,09	56182,94
Almirante Tamandaré do Sul	4300471	-	-	27847,41	31757,29	49133,84	39930,72	23748,38	45319,22	51932,62	61804,41	59517,19	55547,54
Alpestre	4300505	31276,53	34752,72	39126,53	40886,75	61006,95	54954,52	56069,22	63574	78876,1	84930,29	87491,8	107537
Alto Alegre	4300554	9702,852	10672,15	14537,94	16680,09	23352,97	16730,2	13450,43	20091,13	24372,64	30924,83	34216,01	33297,1
Alto Feliz	4300570	12948,37	14603,94	17364,97	17409,69	18804,66	20148,06	21192,95	23420,62	23573,89	23408,59	27992,44	38175,54
Alvorada	4300604	376974,3	424955,3	490581,2	582722,9	637090,9	695045,4	788897,8	852208,7	943154,4	1074664	1204554	1467892
Amaral Ferrador	4300638	16124,67	16873,13	23985,45	26875,45	34821,67	36491,25	32160,2	39456,28	43374,03	49313,78	61772,56	58605,17
Ametista do Sul	4300646	19404,43	21041,6	24030,81	29976,51	36928,59	38783,2	41353,7	44536,02	47420,06	48770,35	53812,46	67605,21
André da Rocha	4300661	8607,112	9472,898	12950,81	12835,6	22925,31	22888,82	15799,93	27623,91	32111,18	41153,86	39286,41	35228,03
Anta Gorda	4300703	28486,63	34686,79	47885,14	56318,88	71394,04	68870,62	55138,7	65859,92	80940,44	91101,81	92958,94	108386,8
Antônio Prado	4300802	102025,3	105410	115941,3	136703,4	160942,3	170996,9	172025	190334,4	196274,8	209272,2	237815,2	279116,1
Arambaré	4300851	24112,15	25529,52	30055,79	32683,72	45296,87	55374,96	39060,12	41688,45	50761,04	62027,64	72358,32	79367,55
Araricá	4300877	17660,68	18799,98	17146,66	21529,86	27014,59	31665,69	35394,29	39048,07	40381,91	45917,11	60713,48	69341,48
Aratiba	4300901	53720,16	76058,54	158178,7	102313,3	160845,2	232744,3	231446	249418,1	288783,9	245181,2	329578,8	382969,9
Arroio do Meio	4301008	199426,4	224258,8	228365,9	287729,7	348722,5	348869,5	402224,6	409539,4	432122,1	436223,8	528966,5	588317,6
Arroio do Padre	4301073	-	-	9112,964	10200,37	13975,69	15009,91	16763,98	19688,74	21120,03	23207,45	26117,94	29488,18
Arroio do Sal	4301057	23888,54	26757,98	29690,3	34049,18	42258,7	45920,78	53637,87	62869,34	67840,98	71808,81	80793,55	97310,38
Arroio dos Ratos	4301107	42110,8	52190,75	55177,55	72124,66	72124,66	87541,73	117942,8	99631,52	104400,9	117132,7	108628,4	121874,4
Arroio do Tigre	4301206	58409,21	66164,6	88062,54	99194,7	118982,4	128393,9	126675,3	155621,8	169347	195746,3	210520,6	202406,1
Arroio Grande	4301305	107563,8	99540,97	126001,8	131112,8	170611,8	204297,1	169924,2	189422,5	209842,5	260428,8	315712,4	349018
Arvorezinha	4301404	37965,12	41252,53	53786	59579,43	62584,92	71112,53	74055,14	90974,61	98540,7	105035,4	121134	125878,5
Augusto Pestana	4301503	41967,63	38000,64	51144,07	55972,84	85903,16	77216,43	64350,26	80274,77	115027	132309	129934,1	149587,3
Áurea	4301552	13928,32	14608,59	18213,01	20416,86	32563,69	30497,28	24276,24	31917,34	42096,9	44762,53	45139,6	52225,57
Bagé	4301602	540219	583963,7	606687,3	676716,8	775673,3	861892,7	909794,9	972091,8	1023342	1117882	1197238	1428995
Balneário Pinhal	4301636	33015,09	36334,65	39997,25	43291,61	53504,88	58516,41	63310,35	70967,6	78018,91	85512,27	93531,43	114511,5
Barão	4301651	22123,8	23334,1	28694,41	31542,99	33263,99	37034,3	40148,84	43974,23	49929,48	59474,7	80984,35	101730,9
Barão de Cotegipe	4301701	30684,23	31797,23	33024,85	35812,41	58238,75	51496,47	50006,7	60359,98	77698,25	86387,2	92871,66	110709,5
Barão do Triunfo	4301750	21385,1	22765,9	29079,6	37454,68	49215,65	59279,87	47368,56	50028,61	52784,23	60916,02	71024,03	66930,8
Barracão	4301800	23363,28	27633,49	34154,87	38704,1	61585,77	44849,49	40421,48	63568,1	83817,15	102501,2	99556,99	118143,4
Barra do Guarita	4301859	8152,391	11370,13	14801,04	12111,47	14218,85	12962,05	14866,23	17305,92	19672,29	21154,05	22052,83	26494,63
Barra do Quaraí	4301875	45904,24	37052,26	47260,82	55820,75	69507,1	84717,47	61884,17	78905,89	63634,39	98352,18	108405,2	134310,9
Barra do Ribeiro	4301909	48357,76	53229,21	58715,54	71823,31	90783,01	101929,8	90664,56	101929,8	108691,7	125592	141488,7	169879,6
Barra do Rio Azul	4301925	9141,821	9640,353	11114,6	11379,15	16393,15	16116,61	16278,67	17915,81	25839,34	32354,1	28704,89	34438,61
Barra Funda	4301958	18101,71	19327	20344,46	23429,31	34600,13	38389,47	36063,57	44619,23	51672,67	57125,82	63511,94	74214,93
Barros Cassal	4302006	36581,86	39076,94	50345,01	58738,04	74722,44	77512,73	70087,82	87712,72	89024,56	99136,2	113025,6	117751
Benjamin Constant do Sul	4302055	6712,186	7038,696	8312,677	8342,853	12922,96	11920,19	11384,87	13185	15190,83	18032,31	17007,98	21882,92
Bento Gonçalves	4302105	1014926	1129440	1191608	1356656	1453355	1696799	1914715	1990020	2204611	2219408	2605711	3103029
Boa Vista das Missões	4302154	12732,93	14838,12	26155,07	23169,19	43613,97	32474,79	17642,1	34183,02	51077,2	57208,41	52220,9	60590,37
Boa Vista do Buricá	4302204	32815,26	40372,38	47587,91	50019,69	59505,8	61191,14	59539,54	67578,17	83186,36	87016,24	84555,14	106443,3
Boa Vista do Cadeado	4302220	-	-	32656,06	30301,86	68480,11	56091,84	23815,84	56840,78	75512,16	99740,69	89093,82	83728,93
Boa Vista do Incra	4302238	-	-	32997,88	36546,28	67197,61	38838,49	21683,99	45564,55	53247,44	72709,52	74581,31	66438,31
Boa Vista do Sul	4302253	13957,15	14823,36	20191,95	24073,83	31223,51	28113,08	27556,14	31792,78	33922,1	32595,69	38001,26	60866,29
Bom Jesus	4302303	60328,77	67591,12	68167,12	74831,16	88496,02	82817,18	90372,52	92868,07	126216,2	135083,7	198817,6	213852,8
Bom Princípio	4302352	62032,93	71423,99	86114,4	97017,8	113089	133803,7	142091,3	156876	16799,87	181458,1	209097	232489,8
Bom Progresso	4302378	11485,52	13333,72	16401,6	17814,91	24615,41	22898,33	16030,65	23575,15	28060,38	25163,23	27870,05	35894,5

ANEXO F – FEE: CONJUNTO DE DADOS SOBRE DEMOGRAFIA

Tabela 11. Trecho do arquivo contendo dados demográficos - FEE

		Demografia																
Municípios	Código	População Homens de 05 a 09 anos																
		1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
Aceguá	4300034	-	-	-	-	195	193	196	210	203	208	210	192	181	165	151	139	145
Água Santa	4300059	209	211	217	222	191	182	177	177	168	164	143	135	134	126	121	118	116
Agudo	4300109	734	735	733	721	722	728	709	694	683	654	600	595	558	530	495	480	450
Ajuricaba	4300208	351	343	335	330	305	309	295	290	283	248	221	214	211	204	222	207	204
Alecrim	4300307	422	410	400	384	358	339	325	305	294	276	254	232	208	180	163	163	148
Alegrete	4300406	3942	3947	3968	3965	3919	3895	3800	3646	3543	3344	3173	3050	2823	2613	2430	2341	2202
Alegria	4300455	254	251	250	247	231	225	204	184	162	152	126	129	128	134	125	127	107
Almirante Tamandaré do Sul	4300471	-	-	-	-	96	100	100	88	73	65	55	53	52	48	42	47	46
Alpestre	4300505	633	590	546	512	450	430	428	414	396	375	333	297	291	290	269	252	245
Alto Alegre	4300554	72	76	82	86	79	76	71	68	65	62	62	54	58	59	58	51	51
Alto Feliz	4300570	110	112	112	114	102	112	105	107	101	102	79	86	84	91	85	83	69
Alvorada	4300604	9284	9518	9751	9943	10070	10276	10312	10326	10364	10215	9905	9544	9071	8556	8169	7964	7705
Amaral Ferrador	4300638	250	249	247	245	262	253	249	254	254	260	266	271	271	258	262	270	263
Ametista do Sul	4300646	388	408	427	439	429	429	429	408	393	357	330	320	311	282	283	283	267
André da Rocha	4300661	45	45	45	46	46	42	40	42	43	42	38	40	37	40	42	43	32
Anta Gorda	4300703	251	253	255	253	245	234	219	195	185	159	151	139	146	141	141	144	147
Antônio Prado	4300802	528	532	532	527	555	550	556	540	523	462	442	401	382	382	376	353	357
Arambaré	4300851	156	167	180	185	181	171	170	174	172	169	169	165	140	123	121	115	105
Araricá	4300877	192	203	217	227	221	234	228	224	223	217	214	211	209	188	183	181	179
Aratiba	4300901	338	335	330	321	299	287	278	270	245	234	216	205	187	185	181	172	148
Arroio do Meio	4301008	755	737	716	698	663	664	632	638	625	629	606	606	564	524	500	479	477

ANEXO G – DENATRAM: CONJUNTO DE DADOS SOBRE FROTA DE VEÍCULOS

Tabela 12. Trecho do arquivo contendo dados de frota de veículos - DENATRAM

UF	MUNICIPIO	TOTAL	AUTOMOVEL	BONDE	CAMINHAO	CAMINHAO TRATOR	CAMINHONETE	CAMIONETA	CHASSI PLATAF	CICLOMOTOR
AC	ACRELANDIA	4166	892	0	213	9	346	34	0	0
AC	ASSIS BRASIL	1105	211	0	19	0	80	10	0	0
AC	BRASILEIA	6209	1566	0	193	48	609	66	0	2
AC	BUJARI	1391	415	0	93	6	181	16	0	3
AC	CAPIXABA	1364	384	0	70	1	156	13	0	1
AC	CRUZEIRO DO SUL	23795	4814	0	662	51	1904	237	1	0
AC	EPITACIOLANDIA	3996	1054	0	190	55	423	45	0	0
AC	FEIJO	2461	420	0	98	1	268	30	0	0
AC	JORDAO	67	9	0	2	0	10	0	0	0
AC	MANCIO LIMA	2541	369	0	75	6	205	15	0	1
AC	MANOEL URBANO	618	144	0	25	0	76	6	0	2
AC	MARECHAL THAUMATURGO	216	8	0	8	0	16	0	0	0
AC	PLACIDO DE CASTRO	3389	937	0	133	7	312	27	0	4
AC	PORTO ACRE	2595	533	0	144	2	249	33	0	2
AC	PORTO WALTER	137	1	0	6	0	11	0	0	0
AC	RIO BRANCO	146497	55739	0	3950	419	14061	3010	8	209
AC	RODRIGUES ALVES	963	106	0	12	1	53	1	0	0
AC	SANTA ROSA DO PURUS	111	12	0	3	0	9	0	0	0
AC	SENA MADUREIRA	5673	1062	0	90	4	601	49	0	8
AC	SENADOR GUIOMARD	4980	1492	0	321	17	538	48	0	2
AC	TARAUACA	3316	413	0	135	4	270	46	0	5
AC	XAPURI	2810	596	0	65	4	239	12	1	1
AL	AGUA BRANCA	3008	579	0	67	2	445	26	0	0
AL	ANADIA	2053	731	0	45	0	98	81	0	0
AL	ARAPIRACA	81354	25456	0	2854	399	5477	1333	1	4