

Desafios Sociotécnicos e Prospecções para Promover Transparência de Dados na Universidade

Breno Bernard Nicolau de França¹, Julio Cesar dos Reis¹, Rodolfo Jardim de Azevedo¹

¹Instituto de Computação – Universidade Estadual de Campinas (UNICAMP)
Av. Albert Einstein, 1251, Cidade Universitária, CEP 13083-852 – Campinas – SP – Brasil

{breno, jreis, rodolfo}@ic.unicamp.br

Abstract. *Data transparency is a key factor for the development of several sectors in society. Transparency not only boosts the understanding regarding what is being developed at public universities, but also shows how public resources have been invested. In this paper, we present results from a study on transparency, based on data collected from two initiatives for disseminating the transparency culture across the UNICAMP. We identified sociotechnical challenges, as well as described an architectural solution to improve transparency processes within the university to promote unrestricted access to public data.*

Resumo. *Transparência de dados é um aspecto chave para o desenvolvimento de diversos setores da sociedade. Em universidades públicas, a transparência potencializa o conhecimento sobre o que é desenvolvido permitindo o entendimento de onde os recursos são investidos. Neste artigo, apresentamos os resultados de um estudo sobre transparência, com base em dados coletados de duas iniciativas para difundir a cultura da transparência na UNICAMP. Identificamos desafios sociotécnicos, bem como apontamos uma solução arquitetural para facilitar os processos associados à transparência dentro da universidade e promover acesso irrestrito e facilitado aos dados públicos.*

1. Introdução

Acesso à informação para todos é essencial para o exercício da cidadania no século XXI. Pessoas necessitam de meios adequados para acessarem dados públicos visando obter informações claras sobre a eficiência das instituições. Isso se torna ainda mais relevante no contexto de instituições públicas, sobre as quais o cidadão demanda esclarecimentos sobre como os impostos recolhidos são aplicados. A transparência de dados desenvolve então um papel chave na garantia de governança pública adequada e combate à corrupção.

Dados governamentais abertos permitem melhor cooperação entre instituições públicas e o cidadão em geral. A abertura de dados não apenas beneficia a democracia por meio da transparência, como também permite melhor eficiência econômica com aprimoramento da gestão pública e disseminação de resultados [Mutuku and Colaco 2012]. Apesar dos benefícios teoricamente esperados, os desafios para a transparência demandam tanto investigações quanto mudanças organizacionais e tecnológicas para garantir acesso igualitário e público às informações governamentais [Jaeger and Bertot 2010].

A literatura tem apresentado diversas iniciativas para abertura de dados governamentais [Attard et al. 2015]. Mais especificamente, o uso de Tecnologias da Informação e Comunicação (TICs) tem indicado benefícios para promover transparência em dados e

umentar a clareza de ações governamentais [Bertot et al. 2010]. Particularmente, Bertot *et al.* [Bertot et al. 2010] mostraram que TICs desenvolvem um impacto relevante para criar mudanças sociais nas atitudes em relação à transparência. Nesse contexto, Böhm *et al.* [Böhm et al. 2012] apresentaram o *GovWILD*, um sistema web para integrar dados governamentais abertos, promovendo perguntas e respostas sobre políticos, campanhas e recursos governamentais.

Em particular, no contexto de universidades, Keßler & Kauppinen [Kessler and Kauppinen 2015] apresentaram uma infraestrutura para a publicação de dados abertos da universidade de Münster, na Alemanha. Eles indicaram que tal infraestrutura desenvolve um papel central para a visibilidade e acessibilidade dos dados produzidos e coletados pela universidade. Diversos projetos tem evidenciado a relevância da transparência e acesso a dados não apenas para gestão pública, mas igualmente para o desenvolvimento da ciência no geral, como a proposta de ciência aberta na medicina que visa a publicação e reuso de dados científicos [Krumholz et al. 2013].

Existem muitos dados de interesse público nos mais diferentes segmentos da sociedade, incluindo universidades públicas. No entanto, partes interessadas usualmente enfrentam dificuldade em encontrar os dados de interesse. Frequentemente, cidadãos buscam responder questões que demandam acesso e integração em fontes fragmentadas de informação. Esse problema é refletido pela utilização de diversos sistemas de informação isolados, que dificilmente interagem e permitem troca de dados de alguma natureza. O fato de não haver uma arquitetura para integração das informações, e que facilite o acesso ao usuário final com refinadas técnicas de consulta, cria barreiras na obtenção de respostas às necessidades de informação dos cidadãos.

Neste artigo, conduzimos um estudo sobre uma iniciativa de transparência de dados na Universidade Estadual de Campinas (UNICAMP). Analisamos os resultados obtidos em uma disciplina ofertada no contexto de graduação sobre transparência de dados e um evento de desenvolvimento de *software* promovido para obter ideias de soluções tecnológicas inovadoras visando transparência de dados na universidade.

Nossos resultados indicam desafios do ponto de vista organizacional e técnico. Dentre os desafios identificados, ressaltamos a falta de cultura sobre transparência tanto de usuários (no caso alunos) quanto de servidores públicos responsáveis por criar, organizar e disseminar os dados. Adicionalmente, o estudo realçou que transparência de dados requer inovações tecnológicas para disponibilizar efetivamente dados para um grande público-alvo. Assim, com base na análise, prospectamos uma arquitetura de *software* como uma infraestrutura para apoiar a disponibilização e acesso aos dados.

O restante deste artigo está organizado da seguinte maneira: a seção 2 apresenta a metodologia de pesquisa com os passos percorridos para obtenção e análise dos dados. Enquanto a seção 3 relata os desafios e a proposta de arquitetura, a seção 4 discute os resultados obtidos. A seção 5 apresenta as considerações finais.

2. Metodologia

2.1. Materiais e Participantes

Este estudo se fundamenta em duas iniciativas desenvolvidas na universidade para promover a cultura de transparência de dados.

Curso de transparência de dados. Uma disciplina motivada por alunos foi ofertada com o objetivo de discutir e entender aspectos, problemas e potenciais soluções sobre transparência na UNICAMP. Participaram, durante um semestre letivo, 51 estudantes de diversos cursos em nível da graduação. Os alunos foram então convidados a refletirem sobre questões de interesse em relação à universidade, dentro dos mais diversos aspectos.

Então, o curso foi organizado com base em três projetos desenvolvidos em grupos ao longo do semestre com os seguintes objetivos e atividades:

- **Projeto 1 (Aquecimento):** Definição de uma questão sobre transparência na universidade e tentativa inicial de respondê-la. O entregável desse projeto deveria ser no formato de um infográfico contendo a(s) pergunta(s) e resposta(s).
- **Projeto 2 (Investigar um tema):** Com base na sugestão de um conjunto de temas sobre diversos aspectos na universidade (*e.g.*, vivência no campus, orçamento, anuário estatístico, vestibular, comunidade, *etc.*), grupos definiram problemáticas que gostariam de investigar. Assim, os grupos deveriam relatar os passos seguidos para obtenção e entendimento dos dados, indicando os problemas encontrados.
- **Projeto 3 (Aprofundamento no tema):** Nesta fase, os grupos se aprofundaram em um aspecto particular do tema. Similar ao Projeto 2, os alunos relataram os meios e dificuldades em se obter e entender os dados sobre a universidade.

Todos os projetos seguiram uma lógica semelhante de entregas, sendo uma definição de problemas a serem investigados (por meio de questões ou temas pré-definidos), a coleta e disponibilização dos dados obtidos para responder às questões, e descrição do caminho (denominado em sala como “Odisseia”) utilizado para se obter às respostas, incluindo dificuldades/desafios. Particularmente, a “Odisseia” relatada pelos alunos foi utilizada para identificar os desafios descritos na Seção 3.1. Como última atividade, alunos foram convidados a criar um vídeo de aproximadamente 5 minutos apresentando os resultados descobertos durante o semestre.

Hackathon de transparência de dados. Neste evento, parcialmente organizado por alunos da UNICAMP, objetivou-se a proposição de soluções de *software* inovadoras com foco em transparência e dados abertos no contexto da universidade; no sentido de soluções para facilitar acesso, visualização e análise desses dados. O evento durou 36 horas e foi constituído de 13 times de alunos/desenvolvedores de *software* de diferentes instituições. Para contextualizar os alunos, diversas palestras com aspectos organizacionais e tecnológicos foram ministradas (*e.g.*, *Big Data* na Indústria, *Linked Data* na Prática, Transparência no Governo). A proposta dos grupos foram avaliadas por um comissão julgadora que deliberou 3 finalistas. Seção 3.2 apresenta os resultados e lições apreendidas.

2.2. Análise de dados

O seguinte procedimento foi conduzido para analisarmos qualitativamente os materiais coletados nos eventos descritos.

1. Exame detalhado do material obtido no curso (resultado do desenvolvimento de trabalhos dos alunos) para identificar problemas recorrentes de acesso aos dados da universidade e classificá-los;
2. Detecção e análise dos problemas endereçados e soluções propostas no *Hackathon* pelas equipes de trabalho;

O entendimento dos problemas em diversas perspectivas informou elementos relevantes para a concepção de componentes essenciais de uma arquitetura de software que permita melhor integração entre *software* existentes na universidade de onde os dados para transparência são providos.

3. Resultados

Apresentamos os desafios identificados no entendimento de transparência de dados na universidade dentro do contexto da disciplina ofertada (Seção 3.1), seguido das proposições obtidas e lições apreendidas no *Hackathon* (Seção 3.2). Como base nesses resultados, concebemos uma arquitetura de *software* para permitir a disponibilização integrada de dados oriundos de diferentes sistemas da universidade (Seção 3.3).

3.1. Desafios Identificados

O primeiro projeto da disciplina consistiu em levantar questões de interesse dos alunos a serem respondidas pelos mesmos através da disponibilização de dados públicos pela universidade. Exemplos de questões levantadas pelos alunos foram:

- *Como que chega o dinheiro na universidade?*
- *Quanto é o ganho e o gasto nos vestibulares?*
- *Quanto é o gasto mensal com bolsa para alunos?*
- *Qual a média de contratação de professores por ano na universidade?*
- *Quantos professores aposentados tem a universidade?*
- *Qual instituto recebe o maior repasse da universidade?*
- *Quantas pessoas almoçam nos restaurantes universitários diariamente?*
- *Qual o número médio de livros extraviados na universidade?*
- *Qual é o investimento mensal destinado para os ônibus internos?*
- *Onde são os pontos de Wi-Fi da universidade?*
- *Quantas vagas para deficientes físicos e idosos a universidade possui?*
- *Quanto é o gasto na segurança do Campus? Esse investimento quando aumentado, tem melhor resultado?*
- *Quantos alunos estrangeiros são matriculados em cursos da universidade?*
- *Qual o consumo médio de energia mensal da universidade?*
- *Quais são os patrimônios da universidade?*

Para os três projetos da disciplina, os empecilhos para chegar às informações foram capturados e classificados. Como eles foram descritos de maneira similar em todos os projetos e recaem sobre problemas comuns, sintetizamos os problemas nesta seção.

Considerando algumas questões e temas, sobretudo financeiros, as informações foram encontradas em menos de uma hora por meio de buscas *online* (via mecanismo de busca do *Google*), anuário estatístico da instituição, e sites institucionais. Nesses casos, mesmo que as informações pudessem ser completamente encontradas de maneira aberta, elas se encontravam em formatos distintos e de difícil manipulação (por exemplo, arquivos PDF e imagens), além de estarem frequentemente fragmentadas em diferentes fontes, necessitando de um esforço adicional para relacioná-las.

A grande maioria das questões e, portanto, os dados necessários para respondê-las precisaram de mais de três (chegando até seis) iterações, ou seja, idas e vindas em

diferentes setores da universidade, para se obter de fato os dados/informações adequados com a finalidade de alcançar as respostas corretas. Diante deste cenário, destacamos os seguintes desafios recorrentes:

Conteúdo de difícil entendimento sem conhecimento específico da área. Além do difícil acesso, os dados não são apresentados de forma adequada para o entendimento pela maior parte da população (leiga em conceitos específicos da universidade). Por exemplo, dados financeiros, questões orçamentárias, e políticas de distribuições de bolsas.

Temporalidade dos dados. Em qualquer caso dentre os observados, os dados refletiam apenas o passado e podiam demandar até 20 dias para obtê-los. Encontramos casos que as informações requisitadas em 2017, mas os alunos solicitantes receberam dados apenas referente ao ano de 2013.

Burocracia no acesso. A forma de acesso aos dados obtidos variou bastante dentro da mesma instituição. O acesso pode ser via informações estáticas abertamente disponíveis em portais da universidade, formulários de pedido de acesso a informação, *e-mails*, ofícios enviados por meio do Diretor de uma unidade, até presencialmente com reuniões agendadas. Informações sobre unidades, salários implementados na organização, puderam ser obtidos por meio de portais *Web*. Informações sobre questões acadêmicas como evasões, ingressantes em vestibular, informações socioeconômicas dos alunos, necessitaram de ofícios e foram recebidas em meios impressos.

Informações incompletas e/ou fragmentadas. As informações solicitadas por vezes estavam incompletas e, quando possível, o acesso à informação completa dependia do acesso à diferentes fontes, incluindo fontes externas à universidade, sendo umas *online* e outras obtidas por meio de ofício. Este foi o caso de informações solicitadas a respeito do patrimônio da universidade e investimento em projetos de pesquisa.

Indisponibilidade dos dados/informações. Algumas tentativas de se obter os dados solicitados foram respondidas, mas sem fornecer os dados/informações de fato. Nesses casos, foi alegado que o setor não possuía os dados ou que eles não estavam disponíveis e/ou permitidos para divulgação.

Conhecimento interno da organização para ter sucesso na solicitação. Existiram casos em que o solicitante precisou saber a qual setor interno da universidade ele deveria pedir os dados. Entretanto, o cidadão comum não tem esse conhecimento de como a universidade esta estruturada. É perfeitamente possível que uma solicitação realizada por um canal de propósito geral da universidade para acesso à informação tenha um resultado diferente da solicitação realizada por um canal específico de um setor interno. Isso evidencia limitações existentes na integração da universidade em relação aos dados e responsabilidades.

Contatos sem resposta. Algumas tentativas de fazer contato com os setores responsáveis simplesmente não foram respondidas. Esse foi um fator que dificultou o acesso à informação, mesmo que fosse necessário retirar pessoalmente algum documento.

Telefone de contato não alcançáveis. Em um caso específico, a única forma de contato encontrada com o setor (outra instituição pública detentora de dados da universidade) foi um número de telefone. Entretanto, o mesmo era bloqueado para ligações originadas de celulares.

3.2. Proposições e Lições Aprendidas no *Hackathon*

No geral, os grupos propuseram soluções diversificadas com objetivos específicos como detecção e visualização de anomalias na folha de pagamentos dos servidores, aplicação interativa com gráficos para apresentação de dados socioeconômicos e financeiros de cada instituto. Focamos na descrição dos objetivos e resultados dos três grupos finalistas avaliados pela comissão julgadora do evento. A motivação dos grupos foi, principalmente, o fato de haver dados de difícil acesso e análise, além de pouco intuitivos para usuários finais. De maneira geral, os grupos visaram disponibilizar os dados de maneira simples.

Visualização interativa de dados. Nesta proposta, criou-se uma solução de visualização interativa em um portal disponibilizando uma busca unificada para aprimorar a recuperação de documentos da universidade.

Chatbot. Esta proposta explorou o conceito de *chatbot*, como um assistente virtual de perguntas e respostas. O objetivo foi obter uma aplicação onde o usuário efetue perguntas sobre a universidade e o sistema de *chatbot* apresente resultados sobre a questão. Foi alcançado um protótipo de um *chatbot* implementado em uma rede social.

Centralização de dados. Esta proposta visou unificar os dados e disponibilizar seu acesso com técnicas de computação em nuvens através de uma única interface de programação. O objetivo foi permitir um aprimoramento na capacidade de cruzamento dos dados de naturezas diversas e facilitar a manipulação dos mesmos. No protótipo, um mapa interativo foi criado que disponibilizou relatórios de unidades da universidade com base nos dados centralizados.

O evento permitiu que os participantes refletissem sobre maneiras inovativas de disponibilizar os dados para um grande público-alvo. Tecnicamente, o evento permitiu aos alunos descobrirem o uso de novas tecnologias e linguagens de programação explorando o tema transparência de dados.

3.3. Proposta de Arquitetura de Software

Os resultados obtidos indicaram que a transparência requer o acesso e análise de dados armazenados em sistemas computacionais distintos. Esses dados estão dispersos em diversas fontes, são representados em modelos de dados distintos, com formatos heterogêneos, e em sistemas implementados com linguagens e por equipes distintas. Adicionalmente, os dados descrevem informação de diversos domínios (vida acadêmica de alunos, vestibular, orçamentário/financeiro, funcionários e docentes, entre outras), sendo por vezes desconexos e não pré-processados para permitir análise e entendimento dos mesmos. Assim, propomos uma arquitetura computacional para permitir o desenvolvimento organizado de uma infraestrutura para disponibilização e acesso aberto aos dados (*cf.* Figura 1).

O componente “*Backoffice*” é acessado pelas unidades da universidade, as quais disponibilizam os dados para acesso, por meio das interfaces desse componente (“*CarregarDocumento*, *CarregarImagem*, e *CarregarDados*”). Ele permite que ambos dados estruturados (*e.g.*, CSV, XML, JSON, SQL) e artefatos como documentos PDF e imagens sejam inseridos na plataforma, por vias distintas. É importante ressaltar que essas interfaces devem estar disponíveis de forma a serem acessadas por usuários e integradas com outros sistemas da universidade. O gerenciador de transformação de dados (“*GerenciadorTransformacoes*”) e o gerenciador de artefatos (“*GerenciadorArtefatos*”) são

responsáveis por essas tarefas. Ao carregar quaisquer informações ao “*Backoffice*” por meio de suas interfaces, metadados devem descrever a unidade da organização detentora dos dados (origem do dado), bem como os domínios nos quais os dados se aplicam (planejamento, financeiro, pesquisa, ensino, etc.).

A arquitetura prevê dois tipos de repositórios. Um para armazenar artefatos como documentos e imagens, por exemplo editais, regulamentações, fotografias e plantas de obras, e outro repositório focado na centralização e relação entre dados. O repositório de dados estruturados necessita considerar modelos compartilhados de dados e permitir uma descrição flexível dos dados para possibilitar análises e descoberta de relações entre os dados de diversos domínios. Base de dados orientada a grafos, ou ainda bancos de dados não relacionais de alto desempenho em consultas, podem acomodar esses requisitos e permitir consultas com alto desempenho e disponibilidade.

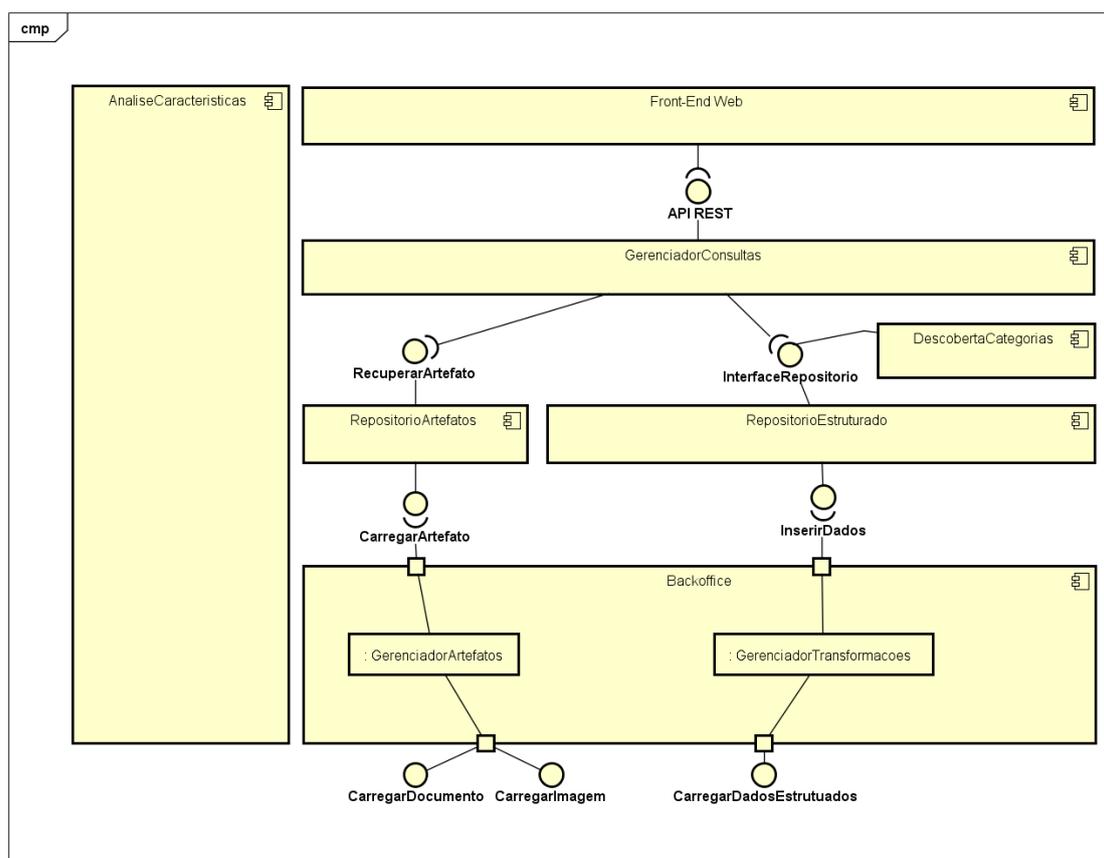
O componente “*GerenciadorConsultas*” é responsável por tratar consultas em ambos repositórios. Consultas podem ser expressas por meio de diferentes facetas e filtros, além de palavras-chaves e categorias de análise pré-determinadas. O componente de descoberta de categorias (“*DescobertaCategorias*”) visa extrair possíveis classes de consulta de interesse do usuário através de técnicas de classificação e agrupamento aplicados sobre os dados do repositório e de consultas submetidas pelos usuários.

O componente de “*Front-End Web*” visa implementar interfaces Web com usuários finais para expressar consultas a dados e múltiplas perspectivas de visualização desses dados (gráficos, tabelações, entre outras). Ainda, é relevante considerar o perfil de acesso e utilização dos usuários de forma a entender como eles utilizam e que tipos de informações são mais acessadas por determinados perfis. Para esse fim, o componente “*AnaliseCaracteristicas*” prevê capturar os dados mais acessados, as visualizações mais exploradas, os domínios de maior interesse dos usuários, de forma a direcionar a atenção para determinados conteúdos ou funcionalidades que necessitem de melhorias.

4. Discussão

O interesse em conhecer melhor a universidade, entendendo como os investimentos são realizados e qual retorno se obtém desses investimentos, perpassa as diferentes áreas da universidade, desde situação orçamentária e infraestrutura até elementos do dia-a-dia no campus. Esse interesse se dá em função do desconhecimento sobre o funcionamento interno da universidade e sua importância para a sociedade, até certo ponto justificáveis pelo tamanho e complexidade da instituição.

O problema caracterizado nos desafios para o acesso aos dados da universidade possui questões técnicas e sociais, do ponto de vista político-organizacional e da cultura da mesma. Do ponto de vista social, por exemplo, é interessante notar que certos dados são negados ou dados como indisponíveis por um setor em seu canal imediato com o cidadão, mas são disponibilizados se contatar as pessoas “certas” dentro desse setor. Isso evidencia ou a falta de conhecimento entre os funcionários de determinados setores sobre quais informações eles detém e quais deveriam estar disponíveis ao público, ou um receio sobre os riscos associados à disponibilização de determinadas informações ao público. Na prática, nota-se haver uma desconfiança, quando se solicita informações, de que alguém fará mal uso das informações ou que poderá prejudicar o indivíduo ou o setor responsável pela informação. Isso justifica, em parte, uma quantidade razoável de negativas ao tentar



powered by Astah

Figura 1. Proposta de Arquitetura

solicitar informações de caráter público, mas que o livre acesso é desconhecido pelas pessoas internas àquelas organizações.

Ainda explorando os aspectos não-técnicos, percebe-se que não há uma cultura de transparência, no sentido de estabelecer quais, por quem e como os dados devem ser disponibilizados à sociedade. Internamente nos órgãos da universidade, parece haver uma alta proteção sobre esses dados, muitas vezes sem fundamento, onde qualquer informação apenas deve ser disponibilizada mediante autorização direta de cargos da alta gestão da universidade como coordenadores, diretores, pró-reitores e reitores.

Percebemos, também, que a utilização de determinados jargões ou termos no ato da solicitação pode influenciar no mal entendimento sobre quais dados se deseja obter, ou ainda recair na situação do setor entender que aquele dado não está sob sua responsabilidade. Um exemplo é quando se pede informações de cunho financeiro a setores/unidades que não são responsáveis pela questão orçamentária da universidade. Nesses casos, as repostas tendem a redirecionar o solicitante a outros setores ou simplesmente responder negativamente à solicitação, mesmo sendo informações específicas daquele setor.

Não é razoável assumir que o cidadão necessite entender a estrutura interna da instituição pública para requerer dados. A transparência deve seguir no sentido inverso, fornecendo um canal de comunicação único com o cidadão. Esse tipo de serviço já é disponibilizado pela UNICAMP. Entretanto, não se pode afirmar que tal serviço funciona de

forma completamente transparente ou facilita a obtenção dos dados, uma vez que grande parte das informações necessitam ser pedidas via formulários.

Entendemos que as iniciativas de transparência ou dados abertos de instituições públicas deveriam, idealmente, ser pró-ativas no sentido de entender as demandas da sociedade e disponibilizar as informações sem a necessidade de solicitações via formulários, ofícios, e outros meios mais burocráticos. Entretanto, a universidade passa por um momento de entendimento desses aspectos e a transição no processo de transparência necessita, também, do apoio tecnológico.

Diferentes órgãos públicos fazem uso de portais de dados abertos, os quais funcionam como uma ferramenta de busca, e repositório de dados sobre diferentes áreas da esfera pública. Entretanto, a proposta desses portais assume que o cidadão conheça a estrutura interna do governo, entendendo por exemplo, quais dados estão associados a que setores. Este estudo detectou algo similar na universidade, onde é necessário conhecer posição geográfica dos institutos ou órgãos formais da instituição para se obter informação, por vezes de forma presencial.

Os aspectos técnicos endereçados na proposta da arquitetura supõem um esforço reduzido das unidades detentoras dos dados para disponibilizá-los, evitando um custo considerável em uma eventual integração sistêmica entre os diversos sistemas utilizados na universidade. Assim, a arquitetura prevê não somente meios para disponibilização dos dados, mas componentes que permitam o melhor entendimento das demandas do cidadão em relação aos dados públicos da universidade, por meio de componentes de análise de dados de uso e disponibilizados nos repositórios. O grande desafio da arquitetura é lidar e definir um modelo compartilhado de dados para permitir consultas mais complexas.

Identificar os problemas existentes e encaminhamentos de soluções de transparência para a universidade tem sido uma preocupação recorrente. Esforços como este estudo podem ajudar no diagnóstico de problemas, ainda que em situações particulares e com a limitação em termos de abrangência das soluções propostas. Alguns dos desafios identificados podem ser entendidos como de percepção geral, mas que nessa oportunidade ficam evidenciados, seja por meio dos relatos das “Odisséias” dos alunos ou pelas inúmeras solicitações formais (ofícios, formulários e *e-mails*) de acesso aos dados. Ainda, entendemos que iniciativas desse tipo disseminam a importância do tema, ajudam a criar uma cultura de transparência na universidade, e desenvolve um papel educacional para com os alunos no entendimento do tema e de aspectos relacionados à universidade.

Do ponto de vista dos alunos, conseguimos notar um maior engajamento e curiosidade à medida em que mais dados eram coletados pelos grupos e apresentados em sala de aula. Enquanto, no início, a maior parte dos alunos estava preocupada com questões orçamentárias, o foco foi sendo transferido para outras questões como resultados acadêmicos (evasão, alunos por curso, *etc.*), uso de recursos pelos hospitais, patrimônios da universidade, organização de setores, *etc.* O *Hackathon* trouxe outra dinâmica ao trabalho, agregando código ao processo e fomentando ligações entre os dados já coletados. Ao final do semestre, tanto a visão sobre a universidade quanto a capacidade de analisar os dados encontrados foram aprimoradas.

Este estudo evidenciou características chaves do problema, mas entendemos que não é possível utilizá-las como um diagnóstico completo do estado da universidade em

relação ao tema de transparência. Os casos enfrentados representam situações particulares, mas que se repetem em diferentes níveis e setores da universidade. Ainda, não temos como afirmar que outras universidades, ou ainda outros órgãos públicos do país enfrentam necessariamente os mesmos problemas em função da observação e coleta de dados localizada. Finalmente, a solução arquitetural encaminhada deve ser amplamente discutida com os setores de TICs da universidade para evoluí-la e entender desafios específicos.

5. Considerações Finais

Transparência de dados na universidade envolve tanto desafios organizacionais quanto tecnológicos com inovações que permitam desenvolver soluções que façam sentido aos usuários finais. Neste artigo, conduzimos um estudo para entender os desafios sociotécnicos para a criação de soluções visando transparência de dados no contexto universitário. Nosso estudo detectou diversas dificuldades para encontrar informações sobre a universidade. Propomos uma arquitetura como meio de organizar o desenvolvimento de *software* de uma plataforma de transparência na universidade. A arquitetura permitirá que dados de diferentes naturezas sejam integrados para a resposta de consultas e filtro de dados relevantes. Trabalhos futuros envolvem investigar meios de implementar e implantar a arquitetura, utilizando-a como peça chave para abertura de dados públicos.

Agradecimentos

Esta pesquisa tem apoio financeiro da Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) (Processo #2017/02325-5) e CNPq.

Referências

- Attard, J., Orlandi, F., Scerri, S., and Auer, S. (2015). A systematic review of open government data initiatives. *Government Information Quarterly*, 32(4):399 – 418.
- Bertot, J. C., Jaeger, P. T., and Grimes, J. M. (2010). Using icts to create a culture of transparency: E-government and social media as openness and anti-corruption tools for societies. *Government Information Quarterly*, 27(3):264 – 271.
- Böhm, C., Freitag, M., Heise, A., Lehmann, C., Mascher, A., Naumann, F., Ercegovic, V., Hernandez, M., Haase, P., and Schmidt, M. (2012). Govwild: Integrating open government data for transparency. In *Proceedings of the 21st International Conference on World Wide Web, WWW '12 Companion*, pages 321–324, New York, NY, USA.
- Jaeger, P. T. and Bertot, J. C. (2010). Transparency and technological change: Ensuring equal and sustained public access to government information. *Government Information Quarterly*, 27(4):371 – 376. Special Issue: Open/Transparent Government.
- Kessler, C. and Kauppinen, T. (2015). Linked open data university of münster. In Springer, Berlin, H., editor, *Infrastructure and Applications*. University of München, Germany.
- Krumholz, H., Ross, J., Gross, C., and et al. (2013). A historic moment for open science: The yale university open data access project and medtronic. *Annals of Internal Medicine*, 158(12):910–911.
- Mutuku, L. N. and Colaco, J. (2012). Increasing kenyan open data consumption: A design thinking approach. In *Proceedings of the 6th International Conference on Theory and Practice of Electronic Governance, ICEGOV '12*, pages 18–21, New York, NY, USA.