

Daniel Ianegitz Vieira

**Uma Plataforma para Disponibilização
Centralizada de Dados Abertos Governamentais
como Suporte para Aplicações no Contexto de
Cidades Inteligentes**

Sorocaba, SP

27 de Julho de 2016

Ianegitz Vieira, Daniel

Uma Plataforma para Disponibilização Centralizada de Dados Abertos Governamentais como Suporte para Aplicações no Contexto de Cidades Inteligentes / Daniel Ianegitz Vieira. -- 2016.

91 f. : 30 cm.

Dissertação (mestrado)-Universidade Federal de São Carlos, campus Sorocaba, Sorocaba

Orientador: Alexandre Alvaro

Banca examinadora: Luciana A. M. Zaina, Vinicius Cardoso Garcia

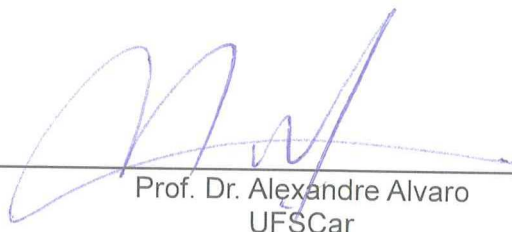
Bibliografia

1. Dados Abertos Governamentais. 2. Cidades Inteligentes. 3. API. I. Orientador. II. Universidade Federal de São Carlos. III. Título.



Folha de Aprovação

Assinaturas dos membros da comissão examinadora que avaliou e aprovou a defesa de dissertação de mestrado do candidato Daniel Ianegtz Vieira, realizada em 27/06/2016:



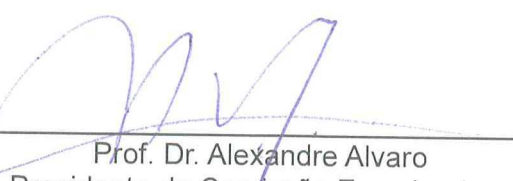
Prof. Dr. Alexandre Alvaro
UFSCar

Prof. Dr. Vinicius Cardoso Garcia
UFPE



Profa. Dra. Luciana Aparecida Martinez Zaina
UFSCar

Certifico que a sessão de defesa foi realizada com a participação à distância do membro Prof. Dr. Vinicius Cardoso Garcia e, depois das arguições e deliberações realizadas, o participante à distância está de acordo com o conteúdo do parecer da comissão examinadora redigido no relatório de defesa do aluno Daniel Ianegtz Vieira.



Prof. Dr. Alexandre Alvaro
Presidente da Comissão Examinadora
UFSCar

Daniel Ianegitz Vieira

Uma Plataforma para Disponibilização Centralizada de Dados Abertos Governamentais como Suporte para Aplicações no Contexto de Cidades Inteligentes

Dissertação de mestrado apresentada ao Programa de Pós-Graduação em Ciência da Computação (PPGCC-So) da Universidade Federal de São Carlos como parte dos requisitos exigidos para a obtenção do título de Mestre em Ciência da Computação. Área de concentração: Engenharia de Software e Redes de Computadores.

Universidade Federal de São Carlos – UFSCar

Centro de Ciências em Gestão e Tecnologia – CCGT

Programa de Pós-Graduação em Ciência da Computação – PPGCC-So

Orientador: Prof. Dr. Alexandre Alvaro

Sorocaba, SP

27 de Julho de 2016

Agradecimentos

Agradeço,

aos meus pais, Maria Inês e Bernardino Valter, pelo dom da vida, o incentivo aos estudos, por prezar a honestidade e integridade acima de tudo e pelo suporte prestado.

ao meu orientador, Prof. Dr. Alexandre Alvaro, por acreditar no meu potencial e aceitar me coordenar, por todo o empenho dispensado e por me guiar durante todo o processo intra e extra-curricular.

aos professores da UFSCar, por se dedicarem arduamente a repassar seus conhecimentos, e, em especial, às Profs. Dr.as Luciana Zaina e Sahudy Montenegro, por serem responsáveis diretas pela realização deste trabalho, bem como por me auxiliarem, aconselharem e não medirem esforços para que eu atingisse os objetivos finais.

aos colegas de curso, por todos os projetos e trabalhos desenvolvidos, conhecimentos adquiridos e agradáveis momentos de lazer e descontração.

a todos os meus amigos, especialmente Thamires Luz, por ser um pilar de sustentação durante todo o percurso, Eline Welter e Johannes Lochter, pela disponibilidade no auxílio de questões técnicas, e Tamires Belinassi, pelo incansável incentivo e apoio prestado.

à UFSCar Sorocaba pela infraestrutura oferecida aos alunos.

à CAPES pelo auxílio financeiro.

e a todos aqueles que contribuíram, direta ou indiretamente, para a realização deste trabalho.

“Tudo o que temos de decidir é o que fazer com o tempo que nos é dado.”
(J. R. R. Tolkien)

Resumo

O conceito de Cidades Inteligentes (CI) visa prover às cidades sistemas dinâmicos e flexíveis, que interagem entre si e auxiliam a manter a estabilidade e escalabilidade de seus recursos e serviços, usualmente utilizando-se de sensores e coletores, com altos custos de implantação e manutenção. Por outro lado, os Dados Abertos Governamentais (DAG) são dados referentes às atividades governamentais, disponibilizados através da Internet. As publicações são fragmentadas em portais distintos e não correlacionados, além de não seguirem um padrão de formato definido. Assim, este trabalho propõe uma plataforma de coleta e disponibilização centralizada de DAG, com o intuito de prover informações consolidadas ao contexto de CI. Ainda, são apresentadas duas validações desta proposta, a primeira promovendo a solidificação da plataforma e a segunda um estudo experimental aplicado no ambiente acadêmico, no qual foram supervisionados oito estudantes de graduação, divididos em dois grupos, durante o desenvolvimento de dois projetos de CI, consistindo em portais de visualização gráfica e comparativa de DAG. O grupo experimental utilizou a plataforma aqui proposta como fonte de DAG, e o grupo controle utilizou-se de DAG coletados diretamente nos portais de transparência governamental. Ao fim do estudo, com o intuito de encontrar benefícios na utilização de uma plataforma centralizada, foram aplicados testes estatísticos às métricas coletadas, e expressivas vantagens em relação à quantidade de linhas de código-fonte e ao tempo de desenvolvimento foram encontradas para o grupo experimental, indicando que a plataforma ofereceu benefícios durante o ciclo de desenvolvimento do projeto.

Palavras-chaves: Cidades Inteligentes; Dados Abertos Governamentais; Plataforma; *Web Service*; Estudo experimental.

Abstract

The Smart City (SC) concept aims to promote dynamic and flexible systems to the cities, that interact with each others and help to maintain the stability and scalability of it's resources and services, usually through sensors and collectors, with high implantation and maintenance costs. On the other hand, the Open Government Data (OGD) are data relationated to the governamental activity, available on the Internet. The publications are fragmented in different and uncorrelated portals, also do not follow a standard pattern format. Thus, this work proposes a platform to collect and distribute OGD in a centralized way, in order to provide consolidated information to the context of SC. Also, two validations of this proposal are presented, the first one promoting the solidification of the platform and the second one being an experimental study applied at the academic environment, were eight graduation students where supervisionated, divided into two groups, during the development of two SC projects, consisting of comparative and graphic visualization OGD portals. The experimental group used the platform proposed here as a OGD source, and the control group used OGD collected directly through the governmental transparency portals. At the end of the study, with the intention to find benefits in the utilization of a centralized platform, statistical tests were applied to the collected metrics, and expressive advantages to the experimental group were found, related to the quantity of source-code lines and the development time, indicating that the platform offered benefits during the cycle of development of the project.

Key-words: Smart Cities; Open Government Data; Platform; Web Service; Experimental Study.

Lista de ilustrações

Figura 1 – Visão geral da metodologia aplicada.	25
Figura 2 – Gráfico de linha do tempo dos artigos.	32
Figura 3 – Plataforma de análise e disponibilização centralizada de Dados Abertos Governamentais.	41
Figura 4 – Diagrama Entidade Relacionamento da plataforma proposta.	44
Figura 5 – Modelo de requisição de dados da API, em visão macro.	45
Figura 6 – Portal de transparência da cidade de Campinas.	48
Figura 7 – Representação parcial dos DAG coletados, em formato CSV.	49
Figura 8 – DAG reorganizados, no formato aceito pela plataforma, em representação parcial.	50
Figura 9 – Resultados em XML sem tratamento.	52
Figura 10 – Gráfico de perfis dos participantes.	62
Figura 11 – Gráfico relativo às competências TAM.	64
Figura 12 – Gráfico geral sobre as concordâncias apontadas em todas as competências.	65
Figura 13 – Gráfico de concordância geral das perspectivas TAM.	67
Figura 14 – Gráficos PUA relativos a Presença, Utilidade e Adequação.	69
Figura 15 – Gráfico de tempo de desenvolvimento.	72
Figura 16 – Codificação de acesso a DAG, Grupo Controle.	73
Figura 17 – Codificação de acesso a DAG, Grupo Experimental.	73
Figura 18 – Gráfico de quantidade de linhas de código.	74
Figura 19 – Boxplot referente às horas de desenvolvimento para acesso aos DAG.	77
Figura 20 – Valores <i>t-test</i> referentes às horas de desenvolvimento para acesso aos DAG.	77
Figura 21 – Boxplot referente às horas de desenvolvimento de visualização gráfica dos DAG.	78
Figura 22 – Valores <i>t-test</i> referentes às horas de desenvolvimento de visualização gráfica dos DAG.	78
Figura 23 – Boxplot referente às horas de desenvolvimento de todo o projeto.	78
Figura 24 – Valores <i>t-test</i> referentes às horas de desenvolvimento de todo o projeto.	79
Figura 25 – Boxplot referente à quantidade de linhas de código.	79
Figura 26 – Valores <i>t-test</i> referentes à quantidade de linhas de código.	79
Figura 27 – Material utilizado na aula sobre <i>web services</i>	91
Figura 28 – Material utilizado na aula sobre DAG.	91
Figura 29 – Material utilizado na aula sobre a plataforma.	94
Figura 30 – Questionário de perfil do participante aplicado aos alunos.	95
Figura 31 – <i>Status report</i> grupo experimental.	97

Lista de tabelas

Tabela 1 – Comparação dos trabalhos relacionados.	39
Tabela 2 – Aferições individuais de perfis dos participantes.	63
Tabela 3 – Aferições totais relativas às competências da plataforma.	64
Tabela 4 – Resultados sob a perspectiva de facilidade de uso percebida.	65
Tabela 5 – Resultados sob a perspectiva de utilidade da técnica.	66
Tabela 6 – Resultados sob a perspectiva de facilidade de compreensão.	66
Tabela 7 – Competências PUA.	68
Tabela 8 – Resultados PUA.	70
Tabela 9 – Tempos de desenvolvimento utilizados pelos grupos.	71
Tabela 10 – Linhas de código utilizadas durante o desenvolvimento.	72
Tabela 11 – Complexidade ciclomática dos métodos utilizados para acesso a DAG.	75

Lista de abreviaturas e siglas

AM	Aprendizado de Máquina
API	<i>Application Programming Interface</i>
CGU	Controladoria-Geral da União
CI	Cidades Inteligentes
Consocial	Conferência Nacional sobre Transparência e Controle Social
COR	Centro de Operações Rio
DAG	Dados Abertos Governamentais
DER	Diagrama Entidade Relacionamento
DIGO	<i>Delivering Information of GOvernment</i>
GC	Grupo Controle
GE	Grupo Experimental
h	Hora(s)
IDE	Ambiente de Desenvolvimento Integrado
ITU	<i>International Telecommunication Union</i>
GQM	<i>Goal/Question/Metrics</i> (Objetivo/Pergunta/Métricas)
MIT	Instituto de Tecnologia de Massachussets
MVC	<i>Model-View-Controller</i>
PRISMA	<i>PiattafoRme cloud Interoperabili per SMArt-government</i>
PUA	Presença-Utilidade-Adequação
RDS	<i>Relational Database Service</i>
Sefti	Secretaria de Fiscalização de Tecnologia da Informação
SGBD	Sistema de Gerenciamento de Banco de Dados
SOAP	<i>Simple Object Access Protocol</i>

t-test	Teste estatístico <i>t de Student</i> para amostras independentes
TAM	Modelo de Aceitação de Tecnologia
TCU	Tribunal de Contas da União
TIC	Tecnologia da Informação e Comunicação
TI	Tecnologia da Informação

Lista de símbolos

\div	Divisão
μ	Média
\times	Multiplicação
α	Significância
Σ	Somatório

Sumário

	Introdução	21
1	DADOS ABERTOS GOVERNAMENTAIS	27
1.1	Conceito	27
1.2	Adoção no Brasil	28
1.2.1	Lei da transparência	28
1.2.2	Motivos	29
1.2.3	Diretrizes	29
1.2.4	Efeitos	30
1.3	Considerações	30
2	REVISÃO BIBLIOGRÁFICA	31
2.1	Linha do tempo	31
2.2	Trabalhos relacionados	37
2.2.1	Análise dos trabalhos	38
2.3	Considerações	39
3	PLATAFORMA DE DADOS ABERTOS GOVERNAMENTAIS	41
3.1	Modelagem	41
3.1.1	Arquitetura	41
3.1.2	Banco de dados	43
3.1.3	API	43
3.2	Desenvolvimento	44
3.2.1	Ferramentas e linguagens	45
3.2.2	<i>Web services</i>	46
3.2.3	Hospedagem	47
3.3	Coleta de dados	47
3.3.1	Importação	48
3.3.2	Organização	49
3.3.3	Validação dos registros	50
3.4	Disponibilização	50
3.4.1	Exemplo de acesso	50
3.5	Considerações	53
4	VALIDAÇÃO DA PLATAFORMA	55
4.1	Validação preliminar	55

4.1.1	Evolução da plataforma	55
4.1.2	Consolidação da plataforma	56
4.2	Estudo experimental	57
4.2.1	Estudo de caso	57
4.2.1.1	Objetivo do estudo	57
4.2.1.2	Questões e métricas	58
4.2.1.3	Objetivo da medição	58
4.2.1.4	Participantes	58
4.2.1.5	Projeto a ser validado	59
4.2.1.6	Treinamento	59
4.2.1.7	Mecanismos de acompanhamento	60
4.2.2	Planejamento	60
4.2.2.1	Seleção de indivíduos	60
4.2.2.2	Seleção do contexto	61
4.2.2.3	Variáveis	61
4.2.3	Operação	61
4.2.3.1	Perfis dos participantes	62
4.2.3.2	Questionário de competências	62
4.2.4	Análise e interpretação dos resultados	63
4.2.4.1	Modelo de aceitação de tecnologia	64
4.2.4.2	Presença-Utilidade-Adequação	68
4.2.4.3	Aplicação do teste estatístico	75
4.3	Considerações	80
	Conclusão	83
	Referências	85
	APÊNDICE A – MATERIAIS UTILIZADOS NAS AULAS.	91
	APÊNDICE B – QUESTIONÁRIO DE PERFIL DO PARTICIPANTE	95
	ANEXO A – STATUS REPORTS	97

Introdução

O crescimento populacional das cidades leva conseqüentemente ao crescimento de demanda residencial, de transporte, saúde pública, energética, coleta de lixo urbano, educação, entre outros inúmeros serviços fornecidos pelo governo e utilizados pela população. Estes serviços podem apresentar uma queda de qualidade quando não acompanham o ritmo acelerado do crescimento demográfico (SUAKANTO et al., 2013). Atualmente, mais da metade da população do mundo vive em cidades, de acordo com a *World Health Organization*¹, e estima-se que em 2050 este número chegue a 70% (PRETZ, 2014b). Para auxiliar na gestão e tomada de decisão, além de suportar as inovações e monitorar o crescimento de grandes cidades, tem-se proposto o uso de Tecnologia da Informação e Comunicação (TIC), surgindo assim o termo Cidades Inteligentes (CI) (SUAKANTO; THERESA, 2011).

Uma das metodologias aplicáveis para oferecer inteligência a uma cidade é através da implantação de uma malha composta de sensores e coletores interligados entre si, com funções de captura de dados referentes a temperatura, poluição do ar, poluição da água, tráfego de veículos, desastres naturais, entre outros fatores de uma grande cidade (SUAKANTO et al., 2013). Sensores inteligentes de monitoramento de energia elétrica têm sido amplamente estudados na área de CI, denominados *Smart Grids* (PORTELA; SLOOTWEG; EEKELEN, 2013; HARTMANN; WOLTER; KRAUSS, 2013).

Apesar dessa tendência relatada em diversos meios, das leis providas pelo estado e pelos inúmeros trabalhos encontrados na literatura na área de CI, ainda existem poucos projetos que já estão em pleno funcionamento ao redor do mundo (CHEN, 2010). Isto se deve, principalmente, aos altos custos envolvidos em infraestrutura e serviços de telecomunicações, da complexidade e tempo despendidos na implementação dos sensores e *hardwares* envolvidos (EVERS, 2013), infraestrutura obsoleta, problemas de crescimento desordenado em novas cidades, falta de planejamento para suportar o crescimento conforme a demanda, entre outros fatores. Este cenário dificulta a implantação de novas tecnologias que envolvem equipamentos nas cidades (DONOLO, 2013).

Em contrapartida, os governos vêm disponibilizando informações a respeito dos serviços prestados à população, como saúde, transporte, educação, segurança, dentre outros. Estes dados, denominados Dados Abertos Governamentais (DAG), são disponibilizados nas mais variadas plataformas e tecnologias, visto que não há uma padronização para publicação entre municípios, estados e federação. Isto acaba por dificultar a coleta, análise e processamento dos dados visando a geração de informação útil para a população ou

¹ <http://www.who.int/about/en/> Acessado em Jun/2016

mesmo para o próprio governo (BRITO et al., 2014; CORRÊA; CORRÊA; SILVA, 2014).

Além de inúmeros formatos, sendo eles compreensíveis por humanos, como PDF², DOC³ e XLS⁴, ou máquina, como JSON⁵, CSV⁶, XML⁷ e HTML⁸ (BRITO et al., 2014), também existem muitos portais de publicação, como o da Câmara dos Deputados⁹, o do Senado Federal¹⁰ e os portais locais referentes a prefeituras dos municípios. Nenhum destes portais citados integram-se ao portal oficial dos Dados Abertos Governamentais Brasileiro. Devido à descentralização dos portais fornecedores de dados, bem como ao formato de publicação dos mesmos, os projetos de *software* que utilizam Dados Abertos Governamentais brasileiros tendem a funcionar em um âmbito restrito aos portais nos quais foram originalmente projetados, não sendo escaláveis a níveis superiores, como estadual ou nacional, ou mesmo a outros municípios (CORRÊA; CORRÊA; SILVA, 2014).

Com isso, a consulta e/ou análise dos dados de forma mais ampla torna-se complexa e, muitas vezes, apenas informações contextualizadas são obtidas, não sendo possível a extração de conhecimento. Adicionalmente, torna-se complexo o desenvolvimento de aplicações escaláveis que aproveitem o potencial dos dados abertos, uma vez que os formatos de divulgação são diversificados e não existe uma centralização ou padronização para agregação dos dados (BRITO et al., 2014), além da burocracia envolvendo modificações nas estruturas governamentais, impossibilitando projetos que ambicionam implementações que envolvam participação governamental em grande escala, como a nível municipal.

Motivação

Três motivações principais foram inicialmente definidas, guiando o desenvolvimento deste trabalho:

1. Apoiar o desenvolvimento de projetos de CI;
2. Reaproveitar estruturas já existentes; e
3. Fornecer um facilitador a desenvolvedores de *software*.

Neste cenário, esta dissertação visa prover uma fonte de dados centralizada, sem a necessidade de implantação de nenhuma nova tecnologia física ou infraestrutura de

² <http://tools.ietf.org/html/rfc3778> Acessado em Jun/2016

³ <https://msdn.microsoft.com/en-us/library/cc313153.aspx> Acessado em Jun/2016

⁴ <https://msdn.microsoft.com/en-us/library/cc313154.aspx> Acessado em Jun/2016

⁵ <http://json.org> Acessado em Jun/2016

⁶ <http://tools.ietf.org/html/rfc4180>

⁷ <http://www.w3.org/XML/> Acessado em Jun/2016

⁸ <http://www.w3.org/html/>

⁹ <http://www2.camara.leg.br/transparencia/dados-abertos> Acessado em Jun/2016

¹⁰ <http://www.senado.gov.br/transparencia/dadosAbertos/dadosAbertos.asp> Acessado em Jun/2016

hardware. É proposto o desenvolvimento de uma plataforma de coleta, processamento e armazenamento dos DAG como fonte de dados para a promoção de aplicações voltadas ao contexto de CI, voltada a desenvolvedores de *software*. Por fim, será apresentada uma validação experimental desta plataforma, no contexto de disciplinas de graduação de Bacharelado em Ciência da Computação da Universidade Federal de São Carlos (UFSCar) - Sorocaba.

Objetivos

Este trabalho propõe uma plataforma para coleta, armazenamento e disponibilização centralizada de DAG atuando como fonte de dados a projetos voltadas ao cenário de CI. Os objetivos que visa-se alcançar são:

- Reduzir ou zerar os custos com coleta de dados;
- Reduzir ou zerar o tempo de coleta de dados;
- Fornecer dados reais;
- Reaproveitar os DAG públicos e gratuito; e
- Prover uma plataforma para distribuição centralizada de DAG para aplicações de CI.

A plataforma compreende os seguintes atributos:

- Funcionais: coleta, organização, armazenamento e distribuição de DAG;
- Escalabilidade: projetada para suportar DAG de diferentes países, estados e municípios, em diferentes granularidades temporais;
- Adaptabilidade: a plataforma foi modelada com foco em projetos de CI, atendendo aos requisitos não planejados que surgiram durante o ciclo de vida de desenvolvimento de *software* utilizando a mesma;
- Robustez: oferece uma documentação clara e concisa sobre sua utilização, além de contar com serviços líderes de mercado para o seu suporte, com relação a tecnologias utilizadas; e
- Validade: a plataforma foi utilizada experimentalmente em projetos de *software* de CI.

Escopo Negativo

Por este projeto focar no desenvolvimento da plataforma de DAG, o escopo negativo deste trabalho é:

- População de Base de Dados: a base de dados foi disponibilizada com DAG suficientes para a utilização em produção da plataforma, em ambiente de desenvolvimento, contudo a coleta de dados não prosseguiu após um limite mínimo atingido, e outras tarefas de desenvolvimento foram priorizadas;
- Coleta de dados automatizada: o foco do trabalho é fornecer uma plataforma centralizadora e, portanto, foram utilizadas coletas de dados manuais a fim de preencher a plataforma, visando a sua validação apenas;
- Organização de registros através de Aprendizado de Máquina (AM). Os registros foram organizados baseados em sua origem, granularidade e finalidade, porém, embora a plataforma não conta com esses serviços, não é restritiva, aceitando DAG oriundos de técnicas de AM; e
- Desenvolvimento de *software* voltado a CI: o objetivo é fornecer DAG através de uma API, atuando como *middleware* neste processo, visando que aplicações possam ser desenvolvidas para o contexto de CI utilizando a plataforma, sem outras fontes de dados. Porém, estudos de caso foram desenvolvidos, a fim de validar a plataforma.

Estrutura do trabalho

Uma visão geral da metodologia aplicada ao desenvolvimento desse trabalho pode ser vista na Figura 1, apresentando todo o processo, desde a motivação inicial, a revisão bibliográfica fundamentando este trabalho, o desenvolvimento e consolidação da plataforma, e, por fim, a validação do mesmo, através de um estudo experimental.

Com o intuito de facilitar a compreensão, este documento está organizado da seguinte maneira:

- Capítulo 1: define e conceitua os principais assuntos para o entendimento do tema deste trabalho. Sendo assim, é apresentado, desde o seu início, todo o processo de regulamentação de DAG no Brasil e no mundo;
- Capítulo 2: é apresentada uma revisão da literatura, apresentando projetos relevantes voltados a DAG e CI. Também são relacionados os trabalhos que se aproximam ou relacionam à plataforma proposta;
- Capítulo 3: apresenta a plataforma desenvolvida, com toda a sua estrutura técnica;

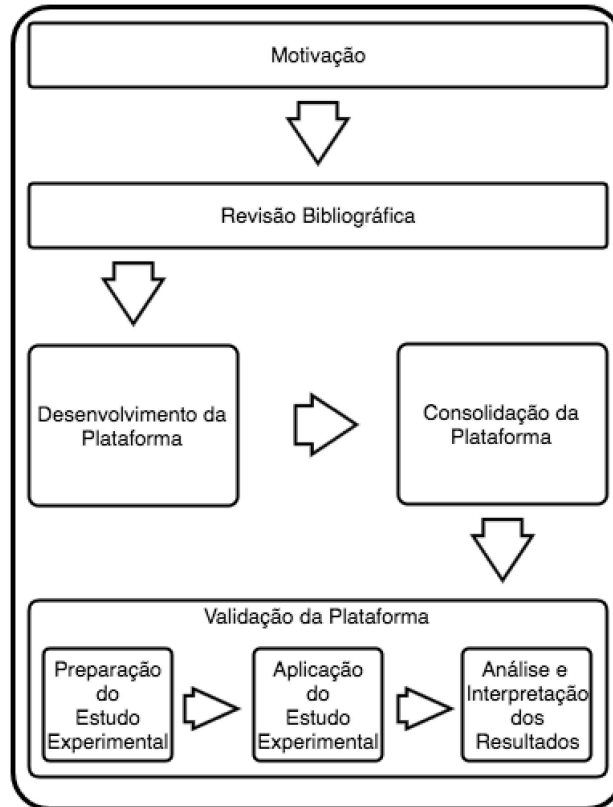


Figura 1 – Visão geral da metodologia aplicada.

- Capítulo 4: descreve a validação do trabalho, com os resultados obtidos e projetos funcionais desenvolvidos a partir da plataforma, fazendo uma comparação com projetos desenvolvidos sem a utilização da mesma, mostrando as vantagens obtidas em tempo de desenvolvimento e linhas de código-fonte; e
- Conclusão: discute os resultados apresentados e as possíveis linhas de estudo para trabalhos futuros.

1 Dados Abertos Governamentais

Este capítulo apresenta a fundamentação teórica para uma melhor compreensão deste trabalho, com uma conceituação sobre Dados Abertos Governamentais, desde suas discussões iniciais e introdução no governo brasileiro até o presente momento, bem como uma visão geral de DAG em escala global.

1.1 Conceito

O conceito de dados abertos governamentais é baseado em 3 leis e 8 princípios ([OPENGOVDATA, 2007](#)). As leis foram formuladas em 2009 pelo especialista canadense em políticas públicas e ativista dos dados abertos David Eaves ([EAVES, 2009](#)), e aceitas pela comunidade e pelo W3C^{1,2}, como seguem:

1. Se o dado não pode ser encontrado e indexado na *Web*, ele não existe;
2. Se não estiver aberto e disponível em formato compreensível por máquina, ele não pode ser reaproveitado; e
3. Se algum dispositivo legal não permitir sua replicação, ele não é útil.

Os 8 princípios foram definidos em uma conferência na Califórnia, Estados Unidos da América ([OPENGOVDATA, 2007](#)), e estão listados abaixo:

1. Completos: todos os dados públicos são disponibilizados. Dados são informações eletronicamente gravadas, incluindo, mas não se limitando a, documentos, bancos de dados, transcrições e gravações audiovisuais. Dados públicos são dados que não estão sujeitos a limitações válidas de privacidade, segurança ou controle de acesso, reguladas por estatutos;
2. Primários: os dados são publicados na forma coletada na fonte, com a mais fina granularidade possível, e não de forma agregada ou transformada;
3. Atuais: os dados são disponibilizados o quão rapidamente seja necessário para preservar o seu valor;
4. Acessíveis: os dados são disponibilizados para o público mais amplo possível e para os propósitos mais variados possíveis;

¹ <https://www.w3.org/Consortium/> Acessado em Jun/2016

² <http://www.w3c.br/divulgacao/pdf/dados-abertos-governamentais.pdf> Acessado em Jun/2016

5. Processáveis por máquina: os dados são razoavelmente estruturados para possibilitar o seu processamento automatizado;
6. Acesso não discriminatório: os dados estão disponíveis a todos, sem que seja necessária identificação ou registro;
7. Formatos não proprietários: os dados estão disponíveis em um formato sobre o qual nenhum ente tenha controle exclusivo; e
8. Livres de licenças: os dados não estão sujeitos a regulações de direitos autorais, marcas, patentes ou segredo industrial. Restrições razoáveis de privacidade, segurança e controle de acesso podem ser permitidas na forma regulada por estatutos.

1.2 Adoção no Brasil

A Parceria de Governos Abertos (*Open Government Partnership*)³ é uma organização não governamental que reúne regras sobre os Dados Abertos Governamentais (DAG). Participam, até o presente momento, 69 países, dos quais o Brasil é membro ativo desde o início, em 2011. Entretanto, o Brasil já coletava e disponibilizava à sua população dados referentes às suas ações, gastos, investimentos e projeções desde 2009, quando foi criado um catálogo central de atividades públicas, para promover e monitorar a atividade governamental (BREITMAN et al., 2012).

1.2.1 Lei da transparência

O Brasil estabeleceu a lei federal do acesso às informações Número 12.527 (CASA CIVIL, 2011), válida para todo o território nacional, com o objetivo de garantir o acesso às informações, conforme afirma o seu artigo oitavo: “É dever dos órgãos e entidades públicas promover, independentemente de requerimentos, a divulgação em local de fácil acesso, no âmbito de suas competências, de informações de interesse coletivo ou geral por eles produzidas ou custodiadas”. Enfatizada no seu parágrafo segundo: “Para cumprimento do disposto no caput, os órgãos e entidades públicas deverão utilizar todos os meios e instrumentos legítimos de que dispuserem, sendo obrigatória a divulgação em sítios oficiais da rede mundial de computadores (Internet)”.

Um projeto de lei tenta corrigir os problemas da Lei da Transparência, como a falta de padronização e a utilização de formatos de publicação que requerem *softwares* fechados e pagos, além de ampliar o alcance da lei para os poderes legislativos e judiciários, criando o Comitê Gestor de Dados Públicos, com o objetivo de disponibilizar para a população os

³ <http://www.opengovpartnership.org/about> Acessado em Jun/2016

dados abertos de todo o âmbito federal em um único local. Atualmente, o PL 7804/2014⁴ aguarda a deliberação na Comissão de Ciência e Tecnologia, Comunicação e Informática (CCTCI).

1.2.2 Motivos

O Tribunal de Contas da União (TCU), por intermédio da Secretaria de Fiscalização de Tecnologia da Informação (Sefti), a fim de conhecer iniciativas de publicação de dados abertos na Administração Pública Federal, e em âmbito internacional, publicou cinco motivos para a abertura de dados na administração pública (TCU, 2015), sendo eles:

1. Porque a sociedade exige mais transparência na gestão pública;
2. Porque a própria sociedade pode contribuir com serviços inovadores ao cidadão;
3. Porque ajuda a aprimorar a qualidade dos dados governamentais;
4. Para viabilizar novos negócios; e
5. Porque é obrigatório por Lei.

1.2.3 Diretrizes

No ano de 2012 foi realizada no Brasil a primeira Conferência Nacional sobre Transparência e Controle Social (Consocial) (CGU, 2012), coordenada pela Controladoria-Geral da União (CGU). Convocada em 2010 e realizada em 2012, mobilizou diretamente mais de 150 mil cidadãos e cerca de 1200 delegados, com o tema “A Sociedade no Acompanhamento e Controle da Gestão Pública”, baseada em 80 propostas (CONSOCIAL, 2012b) e 76 diretrizes votadas e aprovadas (CONSOCIAL, 2012a), dentre as quais, são relevantes ao contexto de DAG, as seguintes:

- E1.D1: aplicar a Lei de Acesso à Informação Pública com severidade e rigor;
- E1.D6: divulgar todas as informações relativas às licitações em tempo real e em formato de dados abertos;
- E1.D8: garantir acesso à tecnologia da informação para toda a população;
- E1.D9: inventariar, catalogar e disponibilizar na Internet todas as bases de dados governamentais;
- E1.D13: prestar contas em tempo real e em formato de dados abertos.

⁴ <http://www.camara.gov.br/proposicoesWeb/fichadetramitacao?idProposicao=620193> Acessado em Jun/2016

- E1.D16: regulamentar e padronizar a publicação na Internet de dados governamentais em formato aberto;
- E2.D2: capacitar a Sociedade Civil para analisar os dados governamentais; e
- E2.D14: tornar a Consocial uma conferência permanente.

Com isso, o portal brasileiro de dados abertos⁵ foi lançado em 2012. Em Janeiro de 2014, o portal contava com 240 conjunto de dados e, dois anos depois, em 2016, são 1.115 conjuntos listados⁶. Paralelamente, algumas cidades criaram os seus próprios portais, como São Paulo, Rio de Janeiro e Recife (BRITO et al., 2014).

1.2.4 Efeitos

Segundo o portal brasileiro de dados abertos, os efeitos dos dados abertos governamentais sobre as políticas públicas são (MOREIRA; MALIN, 2015):

- Inclusão: fornecer dados em formatos padronizados abertos e acessíveis permite que qualquer cidadão utilize qualquer ferramenta de software para adaptá-los às suas necessidades;
- Transparência: informações do setor público abertas e acessíveis melhoram a transparência, pois as partes interessadas podem usá-las da maneira mais adequada ao seu propósito; e
- Responsabilidade: os conjuntos apropriados de dados abertos, devidamente associados, podem oferecer vários pontos de vista sobre o desempenho dos governo no cumprimento de suas metas em políticas públicas.

1.3 Considerações

Embora existam inúmeras iniciativas, diretrizes, recomendações e leis, os governos brasileiros e internacionais ainda estão se adaptando de acordo com as regras preestabelecidas, similares aos princípios de DAG (CORRÊA; CORRÊA; SILVA, 2014), justamente por não existirem definições técnicas rígidas e padrões de divulgação obrigatórios, o que leva a uma fragmentação em inúmeras escalas de diferentes órgãos, além de dificultar tecnicamente o desenvolvimento de projetos escaláveis que tirem proveito de DAG como fontes de dados. Além disso, os governos não são receptivos a implantações e intervenções externas.

⁵ <http://dados.gov.br/sobre/> Acessado em Jun/2016

⁶ <http://dados.gov.br/dataset> Acessado em Jun/2016

2 Revisão Bibliográfica

Analisando a literatura, o cenário atual em relação a trabalhos desenvolvidos em DAG é abrangente, com projetos de *software*¹ sendo desenvolvidos no Brasil² (BRITO et al., 2014) e em todas as partes do mundo³, além de diversas cartilhas e manuais de métodos e procedimentos de desenvolvimento com foco em dados abertos, voltadas a programadores⁴ e analistas de sistemas⁵ (PARK; VANROEKEL, 2013).

Neste sentido, foram realizadas buscas nas bases de dados *IEEE Explorer*⁶, *ACM*⁷, *Google Scholar*⁸, *AIS*⁹ e *SCOPUS*¹⁰ utilizando os seguintes termos e palavras-chave: “*Open Government Data*”, “*Open Governamental Data*”, “*Open Data*”, “Dados Abertos Governamentais”, “Dados Abertos”, “*Smart Cities*”, “*Smart City*”, “*Smart Grid*” e “Cidades Inteligentes”, além de pesquisas combinadas, como: “*Open Government Data Smart City*” e “Dados Abertos Governamentais Cidades Inteligentes”. Com isso, as pesquisas utilizando estas palavras chave retornaram 54 artigos, dentre os quais 41 foram selecionados para leitura por estarem diretamente relacionados a este projeto, enquanto os outros 13 foram excluídos baseado na leitura do resumo. A partir da leitura integral, 28 artigos foram considerados relevantes, e foram, efetivamente utilizados para fundamentar as bases da revisão bibliográfica apresentada neste trabalho, além do apoio de *sites* e artigos *web*.

2.1 Linha do tempo

A revisão bibliográfica é relacionada, por motivos organizacionais, em ordem cronológica, apresentando projetos e estudos relacionados a Dados Abertos Governamentais e também a Cidades Inteligentes, no Brasil e no mundo. O gráfico da Figura 2 representa os artigos aqui utilizados, ao longo do tempo, conforme sua publicação.

No ano de 1997, o governo de Singapura preparou um plano de trabalho integrando a Tecnologia da Informação (TI) ao ambiente educacional, com uma estratégia de ensino e aprendizado baseado em TI em todas as escolas, disponibilizando um computador para cada dois alunos realizarem os seus estudos. Em 1999, iniciaram-se os planejamentos para

¹ <http://dataviva.info/pt/about/> Acessado em Jun/2016

² <http://br.okfn.org/projetos/> Acessado em Jun/2016

³ <http://opengovernmentdata.org/about/> Acessado em Jun/2016

⁴ http://www.w3c.br/pub/Materiais/PublicacoesW3C/manual_dados_abertos_desenvolvedores_web.pdf Acessado em Jun/2016

⁵ <http://dados.gov.br/cartilha-publicacao-dados-abertos/> Acessado em Jun/2016

⁶ <http://ieeexplore.ieee.org/> Acessado em Jun/2016

⁷ <http://dl.acm.org> Acessado em Jun/2016

⁸ <http://scholar.google.com> Acessado em Jun/2016

⁹ <http://aisel.aisnet.org/> Acessado em Jun/2016

¹⁰ <http://www.scopus.com> Acessado em Jun/2016

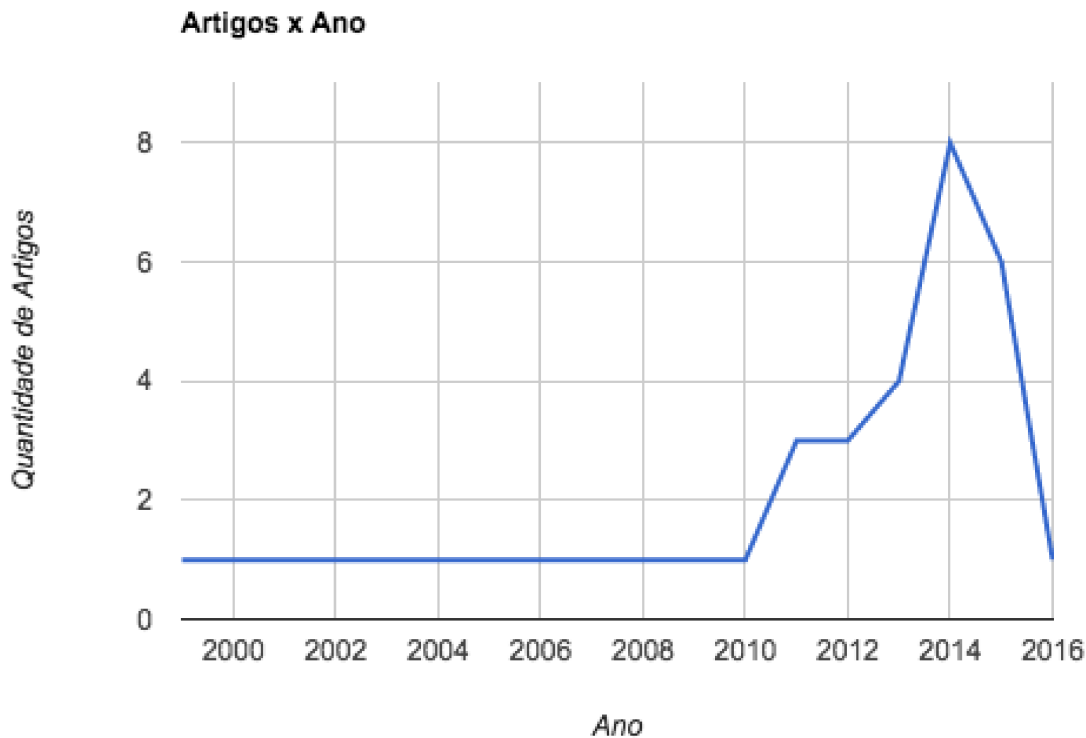


Figura 2 – Gráfico de linha do tempo dos artigos.

se tornar uma ilha inteligente, sendo um dos primeiros países do mundo a possuir uma Infraestrutura Nacional de Informação, na qual cada casa, escola e escritório estariam conectados, promovendo uma vasta gama de transações comerciais ou governamentais, e também promovendo fontes de conhecimento e meios de lazer para a sua população (MAHIZHNAN, 1999).

Neste período, no ano de 2000, na cidade de Southampton, Inglaterra, foi implantado o *SmartCities Card*¹¹ como parte do projeto *SmartPath*, provendo acesso a serviços como transporte público através da Internet. A partir de 2002, o projeto foi expandido, e foi garantida a utilização dos serviços da câmara municipal à população e, para os funcionários públicos, foi fornecido o acesso aos sistemas e registros. Também foi facilitada a comunicação entre laboratórios de patologia e conselhos gerais de medicina, tornando menos burocrática e mais precisa a troca de informações entre os setores (MAHIZHNAN, 2002). Atualmente, o projeto conta com serviços relacionados a pagamentos de impostos municipais, renovação de livros em bibliotecas, coleta de lixo e reciclagem, educação, planejamento, estacionamento e transporte público¹².

Desde 2004 a cidade de Medellín, na Colômbia, vem implantando projetos

¹¹ <https://www.southampton.gov.uk/roads-parking/travel/smartcities-card/> Acessado em Jun/2016

¹² <http://www.southampton.gov.uk/atoz/> Acessado em Jun/2016

centrados em educação, cultura e inclusão social, como o programa Digital Medellín¹³, criado em 2006. Nos anos seguintes foram criados projetos relacionados a CI em diferentes setores, como mobilidade, saúde e segurança, formuladas em um contexto tecnológico local. Em 2012, Medellín foi eleita a cidade mais inovativa do ano¹⁴, devido a bons resultados em emissão de CO₂, construção de espaços culturais e redução de criminalidade. Porém, foi constatado que transformar a sociedade e a economia leva tempo e que a aceitação tecnológica pelas pessoas não é sempre fácil, sendo necessárias campanhas pedagógicas em diferentes meios de comunicação, além da continuidade das políticas e iniciativas públicas de longo prazo (USECHE; SILVA; VILAFANE, 2013).

Posteriormente, em 2009, o governo dos Estados Unidos anunciou um estímulo federal de 3,4 bilhões de dólares para 100 projetos relacionados a CI e *Smart Grids*, como parte do Ato de Recuperação e Reinvestimento Americano. Além do incentivo público, gerou-se uma expectativa de investimento na ordem de 4,7 bilhões de dólares por empresas privadas no setor (CHEN, 2010). A cidade de Seattle já possui um projeto de implementação de circuitos de sensores em andamento (ALAWADHI; SCHOLL, 2013). Um projeto semelhante está sendo desenvolvido em Bandung, na Indonésia (SUAKANTO et al., 2013), e um plano estratégico/operacional para toda a Europa está sendo debatido^{15,16}.

Muitas outras cidades ao redor do mundo também estão desenvolvendo os seus projetos em parcerias com empresas privadas. A IBM vem realizando, desde 2010, o *Smarter Cities Challenge*¹⁷, auxiliando 130 cidades, com contribuições de mais de 66 milhões de dólares. Dentre todas as cidades auxiliadas são destacadas Barcelona - Espanha, Da Nang - Vietnam e Rio de Janeiro - Brasil (MARCA, 2014; PRETZ, 2014b). Em Barcelona desenvolvem-se projetos relacionados a coleta de lixo com sensores indicando quando devem ser esvaziados os contêineres, iluminação inteligente baseada em sensores de presença de movimento, além de projetos de reestruturação de rotas de ônibus urbanos. Em Da Nang reduziu-se o congestionamento instalando controladores de tráfego que utilizam técnicas de *big data* e análises preditivas, além de projetos para o transporte público e coleta de água, com medidores indicando a qualidade da água em tempo real.

No Rio de Janeiro inaugurou-se, em 2010, o Centro de Operações Rio (COR)¹⁸, um centro de gerenciamento de informações públicas para a cidade. O centro conta com interconexões de informações de vários departamentos e agências, melhorando assim o tempo de resposta da cidade a vários tipos de incidentes. As operações são unificadas em uma plataforma para consolidação de sistemas de dados urbanos, para visualização, análise

¹³ <http://www.mdeinteligente.co/> Acessado em Jun/2016

¹⁴ <http://online.wsj.com/ad/cityoftheyear> Acessado em Jun/2016

¹⁵ http://ec.europa.eu/eip/smartcities/files/sip_final_en.pdf Acessado em Jun/2016

¹⁶ http://ec.europa.eu/eip/smartcities/files/operational-implementation-plan-oip-v2_en.pdf Acessado em Jun/2016

¹⁷ <https://smartercitieschallenge.org/about> Acessado em Jun/2016

¹⁸ <http://www.rio.rj.gov.br/web/corio> Acessado em Jun/2016

e monitoramento, permitindo às lideranças da cidade tomarem decisões baseadas em informações em tempo real. Atualmente, o centro conta com 560 câmeras que monitoram todos os pontos da cidade, auxiliando em momentos de crise (ROCHE et al., 2012).

Guadalajara, no México, foi escolhida para o desenvolvimento do projeto chamado *Ciudad Creativa Digital*¹⁹, a partir de uma parceria envolvendo o IEEE, engenheiros do Instituto de Tecnologia de Massachussets (MIT)²⁰, o *International Telecommunication Union* (ITU)²¹ e estudantes de engenharia de universidades locais. O time foca inicialmente em oito áreas: transporte público e estacionamento, coleta de lixo, segurança, telecomunicações, economia, meio ambiente, operações governamentais e o quanto os cidadãos desejam se envolver para melhorar as suas comunidades (PRETZ, 2014a), visando alimentar a inovação e a produtividade através do talento, fortalecer a economia do México com melhores empregos e maior competitividade das indústrias.

A cidade de Santander, na Espanha, iniciou o projeto para se tornar uma cidade inteligente em 2010, com um custo inicial de 9 milhões de euros, financiados em grande parte pela União Européia. O objetivo foi desenvolver um protótipo de Cidade Inteligente para controle de vagas de estacionamento no centro da cidade, com 10.000 sensores e uma equipe de 20 pesquisadores e programadores. Atualmente, o projeto está em expansão, iniciando cobertura a outros serviços, como iluminação pública e coleta de lixo²² (EVERS, 2013; NEWCOMBE, 2014).

Também em 2010 foi conduzido um dos primeiros experimentos relacionados a Dados Abertos Governamentais no Brasil. Foi convertida uma subseção do banco de dados relacional do portal DadosGov em *datasets* de triplas, com formato RDF²³. Porém nenhuma ferramenta oferecia o suporte necessário durante o estágio de modelagem conceitual, como a ferramenta *Triplify*²⁴. Para solucionar esse problema foi utilizada uma ferramenta específica para o cenário envolvido, chamada *StdTrip* (SALAS et al., 2011). Apesar deste esforço, a maioria das triplas resultantes usou um vocabulário RDF especialmente criado para descrever questões específicas do governo brasileiro. O uso de tal contextualização, embora localmente relevante, foi muito restritivo (BREITMAN et al., 2012).

No ano de 2011 foi realizado um estudo na Alemanha, no qual dados estatísticos coletados de diferentes instituições governamentais foram formalmente e semanticamente representados. Também foram aplicados métodos de combinação e visualização de resultados de maneira inteligível a máquinas e humanos, visando atingir a integração dos dados para a utilização dos cidadãos, com conhecimento técnico ou não, para que os mesmos descubram

¹⁹ <http://ccdguadalajara.com/> Acessado em Jun/2016

²⁰ <http://web.mit.edu/aboutmit/> Acessado em Jun/2016

²¹ <http://www.itu.int/en/about/Pages/default.aspx> Acessado em Jun/2016

²² <http://www.smartsantander.eu> Acessado em Jun/2016

²³ <http://www.w3.org/RDF/> Acessado em Jun/2016

²⁴ <https://www.w3.org/2001/sw/wiki/Triplify> Acessado em Jun/2016

os dados estruturados, os interpretem, utilizem e reutilizem. O trabalho listou os principais problemas da divulgação dos dados, como a falta de semântica para explicar o que o dado descreve e falta de inteligibilidade para a comunidade de usuários que os utilizaria. Por isso, foi proposta uma plataforma para prover aos usuários finais uma ferramenta online de visualização gráfica de dados, utilizando a *API Google Visualization*²⁵, na qual os dados eram convertidos de múltiplos formatos como documentos de texto, planilhas e arquivos XML para triplas RDF, seguindo os princípios de *Linked Data* (HOXHA; BRAHAJ, 2011).

Em dezembro de 2012 e fevereiro de 2013 foram realizadas duas análises de portais de Dados Abertos Governamentais de 35 países (SAYOGO; PARDO; COOK, 2014). Foi constatado que 51% dos portais divulgavam seus dados utilizando algum tipo de formato proprietário, enquanto 49% dos portais divulgavam dados nos formatos *Linked Data* exigidos pelas normas de DAG.

Entre setembro e outubro de 2013, uma pesquisa similar (CORRÊA; CORRÊA; SILVA, 2014) foi realizada em 20 municípios brasileiros, com população entre 11.000 e 1.100.000 habitantes. Foram coletadas informações referentes aos dados disponibilizados publicamente pelos portais locais. Foi constatado que 68% dos dados eram divulgados em formato HTML, 22% em formato PDF e apenas 10% em *Linked Data*, formato aceito pelas normas de DAG.

A cidade do Rio de Janeiro promoveu, em agosto de 2013, um desafio de aplicações para dispositivos móveis, RioApps²⁶, com o objetivo de promover a cidade, a vida de seus moradores, os negócios e o turismo, através de aplicações de *software*. Foi utilizado o portal de dados abertos da cidade, *RioDatamine*, contendo 46 *datasets*. Foram submetidas 264 aplicações em 16 categorias diferentes (BRITO et al., 2014). O concurso continua a ser realizado anualmente. Em 2016 foi realizada a competição “Rio Apps Cidade Olímpica”²⁷, onde foram submetidas 1009 propostas em 18 categorias, com idéias próprias ou baseadas nas mais de 5300 ideias previamente inscritas nas três últimas edições do “Rio Ideias”²⁸, que precedem o “Rio Apps”. Um dos critérios de julgamento foi o potencial impacto sobre os moradores e visitantes do Rio de Janeiro, além da utilização das informações sobre a cidade disponíveis no portal de dados abertos DataRio²⁹.

Em 2014 foi proposto no Brasil um levantamento de iniciativas e soluções de informática médica que otimizam o atendimento médico à população, otimizando processos e gerenciando recursos nas CI que utilizam informática médica como um de seus domínios. Através de processos de triagem foram estudados 76 artigos relevantes sobre o tema, visando demonstrar soluções e propostas relevantes à CI, como sistemas que controlam a pressão

²⁵ <https://developers.google.com/chart/interactive/docs/reference> Acessado em Jun/2016

²⁶ <http://rioapps.com.br/> Acessado em Jun/2016

²⁷ <http://portalrioapps.com.br/> Acessado em Jun/2016

²⁸ <http://ideias.rioapps.com.br/> Acessado em Jun/2016

²⁹ <http://data.rio/about> Acessado em Jun/2016

arterial e sistemas inteligentes para ambulâncias utilizando sensores que possibilitem o pronto atendimento dos pacientes (AFONSO et al., 2014).

Em março de 2014 a cidade de Recife promoveu um *workshop* para desenvolvedores e idealizadores de projetos conceituais envolvendo seus bancos de dados abertos governamentais, com o objetivo de promover o acesso e a utilização de seus registros pela população. Foram submetidos cinquenta e sete projetos, sendo eles vinte e três *websites* ou aplicativos para dispositivos móveis e trinta e quatro projetos conceituais. Também foram utilizadas vinte e três bases de dados abertos³⁰.

Foi proposto, no ano de 2015, um modelo de maturidade para CI brasileiras, chamado Br-SCMM (*Brazilian Smart Cities Maturity Model*), explorando os domínios e indicadores para melhoria de todos os processos gerenciais, através da criação de um modelo composto de dez áreas chamadas de “Domínios Básicos”, cada uma com seu respectivo “Indicador Base”. O modelo usa os domínios e indicadores para medir o nível base em uma escala, identificando as possíveis áreas que necessitam melhorias. Os níveis são divididos em cinco categorias, formando a palavra *SMART* (*Simplified, Managed, Applied, Measured e Transformed*). Todas as capitais brasileiras foram medidas utilizando a categoria *S* do modelo. O modelo está sendo desenvolvido em parceria com três universidades brasileiras e os municípios em que estão localizadas estas instituições, onde os níveis poderão ser testados assim que implementados (AFONSO et al., 2015b).

Um estudo, realizado no ano de 2015, promoveu a comparação entre os domínios e indicadores utilizados para categorizar CI, baseado no agrupamento de similaridades, oferecendo, assim, uma maior compreensão das possibilidades de otimização de recursos. O estudo identificou similaridades e diferenças entre as capitais através de *clusters* que agruparam os indicadores, oferecendo bases para a tomada de decisões nos governos municipais, além de oferecer compreensão entre particularidades e potencialidades para a criação de um *ranking* nacional de CI, através de dados públicos e ferramentas de cálculos e gerações de gráficos (AFONSO et al., 2015a).

Na cidade de Sorocaba, em São Paulo, Brasil, é realizado o *Startup Sorocaba*³¹, promovendo palestras, *workshops*, cursos, consultorias, mentorias e treinamentos relacionados a projetos inovadores, com o intuito de desenvolver a cultura do empreendedorismo na região. O evento é realizado desde 2014, e, no ano de 2016, em sua quinta edição, o *Startup Weekend Smart Cities*³² abordou o tema CI. Foram reunidos mais de cem participantes, entre empreendedores, estudantes, *designers*, desenvolvedores, arquitetos e engenheiros, além de pessoas interessadas em contribuir com troca de experiências e aprendizados sobre o tema CI. Treze ideias foram trabalhadas durante o evento, dentre

³⁰ <http://www.cidadaointeligente.rec.br> Acessado em Jun/2016

³¹ <http://startupsorocaba.com/sobre/> Acessado em Jun/2016

³² <http://startupsorocaba.com/startup-sorocaba-smart-cities-e-tema-do-proximo-startup-weekend-em-sorocaba/> Acessado em Jun/2016

as trinta e oito propostas iniciais apresentadas. Foram desenvolvidos projetos nos mais variados temas, como soluções para estacionamento de bicicletas, compartilhamento de créditos para utilização de energia solar, vigilância comunitária, compartilhamento de banheiros para oferecimento de banhos, educação financeira sustentável e informações sobre espaços públicos através de *QR Code*³³.

No ano de 2016 foram conduzidas pesquisas em quatro cidades do norte europeu, com o intuito de verificar se o acesso aos dados é neutro ou visto como neutro, examinar como os dados são gerenciados e movidos através das cidades, trazer desafios envolvendo cidadãos e dados *crowdsorce*, e, finalmente, considerar como a burocracia natural das cidades impacta no uso dos dados. Durante as pesquisas foram identificadas necessidades de entender a cidade como um ator no desenvolvimento de novas tecnologias, serviços e suporte, visto que a cidade é tanto um facilitador quanto um guardião dos dados, sendo ela não simplesmente um pano de fundo de inovação tecnológica, mas um parceiro ativo, embora complexo (MCMILLAN et al., 2016).

2.2 Trabalhos relacionados

No Brasil, no ano de 2011, foi proposto um projeto denominado *Delivering Information of Government* (DIGO), com o intuito de definir uma arquitetura padrão para os portais governamentais divulgarem os seus dados entre as suas diferentes fontes de dados. O objetivo foi de minimizar o conflito de dados gerados por diferentes modelos de divulgação e, assim, permitir aos cidadãos combinar informações de diferentes portais e produzir novas informações e aplicações (MACHADO; OLIVEIRA, 2011). Entretanto, não foram encontrados trabalhos que deram continuidade ao desenvolvimento desta arquitetura ou interesse governamental em implantar a mesma.

Em 2012 foi proposta uma plataforma relacionada a mobilidade em ambientes de CI consumindo Dados Abertos (TCHOLTCHEV et al., 2012), não sendo estes, necessariamente, governamentais. A plataforma utilizaria também serviços de computação em nuvem e provedores de rede, com o objetivo de locar e reservar veículos elétricos, além de transmitir informações referentes a localização do veículo, do requisitante e dos pontos de alimentação de energia durante o trajeto. A proposta visa ainda investigar dados pessoais alavancando o espaço móvel personalizado, para eficiência energética em transporte, através de uma nuvem de dados endereçando aspectos como confiança, confidencialidade e integridade dos dados, fornecida pelos próprios usuários do sistema.

Uma análise empírica foi realizada no Brasil, com o intuito de verificar o impacto do uso de aplicações baseadas em DAG, através de entrevistas conduzidas com visitantes

³³ <http://startupsorocaba.com/conheca-os-projetos-vencedores-do-startup-weekend-smart-cities-sorocaba/> Acessado em Jun/2016

do *website* “Meu Congresso Nacional”³⁴, durante as eleições do ano de 2014. Chegou-se à conclusão de que as pessoas consideram tais ferramentas mais úteis que sites oficiais governamentais de transparência, auxiliando na redução da corrupção e influenciando na escolha do candidato (BRITO et al., 2015).

Também no Brasil, em 2015, foi desenvolvida uma arquitetura de referência para plataformas de *crowdsensing* em CI. O projeto foca em conexão e integração de sensores através de filas de mensagens e utilização de sensores humanos, além de processamentos e análises em tempo real. Foi apresentado um aplicativo *Android*, o qual apresenta, em tempo real, informações aos usuários do sistema como acidentes ou risco de colisões (DINIZ; SILVA; GAMA, 2015).

Visando contribuir com a pesquisa de processamento heterogêneo de dados no padrão RDF, em 2015 foi proposta uma ferramenta operante em nuvem chamada *ConvertView*, que possibilita aos usuários a conversão de arquivos XML em arquivos RDF, para a geração de dados semânticos. Várias combinações de conjuntos de dados (*datasets*) podem ser selecionadas, oriundas de arquivos XML, e a ferramenta mapeia e apresenta os termos distintos presentes no mesmo, gerando um arquivo RDF, além de possibilitar a pesquisa nos dados através da linguagem SPARQL (AFONSO et al., 2015c).

Ainda em 2015, na cidade de Catânia, Itália, foi proposto um modelo de dados para CI, capaz de integrar dados de múltiplos domínios, como informações geográficas, transporte público, manutenção de rodovias, entre outros, facilitando a observação e interação dos cidadãos com a administração pública através de um modelo de dados semântico, como parte do projeto chamado *PiattafoRme cloud Interoperabili per SMArt-government* (PRISMA)³⁵. Embora ambas as plataformas se aproximem nas etapas de extração e, em menor grau, na etapa de divulgação dos dados, os objetivos se distanciam, sendo completamente distintos. O modelo de Catânia é voltada a divulgar geolocalização e informações sobre locais públicos, enquanto a plataforma aqui descrita foca em escalabilidade de informações relacionadas a gastos públicos, sendo este o seu grande diferencial (CONSOLI et al., 2015).

2.2.1 Análise dos trabalhos

Para uma análise dos trabalhos relacionados, sendo possível encontrar suas contribuições e aproximações com a plataforma proposta, cinco requisitos essenciais foram definidos, sendo eles:

1. Fornecer suporte a projetos de CI;
2. Reutilizar estruturas já existentes;

³⁴ <http://meucongressonacional.com/sobre> Acessado em Jun/2016

³⁵ <http://www.ponrec.it/open-data/progetti/scheda-progetto?ProgettoID=5830> Acessado em Jun/2016

3. Prover ferramentas para desenvolvedores de *software*;
4. Oferecer dados reais; e
5. Não necessitar de intervenção governamental.

Uma matriz comparando os trabalhos é apresentada na Tabela 1, onde são feitas as análises individuais em relação aos requisitos que a plataforma proposta neste trabalho atende completamente.

Tabela 1 – Comparação dos trabalhos relacionados.

Trabalho	Requisitos				
	1	2	3	4	5
DIGO	P	N	P	P	N
Mobilidade em ambientes de CI	P	P	P	P	C
Meu Congresso Nacional	N	C	P	C	C
Arquitetura de referência para plataformas de <i>crowdsensing</i> em CI	C	P	C	P	C
<i>ConvertView</i>	C	P	C	N	C
PRISMA	N	C	N	C	P
Plataforma Proposta	C	C	C	C	C

Legenda - N = Não Atende; P = Atende Parcialmente; C = Atende Completamente

Como pôde ser analisado nos trabalhos apresentados, ainda não existem propostas que atendam completamente as necessidades às quais este trabalho investiga, sendo elas maneiras de realizar a captura, análise, processamento e armazenamento de DAG de municípios brasileiros visando disponibilizar os dados de maneira centralizada, para que possam ser consumidos por fontes diversas no âmbito de desenvolvedores de *software* voltados ao desenvolvimento de aplicações no cenário de CI, utilizando os dados já existentes, abertos e gratuitos, disponibilizados pelo governo.

2.3 Considerações

Pôde ser notado que trabalhos visando modificações nas estruturas governamentais tendem a enfrentar problemas durante suas fases de implementação, visto que os governos não são abertos a intervenções externas, devido a grande burocracia ou até mesmo fragmentações diversas, impossibilitando um padrão a ser encontrado nos portais governamentais. Também percebe-se que soluções visando o usuário final são muito restritivas ao contexto local às quais foram desenvolvidas. Por fim, foram analisados trabalhos relacionados a esta proposta, onde nenhum se aproximou dos objetivos aqui buscados, oferecendo suporte a desenvolvedores de *software* voltados a projetos de CI, sem necessidade de intervenção governamental. Com isso, a proposta mostra-se única, atuando em um segmento altamente relevante, relativamente novo e pouco explorado, contribuindo com a comunidade científica de uma maneira inédita.

3 Plataforma de Dados Abertos Governamentais

Este capítulo descreve a plataforma de DAG proposta, concebida através de quatro etapas distintas, a saber: (i) modelagem, (ii) desenvolvimento, (iii) coleta de dados, e (iv) disponibilização dos dados.

Os objetivos principais da plataforma são a análise e disponibilização de DAG de maneira centralizada, para serem consumidos por aplicações ou serviços no âmbito de CI. A arquitetura também é passível de prover DAG para visualizações agregadas e comparativos, com granularidade mensal ou anual, relativos a valores envolvidos nos setores públicos de diferentes municípios.

3.1 Modelagem

Com finalidade organizacional, a modelagem da plataforma foi dividida em três partes: (i) arquitetura, (ii) banco de dados e (iii) API, apresentadas nas subseções a seguir.

3.1.1 Arquitetura

A Figura 3 fornece uma visão macro da plataforma e abaixo são apresentadas as três fases que compõem a modelagem da plataforma, sendo elas:

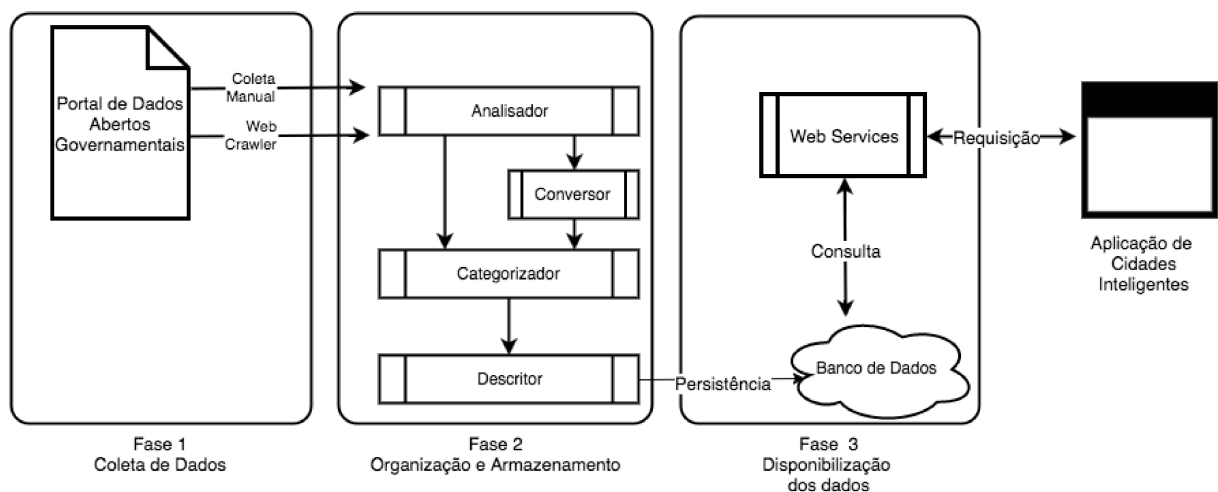


Figura 3 – Plataforma de análise e disponibilização centralizada de Dados Abertos Governamentais.

1. Coleta de dados: pode-se coletar registros de inúmeros portais de dados abertos governamentais do Brasil ou de outros países, utilizando técnicas de coleta manuais,

ou mesmo automatizadas, conhecidas como *crawlers*¹. Inicialmente, após definidos os portais para obtenção dos dados, esta proposta definiu a camada de dados como genérica, com o intuito de facilitar o desenvolvimento de novos conectores, ao invés de coletar dados via *crawlers*, *JSON*, *Web Services*², dentre outras tecnologias de integração. A estrutura do banco de dados foi definida baseada nas análises e categorizações apuradas durante o processo de coleta de dados;

2. Organização e Armazenamento: os dados são organizados e padronizados para futura persistência e disponibilização. Para isso, passaram por uma série de critérios e métodos, chamados aqui de Analisador, Conversor, Categorizador e Descritor, através dos seguintes passos:

- Analisador: é o método de verificar o conteúdo dos dados, quanto à sua semântica e sintaxe, para posterior categorização dos mesmos, referentes a todas as atividades governamentais que possam estar envolvidos, além de reunir informações temporais e geográficas às quais os dados pertencem. Também foi verificado o formato no qual os dados foram publicados;
- Conversor: método não obrigatório, onde modificou-se ou não a estrutura do registro, para que o mesmo se adeque à plataforma, dependendo do formato no qual foram disponibilizados em sua origem. O formato adotado para persistência foi o modelo relacional, hospedado em nuvem (*cloud computing*), normalizado na terceira forma normal, segundo as normas de boas práticas de estruturas de bancos de dados relacionais (BEERI; BERNSTEIN; GOODMAN, 1978);
- Categorizador: define a qual categoria de serviço, gasto ou investimento público o dado pertence, para assim ser possível a correlação de dados oriundos de diferentes origens ou portais governamentais; e
- Descritor: registra informações de metadados, que podem ser úteis tanto para a organização conceitual dos registros na etapa de modelagem do banco de dados quanto para os usuários finais entenderem ao que se referem as informações ali contidas, atuando como um Dicionário de Dados.

3. Disponibilização dos dados: tem o objetivo de disponibilizar, de maneira única, padronizada e centralizada, os DAG coletados em diferentes localidades e em diferentes formatos para aplicações de CI. A disponibilização é pública, através da Internet, inicialmente, por meio de uma *Application Programming Interface* (API) em *Web Services*, sendo possível disponibilizar em outros formatos conforme necessidades dos desenvolvedores, como formatos JSON e RDF.

¹ http://www.w3.org/wiki/Write_Web_Crawler Acessado em Jun/2016

² <http://www.w3.org/TR/ws-arch/> Acessado em Jun/2016

3.1.2 Banco de dados

A estrutura de Banco de Dados, conforme Diagrama Entidade Relacionamento (DER) apresentado na Figura 4, foi modelada a partir de todas as premissas da plataforma, compreendendo os diversos cenários de DAG a serem disponibilizados nos portais governamentais, bem como as necessidades dos desenvolvedores envolvidos em projetos de CI.

O modelo centraliza a despesa como entidade principal, visto que a mesma deve ser divulgada, por lei, em todo o território brasileiro. Os relacionamentos são criados com entidades externas, embora possuam no corpo da despesa também as suas descrições, visto que os portais muitas vezes divulgam apenas a despesa com todas as descrições internalizadas, sem nenhum relacionamento. Nestes casos, o mesmo deve ser criado manualmente, via *scripts* ou automatizadamente, conforme relatado na subseção 3.1.1, em “Organização e Armazenamento”, método “Conversor”. Dessa maneira, a modelagem atende a ambos os cenários indistintamente, padronizando-os internamente e distribuindo DAG de maneira uniformizada.

3.1.3 API

Uma visão geral do funcionamento da API é apresentada na Figura 5. Nela vemos aplicações distintas de CI fazendo requisições à plataforma, e a mesma respondendo às requisições com os DAG. O conteúdo dos DAG e a sua forma de coleta são transparentes à aplicação, não sendo necessários posteriores ajustes ou novas fontes de dados para uma escalabilidade a registros de outros municípios, uma vez que estes sejam disponibilizados pela plataforma.

Duas chamadas de requisição à API foram modeladas, compreendendo todos os cenários para disponibilização de DAG a projetos de CI. São elas:

- Lista de registros: é a chamada principal da plataforma. Nela são entregues os DAG detalhados, com todas as informações disponíveis, apresentados em formato XML. Os filtros são opcionais, podendo estes serem combinados entre si indistintamente, retornando as despesas que atendam à solicitação, ou suprimidos, retornando uma lista de registros menos específica. As consultas são passíveis de paginação, a qual também é opcional;
- Consulta agregada: é uma chamada que retorna apenas a sumarização do valor das despesas contidas nos filtros opcionais, os quais funcionam de maneira idêntica à consulta detalhada.

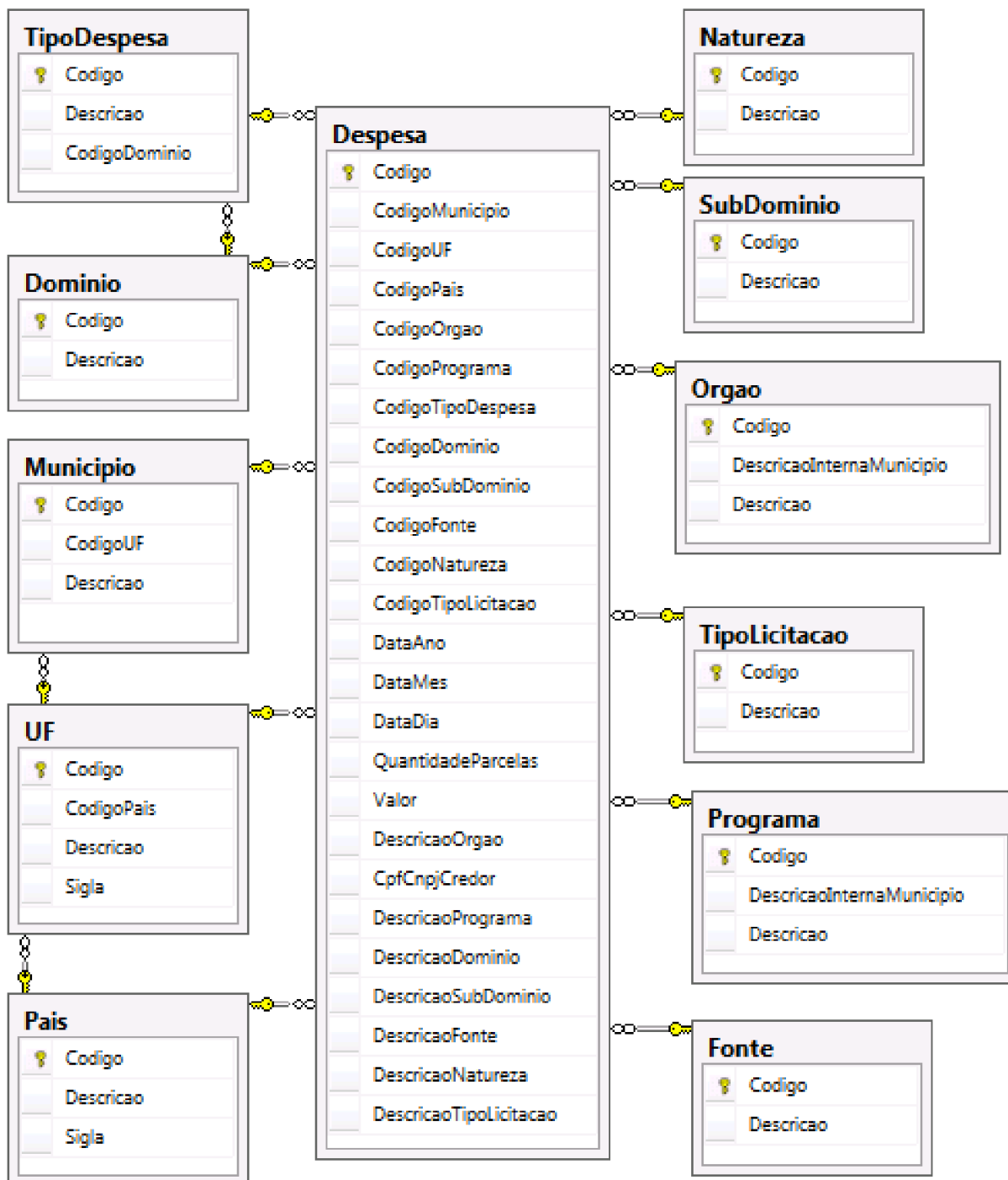


Figura 4 – Diagrama Entidade Relacionamento da plataforma proposta.

3.2 Desenvolvimento

A etapa de desenvolvimento da plataforma foi realizada durante o ano de 2014 e o primeiro semestre de 2015, considerando iterações com alunos de graduação. Após este período, a plataforma ficou sólida e estável o suficiente para utilização em produção. Ao longo do segundo semestre do mesmo ano, já com a plataforma estável e imutável, dois

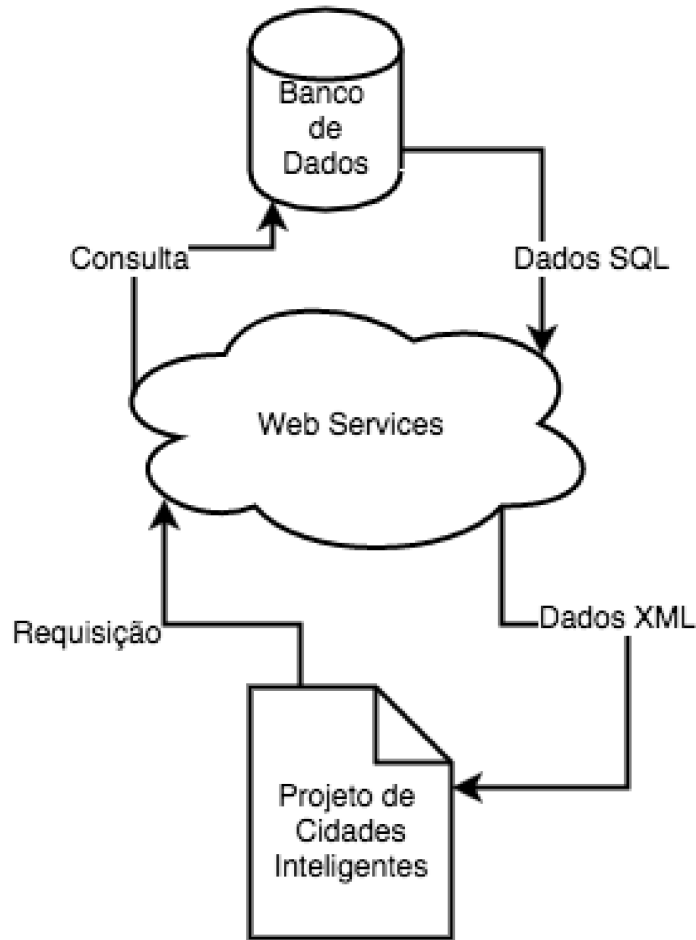


Figura 5 – Modelo de requisição de dados da API, em visão macro.

grupos de desenvolvedores realizaram o desenvolvimento do projeto utilizando-a novamente como fonte de dados, sendo estes projetos considerados na validação experimental.

3.2.1 Ferramentas e linguagens

As linguagens de programação *Java* ³, através do Ambiente de Desenvolvimento Integrado (IDE) *Netbeans* 8.0⁵ e *Visual C#* 5.0⁶ utilizando a IDE *Visual Studio Express 2013 for Web* ⁷ em *.NET Framework* 4.5⁸ foram utilizadas para o desenvolvimento da plataforma.

Para a persistência dos DAG, inicialmente, foi utilizado o *Microsoft SQL Server 2014 Express Edition* ⁹ e o seu Sistema de Gerenciamento de Banco de Dados (SGBD) equivalente. Durante o semestre de desenvolvimento, a plataforma foi vinculada ao banco de dados

³ ⁴ Acessado em Jun/2016

⁵ <https://netbeans.org/community/releases/80/> Acessado em Jun/2016

⁶ <https://www.microsoft.com/en-us/download/details.aspx?id=7029> Acessado em Jun/2016

⁷ <https://www.microsoft.com/en-US/download/details.aspx?id=44912> Acessado em Jun/2016

⁸ <https://www.microsoft.com/pt-br/download/details.aspx?id=30653> Acessado em Jun/2016

⁹ <https://msdn.microsoft.com/sqlserver2014expressresources.aspx> Acessado em Jun/2016

*PostgreSQL*¹⁰, para fins comparativos, não apresentando eventuais incompatibilidades. A ferramenta *SQL Server Import and Export Wizard*¹¹ serviu como suporte para a integração dos registros importados dos portais de DAG governamentais ao Banco de Dados *Sql Server*.

Os sistemas operacionais *Windows 7 Professional*¹² e *Mac Os X*, nas versões 10.9 *Mavericks*¹³, 10.10 *Yosemite*¹⁴ e 10.11 *El Capitan*¹⁵ serviram como suporte às ferramentas.

3.2.2 Web services

Os *Web Services* foram desenvolvidos utilizando a arquitetura *Simple Object Access Protocol* (SOAP)¹⁶, para garantir a integridade e confiabilidade de acesso aos DAG, além da robustez das chamadas e imutabilidade organizacional, uma vez que os métodos de requisição de DAG são rigidamente definidos e não sofrerão alterações estruturais, sendo essa uma das premissas do projeto como um todo.

Para incorporar a API a um projeto de CI, basta informar o WSDL Acessado em Jun/2016, e todo o acesso já estará disponível, restando fazer as chamadas via código. Para evitar quaisquer incompatibilidades, os métodos de requisição retornam uma lista de registros em XML no formato texto (*string*), bem como todos os parâmetros opcionais de entrada, para filtragem, também são em formato *string*. Todos os arquivos podem ser encontrados publicamente na Internet¹⁷.

As duas chamadas à API receberam as seguintes nomenclaturas:

- “getListaDespesa”: para a chamada referente à lista detalhada de registros, com todas as informações dos DAG disponíveis; e
- “getTotalDespesa”: para a chamada referente à sumarização das despesas.

Para ambas as chamadas foram definidos filtros opcionais, que podem ser utilizados ou suprimidos, sendo eles: Cidade, Ano, Mês, Domínio, Subdomínio, Natureza da Despesa, Fonte da Despesa e Tipo de Licitação, além da paginação. Caso não se deseje utilizar um dos parâmetros, são considerados suprimidos quando informados em branco, vazios ou nulos (*null*).

¹⁰ <https://www.postgresql.org/> Acessado em Jun/2016

¹¹ [https://msdn.microsoft.com/en-us/library/ms141209\(v=sql.120\).aspx](https://msdn.microsoft.com/en-us/library/ms141209(v=sql.120).aspx) Acessado em Jun/2016

¹² <https://www.microsoft.com/en-us/software-download/windows7> Acessado em Jun/2016

¹³ <https://support.apple.com/kb/SP702> Acessado em Jun/2016

¹⁴ <https://support.apple.com/kb/SP711> Acessado em Jun/2016

¹⁵ <https://support.apple.com/kb/SP728> Acessado em Jun/2016

¹⁶ <http://www.w3.org/TR/soap/> Acessado em Jun/2016

¹⁷ <https://github.com/DadosAbertosGovernamentais/Arquivos> Acessado em Ago/2016

Ao incorporar a API a um projeto de CI, os métodos “getTotalDespesa” e “getListadespesa” são listados imediatamente, sem restrições. Um exemplo de acesso é apresentado na subseção 3.4.1.

3.2.3 Hospedagem

A plataforma foi hospedada, inicialmente, utilizando os serviços de computação em nuvem *Microsoft Azure*¹⁸, devido ao fato de serem integrados com a IDE *Visual Studio* e oferecerem um plano inicial gratuito. Porém, esta abordagem logo se provou inviável e foi descartada, uma vez que os desenvolvedores rapidamente consumiram todos os recursos sem custos ofertados durante os testes iniciais da validação da plataforma.

Portanto, uma nova abordagem era necessária, e, após estudos juntamente ao orientador e os desenvolvedores, além de inúmeros testes, foram definidos os serviços *Amazon AWS*¹⁹, por oferecerem propostas de hospedagem baseadas em horas de uso, com um número satisfatório de horas iniciais grátis. Com isso, foram utilizados os planos *Amazon Relational Database Service (RDS)*²⁰ para hospedagem do banco de dados, e *Elastic Beanstalk*²¹ para hospedagem dos *Web Services*. O WSDL²² encontra-se disponível para acesso através de URL pública.

3.3 Coleta de dados

Para iniciar esta etapa, foram analisados portais de divulgação de DAG de algumas cidades do Estado de São Paulo, a saber: Campinas, Ribeirão Preto, São José dos Campos, Sorocaba e a cidade de São Paulo. Embora o objetivo principal deste trabalho não seja especificamente a coleta de dados, esta tarefa é essencial para a utilização da plataforma como um todo, sendo ela a única etapa que deve ser expandida após a finalização do desenvolvimento, enquanto as outras etapas são imutáveis, além de ser necessária para a validação do projeto como um todo.

Após análise da qualidade dos dados disponibilizados, bem como seus metadados, foram definidas as cidades de Campinas²³, Ribeirão Preto²⁴ e São José dos Campos²⁵ para inicializar a coleta dos dados. Isto se deve ao fato dos dados estarem estruturados e serem disponibilizadas informações com granularidade mensal, não agregadas e disponíveis em formato CSV.

¹⁸ <https://azure.microsoft.com/en-us/services/> Acessado em Jun/2016

¹⁹ <https://aws.amazon.com/about-aws/> Acessado em Jun/2016

²⁰ <https://aws.amazon.com/rds/> Acessado em Jun/2016

²¹ <http://docs.aws.amazon.com/elasticbeanstalk/latest/dg/Welcome.html> Acessado em Jun/2016

²² <http://www.w3.org/TR/wsdl> Acessado em Jun/2016

²³ <http://transparencia.campinas.sp.gov.br> Acessado em Jun/2016

²⁴ <http://www.ribeiraopreto.sp.gov.br/transparencia/> Acessado em Jun/2016

²⁵ http://www.sjc.sp.gov.br/servicos/portal_da_transparencia.aspx Acessado em Jun/2016

Apesar da qualidade com que é realizada a disponibilização dos dados das cidades acima, foram necessários ajustes manuais em todos os processos descritos no modelo da plataforma, como forma de adequação à modelagem proposta. Com isso, foram utilizadas técnicas manuais para analisar, converter, categorizar e descrever os dados coletados, sendo possível o posterior desenvolvimento de métodos computacionais automatizados para realização de tais tarefas.

3.3.1 Importação

A intenção é de realizar a importação de forma incremental, assim sendo possível o cruzamento de dados de municípios e a geração de informação relevante tanto para o poder público quanto para os cidadãos das cidades envolvidas, além dos desenvolvedores possuírem uma plataforma centralizada rica em conjuntos de dados temporais.

Um exemplo de acesso para coleta de dados através do portal governamental de divulgação de DAG da cidade de Campinas é apresentado na Figura 6, enquanto os dados em CSV, em seu estado bruto, coletados neste portal, são representados parcialmente na Figura 7.

Após a coleta, duas validações estruturais foram realizadas, através da ferramenta *SQL Server Import and Export Wizard* e de um *script* com instruções *COPY*²⁶ do *PostgreSQL*. Posteriormente, os dados foram exportados e disponibilizados para uso na disciplina de Banco de Dados, pertencente à grade curricular da UFSCar, e foram então incorporados ao banco de dados *PostgreSQL* pelo professor, passando por uma nova validação estrutural.



The screenshot shows a web interface with four buttons: FILTRAR, IMPRIMIR, EXPORTAR PDF, and EXPORTAR CSV. Below the buttons, it states 'Total de registros encontrados: 3261 - Exibindo página 1 de 164' with navigation links '<< Anterior | 1 2 3 4 5 6 7 8 9 10 até 164 | Próxima >>'. The main content is a table with the following columns: Órgão, CNPJ ou CPF/Credor, Função, Subfunção, Programa, Fonte, Natureza despesa, Valor Pago Acumulado, and Procedimento/Tipo de Licitação. The table contains six rows of data.

Órgão	CNPJ ou CPF/Credor	Função	Subfunção	Programa	Fonte	Natureza despesa	Valor Pago Acumulado	Procedimento/Tipo de Licitação
ENCARGOS GERAIS DO MUNICIPIO	00360305000104 - CAIXA ECONÔMICA FEDERAL - CEF	28 - Encargos Especiais	844 - Serviço Da Dívida Externa	4009 - Manutenção dos Serviços	01100000 - Geral Total	32902100 - JUROS SOBRE A DIVIDA POR CONTRATO	R\$ 0,00	Serviços/Outros
ENCARGOS GERAIS DO MUNICIPIO	00394480000141 - MINISTÉRIO DA FAZENDA	28 - Encargos Especiais	843 - Serviço Da Dívida Interna	4009 - Manutenção dos Serviços	01100000 - Geral Total	46907700 - PRINCIPAL CORRIG.DA DIV.CONTRAT.REFINANCIADO	R\$ 134.580,08	Serviços/Outros
ENCARGOS GERAIS DO MUNICIPIO	51885242000140 - PREFEITURA MUNICIPAL DE CAMPINAS	28 - Encargos Especiais	846 - Outros Encargos Especiais	4009 - Manutenção dos Serviços	01100000 - Geral Total	33909300 - INDENIZACOES E RESTITUICOES	R\$ 10.807,90	Folha de Pessoal - Automático
ENCARGOS GERAIS DO MUNICIPIO	29979038000140 - INSTITUTO NACIONAL DE SEGURIDADE SOCIAL- INSS	28 - Encargos Especiais	843 - Serviço Da Dívida Interna	4009 - Manutenção dos Serviços	01100000 - Geral Total	46907700 - PRINCIPAL CORRIG.DA DIV.CONTRAT.REFINANCIADO	R\$ 805.279,74	Serviços/Outros
ENCARGOS GERAIS DO MUNICIPIO	00394480000141 - MINISTÉRIO DA FAZENDA	28 - Encargos Especiais	843 - Serviço Da Dívida Interna	4009 - Manutenção dos Serviços	01100000 - Geral Total	32902200 - OUTROS ENCARGOS S/A DIVIDA POR CONTRATO	R\$ 14.886,50	Serviços/Outros
ENCARGOS GERAIS DO MUNICIPIO	00360305000104 - CAIXA ECONÔMICA FEDERAL - CEF	28 - Encargos Especiais	846 - Outros Encargos Especiais	4009 - Manutenção dos Serviços	01100000 - Geral Total	32902100 - JUROS SOBRE A DIVIDA POR CONTRATO	R\$ 290.169,61	Serviços/Outros

Figura 6 – Portal de transparência da cidade de Campinas.

²⁶ <https://www.postgresql.org/docs/current/static/sql-copy.html> Acessado em Jun/2016

	A	B	C	D	E	F	G	H	I	J	K
1	Orgao	CNPJ ou CPF/Cre Funcao		Subfuncao	Programa	Fonte	Natureza despesa	Valor Pago Acumul	Procedimento/Tipo de Licitacao		
2	ENCARGO 06916689000185	28 - Encargos Especiais 843 - Servico Da Divida Int 4009 - Manutenca	01100000 - Geral Tot 32912200 - OUTROS ENCAF					R\$ 3.077.983	71 Servicos/Outros		
3	ENCARGO 51174001000193	28 - Encargos Especiais 846 - Outros Encargos Espi 4009 - Manutenca	01100000 - Geral Tot 33909100 - SENTENCAS JU					R\$ 83	0 Servicos/Outros		
4	ENCARGO 00000000000191	28 - Encargos Especiais 841 - Refinanciamento Da 4009 - Manutenca	01100000 - Geral Tot 46907700 - PRINCIPAL COR					R\$ 1.178.226	78 Servicos/Outros		
5	ENCARGO 51174001000193	28 - Encargos Especiais 846 - Outros Encargos Espi 4009 - Manutenca	01100000 - Geral Tot 33909100 - SENTENCAS JU					R\$ 82.100	0 Servicos/Outros		
6	ENCARGO 00394460000141	28 - Encargos Especiais 843 - Servico Da Divida Int 4009 - Manutenca	01100000 - Geral Tot 32902200 - OUTROS ENCAF					R\$ 29.735	31 Servicos/Outros		
7	ENCARGO 90400888000142	28 - Encargos Especiais 846 - Outros Encargos Espi 4009 - Manutenca	01100000 - Geral Tot 32902300 - JUROS					R\$ 98.223	R\$ 191.310	7 Servicos/	
8	ENCARGO 06916689000185	28 - Encargos Especiais 843 - Servico Da Divida Int 4009 - Manutenca	01100000 - Geral Tot 46917100 - PRINC.DIVID.CC					R\$ 2.194.510	80 Servicos/Outros		
9	ENCARGO 06916689000185	28 - Encargos Especiais 845 - Outras Transferenci 4009 - Manutenca	01100000 - Geral Tot 33913900 - OUTROS SERV.					R\$ 870.000	0 Servicos/Outros		
10	ENCARGO 51174001000193	28 - Encargos Especiais 846 - Outros Encargos Espi 4009 - Manutenca	01100000 - Geral Tot 33909100 - SENTENCAS JU					R\$ 0	0 Servicos/Outros		

Figura 7 – Representação parcial dos DAG coletados, em formato CSV.

3.3.2 Organização

Os registros são organizados em oito categorias, porém, devido a divergências e despadronização de publicação de DAG em território brasileiro, estas podem ou não serem divulgadas pelos portais de DAG, sendo elas:

- Domínio: a pasta base, como Saúde, Educação, Transporte ou Saneamento;
- Subdomínio: as divisões organizacionais internas dos domínios, como assistência hospitalar, educação infantil, transporte especial ou coleta de lixo urbano;
- Natureza da Despesa: onde propriamente foi empenhado o recurso, como material de consumo, recursos humanos ou obras. Podem pertencer a qualquer Domínio e/ou Subdomínio;
- Fonte: quem provê o recurso, como convênios, recursos próprios ou fundos diversos;
- Tipo de Licitação: a maneira como foi licitado o recurso, como adiantamento, contrato de locação ou contrato de obras;
- Órgão: quem licitou o recurso, como Câmara Municipal, Gabinete de Secretaria ou Fundo de Apoio;
- Programa: especificação destinada a descrever onde foi empenhado o recurso, como reforma ministerial, construção de praça pública ou modernização de escolas; e
- Tipo Despesa: mecanismo interno relativo à plataforma, agindo como *filler* ao importar os registros.

Os dados, após a coleta, foram reorganizados em um formato acessível à plataforma, conforme Figura 8. A estrutura segue o modelo proposto no DER, contando com os seguintes campos, com esta exata nomenclatura e ordenação: Codigo, CodigoMunicípio, CodigoUF, CodigoPais, CodigoOrgao, CodigoPrograma, CodigoTipoDespesa, CodigoDominio, CodigoSubDominio, CodigoFonte, CodigoNatureza, CodigoTipoLicitacao, DataAno, DataMes, DataDia, QuantidadeParcelas, Valor, DescricaoOrgao, CpfCnpjCredor, DescricaoPrograma, DescricaoDominio, DescricaoSubDominio, DescricaoFonte, DescricaoNatureza

e DescricaoTipoLicitacao. Embora não obrigatórios, visto que nem todos os portais divulgam todos estes detalhamentos, quanto mais informações forem coletadas sobre o registro, mais rico e informativo ele será, tornando a comparação cada vez mais útil aos usuários finais.

Codig	Codigol	Codigop	Codigoq	Codigor	Codigos	Codigot	Codigou	Codigov	Codigow	Codigox	Codigoy	DataAno	DataMes	DataQua	Valor	Descricao	CpfCnpjCredor	DescricaoProgr
1	1	1	1	1	4	2	2	15	5	7	2014	1		4	20	FUNDACAO MU:4611985500013		
2	1	1	1	1	6	2	10	17	5	7	2014	1		7	88	FUNDACAO MU:3353048601256		
3	1	1	1	1	3	2	5	15	4	7	2014	1		27	57	FUNDACAO MU:3358275000017		
4	1	1	1	1	6	2	10	13	14	1	2014	1		35	0	FUNDACAO MU:2997903600014		
5	1	1	1	1	4	2	2	15	5	7	2014	1		36	33	FUNDACAO MU:3305019600018		
6	1	1	1	1	4	2	2	15	2	7	2014	1		53	42	FUNDACAO MU:82062455887 -		
7	1	1	1	1	3	2	5	15	5	7	2014	1		97	28	FUNDACAO MU:3082293600016		
8	1	1	1	1	4	2	2	15	2	7	2014	1		128	60	FUNDACAO MU:04619749834 -		
9	1	1	1	1	3	2	5	15	17	10	2014	1		137	0	FUNDACAO MU:24657950827 -		
10	1	1	1	1	3	2	5	15	5	10	2014	1		239	0	FUNDACAO MU:10783253834 -		
11	1	1	1	1	4	2	2	15	2	7	2014	1		240	79	FUNDACAO MU:10577546848 -		
12	1	1	1	1	4	2	2	15	2	7	2014	1		240	79	FUNDACAO MU:28076842115 -		

Figura 8 – DAG reorganizados, no formato aceito pela plataforma, em representação parcial.

A base de dados conta com mais de 35.000 registros armazenados, relacionados aos setores públicos de Saúde, Educação, Transporte, Saneamento, Urbanização, Cultura e Segurança, organizados com granularidade mensal, e todos valorizados.

3.3.3 Validação dos registros

Com os registros coletados e organizados, foi necessária uma validação prévia de sua qualidade, para que, assim, a plataforma provesse dados coesos e concisos.

Os metadados e a consistência estrutural foram avaliados durante o primeiro semestre de 2015, quando professor e alunos testaram incessantemente a base de dados, com consultas simples e complexas, além de criações de índices, junções, ordenações, visões e procedimentos, com o intuito de encontrar problemas e violações. O estudo da disciplina foi baseado inteiramente na utilização dos registros contidos nesta base de dados durante o semestre letivo.

Ao encerramento do ciclo a base de dados foi aprovada, validando, assim, a sua usabilidade e integridade, sendo, finalmente, incorporada à plataforma e disponibilizada para uso em produção.

3.4 Disponibilização

Foi definida uma API baseada em XML oriundos de *Web Services* para a divulgação dos dados. A plataforma é projetada para ser hospedada em nuvem, sem restrições de uso.

3.4.1 Exemplo de acesso

Um código de exemplo para acesso e uso da *API* foi desenvolvido em linguagem de programação *Java*, que pode ser visto abaixo. O objetivo é de tornar muito simples e

padronizado o acesso, espelhando outras APIs de mercado já estabelecidas, como *Facebook*²⁷, *Twitter*²⁸ e *Instagram*²⁹.

Algoritmo 3.1 – Exemplo de acesso à API

```
public static void main(String [] args) {

    //Instancia do Web Service
    TransparenciaWS t = new TransparenciaWS ();
    TransparenciaWSSoap ws = t.getTransparenciaWSSoap ();

    //Chamadas a API

    //Sumarizacao de registros com filtros opcionais
    String Total = ws.getTotalDespesa
        ("", "2015", "07", "", "", "", "", "", "", "");

    //Lista de registros com filtros opcionais
    String Lista = ws.getListadespesa
        ("Campinas", "2014", "", "", "", "", "", "", "", "");

    //Exibicao das chamadas
    System.out.println ("Valor Total: " + Total);
    System.out.println ("Lista de registros em XML: " + Lista);
}
```

O exemplo apresenta as duas chamadas à API, utilizando parâmetros opcionais. O método “getTotalDespesa” retorna a soma valorizada dos DAG referentes ao mês 07 do ano de 2015. Como não foi informada a cidade, o método retornará o valor agregado de todas as cidades incorporadas à plataforma no momento da requisição. Já o método “getListadespesa” retorna uma lista detalhada de DAG, referentes a todos os meses do ano de 2014 e apenas à cidade de Campinas, visto que ela foi informada no parâmetro. Os resultados da invocação de ambos os métodos são reproduzidos parcialmente na Figura 9, apresentados em XML puro, sem tratamento, para posterior incorporação a projetos de CI.

Desta forma, aplicações de CI podem consumir DAG providos pela plataforma de maneira transparente, com inúmeras possibilidades, desde uma simples visualização

²⁷ <https://developers.facebook.com> Acessado em Jun/2016

²⁸ <https://dev.twitter.com> Acessado em Jun/2016

²⁹ <https://instagram.com/developer/> Acessado em Jun/2016

```

Valor Total: 27351151,77
Lista de Registros em XML: <?xml version="1.0" encoding="utf-8"?>
<DataSet>
<xs:schema id="NewDataSet" xmlns="" xmlns:xs="http://www.w3.org/2001/XMLSchema" xmlns:msdata="u
<xs:element name="NewDataSet" msdata:IsDataSet="true" msdata:UseCurrentLocale="true">
  <xs:complexType>
    <xs:choice minOccurs="0" maxOccurs="unbounded">
      <xs:element name="Despesa">
        <xs:complexType>
          <xs:sequence>
            <xs:element name="Codigo" type="xs:int" minOccurs="0" />
            <xs:element name="CodigoMunicipio" type="xs:int" minOccurs="0" />
            <xs:element name="CodigoUF" type="xs:int" minOccurs="0" />
            <xs:element name="CodigoPais" type="xs:int" minOccurs="0" />
            <xs:element name="CodigoOrgao" type="xs:int" minOccurs="0" />
            <xs:element name="CodigoPrograma" type="xs:int" minOccurs="0" />
            <xs:element name="CodigoTipoDespesa" type="xs:int" minOccurs="0" />
            <xs:element name="CodigoDominio" type="xs:int" minOccurs="0" />
            <xs:element name="CodigoSubDominio" type="xs:int" minOccurs="0" />
            <xs:element name="CodigoFonte" type="xs:int" minOccurs="0" />
            <xs:element name="CodigoNatureza" type="xs:int" minOccurs="0" />
            <xs:element name="CodigoTipoLicitacao" type="xs:int" minOccurs="0" />
            <xs:element name="DataAno" type="xs:int" minOccurs="0" />
            <xs:element name="DataMes" type="xs:int" minOccurs="0" />
            <xs:element name="DataDia" type="xs:int" minOccurs="0" />
            <xs:element name="QuantidadeParcelas" type="xs:int" minOccurs="0" />
            <xs:element name="Valor" type="xs:decimal" minOccurs="0" />
            <xs:element name="DescricaoOrgao" type="xs:string" minOccurs="0" />
            <xs:element name="CpfCnpjCredor" type="xs:string" minOccurs="0" />
            <xs:element name="DescricaoPrograma" type="xs:string" minOccurs="0" />
            <xs:element name="DescricaoDominio" type="xs:string" minOccurs="0" />
            <xs:element name="DescricaoSubDominio" type="xs:string" minOccurs="0" />
            <xs:element name="DescricaoFonte" type="xs:string" minOccurs="0" />
            <xs:element name="DescricaoNatureza" type="xs:string" minOccurs="0" />
            <xs:element name="DescricaoTipoLicitacao" type="xs:string" minOccurs="0" />
          </xs:sequence>
        </xs:complexType>
      </xs:element>
    </xs:choice>
  </xs:complexType>
</xs:element>
</xs:schema>
<diffgr:diffgram xmlns:msdata="urn:schemas-microsoft-com:xml-msdata" xmlns:diffgr="urn:schemas-mic
<NewDataSet>
  <Despesa diffgr:id="Despesa1" msdata:rowOrder="0">
    <Codigo>44</Codigo>
    <CodigoMunicipio>1</CodigoMunicipio>
    <CodigoUF>1</CodigoUF>
    <CodigoPais>1</CodigoPais>
    <CodigoOrgao>1</CodigoOrgao>
    <CodigoPrograma>6</CodigoPrograma>
    <CodigoDominio>2</CodigoDominio>
    <CodigoSubDominio>10</CodigoSubDominio>
    <CodigoFonte>13</CodigoFonte>
    <CodigoNatureza>4</CodigoNatureza>
    <CodigoTipoLicitacao>7</CodigoTipoLicitacao>
    <DataAno>2014</DataAno>
    <DataMes>2</DataMes>
    <Valor>4.58</Valor>
  </Despesa>
  .

```

Figura 9 – Resultados em XML sem tratamento.

de gastos sumarizados até complexas visualizações gráficas tridimensionais, comparando investimentos em setores através do tempo e seus reflexos em outros setores públicos.

3.5 Considerações

Como pôde ser visto, o acesso à API é simplificado, e o poder do cruzamento dos dados de forma temporal pode dar a visibilidade de informações relevantes como: o investimento em educação e o investimento em segurança trouxe melhoria de vida para a população de Campinas no ano de 2014? Ou ainda, quando investiu-se em saneamento pôde-se notar uma melhoria no nível de saúde da população de Campinas? Estes são apenas alguns exemplos do nível de informação que pode ser gerado pela plataforma. Assim, os sistemas de informação que irão consumir tais dados poderão trabalhar com o cruzamento de informações e trazer inúmeras possibilidades de visualização de dados. Com isso, a geração de informações ricas podem transformar a maneira como é feito o planejamento das políticas públicas bem como aumentar o grau de entendimento e participação popular.

Por fim, a intenção é de unir os mais variados tipos de dados na plataforma, de maneira a gerar cada vez mais informação de qualidade para as aplicações que irão consumir a API, sem que estas necessitem, obrigatoriamente, alterar seus projetos para escalar a abrangência a DAG de novas cidades, conforme a plataforma for incrementada ao longo do tempo.

4 Validação da plataforma

Neste capítulo são apresentadas as duas validações pelas quais este trabalho foi submetido. Inicialmente, uma validação parcial foi realizada, e, posteriormente, a plataforma passou por um estudo experimental. Os estudos seguiram a metodologia de engenharia de *software* experimental (WOHLIN, 2000; TRAVASSOS; GUROV; AMARAL, 2002; WOHLIN, 2014).

4.1 Validação preliminar

No primeiro semestre letivo de 2015 foi realizada uma validação inicial da proposta, com o intuito de evoluir as interfaces da plataforma, baseado nas necessidades dos seus usuários, participantes do estudo, através de *feedbacks*. Assim, a validação foi realizada em parceria com três disciplinas do curso de Bacharelado em Ciências da Computação da UFSCar, sendo elas Banco de Dados, Desenvolvimento *Web* e Gestão de Projetos. Ao longo desse período foram acompanhados doze alunos de graduação, divididos em quatro grupos, onde dois grupos utilizaram a plataforma para desenvolver o seu projeto, uma quantidade suficiente para testar a efetividade da mesma (NIELSEN, 2000). Embora se tenham aplicados previamente questionários a fim de verificar os conhecimentos específicos relacionados a linguagens de programação, banco de dados e *web services*, os graduandos tiveram liberdade para formar os grupos indistintamente, baseados em seus critérios e afinidades pessoais. Para a execução do estudo, dois grupos desenvolveram seus projetos utilizando dados oriundos da plataforma proposta, enquanto os outros dois grupos interagiram diretamente com os dados abertos disponibilizados pelos portais governamentais.

Uma vez que o desenvolvimento da plataforma ocorreu paralelamente à sua utilização e validação, foram verificadas instabilidades e necessidades de alterações e manutenções durante o decorrer do semestre. Com isso, os esforços foram direcionados para a adequação da plataforma às necessidades dos desenvolvedores, as quais surgiram durante o percurso de desenvolvimento da mesma.

4.1.1 Evolução da plataforma

A plataforma absorveu os *feedbacks* relatados semanalmente, além de constantes reportes via e-mail, conforme os grupos prosseguiram no desenvolvimento de seus respectivos projetos, e assim evoluiu e se moldou às necessidades ocasionais, alcançando, ao fim do semestre, o seu objetivo final. Os desenvolvedores igualmente foram capazes de atingir suas metas, o desenvolvimento de um portal de visualização e comparação de DAG utilizando

a plataforma como fonte única de dados.

Ao fim dos projetos, foram aplicados questionários e entrevistas com os participantes, com o intuito de evoluir a plataforma exercendo melhorias necessárias à mesma, segundo as necessidades, sugestões e críticas providas por seus próprios usuários. Estas foram analisadas e, uma vez provadas necessárias, adicionadas ao escopo do projeto e implementadas posteriormente.

4.1.2 Consolidação da plataforma

O primeiro semestre de 2015 serviu como uma base sólida para tornar a plataforma mais robusta. A evolução da plataforma foi constante, e ao fim do semestre estava sólida o suficiente para não necessitar de ajustes e atender todas as necessidades dos desenvolvedores. Visando validar a plataforma, iniciando um novo ciclo de desenvolvimento, os seguintes passos foram tomados:

- Planejamento do experimento: realização de um planejamento detalhado de como será o experimento, assim o acompanhamento durante o desenvolvimento do projeto não será afetado;
- Alterações no escopo do projeto: escopo do projeto definido e imutável, evitando assim alterações no mesmo e reinício do experimento;
- Desenvolvimento e testes da plataforma: o desenvolvimento e os testes de uso do *software* proposto para experimentação foram realizados antes de iniciar a sua validação;
- Ferramentas de suporte: foram definidas previamente as ferramentas que suportarão o acompanhamento do estudo experimental;
- Seleção dos participantes do experimento: selecionar com cautela os participantes de um estudo experimental para que se possa garantir um amplo alcance do objeto que se pretende avaliar; e
- Treinamento: planejar o treinamento dos participantes do estudo para que estejam aptos a executar o experimento com conhecimentos prévios e equilibrados.

Assim, após a evolução da plataforma durante todo o semestre, foi possível o início de um estudo experimental inteiramente planejado, com a plataforma finalizada, estável, sólida, imutável e já utilizada anteriormente com sucesso.

4.2 Estudo experimental

Durante o segundo semestre de 2015, já com a plataforma consolidada a plataforma passou por um estudo experimental onde, semanalmente, foram acompanhados oito desenvolvedores de *software* com ênfase em aplicações de CI utilizando DAG. Os desenvolvedores foram divididos em dois grupos, chamados aqui de *Grupo Experimental (GE)*, o qual utilizou a plataforma como fonte de DAG, e *Grupo Controle (GC)*, que utilizou-se de dados coletados diretamente nos portais governamentais. O estudo foi dividido em quatro etapas, (I) estudo de caso, (II) planejamento, (III) operação e (IV) análise e interpretação dos resultados.

4.2.1 Estudo de caso

Este estudo tem como objetivo avaliar a plataforma de DAG centralizada, verificando os possíveis ganhos de produtividade em relação a tempo de desenvolvimento e quantidade de linhas de código para desenvolvedores de *software* envolvidos em projetos de CI voltados a comparação e visualização tridimensional de DAG. Para isso, foram analisadas todas as competências ou funcionalidades oferecidas pela plataforma, utilizando o modelo *Goal/Question/Metrics*, Objetivo/Pergunta/Métricas (GQM) (BASILI; ROMBACH, 1988). Foram aplicados questionários iniciais para uniformização dos grupos baseados em conhecimentos prévios, além de questionários aplicados durante o semestre, para coletar informações referentes às competências da plataforma.

Para medir o grau de concordância e o nível de aceitação quanto a facilidade de uso percebida, utilidade da técnica e facilidade de compreensão, utilizou-se o *Technology Acceptance Model*, ou Modelo de Aceitação de Tecnologia (TAM) (DAVIS; BAGOZZI; WARSHAW, 1989), apresentado na Subseção 4.2.4.1. A metodologia Presença-Utilidade-Adequação (PUA) (TRAVASSOS; GUROV; AMARAL, 2002), descrita na Subseção 4.2.4.2, foi aplicada conforme os projetos avançavam na utilização das competências respectivas.

Por fim, foi aplicado o teste estatístico *t de Student* para amostras independentes (*t-test*) (WINTER, 2013), uma vez que tratamentos diferentes foram aplicados a cada um dos grupos, com uma variável independente (DYBÅA; KAMPENESA; SJØBERGA, 2006), referente à maneira de acessar os DAG, e os seus resultados são mostrados na Subseção 4.2.4.3.

4.2.1.1 Objetivo do estudo

Analisar as competências da plataforma de DAG.

Com o propósito de avaliação.

Com respeito a efetividade.

Do ponto de vista de desenvolvedores de *software*.

No contexto de estudantes do curso de Ciências da Computação.

4.2.1.2 Questões e métricas

As questões às quais este estudo investiga, bem como as métricas utilizadas para tal, são:

Q1: Existem ganhos de tempo total de desenvolvimento utilizando a plataforma?

Métrica: Tempo total gasto pelos grupos para desenvolver o projeto.

Q2: Existiu diminuição de linhas de código escritas utilizando a plataforma?

Métrica: Quantidade de linhas de código utilizadas pelo grupo para desenvolver o projeto.

4.2.1.3 Objetivo da medição

Tendo como base a plataforma de DAG, caracterizar:

1. Quais são as funcionalidades que os desenvolvedores receberam:
 - Quais são as funcionalidades oferecidas pela plataforma que os desenvolvedores consideram úteis para projetos de CI consumindo DAG;
 - Quais são as funcionalidades oferecidas pela plataforma que os desenvolvedores consideram inúteis para projetos de CI consumindo DAG.
2. Quais são as funcionalidades que os desenvolvedores receberam de maneira inadequada:
 - Quais são as funcionalidades que precisam de um melhor detalhamento;
 - Quais são as funcionalidades que oferecem um detalhamento excessivo.
3. Quais são as funcionalidades que os desenvolvedores gostariam de receber além das oferecidas pela plataforma.

4.2.1.4 Participantes

Participaram do experimento dois grupos compostos por quatro alunos cada, estudantes do curso de Ciências da Computação da UFSCar, matriculados na disciplina de Projeto e Desenvolvimento de *Software*. Apesar da plataforma já ter sido submetida a uma validação funcional anteriormente, a validação de aceitação ainda não havia sido aplicada, portanto o número de participantes do estudo, oito, é considerado adequado (NIELSEN, 2012).

4.2.1.5 Projeto a ser validado

O projeto consistiu em um portal de comparação e visualização de DAG, tendo como público alvo usuários interessados em consumir, comparar e ainda visualizar graficamente dados de diferentes setores públicos. Os requisitos do projeto foram:

- Comparação: a visualização deve permitir comparações de diferentes domínios em diferentes cidades, mostrando quais reflexos positivos ou negativos podem trazer um aumento ou redução de investimentos em um setor específico em relação aos outros, como, por exemplo, quais reflexos podem ter na saúde o aumento de investimentos em saneamento básico;
- Visualizações gráficas: o sistema deve considerar todas as entidades disponíveis para DAG e deve permitir, ao menos, três tipos diferentes de visualizações gráficas;
- Realce: os gráficos devem oferecer realce (*highligh*) de ao menos uma informação relevante no período, como picos de investimentos, ou cortes nos mesmos;
- Gráficos elaborados: deve possibilitar a visualização de gráficos simples para usuários sem cadastro, apenas com informações agregadas, e gráficos elaborados, como detalhamentos mensais por setor, através de cadastro e/ou *login* no sistema; e
- Persistência: deve permitir que gráficos criados pelo usuário sejam salvos, para visualizações posteriores.

4.2.1.6 Treinamento

Para o aquecimento dos participantes, antes do início do projeto foram ministradas três aulas presenciais aos alunos, todas apresentando conceitos teóricos e práticos.

A primeira aula abordou os assuntos relacionados a APIs e *web services*, utilizando a linguagem Java. Os alunos, individualmente, desenvolveram um exercício prático criando dois programas simples, um servidor *web service*, utilizando a arquitetura SOAP, e um cliente que acessa esse servidor através de uma chamada com dois parâmetros *web* no formato de números inteiros, e recebe desse servidor uma resposta em XML com a soma destes números. Ao fim da aula, todos os alunos foram capazes de criar uma nova chamada no servidor e utilizar o cliente para acessar esse novo método, onde tiveram liberdade de programar o método que mais lhes aprouveram, visto que o exercício era para fixação do conceito de cliente e servidor utilizando *web services*.

Na segunda aula, sobre DAG, foram explanados os conceitos teóricos como a Lei da Transparência, e alguns portais relevantes foram apresentados, onde os alunos puderam atuar, na prática, realizando consultas e percebendo as diferenças entre os portais. Foram apresentados os portais das cidades de Campinas, São José dos Campos e São Paulo, e

os alunos realizaram o *download* de alguns DAG, além de realizarem manipulações nos arquivos obtidos, com finalidade pedagógica apenas.

A última aula foi relativa à plataforma, sendo esta somente para o GE, que efetivamente a utilizamos. Foi apresentada a motivação, a modelagem, o funcionamento e, ao fim desta aula, um exemplo simples de acesso à plataforma foi aplicado para todos os participantes, onde os mesmos desenvolveram, individualmente, um programa cliente em Java, que acessava a plataforma e realizava ambas as chamadas disponíveis, finalizando com a exibição simples dos DAG na tela, em formato textual, sem quaisquer tratamentos nos mesmos.

As aulas foram realizadas durante o horário letivo, nos laboratórios usuais da disciplina, com duas horas de duração cada. Os materiais utilizados nessas aulas são reproduzidos parcialmente no Apêndice A, onde constam também os exemplos de programação ministrados, em linguagem Java.

O contato por *e-mail* foi disponibilizado, e ao longo do semestre foram respondidas mensagens relacionadas à plataforma ou a DAG, além de explicações presenciais durante as reuniões de acompanhamento semanais, para ambos os grupos.

4.2.1.7 Mecanismos de acompanhamento

O acompanhamento do projeto foi realizado presencialmente, junto aos alunos, onde, semanalmente, em horário letivo e no laboratório da disciplina, eram reportadas as tarefas desenvolvidas, o tempo utilizado para realizar as mesmas, em horas, e também a quantidade de linhas de código utilizadas. Os alunos também desenvolveram um portal para hospedar a documentação semanal, chamada de *status report*. Um exemplo da documentação é ilustrado no Anexo A.

4.2.2 Planejamento

O estudo foi idealizado em Julho de 2015, e assim foi possível iniciar o segundo semestre letivo executando o planejamento definido ao longo do período de sua aplicação.

4.2.2.1 Seleção de indivíduos

Foram selecionados dois grupos de estudantes de graduação da UFSCar, participantes do projeto de “Fábrica de *Software*” e interessados em desenvolvimento de portais de comparação e visualização tridimensional de DAG. Foram aplicados questionários para caracterizar os perfis dos participantes, conforme Apêndice B. Baseado neste questionário, dois grupos foram formados, representando, respectivamente, os alunos que projetaram seus portais utilizando a plataforma aqui proposta (GE) e os alunos que coletaram dados diretamente dos *websites* de transparência governamental (GC).

4.2.2.2 Seleção do contexto

Processo: *offline*.

Participantes: alunos de graduação.

Realidade: problema real.

Generalidade: específico.

O processo é *offline*, pois o acompanhamento foi realizado semanalmente, revisando o que foi desenvolvido neste período. Os participantes são os alunos do curso. O estudo reflete um problema real, com dados não simulados, coletados diretamente dos portais de dados abertos das cidades definidas ou oriundos da plataforma, os quais passaram por severas validações. O contexto possui caráter específico, pois se aplica à plataforma desenvolvida e a projetos de visualização de DAG.

4.2.2.3 Variáveis

Variável dependente: produtividade do projeto.

Variáveis independentes: metodologia de desenvolvimento, projeto desenvolvido, tecnologias utilizadas e ambiente utilizado. Como fator foi estabelecida a variável metodologia de desenvolvimento, por receber influência dos projetos desenvolvidos utilizando as mesmas métricas, cujos efeitos podem ser observados na variável dependente, enquanto as demais variáveis se mantiveram constantes.

Projeto Desenvolvido: portal de visualização de DAG.

Tecnologias utilizadas: Java, HTML, CSS¹, JS(JQuery)², JSP³/Servlet⁴, JDBC⁵, XML e *web services*.

Ambiente utilizado: laboratórios de informática da UFSCar.

4.2.3 Operação

Ao iniciar a operação do estudo, os grupos de desenvolvedores foram formados, e os mesmos foram acompanhados durante o ciclo de desenvolvimento dos projetos. Ao longo desse período foram coletadas as métricas de tempo de desenvolvimento, linhas de código e, ainda, foram aplicados os questionários de competências relativas à plataforma.

¹ <https://www.w3.org/Style/CSS/> Acessado em Jun/2016

² <https://jquery.com/> Acessado em Jun/2016

³ <https://jsp.java.net/> Acessado em Jun/2016

⁴ <http://docs.oracle.com/javase/6/api/javax/servlet/Servlet.html> Acessado em Jun/2016

⁵ <http://www.oracle.com/technetwork/java/overview-141217.html> Acessado em Jun/2016

4.2.3.1 Perfis dos participantes

Os participantes foram categorizados baseados em seus perfis de conhecimentos tecnológicos prévios, através de questionários de perfis de participante. Os resultados da aplicação do questionário são apresentados na Figura 10.

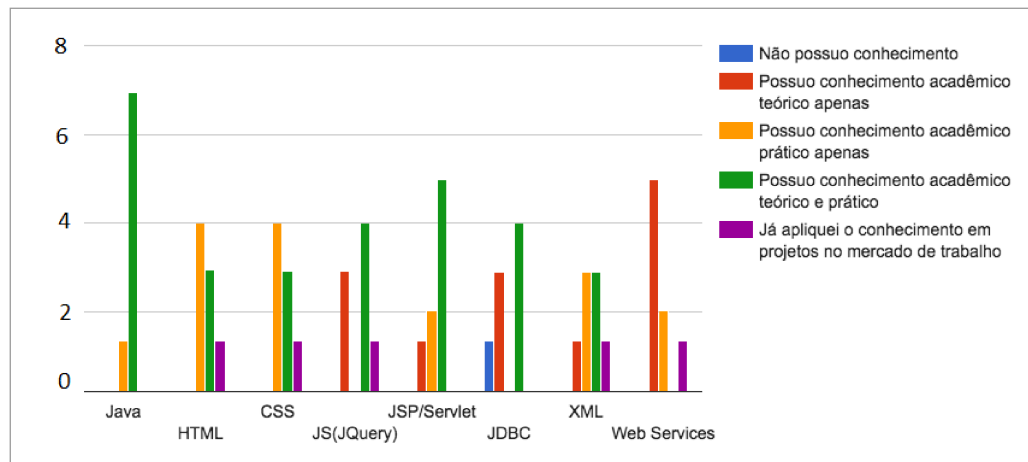


Figura 10 – Gráfico de perfis dos participantes.

Assim, os dois grupos foram formados de maneira homogênea, buscando separar os membros em equipes distintas sempre que dois ou mais indivíduos afirmavam possuir conhecimento prévio de uma mesma tecnologia em níveis semelhantes. A formação dos grupos de participantes é representada individualmente na Tabela 2, apontando os conhecimentos individuais em cada uma das tecnologias utilizadas nos projetos. Os nomes dos participantes foram omitidos. Os apontamentos numéricos equivalem a:

0. Não possuo conhecimento;
1. Possuo conhecimento acadêmico teórico apenas;
2. Possuo conhecimento acadêmico prático apenas;
3. Possuo conhecimento acadêmico teórico e prático; e
4. Já apliquei o conhecimento em projetos no mercado de trabalho.

4.2.3.2 Questionário de competências

Um questionário de competências foi proposto, presumindo as capacidades oferecidas pela plataforma durante o percurso do desenvolvimento. As seguintes competências foram definidas para avaliação da plataforma, sendo C1 a C8 relacionadas ao TAM, e C9 a C13 relacionadas às competência PUA. O questionário foi aplicado ao longo do semestre, conforme os projetos foram sendo desenvolvidos e as competências foram sendo utilizadas,

Tabela 2 – Aferições individuais de perfis dos participantes.

Participante	Java	HTML	CSS	JS JQuery	JSP	JDBC	XML	Web Services
GC1	3	4	4	4	3	1	4	4
GC2	3	2	2	3	1	0	3	1
GC3	3	3	3	1	1	3	1	0
GC4	3	2	2	1	3	3	2	1
GE1	2	2	2	1	2	1	2	2
GE2	3	3	3	3	3	3	3	1
GE3	3	3	3	1	1	1	0	1
GE4	3	3	3	3	3	3	1	1

e as métricas foram coletadas durante as reuniões semanais de reportes e *feedbacks*. As competências definidas foram:

- C1: acesso à ferramenta;
- C2: disponibilidade da ferramenta;
- C3: robustez da ferramenta;
- C4: capacidade de prover dados sem conhecimento prévio de portais municipais;
- C5: capacidade de prover dados sem conhecimento prévio de estruturas de bancos de dados;
- C6: capacidade de prover dados sem conhecimento prévio de comandos SQL;
- C7: flexibilidade da ferramenta;
- C8: dinamismo da ferramenta;
- C9: pesquisas de dados agregados;
- C10: pesquisas de dados com filtros;
- C11: pesquisas de dados de diferentes cidades;
- C12: pesquisas de dados em relação ao tempo; e
- C13: comparação de dados entre municípios.

4.2.4 Análise e interpretação dos resultados

Esta subseção apresenta os resultados analisados e interpretados, após a aplicação dos modelos TAM, PUA e do teste estatístico *t-test*.

4.2.4.1 Modelo de aceitação de tecnologia

Ao fim do projeto, todos os *feedbacks* coletados referentes às competências C1 a C8, relacionadas ao TAM, foram reunidos, e são aqui representados na Figura 11.

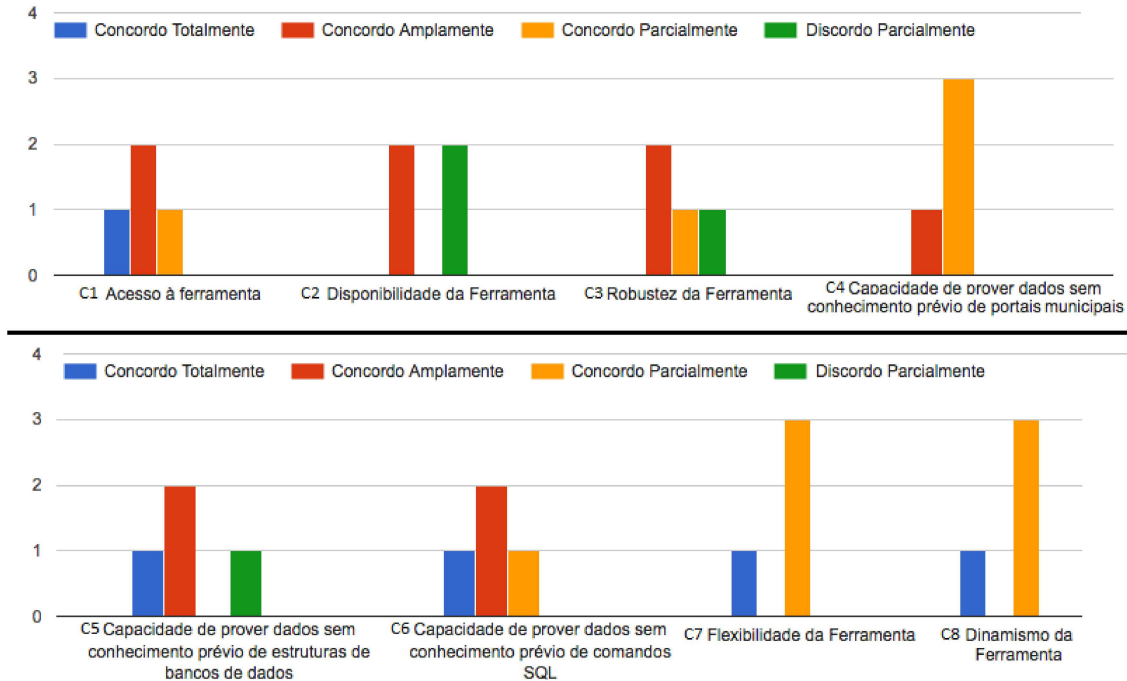


Figura 11 – Gráfico relativo às competências TAM.

Para medir o percentual de concordância das competências foi utilizada a *escala Likert* de seis pontos (LIKERT, 1932). Analisando as aferições indistintamente entre todas as competências, foi aplicada a equação

$$\left(cc = \frac{\sum p}{\sum t} \times 100 \right) \quad (4.1)$$

onde p representa as pontuações aferidas e t o total de pontos máximo, equivalente a 24. Os resultados gerais obtidos são apresentados na Tabela 3, e uma visão completa de todas as aferições é apresentada no gráfico da Figura 12.

Tabela 3 – Aferições totais relativas às competências da plataforma.

Valor	Opção	Aferições
1	“Discordo Totalmente”	00,00%
2	“Discordo Amplamente”	00,00%
3	“Discordo Parcialmente”	12,50%
4	“Concordo Parcialmente”	37,50%
5	“Concordo Amplamente”	34,37%
6	“Concordo Totalmente”	15,63%

Pode-se observar que nenhuma aferição “Discordo Totalmente” ou “Discordo Amplamente” foi apontada. Também nota-se que o maior apontamento foi obtido em “Concordo

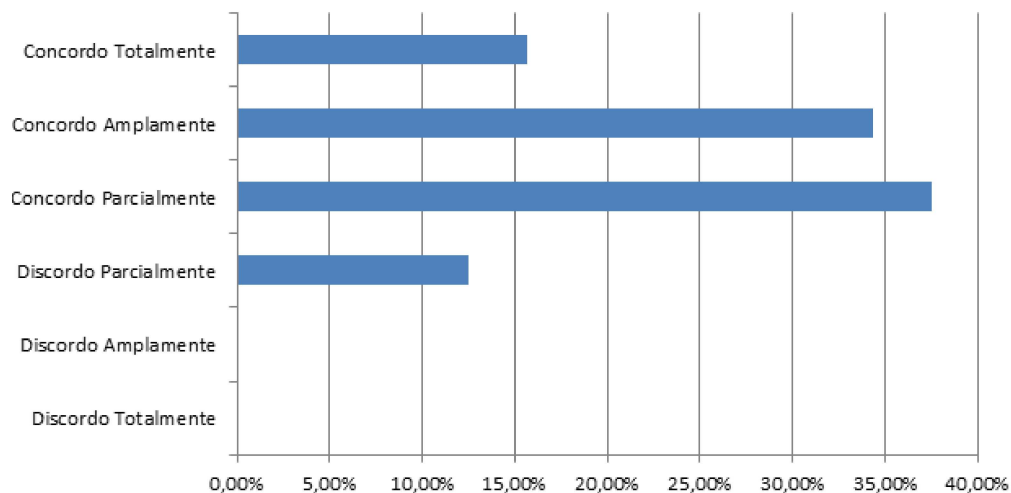


Figura 12 – Gráfico geral sobre as concordâncias apontadas em todas as competências.

Parcialmente”, uma vez que os alunos gostaram das competências oferecidas, mas gostaria que elas fornecessem facilitadores não planejados no escopo do projeto, como manipulações de dados. Esses aferimentos serão discutidos individualmente nas subseções Facilidade de uso percebida, Utilidade da técnica e Facilidade de compreensão, a seguir.

Facilidade de uso percebida

A equação 4.1 foi aplicada aos apontamentos obtidos em relação à perspectiva de facilidade de uso percebida, presentes nas competências C1, C2 e C3, e os resultados, considerando o peso individual de cada apontamento, são apresentados na Tabela 4.

Tabela 4 – Resultados sob a perspectiva de facilidade de uso percebida.

Comp.	Somatório	Média	Mediana	Moda	Aceitação	Concordância
C1	20	5,00	5,00	5	5,00	83,33%
C2	16	4,00	4,00	{3,5}	4,00	66,67%
C3	17	4,25	4,50	5	4,25	70,83%

Ao analisar os resultados, verifica-se que o maior percentual de concordância foi encontrado na competência C1, relacionada ao acesso à ferramenta, provando que o mesmo é democrático e desburocratizado, sendo simples e eficaz. O menor percentual encontrado pertence a C2, relativo a disponibilidade, devido à plataforma possuir sua hospedagem terceirizada e depender da tecnologia de Internet, causando ocasionais atrasos (*delays*) nas requisições, dependendo das condições da conexão, ou se tornando completamente indisponível caso o acesso à Internet esteja intermitente ou interrompido, refletindo também na robustez da ferramenta, na competência C3.

Utilidade da técnica

Em relação às competências analisadas sob a perspectiva de utilidade da técnica, após aplicar a equação 4.1, calculou-se os níveis de aceitação e concordância, e os resultados são apresentados na Tabela 5.

Tabela 5 – Resultados sob a perspectiva de utilidade da técnica.

Comp.	Somatório	Média	Mediana	Moda	Aceitação	Concordância
C4	17	4,25	4	4	4,25	70,83%
C5	19	4,75	5	5	4,75	79,17%
C6	20	5,00	5	5	5,00	83,33%

Como pode ser visto, a competência C6, capacidade de prover dados sem conhecimento prévio de comandos SQL, obteve o maior índice de concordância, sendo essa característica uma das principais vantagens da plataforma, visto que não são necessárias consultas SQL para obter os DAG ou quaisquer mecanismos são necessários para a sua persistência, sendo essas etapas transparentes ao desenvolvedor, mostrando-se extremamente útil. A menor concordância proveio de C4, capacidade de prover dados sem conhecimento prévio de portais municipais. Embora a plataforma ofereça os DAG centralizadamente e uma aula sobre o tema foi lecionada, os alunos visitaram portais para abstrair e aplicar em seus projetos ideias sobre transparência governamental, relacionadas a organização e visualização de DAG. Ainda assim, a competência obteve uma forte aceitação, atingindo 70,83% no geral. A competência C5, capacidade de prover dados sem conhecimento prévio de estruturas de bancos de dados, obteve uma concordância intermediária, uma vez que o projeto da disciplina exigia armazenar gráficos de usuário em bancos de dados, sendo esta obrigação, porém, alheia à plataforma.

Facilidade de compreensão

As competências analisadas sob a perspectiva de facilidade de compreensão foram analisadas através da equação 4.1, e os níveis de aceitação e concordância são apresentados na Tabela 5.

Tabela 6 – Resultados sob a perspectiva de facilidade de compreensão.

Comp.	Somatório	Média	Mediana	Moda	Aceitação	Concordância
C7	18	4,50	4,00	4	4,50	75,00%
C8	18	4,50	4,00	4	4,50	75,00%

As competências C7 e C8 obtiveram um percentual de concordância idêntico, 75,00%. São competências relacionadas respectivamente a flexibilidade e dinamismo, sendo essas características extrínsecas à plataforma, uma vez que a mesma é imutável por definição. Portanto, a facilidade de compreensão sob essas perspectivas foi atingida, uma vez que os alunos apontaram um alto nível de concordância.

Concordância geral do TAM

Para calcular o percentual de concordância de cada perspectiva, foi aplicada a equação

$$cp = \frac{\sum a}{q} \quad (4.2)$$

onde a = aceitação, q = quantidade de competências e v = valor máximo das competências, equivalente a 6.

Após a aplicação do TAM e da realização de análises individuais de cada competência, foi aplicada a equação 4.2 em todas as perspectivas, encontrando o grau de concordância geral referente a facilidade de uso percebida, utilidade da técnica e facilidade de compreensão. A Figura 13 apresenta uma representação gráfica dos resultados atingidos pela plataforma em cada uma das perspectivas.

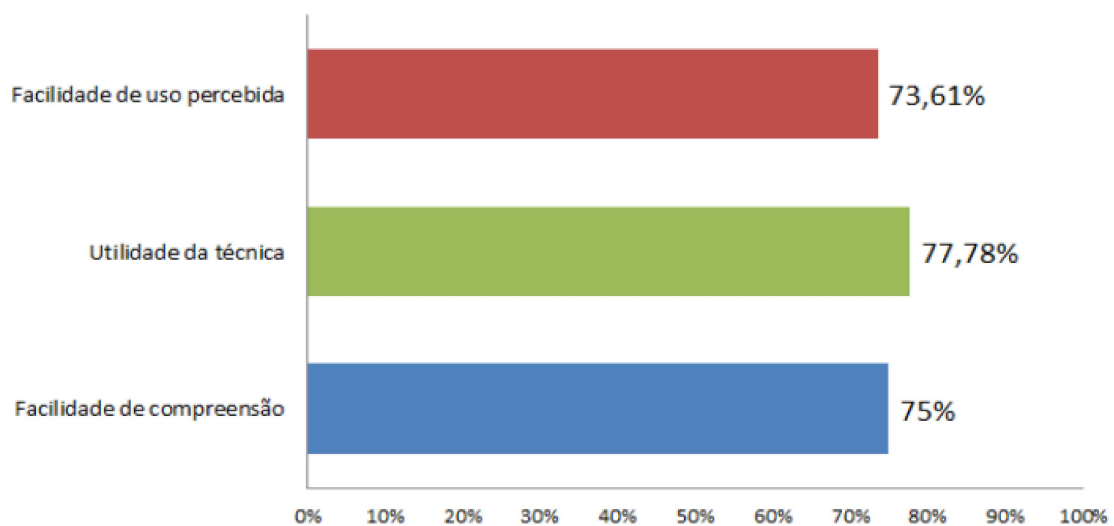


Figura 13 – Gráfico de concordância geral das perspectivas TAM.

Conforme apontamento gráfico, é verificado que o menor percentual de concordância obtido foi de 73,61%, relacionado à facilidade de uso percebida. Embora seja um percentual alto, nota-se que uma nova tecnologia e um novo paradigma de desenvolvimento por muitas vezes acaba encontrando uma certa dificuldade inicial de adaptação. O maior percentual de concordância foi obtido em relação à utilidade da técnica, com 77,78%, mostrando que a plataforma é extremamente útil ao projeto que os desenvolvedores participaram. Os apontamentos relacionados a facilidade de compreensão refletem a pouca experiência dos participantes em relação a DAG e também à plataforma, embora a concordância seja relevante, atingindo um percentual de 75,00%.

Desta maneira, o TAM promove altos índices de aceitação e concordância em relação às competências da plataforma, sob as três perspectivas analisadas.

4.2.4.2 Presença-Utilidade-Adequação

Para cada competência da plataforma foram definidas três possibilidades, referentes a Presença-Utilidade-Adequação (PUA). As competências C9 a C13 foram então avaliadas utilizando a Tabela 7 ao longo do desenvolvimento do projeto, sendo elas:

Presença (P)

- 1 - Funcionalidade não existente.
- 2 - Atende Parcialmente.
- 3 - Atende Completamente.

Utilidade (U)

- 1 - Funcionalidade não se demonstrou útil.
- 2 - Funcionalidade provavelmente útil, mas não foi utilizada.
- 3 - Funcionalidade útil.

Adequação de nível de detalhamento (A)

- 1 - Funcionalidade muito simples.
- 2 - Funcionalidade muito complexa.
- 3 - Funcionalidade não precisa de ajustes.

Tabela 7 – Competências PUA.

Comb.	P	U	A	Implicação das combinações
1	0	0	0	Não é existente, não é util, modificação não é necessária
2	0	0	1	Não é existente, não é util, modificação é necessária
3	0	1	0	Não existente, é util, modificação não é necessária
4	0	1	1	Não existente, é util, modificação é necessária
5	1	0	0	É existente, não é util, modificação não é necessária
6	1	0	1	É existente, não é util, modificação é necessária
7	1	1	0	É existente, é util, modificação não é necessária
8	1	1	1	É existente, é util, modificação é necessária

Análise qualitativa

Para encontrar a utilidade funcional da plataforma, se propôs a análise qualitativa, utilizando as combinações propostas na Tabela 7. Essa análise deve apresentar a lista de ferramentas que foram consideradas úteis pelos alunos, considerando como valor PUA = {X,1,X}, presentes nas combinações de número 3, 4, 7 e 8, sendo críticas as combinações 3 e 4, visto que as competências seriam úteis e não existentes na plataforma, com valor PUA = {0,1,X}. As combinações com valor PUA = {1,1,0} são consideradas ideais, visto que são competências existentes, úteis e não necessitam de modificação.

Competências PUA

Os questionários PUA foram aplicados ao longo do desenvolvimento dos projetos, e os resultados de suas aplicações foram coletados, e são aqui apresentados, reunidos na Figura 14.

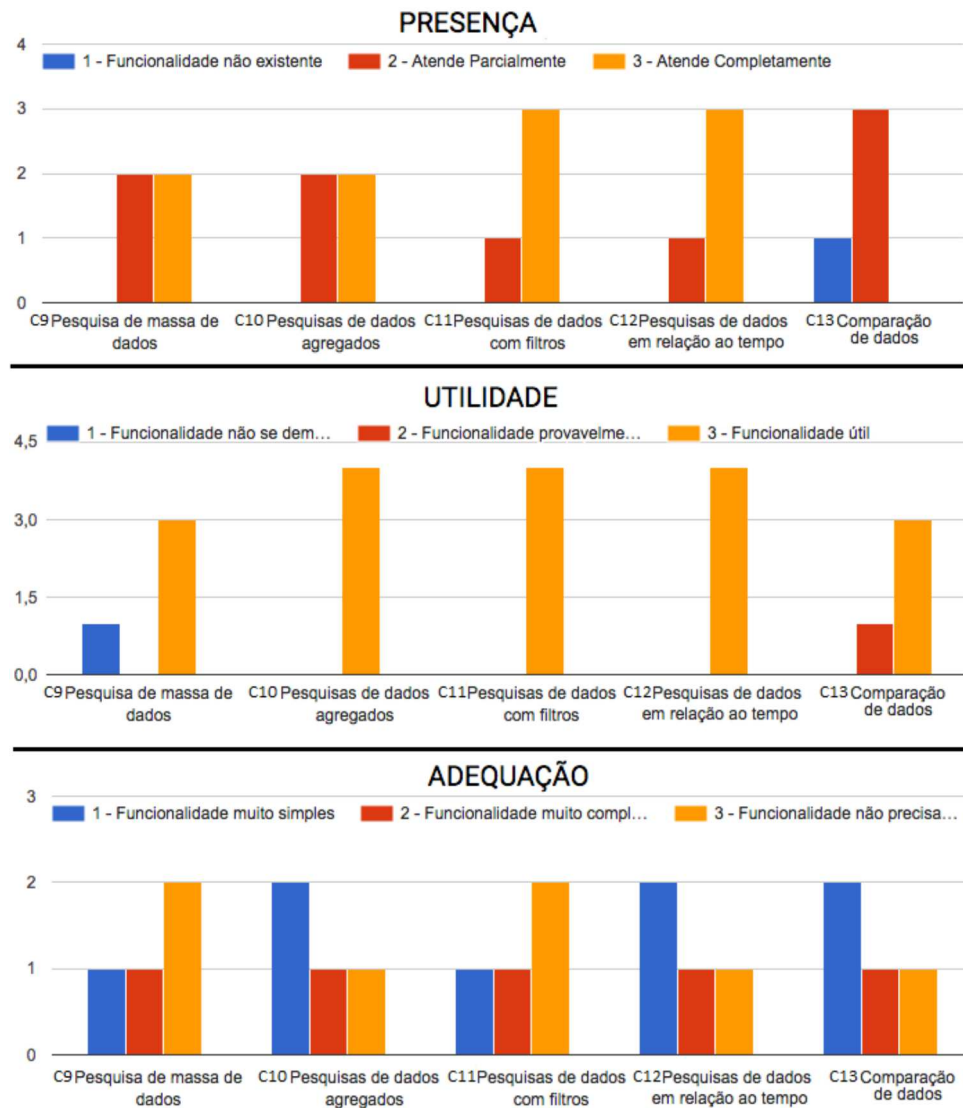


Figura 14 – Gráficos PUA relativos a Presença, Utilidade e Adequação.

Após coletar os apontamentos da aplicação dos questionários PUA, os resultados foram reunidos na Tabela 8.

Como pode ser visto, apenas a competência C13, comparação de dados entre municípios, obteve o valor $PUA = \{0,1,X\}$, onde somente um aluno considerou que esta funcionalidade não é existente e é provavelmente útil, mas não foi utilizada, enquanto os outros três alunos afirmaram ser uma competência útil e que atende parcialmente. Isto se deve ao fato da plataforma apenas fornece os dados em estado cru, sem qualquer tipo de elaboração, cabendo, então, esta tarefa de comparar registros, às aplicações

Tabela 8 – Resultados PUA.

Comp.	Competência	Presença			Utilidade			Adequação		
		1	2	3	1	2	3	1	2	3
C9	Pesquisas de dados agregados	0	2	2	1	0	3	1	1	2
C10	Pesquisas de dados com filtros	0	2	2	0	0	4	2	1	1
C11	Pesquisas de dados de diferentes cidades	0	1	3	0	0	4	1	1	2
C12	Pesquisas de dados em relação ao tempo	0	1	3	0	0	4	2	1	1
C13	Comparação de dados entre municípios	1	3	0	0	1	3	2	1	1

que vão consumir os DAG da plataforma. A competência C11, pesquisas de dados de diferentes cidades, obteve uma forte aceitação, com três alunos apontando que ela é uma funcionalidade que atende completamente e todos eles apontando que ela foi utilizada, provando, assim, que a plataforma promove DAG relevantes de diferentes cidades, restando às aplicações a motivação de como melhor aproveitar as informações providas.

As competências C9 e C10, respectivas a dados agregados e dados com filtros, obtiveram uma aceitação intermediária, enquanto a competência C12, pesquisas de dados em relação ao tempo foi bem aceita, onde todos os desenvolvedores a apontaram como útil, e apenas um deles apontou que a mesma atende parcialmente, enquanto os outros apontaram que atende completamente.

Em relação a adequação, duas competências, C9 e C11, foram apontadas pela maioria dos usuários como funcionalidades que não precisam de ajustes. Nenhuma funcionalidade foi apontada como muito complexa por mais de um usuário, enquanto as competências C10, C12 e C13 foram apontadas como sendo muito simples. Em *feedbacks*, ao serem questionados sobre o apontamento das funcionalidades como muito simples, os alunos, em sua maioria, eram da opinião de que as funcionalidades poderiam retornar os dados pré-trabalhados, como ordenações alfabéticas ou por valores de despesa, além de agrupamentos diversos. Estas funcionalidades, porém, foram planejadas para serem realizadas na camada de desenvolvimento dos portais, devido ao fato da plataforma perder performance e agregar funcionalidades específicas, enquanto o escopo original é de abrangência indistinta e generalizada.

Todas as competências obtiveram o valor $PUA = \{1,1,0\}$, considerado ideal, por ao menos um aluno, onde a funcionalidade é existente, é útil e não precisa de modificações, e as competências C9 e C11 obtiveram este conceito pela maioria dos alunos. Assim, todas as competências da plataforma foram avaliadas positivamente, quando não pela totalidade, ao menos pela maioria de seus usuários, não sendo encontradas competências inexistentes ou desnecessárias, e a plataforma mostrou-se, assim, apta a atender às necessidades dos desenvolvedores de *software* envolvidos.

Tempo de programação

O ciclo de desenvolvimento dos projetos de ambos os grupos foi acompanhado durante o semestre, e todas as tarefas de codificação foram relatadas semanalmente, utilizando a unidade de medida hora (h) como padrão. Foi realizado um agrupamento das tarefas conforme apontamento dos grupos em seus *status report* semanais, durante as três fases de desenvolvimento relacionado a DAG, (a) acesso, (b) manipulação e (c) visualização gráfica. Os resultados são apresentados na Tabela 9.

Tabela 9 – Tempos de desenvolvimento utilizados pelos grupos.

Grupo	Fase	Tarefa	Tempo
GC	Acesso	Coletar DAG e modelar BD	8 h
		Preparar e importar CSVs e inserir nova cidade	9 h
	Manipulação	Visão (<i>View</i>) dos usuários	6 h
		Visão (<i>View</i>) dos gráficos	6 h
	Visualização	Desenvolvimento dos gráficos	7 h
		Exibir grade (<i>grid</i>)	3 h
Salvar gráficos		6 h	
GE	Acesso	Criar as funções de chamada da API	6 h
		Finalizar a criação das funções de chamada da API	4 h
	Manipulação	Desenvolver a comparação entre cidades	5 h
		Desenvolvimento da paginação	7 h
	Visualização	Desenvolver a visualização gráfica dos dados	6 h
		Finalizar a visualização de dados detalhados	5 h
Salvar gráficos de usuário		3 h	

Análise do tempo de programação

Conforme apontamento na Tabela 9, e levantamento no gráfico da Figura 15, o GC necessitou de 17 horas para realizar o acesso aos dados, enquanto a GE finalizou o acesso aos DAG em 10 horas, obtendo assim uma vantagem temporal de 41,18% em relação ao grupo que se utilizou de dados coletados diretamente nos portais governamentais.

Para a manipulação dos DAG, ambos os grupos necessitaram de 12 horas de desenvolvimento e duas tarefas distintas, não houve, assim, diferenças em tempo para esta fase, visto que os dados já estavam incorporados ao projeto e a manipulação dos mesmos foi realizada em linguagens de programação e ambientes presentes na grade curricular e, portanto, confortáveis aos alunos, além dos mesmos estarem equilibradamente dispostos em equipes com níveis de conhecimento muito próximos entre si.

A última fase, relacionada a formulação das visualizações gráficas, custou 16 horas de desenvolvimento ao GC, enquanto o GE a finalizou em 14 horas e, portanto, obteve um desempenho 12,50% superior em relação ao tempo.

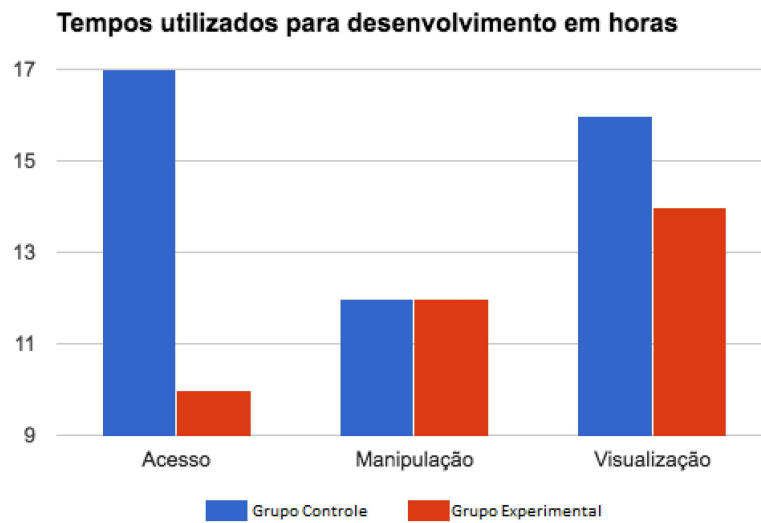


Figura 15 – Gráfico de tempo de desenvolvimento.

Ao fim do projeto, analisando o desenvolvimento como um todo, o GC necessitou de 45 horas de programação relacionada a DAG, ao passo que o GE finalizou as mesmas tarefas em 36 horas, uma vantagem temporal de 20,00%, embora uma plataforma centralizada seja mais relevante nas etapas iniciais de acesso a dados, não influenciando nas outras etapas, após os dados já estarem incorporados ao projeto.

Quantidade de linhas de código-fonte

Ao fim dos projetos, todos os arquivos de código-fonte foram enviados para análise. Aqui foram consideradas apenas as linhas que envolvem diretamente acesso a DAG, não sendo consideradas linhas de código envolvidas em outras etapas do desenvolvimento do projeto, como a manipulação dos dados, visto que diferentes soluções foram utilizadas pelos grupos, como pacotes prontos de folhas de estilos, ou ainda os mesmos foram escritos manualmente pelos desenvolvedores. Os valores registrados são representados na Tabela 10, onde também é apontada a quantidade de linhas de código utilizada no desenvolvimento da API.

Tabela 10 – Linhas de código utilizadas durante o desenvolvimento.

Grupo	Tarefa	Linhas
GC	Acessar os DAG	387
	Desenvolver as visualizações gráficas	509
GE	Acessar os DAG	90
	Desenvolver as visualizações gráficas	154
API	Chamadas da API	341

Os arquivos de código-fonte relacionados a acesso a DAG são representados parcialmente nas Figuras 16 e 17, relacionadas ao GC e GE, respectivamente.


```

479     $data = $this->formatPeriod($request['periodo']);
480     $graphic = new Graphic([
481         'name' => $request['name'],
482         'tipo' => 'totalCities',
483         'dataini' => $this->formatDateToBD($data[0]),
484         'datafim' => $this->formatDateToBD($data[1]),
485     ]);
486     Auth::user()->graphics()->save($graphic);
487     return redirect('user');
488 }
489 public function detailsCity() {
490     $request = Request::all();
491     if(!isset($request['page'])){
492         $data = $this->formatPeriod($request['periodo']);
493         $request['dataini'] = $data[0];
494         $request['datafim'] = $data[1];
495         Session::put('cidade', $request['cidade']);
496         Session::put('periodo', $request['periodo']);
497         Session::put('dataini', $data[0]);
498         Session::put('datafim', $data[1]);
499         Session::put('titulo', $request['titulo']." (detalhes)");
500     }
501     $response = Dag::selectRaw('credor_secretaria, funcao, subfuncao, fonte_de_recurso,
502         valor_pago_acumulado, natureza_da_despesa')
503         ->where('cidade', Session::get('cidade'))->whereBetween('mes_ano', [Session::get('dataini'),
504             Session::get('datafim')])
505         ->orderBy('valor_pago_acumulado', 'desc')->orderBy('credor_secretaria')->orderBy('funcao')->
506         orderBy('subfuncao')->paginate('7');
507     foreach($response as $v) {
508         $v['valor_pago_acumulado'] = $this->formatNumber($v['valor_pago_acumulado']);
509     }
510     return view('pages.detailsCity')->with('response', $response);
511 }

```

Figura 16 – Codificação de acesso a DAG, Grupo Controle.

```

63     :domain => params[:domain].present? ? params[:domain][index] : nil,
64     :subdomain => params[:subdomain].present? ? params[:subdomain][index] : nil,
65     :nature => params[:nature].present? ? params[:nature][index] : nil
66 }
67 end
68
69 def call_api(f, city, year, month, domain, subdomain, nature)
70     # create a client for the service
71     client = Savon.client(wSDL: 'http://transparenciaws.elasticbeanstalk.com/TransparenciaWS.asmx?WSDL')
72
73     fun = if f == 1
74         :get_total_despesa
75     else
76         :get_lista_despesa
77     end
78
79     response = client.call(fun, message: {
80         wNomeCidade: city,
81         wAno: year,
82         wMes: month,
83         wDominio: domain,
84         wSubDominio: subdomain,
85         wNatureza: nature
86     })
87
88     response.body
89 end
90 end

```

Figura 17 – Codificação de acesso a DAG, Grupo Experimental.

Análise da quantidade de linhas de código-fonte

Conforme relatado na Tabela 10, ambos os grupos utilizaram-se de dois arquivos para importar os DAG aos seus projetos, pertencentes à camada *controller* da arquitetura

Model-View-Controller (MVC)⁶, sendo esta uma das exigências da disciplina para o desenvolvimento do projeto.

O GC necessitou codificar 387 linhas para acessar os DAG e outras 509 linhas para fornecer informações às visualizações gráficas, totalizando 896 linhas de código, entre criações de visões, procedimentos armazenados e comandos SQL diversos. O GE, por sua vez, necessitou de apenas 90 linhas de código para acesso aos DAG e 154 linhas para formular os gráficos, em um total de 244 linhas.

Nitidamente, o GE obteve uma vantagem expressiva, necessitando 76,75% menos codificação para acessar os DAG e 69,75% menos linhas para alimentar os gráficos, com um desempenho geral 72,77% superior em relação ao GC, conforme aponta o gráfico da Figura 18, não se levando em conta as linhas de código da API.

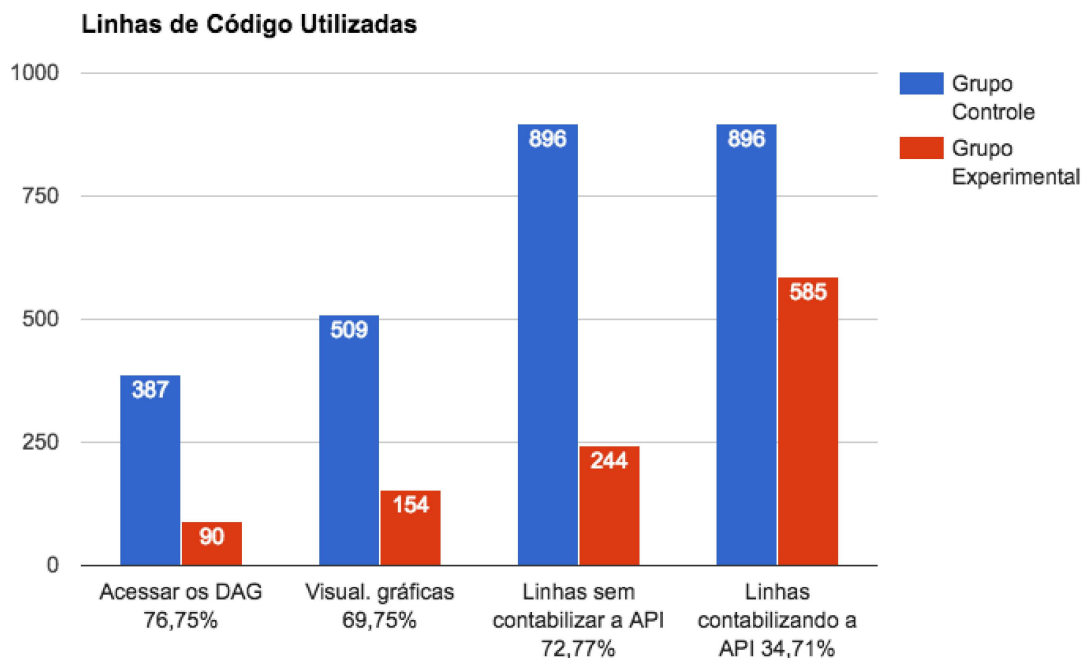


Figura 18 – Gráfico de quantidade de linhas de código.

Por fim, ao analisar as linhas de código do grupo experimental, somando com as linhas de código da API, obtemos um total de 585 linhas, um valor ainda assim 34,71% inferior, em comparação com as 896 linhas utilizadas pelo grupo controle.

Complexidade ciclomática

Os arquivos de código-fonte relacionados a acesso a DAG de ambos os grupos foram submetidos a uma análise de complexidade ciclomática (MCCABE, 1976), e os valores encontrados são apresentados na Tabela 11. Analisando a tabela verifica-se que o grupo experimental utilizou apenas três métodos passíveis de calcular sua complexidade,

⁶ <http://grove.cs.jmu.edu/groverf/papers/WEBIST2011.pdf> Acessado em Jun/2016

atingindo uma complexidade máxima de valor 9 em apenas um método, enquanto o grupo controle utilizou vinte e quatro métodos calculáveis, atingindo o valor de complexidade máxima 8 em três métodos distintos. Pode-se concluir que uma plataforma centralizada aumenta ligeiramente a complexidade ciclomática máxima, mas diminui drasticamente a quantidade de métodos necessários.

Tabela 11 – Complexidade ciclomática dos métodos utilizados para acesso a DAG.

Grupo	Método	Compl.
GE	get_lista_despesa	9
	get_params	7
	get_total_despesa	5
GC	searchbyfunction	8
	searchbyfunctionnormalized	8
	functiongraphic	6
	functionnormalizedgraphic	6
	mygraphics	6
	searchbycity	6
	selectgraphic	6
	searchbytotalcities	5
	citygraphic	4
	detailscity	3
	detailsfunction	3
	formatnumber	3
	totalcitiesgraphic	3
	detailstotalcity	1
	erasegraphic	1
	formatdatetobd	1
	formatdatetouser	1
	formatjson	1
	formatperiod	1
	formattitulofuncao	1
storecitygraphic	1	
storefunctiongraphic	1	
storefunctionnormalizedgraphic	1	
storetotalcitiesgraphic	1	

4.2.4.3 Aplicação do teste estatístico

Após a identificação dos índices de aceitação e concordância das competências da plataforma por seus usuários, e da verificação da presença, utilidade e adequação de suas ferramentas, foram realizados testes estatísticos, com a finalidade de verificar significativas diferenças no desempenho obtido entre os grupos, medidos em tempo de desenvolvimento e quantidade de linhas de código.

A ferramenta *Minitab* 17⁷ foi utilizada como suporte para os cálculos estatísticos e

⁷ <http://support.minitab.com/en-us/minitab/17/> Acessado em Jun/2016

criação de gráficos. Devido às amostras serem reduzidas, contando com medidas de tempo e linhas de código, contabilizadas em tarefas distintas e realizadas por especialistas na área de desenvolvimento de *software*, em um trabalho em grupo, a confiança foi definida em 90%, com significância $\alpha = 0,10$ (WOHLIN, 2000).

Análise das hipóteses

Duas hipóteses foram definidas hipóteses, onde pode não haver diferença significativa entre o uso da plataforma em relação ao projetos que contou com DAG coletados diretamente nos portais governamentais, ou a plataforma pode se mostrar mais eficiente em relação a tempo ou linhas de código. As hipóteses são:

- **Hipótese Nula (H_0)** - Não houve diferença significativa nas métricas de tempo e quantidade de linhas de código na comparação entre os projetos dos grupos que utilizaram a plataforma e os que não utilizaram a mesma.

$$H_0 : \mu_1 = \mu_2 \quad (4.3)$$

- **Hipótese Alternativa (H_A)** - Houve diferença significativa nas métricas de tempo e quantidade de linhas de código na comparação entre os projetos dos grupos que utilizaram a plataforma e os que não utilizaram a mesma.

$$H_A : \mu_1 \neq \mu_2 \quad (4.4)$$

Verificação das hipóteses

Para a verificação das hipóteses, o teste estatístico *t-test* foi aplicado nos valores apresentados nas Tabelas 9 e 10. Os *boxplots* apresentados na Figura 19 comparam os tempos de desenvolvimento de acesso a DAG entre os grupos. Pode-se observar que os resultados dos grupos não se sobrepõem, o que indica que o resultado do teste é significativo, com o valor de $P = 0,047$ (*P-Value*), como mostrado na Figura 20, refutando, assim, a hipótese nula (H_0), e aceitando-se a hipótese alternativa (H_A), onde houve diferença significativa na atividade desenvolvida para os grupos que utilizaram a plataforma.

O próximo teste, mostrado na Figura 21 é relacionado às horas de desenvolvimento de visualizações gráficas, uma vez que a manipulação dos DAG obteve os mesmos valores para ambos os grupos, impossibilitando a aplicação do *t-test*. Analisando a Figura 22, nota-se que $P-Value = 0,317$ e portanto, não obtendo uma significância entre as equipes, aceitando-se, assim a hipótese nula (H_0). Isto deve-se ao fato de que, embora a plataforma auxilie no acesso aos DAG, as visualizações gráficas envolvem processos de manipulação de dados internamente após os mesmos estarem incorporados ao projeto, portanto não oferecendo uma substancial vantagem, o que é comprovado no teste estatístico.

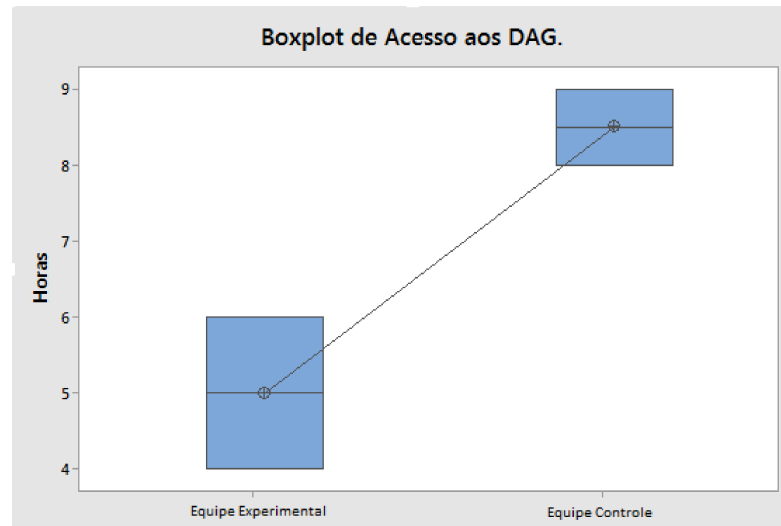


Figura 19 – Boxplot referente às horas de desenvolvimento para acesso aos DAG.

Teste T para Duas Amostras e IC: Acesso Equipe Portal; Acesso Equipe Plataforma

Teste T para 2 amostras para Acesso Equipe Portal vs Acesso Equipe Plataforma

	N	Média	DesvPad	EP Média
Acesso Equipe Portal	2	8,500	0,707	0,50
Acesso Equipe Plataforma	2	5,00	1,41	1,0

Diferença = μ (Acesso Equipe Portal) - μ (Acesso Equipe Plataforma)
 Estimativa para a diferença: 3,50
 Limite inferior de 90% para a diferença: 1,39
 Teste T de diferença = 0,1 (vs >): Valor-T= 3,04 Valor-P = 0,047 GL = 2
 Ambos usam DesvPad Combinado = 1,1180

Figura 20 – Valores *t-test* referentes às horas de desenvolvimento para acesso aos DAG.

Finalmente, foi aplicado o *t-test* em todos os valores, referentes ao ciclo completo de desenvolvimento. Os resultados do *boxplot* podem ser vistos na Figura 23, onde nota-se uma ligeira sobreposição dos valores, verificando-se que, embora próxima de seu valor máximo de significância, a diferença obtida foi suficiente para refutar a hipótese nula (H_0), com *P-Value* = 0,071, conforme representado na Figura 24. Portanto a hipótese alternativa (H_A) é aceita, afirmando, assim, que houve diferença significativa na atividade, onde os grupos que utilizaram a plataforma obtiveram um resultado melhor que os grupos que não utilizaram a mesma.

Com relação às linhas de código, o *t-test* obteve uma diferença significativa, com *P-Value* = 0,021, sendo a melhor diferença encontrada entre todas as aplicações estatísticas, validando, assim, a hipótese alternativa (H_A), e compulsoriamente rejeitando a hipótese nula (H_0). Os resultados do teste podem ser visualizados nas Figuras 25 e 26.

Assim, é comprovada a vantagem de se utilizar uma plataforma centralizada de

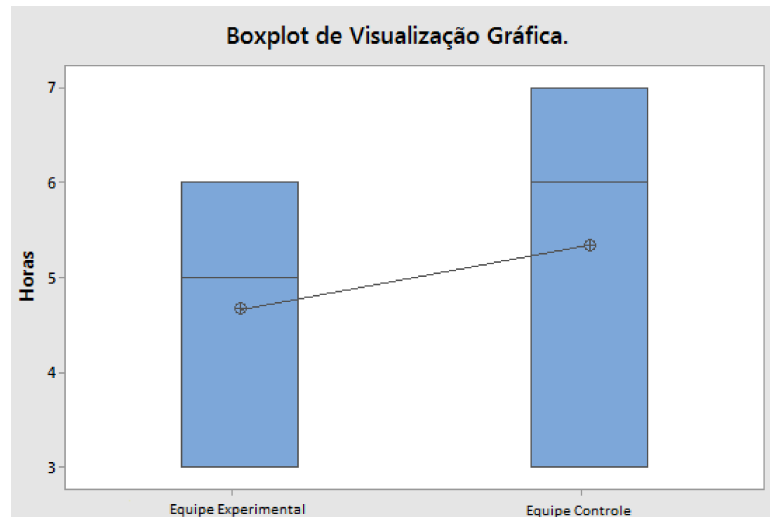


Figura 21 – Boxplot referente às horas de desenvolvimento de visualização gráfica dos DAG.

Teste T para Duas Amostras e IC: Vis. Gráfica Equipe Plataforma; Vis. Gráfica Equipe Portal

Teste T para 2 amostras para Vis. Gráfica Equipe Plataforma vs Vis. Gráfica Equipe Portal

	N	Média	DesvPad	EP Média
Vis. Gráfica Equipe Plat	3	4,67	1,53	0,88
Vis. Gráfica Equipe Port	3	5,33	2,08	1,2

Diferença = μ (Vis. Gráfica Equipe Plataforma) - μ (Vis. Gráfica Equipe Portal)
 Estimativa para a diferença: -0,67
 Limite superior de 90% para a diferença: 1,62
 Teste T de diferença = 0,1 (vs <): Valor-T= -0,51 Valor-P = 0,317 GL = 4
 Ambos usam DesvPad Combinado = 1,8257

Figura 22 – Valores *t-test* referentes às horas de desenvolvimento de visualização gráfica dos DAG.

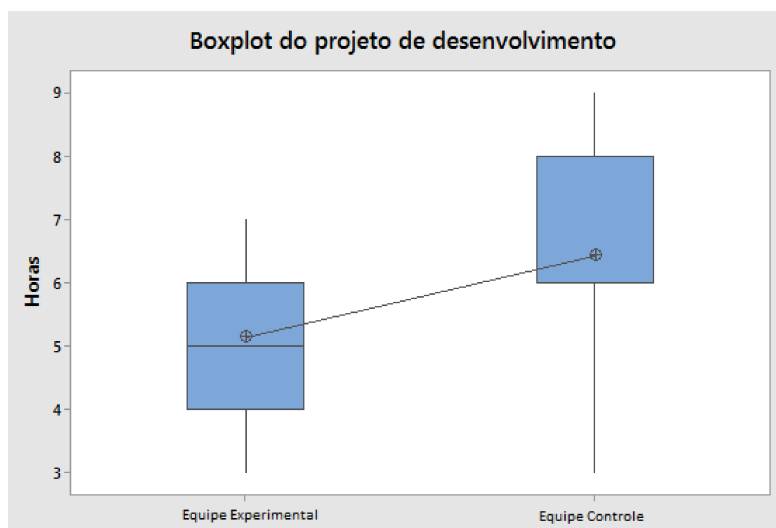


Figura 23 – Boxplot referente às horas de desenvolvimento de todo o projeto.

Teste T para Duas Amostras e IC: Equipe Plataforma; Equipe Portal

Teste T para 2 amostras para Equipe Plataforma vs Equipe Portal

	N	Média	DesvPad	EP Média
Equipe Plataforma	7	5,14	1,35	0,51
Equipe Portal	7	6,43	1,90	0,72

Diferença = μ (Equipe Plataforma) - μ (Equipe Portal)

Estimativa para a diferença: -1,286

Limite superior de 90% para a diferença: -0,091

Teste T de diferença = 0,1 (vs <): Valor-T= -1,57 Valor-P = 0,071 GL = 12

Ambos usam DesvPad Combinado = 1,6475

Figura 24 – Valores *t-test* referentes às horas de desenvolvimento de todo o projeto.

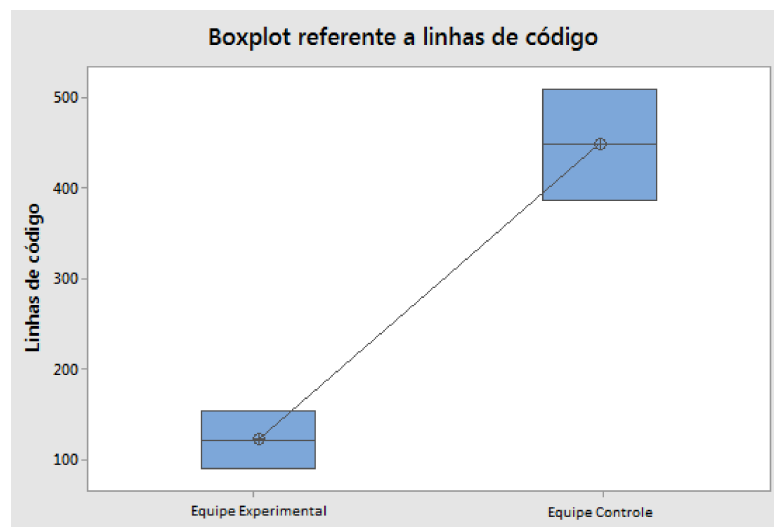


Figura 25 – Boxplot referente à quantidade de linhas de código.

Teste T para Duas Amostras e IC: Equipe Plataforma; Equipe Portal

Teste T para 2 amostras para Equipe Plataforma vs Equipe Portal

	N	Média	DesvPad	EP Média
Equipe Plataforma	2	122,0	45,3	32
Equipe Portal	2	448,0	86,3	61

Diferença = μ (Equipe Plataforma) - μ (Equipe Portal)

Estimativa para a diferença: -326,0

Limite superior de 90% para a diferença: -196,1

Teste T de diferença = 0,1 (vs <): Valor-T= -4,73 Valor-P = 0,021 GL = 2

Ambos usam DesvPad Combinado = 68,8840

Figura 26 – Valores *t-test* referentes à quantidade de linhas de código.

DAG para o desenvolvimento de projetos de CI em relação à ganhos temporais para acesso aos registros, o que reflete na contabilização final de horas do projeto, oferecendo também benefícios que, embora menores em relação à diferença obtida no acesso, ainda assim é significativa. A vantagem em relação a linhas de código também foi obtida. Embora apontem para uma correlação entre ambos os fatores, são necessários estudos para comprovar o

relacionamento entre diminuição de quantidade de linhas de código e seus reflexos no tempo de desenvolvimento, e vice-versa.

Ameaças à validade do estudo

Validade Interna: os participantes são estudantes regulares e desenvolveram projetos voltados ao público alvo interessado em consumo de DAG, com representatividade no mercado de *Software* Comercial. Os grupos foram definidos com base em conhecimentos prévios, por meio da aplicação de um questionário de uniformização. Uma possível ameaça seriam atritos entre os membros, visto que a formação dos grupos se deu com base em seus conhecimentos prévios e não em suas afinidades pessoais. Porém os projetos foram desenvolvidos e nenhum relato de problemas de relacionamento foi mencionado, embora o acompanhamento não tenha sido realizado “*in loco*”. Devido a este fato, também poderiam surgir ameaças de evasão ou abandono, porém, devido ao fato do projeto estar atrelado a uma disciplina obrigatória de graduação, esta ameaça não foi encontrada.

Validade de Conclusão: a aplicação do *t-test* para amostras independentes se mostrou o teste estatístico mais adequado ao cenário, visto que os grupos utilizaram diferentes abordagens para desenvolver o projeto. Uma possível ameaça à validade de conclusão seria o confronto com outros testes, uma vez que não foram aplicados diferentes testes após a definição inicial do *t-test* para amostras independentes.

Validade de Construção: todo o processo do desenvolvimento de *software* foi acompanhado e documentado, em todas as etapas, considerando apenas a opinião dos usuários em formulários respondidos pela Internet, para assim evitar a ameaça de influência na pesquisa. Porém, o acompanhamento foi *offline*, e não puderam ser confirmados os apontamentos de horas dedicadas ao desenvolvimento dos projetos.

Validade Externa: os participantes são considerados influentes para o ambiente de desenvolvedores de *software*, por estarem em estágios avançados nas disciplinas oferecidas na universidade e utilizarem ferramentas representativas no mundo corporativo. Com isso, a validade externa pode ser considerada abrangente, podendo ser replicada em ambientes similares, como grupos de desenvolvedores atuantes em *startups* ou ainda, no mercado corporativo, replicando-se as condições do estudo. O fato da grande maioria dos participantes do estudo nunca terem atuado profissionalmente pode afetar a comparação com um projeto construído por profissionais.

4.3 Considerações

Este capítulo apresentou duas validações à plataforma de DAG, ambas em contexto acadêmico, com a participação de alunos de graduação. A primeira validação, embora não se tenham obtido métricas importantes para estudos estatísticos, serviu como base

para a consolidação da plataforma e para a formalização da próxima validação, já com a plataforma estável e imutável, tornando possível a aplicação do estudo experimental. Nele, foram coletadas métricas de tempo de desenvolvimento e quantidade de linhas de código-fonte escritas, além da aplicação de questionários de competências. Através do modelo de aceitação de tecnologia pode-se verificar que a plataforma atingiu 73,61% de concordância quanto à facilidade de uso percebida, 77,78% de concordância quanto à utilidade da técnica e 75,00% de concordância quanto à facilidade de compreensão. A partir das análises de competências, a plataforma foi aprovada por seus usuários, não obtendo nenhuma competência proposta considerada como inútil, não utilizada ou necessitando de alterações. Por fim, após aplicação de testes estatísticos, constatou-se que foram obtidos 20,00% de redução temporal de desenvolvimento e 72,77% de redução de linhas de código-fonte utilizado para atingir os mesmos objetivos, através da utilização da plataforma. Assim, comprova-se a eficiência em utilizar um modelo centralizado de distribuição de DAG para utilização em aplicações no âmbito de CI.

Conclusão

No contexto de Cidades Inteligentes, a utilização de Dados Abertos Governamentais, distribuídos gratuitamente, para alimentar seus processos e aplicações sem a necessidade de custosos equipamentos e sensores, tende a se tornar uma demanda cada vez mais relevante. Neste sentido, foi apresentada uma plataforma para análise e disponibilização centralizada de DAG a projetos de CI. A plataforma é passível de expansão e disponibilização de dados de inúmeros municípios, além de ser projetada visando suportar métodos computacionais para coleta e organização de dados sem interferência humana no processo.

Ainda, foram apresentadas duas validações experimentais, a primeira servindo como base para a consolidação e a estabilização da plataforma, e a segunda apresentando estudo experimental completo, com comprovações estatísticas das vantagens da utilização da plataforma em relação à um projeto desenvolvido de maneira usual, com dados coletados diretamente nos portais de transparência. Com isso, o objetivo principal deste trabalho foi alcançado, provendo uma plataforma funcional, escalável, adaptada às necessidades, robusta e validada, suprimindo uma necessidade que não foi encontrada em nenhum trabalho relacionado.

Trabalhos futuros

Em relação aos possíveis trabalhos futuros, sugere-se:

- Desenvolvimento de um *crawler* parametrizável: desta maneira podem-se obter dados dos diversos portais que não disponibilizam DAG em formatos amigáveis ou legíveis humanamente para importação na plataforma;
- Desenvolvimento de caracterizadores automáticos de dados: utilizando técnicas de aprendizado de máquina para classificar as informações já coletadas, sendo possível a automatização desta tarefa;
- Integração com a ferramenta *ConvertView*: Dadas as características da ferramenta *ConvertView*, que recebe um XML como parâmetro de entrada e o converte em arquivo RDF, uma integração de ambas as plataformas pode ser estudada, gerando grandes e promissoras possibilidades, como uma maior documentação e extensão dos formatos de distribuição;
- Desenvolver a API utilizando REST: desenvolver e testar as possíveis vantagens da abordagem REST nos *web services*;

- Reteste da plataforma com os mesmos usuários: um novo estudo experimental aplicado aos mesmos usuários em um novo projeto pode se mostrar extremamente vantajoso, uma vez que os usuários já são confortáveis com a utilização da mesma, obtendo assim ainda mais benefícios;
- Estudo experimental aplicado a desenvolvedores individuais: os estudos experimentais apresentados neste trabalho foram idealizados para grupos de desenvolvedores, portanto se desconhece as possíveis vantagens ou desvantagens de sua utilização por um único indivíduo;
- Pesquisas referentes a produtividade com respeito a linhas de código por tempo de desenvolvimento: embora indiquem correlação, não foram estudados os possíveis efeitos dessas duas variáveis e sua aplicação nos projetos, utilizando ou não a plataforma; e
- Testes de performance: não foram realizados testes referentes à performance de ambos os projetos durante o estudo experimental.

Produção científica

Ao longo do projeto, duas submissões foram realizadas. A primeira ainda durante o estágio de desenvolvimento da plataforma, o que implicou em sua reprovação. Porém os *feedbacks* dos revisores foram promissores, o que motivou uma segunda submissão, a uma revista científica (*scientific journal*) em edição especial sobre governo eletrônico. Devido ao período de submissões de artigos da edição conflitar com o encerramento do semestre letivo, o artigo foi submetido sem o estudo experimental, visto que ainda estavam sendo coletadas as métricas para o desenvolvimento do mesmo, o que, novamente, implicou na reprova do artigo. Uma nova submissão está sendo estudada, pela primeira vez contando com a plataforma e o estudo experimental finalizados.

Submissões

- Daniel I. Vieira, Alexandre Alvaro. *Uma Plataforma para Disponibilização Centralizada de Dados Abertos Governamentais como Suporte para Aplicações no Contexto de Cidades Inteligentes*. iSys - Revista Brasileira de Sistemas de Informação (Edição especial sobre Governo Eletrônico, v.9, n.1), 2016.
- Daniel I. Vieira, Alexandre Alvaro. *Em Direção a Uma Plataforma para Disponibilização Centralizada de Dados Abertos Governamentais como Suporte para Aplicações no Contexto de Cidades Inteligentes*. SBSI - XI Simpósio Brasileiro de Sistemas de Informação (Sistemas de Informação de Governo), 2015.

Referências

- AFONSO, R. A. et al. Mapeamento sistemático de informática médica para cidades inteligentes. *XIV Congresso Brasileiro em Informática em Saúde – CBIS*, 2014. Citado na página 36.
- AFONSO, R. A. et al. Smartcluster: Utilizando dados públicos para agrupar cidades inteligentes por domínios. *XI Brazilian Symposium on Information System, Goiânia, GO*, p. 699–702, 2015. Citado na página 36.
- AFONSO, R. A. et al. Br-scmm: Modelo brasileiro de maturidade para cidades inteligentes. *IX Simpósio Brasileiro de Sistemas de Informação (SBSI), 2013, João Pessoa/PB*, p. 511–516, 2015. Citado na página 36.
- AFONSO, R. A. et al. Convertview - a tool for conversion and visualization of open heterogenic governmental data to the rdf standard. *Proceedings of the 11th International Conference on Web Information Systems and Technologies - Lisbon, Portugal*, p. 253–258, 2015. Citado na página 38.
- ALAWADHI, S.; SCHOLL, H. J. Aspirations and realizations: The smart city of seattle. *46th Hawaii International Conference on System Sciences*, p. 1695–1703, 2013. Citado na página 33.
- BASILI, V. R.; ROMBACH, H. D. The tame project: towards improvement-oriented software environments. *IEEE Transactions on Software Engineering (Volume:14 , Issue: 6)*, p. 758 – 773, 1988. ISSN: 0098-5589. Citado na página 57.
- BEERI, C.; BERNSTEIN, P. A.; GOODMAN, N. A sophisticate’s introduction to database normalization theory. *Proceedings of the fourth international conference on Very Large Data Bases-Volume 4*, p. 113–124, 1978. VLDB Endowment. Citado na página 42.
- BREITMAN, K. et al. Open government data in brazil. *IEEE Computer Society*, p. 45–49, 2012. Citado 2 vezes nas páginas 28 e 34.
- BRITO, K. d. S. et al. Brazilian government open data: Implementation, challenges, and potential opportunities. *Proceedings of the 15th Annual International Conference on Digital Government Research*, p. 11–16, 2014. Citado 4 vezes nas páginas 22, 30, 31 e 35.
- BRITO, K. d. S. et al. Assessing the benefits of open government data: the case of meu congresso nacional in brazilian elections 2014. *Proceedings of the 16th Annual International Conference on Digital Government Research*, p. 89–96, 2015. Citado na página 38.
- BRITO, K. d. S. et al. Using parliamentary brazilian open data to improve transparency and public participation in brazil. *Proceedings of the 15th Annual International Conference on Digital Government Research*, p. 171–177, 2014. Citado na página 22.
- CASA CIVIL. *LEI Nº 12.527, DE 18 DE NOVEMBRO DE 2011*. 2011. <http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/l12527.htm> Acessado em Jun/2016. Citado na página 28.

- CGU. *Conferência Nacional sobre Transparência e Controle Social*. 2012. <<http://www.cgu.gov.br/assuntos/controle-social/consocial>> Acessado em Jun/2016. Citado na página 29.
- CHEN, T. M. Smart grids, smart cities need better networks. *IEEE Network*, p. 2–3, 2010. Editor's Note. Citado 2 vezes nas páginas 21 e 33.
- CONSOCIAL. *Diretrizes*. 2012. <<http://aeppsp.org.br/msc/consocial-diretrizes.pdf>> Acessado em Jun/2016. Citado na página 29.
- CONSOCIAL. *Propostas*. 2012. <http://www.cgu.gov.br/assuntos/controle-social/consocial/arquivos/lista80_propostas_finais.pdf> Acessado em Jun/2016. Citado na página 29.
- CONSOLI, S. et al. A smart city data model based on semantics best practice and principles. *Proceeding WWW '15 Companion Proceedings of the 24th International Conference on World Wide Web*, p. 1395–1400, 2015. WWW International World Wide Web Conference. Citado na página 38.
- CORRÊA, A. S.; CORRÊA, P. L. P.; SILVA, F. S. C. Transparency portals versus open government data: an assessment of openness in brazilian municipalities. *Proceedings of the 15th Annual International Conference on Digital Government Research*, p. 178–185, 2014. Citado 3 vezes nas páginas 22, 30 e 35.
- DAVIS, F. D.; BAGOZZI, R. P.; WARSHAW, P. R. User acceptance of computer technology: a comparison of two theoretical models. *INFORMS Institute for Operations Research and the Management Sciences (INFORMS), Linthicum, Maryland, USA*, p. 982 – 1003, 1989. ISSN: 0025-1909. Citado na página 57.
- DINIZ, H. B. M.; SILVA, E. G. F.; GAMA, K. S. A reference architecture for a crowdsensing platform in smart cities. *SBSI 2015 Proceedings of the annual conference on Brazilian Symposium on Information Systems: Information Systems: A Computer Socio-Technical Perspective - Volume 1*, p. 13, 2015. Brazilian Computer Society Porto Alegre, Brazil, Brazil ©2015. Citado na página 38.
- DONOLO, R. M. How to achieve smart cities through smart communication and representation of urban data. *29th Urban Data Management Symposium*, p. 83–86, 2013. London. Citado na página 21.
- DYBÅA, T.; KAMPENESA, V. B.; SJØBERGA, D. I. K. User acceptance of computer technology: a comparison of two theoretical models. *Information and Software Technology Volume 48, Issue 8*, p. 745–755, 2006. Elsevier. Citado na página 57.
- EAVES, D. *Open Activist*. 2009. <<http://eaves.ca/about-david/>> Acessado em Jun/2016. Citado na página 27.
- EVERS, M. *Living Lab: Urban Planning Goes Digital in Spanish 'Smart City'*. 2013. <<http://www.spiegel.de/international/world/santander-a-digital-smart-city-prototype-in-spain-a-888480.html>>. Acessado em Jun/2016. Citado 2 vezes nas páginas 21 e 34.

- HARTMANN, D.; WOLTER, K.; KRAUSS, T. Ict resilience simulations in small confined smart distribution grids. *Proceedings of 2013 35th International Telecommunications Energy Conference 'Smart Power and Efficiency' (INTELEC)*, p. 1–6, 2013. Hamburg. Citado na página 21.
- HOXHA, J.; BRAHAJ, A. Open government data on the web: A semantic approach. *Emerging Intelligent Data and Web Technologies (EIDWT), 2011 International Conference on*, p. 107–113, 2011. Citado na página 35.
- LIKERT, R. *A technique for the measurement of attitudes*. 140. ed. [S.l.]: The Science Press, 1932. Citado na página 64.
- MACHADO, A. L.; OLIVEIRA, J. M. Parente de. Digo: An open data architecture for e-government. *15th IEEE International Enterprise Distributed Object Computing Conference Workshops (EDOCW)*, p. 448–456, 2011. Citado na página 37.
- MAHIZHNAN, A. Smart cities the singapore case. *The International Journal of Urban Policy and Planning*, p. 13–18, 1999. Volume 16 Issue 1. Citado na página 32.
- MAHIZHNAN, A. Southampton goes live with smart city services. *Card Technology Today*, p. 06, 2002. Elsevier. Citado na página 32.
- MARCA, R. B. d. *Smarter Cities... and Wiser Ones?* 2014. <<http://theinstitute.ieee.org/opinions/presidents-column/smarter-cities-and-wiser-ones>> Acessado em Jun/2016. Citado na página 33.
- MCCABE, T. J. A complexity measure. *IEEE Transactions on Software Engineering archive Volume 2 Issue 4, July 1976*, p. 308–320, 1976. Citado na página 74.
- MCMILLAN, D. et al. Data and the city. *Proceeding CHI '16 Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, p. 2933–2944, 2016. ACM New York, NY, USA ©2016. Citado na página 37.
- MOREIRA, D. L. J.; MALIN, A. M. B. Panorama sobre a utilização de dados governamentais abertos no Brasil: Um estudo a partir dos aplicativos desenvolvidos. *XVI Encontro Nacional de Pesquisa em Ciência da Informação (XVI ENANCIB)*, 2015. ISSN 2177-3688. Citado na página 30.
- NEWCOMBE, T. *Santander: The Smartest Smart City*. 2014. <<http://www.governing.com/topics/urban/gov-santander-spain-smart-city.html>>. Acessado em Jun/2016. Citado na página 34.
- NIELSEN, J. *Why You Only Need to Test with 5 Users*. 2000. <<https://www.nngroup.com/articles/why-you-only-need-to-test-with-5-users/>> Acessado em Jun/2016. Citado na página 55.
- NIELSEN, J. *How Many Test Users in a Usability Study?* 2012. <<https://www.nngroup.com/articles/how-many-test-users/>> Acessado em Jun/2016. Citado na página 58.
- OPENGOVDATA. *The Annotated 8 Principles of Open Government Data*. 2007. <<https://opengovdata.org/>> Acessado em Jun/2016. Citado na página 27.
- OPENGOVDATA. *Open Government Data Principles*. 2007. <https://public.resource.org/8_principles.html> Acessado em Jun/2016. Citado na página 27.

- PARK, T.; VANROEKEL, S. *Project Open Data*. 2013. <<https://www.whitehouse.gov/blog/2013/05/16/introducing-project-open-data>> Acessado em Jun/2016. Citado na página 31.
- PORTELA, C.; SLOOTWEG, H.; EEKELEN, M. V. Structuring the design space of the ict architecture for the smart grid. *IEEE Recent Advances in Intelligent Computational Systems (RAICS)*, p. 110–115, 2013. Trivandrum. Citado na página 21.
- PRETZ, K. *Guadalajara: Smart City of the Near Future*. 2014. <<http://theinstitute.ieee.org/technology-focus/technology-topic/guadalajara-smart-city-of-the-near-future>> Acessado em Jun/2016. Citado na página 34.
- PRETZ, K. *An Urban Reality: Smart Cities*. 2014. <<http://theinstitute.ieee.org/technology-focus/technology-topic/an-urban-reality-smart-cities>>. Acessado em Jun/2016. Citado 2 vezes nas páginas 21 e 33.
- ROCHE, S. et al. Are ‘smart cities’ smart enough? *Global Geospatial Conference*, 2012. Citado na página 34.
- SALAS, P. et al. Stdtrip: Promoting the reuse of standard vocabularies in open government data. *Linking Government Data*, p. 113–133, 2011. Citado na página 34.
- SAYOGO, D. S.; PARDO, T. A.; COOK, M. A framework for benchmarking open government data efforts. *2014 47th Hawaii International Conference on System Sciences (HICSS)*, p. 1896–1905, 2014. Citado na página 35.
- SUAKANTO, S. et al. Smart city dashboard for integrating various data of sensor networks. *IEEE International Conference on ICT for Smart Society (ICISS)*, p. 1–5, 2013. Jakarta. Citado 2 vezes nas páginas 21 e 33.
- SUAKANTO, T. N.; THERESA, A. P. Smart city as urban innovation: focusing on management, policy, and context. *ICEGOV '11 Proceedings of the 5th International Conference on Theory and Practice of Electronic Governance*, p. 185–194, 2011. Citado na página 21.
- TCHOLTCHIEV, N. et al. On the interplay of open data, cloud services and network providers towards electric mobility in smart cities. *LCN Workshops*, p. 860–867, 2012. Citado na página 37.
- TCU. *Cinco Motivos para a Abertura de Dados na Administração Pública*. 2015. <<portal.tcu.gov.br/biblioteca-digital/cinco-motivos-para-a-abertura-de-dados-na-administracao-publica.htm>> Acessado em Jun/2016. Citado na página 29.
- TRAVASSOS, G. H.; GUROV, D.; AMARAL, E. A. G. Introdução a engenharia de software experimental. *Relatório Técnico RT-ES-590/02*, 2002. Programa de Engenharia de Sistemas de Computação, COPPE, UFRJ. Citado 2 vezes nas páginas 55 e 57.
- USECHE, M. P.; SILVA, J. C. N.; VILAFANE, C. Medellin colombia: a case of smart city. *13 Proceedings of the 7th International Conference on Theory and Practice of Electronic Governance (ICEGOV)*, p. 231–233, 2013. ACM New York, NY, USA ©2013. Citado na página 33.

WINTER, J. d. Using the student's t-test with extremely small sample sizes. *Practical Assessment, Research & Evaluation - Volume 18, Number 10*, p. 1–12, 2013. ISSN 1531-7714. Citado na página 57.

WOHLIN, C. Guidelines for snowballing in systematic literature studies and a replication in software engineering. *EASE '14 Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering*, p. Article No. 38, 2014. ACM New York, NY, USA ©2014. Citado na página 55.

WOHLIN, C. e. a. Experimentation in software engineering: an introduction. *Kluwer Academic Publishers Norwell, MA, USA ©2000*, 2000. ISBN:0-7923-8682-5. Citado 2 vezes nas páginas 55 e 76.

APÊNDICE A – Materiais utilizados nas aulas.

Protocolos

- *SOAP (Simple Object Access Protocol)*
– *Protocolo de comunicação baseado em XML*
- *WSDL (Web Services Definition Language)*
– *Descreve o Web Service e suas operações e métodos, baseado em XML*
- *Protocolo rígido, com estruturação definida*

Figura 27 – Material utilizado na aula sobre *web services*.

Dados Abertos Governamentais

- Problemas
 - Publicação Descentralizada
 - Inúmeros portais, municipais, estaduais e federais
 - Câmara de Vereadores, dos Deputados, Senado
 - Inúmeros formatos de dados
 - Não existem leis relacionadas a aspectos técnicos de publicação, apenas diretrizes

Figura 28 – Material utilizado na aula sobre DAG.

Algoritmo A.1 – Aula 1, exercício servidor *web service*

```
import javax.jws.WebService;
import javax.jws.WebMethod;
import javax.jws.WebParam;

@WebService()
public class CalculadoraWS {
    @WebMethod(operationName = "soma")

    //Recebimento dos parametros web
    public int soma(@WebParam(name = "n1") int n1,
        @WebParam(name = "n2") int n2) {

        //Calculo dos parametros
        return n1 + n2;
    }
}
```

Algoritmo A.2 – Aula 1, exercício cliente *web service*

```
public static void main(String [] args) {

    //Instancia do Servidor
    org.br.CalculadoraWS_Service service = new
        org.br.CalculadoraWS_Service();
    org.br.CalculadoraWS ws =
        service.getCalculadoraWSPort();

    try {
        //Chamada ao Servidor
        int resultado = ws.soma(2,3);

        //Exibicao do resultado na tela
        System.out.println(“Soma WS = ” + resultado);
    } catch (Exception ex){
        System.out.println(“Exception: ” + ex);
    }
}
```

Algoritmo A.3 – Aula 3, exercício de acesso prático à plataforma

```
package clientetransparenciawsconsole;

public class ClienteTransparenciaWSConsole {

    public static void main(String[] args) {
        System.out.println("Lista de Registros em XML" +
            getListaDespesa
                ("Campinas", "2014", "02", "", "", "", "", ""));
    }

    private static String getListaDespesa
        (java.lang.String wNomeCidade, java.lang.String wAno,
         java.lang.String wMes, java.lang.String wDominio,
         java.lang.String wSubDominio,
         java.lang.String wNatureza,
         java.lang.String wFonte,
         java.lang.String wTipoLicitacao) {

        org.tempuri.TransparenciaWS service = new
            org.tempuri.TransparenciaWS();

        org.tempuri.TransparenciaWSSoap port =
            service.getTransparenciaWSSoap();

        return port.getListaDespesa(wNomeCidade, wAno, wMes,
            wDominio, wSubDominio, wNatureza,
            wFonte, wTipoLicitacao);
    }
}
```

Api de Acesso aos dados de Transparência.

O Endereço do WSDL do Web Service encontra-se hospedado em:

www.transparenciaws.elasticbeanstalk.com/TransparenciaWS.asmx?WSDL

O Web Service contém atualmente 2 métodos genéricos:

getTotalDespesa – Devolve em formato **string** o valor total somado das Despesas, baseado nos parâmetros opcionais da consulta.

getListaDespesa – Devolve em formato **string** uma lista de registros em XML, baseado nos parâmetros opcionais da consulta.

Ambas as consultas possuem os mesmos parâmetros, todos eles do tipo string, iniciando com a letra w para indicar parâmetro web:

(string wNomeCidade, string wAno, string wMes, string wDominio, string wSubDominio, string wNatureza, string wFonte, string wTipoLicitacao, string wInicioRegistros, string wQuantidadeRegistros)

wInicioRegistros: Indica em qual registro a lista vai começar a ser exibida

wQuantidadeRegistros: Indica quantos registros formarão o quadro de registros

Todos os parâmetros são opcionais, podendo ser combinados entre eles.

Caso não se deseje utilizar um parâmetro, pode-se deixar o mesmo em branco "" ou passar um valor **null**.

Os métodos filtrarão somente os parâmetros que forem utilizados.

Exemplos de consulta:

// Retorna uma lista com todos os registros da base

```
System.out.println("Lista em XML: " + getListaDespesa("", "", "", "", "", "", "", ""));
```

//Retorna a soma de todas as despesas da base

```
System.out.println("Valor Total das despesas: " + getTotalDespesa("", "", "", "", "", "", ""));
```

//Retorna uma lista com todos os registros da Cidade de Campinas, no ano de 2014 e no mês 03

```
System.out.println("Lista em XML: " + getListaDespesa("Campinas", "2014", "03", "", "", "", "", ""));
```

// Retorna a soma das despesas da Cidade de Campinas, no ano de 2014, no mês 07, relacionadas ao Domínio 12 – EDUCACAO, Sub Domínio 365 - EDUCACAO INFANTIL e Natureza 44905100 - OBRAS E INSTALACOES

```
System.out.println("Valor Total das despesas: " + getTotalDespesa("Campinas", "2014", "07", "12 - EDUCACAO", "365 - EDUCACAO INFANTIL", "44905100 - OBRAS E INSTALACOES", "", "", ""));
```

//Retorna uma lista com todos os registros relacionados ao Tipo de Licitação Contrato de Locação

```
System.out.println("Lista em XML: " + getListaDespesa("", "", "", "", "", "", "Contrato de Locação", ""));
```

APÊNDICE B – Questionário de perfil do participante

Questionário de Perfil do Participante

Este formulário tem como objetivo coletar informações sobre os participantes das atividades a serem desenvolvidas durante o semestre

***Obrigatório**

Registro acadêmico *

Nome *

Informe sua faixa etária: *

Menos de 20 anos
 Entre 20 e 25 anos
 Entre 26 e 30 anos
 Acima de 30 anos

Você atua ou já atuou no mercado de trabalho de T.I. como Analista e/ou Desenvolvedor Web? *

Nunca atuei profissionalmente no mercado de trabalho de T.I. como Analista e/ou Desenvolvedor Web
 Sim, por menos de 1 ano
 Sim, entre 1 e 2 anos
 Sim, entre 2 e 5 anos
 Sim, por mais de 5 anos

Qual das opções abaixo melhor descreve a sua atuação no momento *

Estudante de graduação apenas
 Estudante de pós-graduação/mestrado/doutorado apenas
 Estudante de graduação e profissional de T.I.
 Estudante de pós-graduação/mestrado/doutorado e profissional de T.I.
 Profissional de T.I. apenas

Aponte o seu nível de experiência nas tecnologias abaixo *

	Não possuo conhecimento	Possuo conhecimento acadêmico teórico apenas	Possuo conhecimento acadêmico prático apenas	Possuo conhecimento acadêmico teórico e prático	Já apliquei o conhecimento em projetos no mercado de trabalho
Java	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
XHTML/HTML	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
CSS	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
JS(jQuery)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
JSP/Servlet	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
JDBC	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
XML	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Web Services	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Figura 30 – Questionário de perfil do participante aplicado aos alunos.

ANEXO A – *Status reports*



Atividades realizadas

- Eduardo Romão
Não realizou nenhum desenvolvimento
0h de desenvolvimento - 1h de gerenciamento
- Hélio Nakayama
Não realizou nenhum desenvolvimento
0h de desenvolvimento - 1h de gerenciamento
- Mateus Craveiro
Desenvolveu a exibição de resultados para uma consulta simples
2h de desenvolvimento - 1h de gerenciamento
- Thiago Wittmann
Finalizou a criação das funções de chamada da API
4h de desenvolvimento - 1h de gerenciamento

Figura 31 – *Status report* grupo experimental.

Tarefas	Planejado	Real	Tempo
Salvar gráficos	Concluir Tarefa	Tarefa Concluída	7 h
Exibir Grid	Concluir Tarefa	Tarefa Concluída	3h
Validar site	Concluir Tarefa	Tarefa Concluída	2h

Principais problemas

Problemas e variações	Comentários
Encontrar pessoas para validar o site	Procuramos alunos e professores na UFSCar.

Figura 32 – *Status report* grupo controle.