

UM MODELO CONCEITUAL PARA USO DE *BIG DATA* E *OPEN DATA* NAS *SMART CITIES*

VINICIUS BARRETO KLEIN

Mestre em Engenharia e Gestão do Conhecimento
Professor do IFC Brusque - Instituto Federal Catarinense Campus Brusque
vinibk@gmail.com / vinicius.klein@brusque.ifc.edu.br

JOSÉ LEOMAR TODESCO

Doutor em Engenharia de Produção
Professor da UFSC – Universidade Federal de Santa Catarina
tite@egc.ufsc.br

RESUMO

Objetivo: Vivemos um contexto onde a sociedade produz constantemente um alto volume de dados, em distintas fontes, formatos e esquemas. Este fato corresponde ao fenômeno *big data*. Contribuindo com este fenômeno, o movimento Dados Abertos também adiciona seus dados a este volume. Os dados *big data* e os dados abertos podem servir de insumo para a geração de conhecimento e as *smart cities* (cidades inteligentes) podem se beneficiar deste processo. Neste contexto, o objetivo desta pesquisa é propor um Modelo Conceitual para identificar e classificar as principais fontes, mecanismos e atores das *smart cities*, visando apoiar pesquisas e a produção de conhecimento para estas cidades.

Design/Metodologia/Abordagem: O artigo apresenta uma pesquisa tecnológica, que produz como resultado final um artefato funcional. Também foi realizada como procedimento uma pesquisa exploratória e aplicado os métodos CESM e OntoKEM.

Resultados: Como resultado o modelo conceitual identificou os principais atores, fontes de dados e necessidades de informação das *smart cities*. As principais fontes de dados foram ainda classificadas e suas características apontadas. Foi também construído um conjunto de artefatos que auxiliam o uso do modelo proposto.

Originalidade/valor: Sua originalidade é identificada no fato de auxiliar a tarefa de identificação de fontes de dados para *smart cities*, passo inicial para a geração de conhecimento para estas cidades. Também se destacam os métodos utilizados para construção do modelo conceitual: Ontokem e CESM. Seu principal valor está em apoiar pesquisadores e projetos que visem utilizar estas fontes de dados para gerar conhecimento para eventuais cidades.

Palavras-chave: *Big data. Open data. Smart cities. Modelo conceitual.*

A CONCEPTUAL MODEL TO USE BIG DATA AND OPEN DATA FOR SMART CITIES

ABSTRACT

Purpose: We live in a context where society constantly produces a high volume of data, in different sources, formats and schemes. This fact corresponds to the big data phenomenon. Contributing to this phenomenon, the Open Data movement also adds new data sources. Big data and open data can serve as input for knowledge generation and smart cities can benefit from this process. In this context, the objective of this research is to propose a Conceptual Model to identify the main sources, mechanisms and actors of the smart cities, in order to support research projects and knowledge generation for these cities.

Design/Methodology/Approach: The article presents a technological research, which produces as a final result a functional artifact. An exploratory research was also carried out as a procedure and the CESM and OntoKEM methods were applied.

Results: The conceptual model identified the main actors, data sources and information needs of smart cities. The main data sources were still classified and their main characteristics pointed out. A set of artifacts was also built to help the use of the proposed model.

Originality/value: Its originality is identified in helping to find and to know the data sources for smart cities, which is an initial step in the generation of knowledge process. Also as originality can be pointed out the methods used to construct the conceptual model: Ontokem and CESM. Its main value is in supporting researchers and projects that aim to use these data sources to generate knowledge for smart cities.

Keywords: Big data. Open data. Smart cities. Conceptual model.

I INTRODUÇÃO

Smart cities são cidades que utilizam TICs (tecnologias da informação e comunicação) e ações governamentais para melhor gerenciar seus principais recursos e serviços, priorizando a qualidade de vida e o desenvolvimento econômico e sustentável (GIFFINGER et al., 2007; VILAJOSANA et al., 2013; NEIROTTI et al., 2014). O adjetivo *smart* caracteriza iniciativas de tornar as cidades atuais mais competitivas economicamente, humanas e sustentáveis, aptas a resolver diversos problemas urbanos atuais, como crises econômicas, trânsito, alta poluição dentre outros (GIFFINGER et al., 2007, NAM; PARDO, 2011). Para resolver estes problemas, este conceito incentiva o uso de TICs e ações governamentais (CARAGLIU; BO; NIJKAMP, 2011). Este processo produz um alto volume de dados nestas cidades, e o estudo do fenômeno *big data* pode auxiliar o gerenciamento destes dados.

Beyer e Laney (2012) afirmam que *big data* exige novas formas de tratamento tecnológico. Isso se deve principalmente a seus altos volumes, velocidades de produção, e por possuir diversos formatos e modelos distintos (DAVENPORT, 2014; O'REILLY, 2012; BEYER; LANEY, 2012; FAN; BIFET, 2012). Segundo Fan e Bifet (2012), este fenômeno representará uma fronteira a ser vencida por pesquisas científicas e aplicações de negócio. Sua mineração possui grande potencial para a descoberta de novos conhecimentos e geração de *insights*. Esta criação de novos conhecimentos pode auxiliar as *smart cities* em seu objetivo de melhorar seus processos. Além disto, as bases de dados *big data* destas cidades podem ser utilizados por diversos setores da sociedade. Por exemplo, o meio empresarial pode gerar inovação e desenvolvimento econômico, o meio acadêmico pode produzir novos conhecimentos a partir destes dados, dentre outras aplicações. Sendo assim, o objetivo desta pesquisa é aplicar conceitos da área de qualidade de dados, engenharia do conhecimento e construção de modelos conceituais para propor um modelo que identifica as principais fontes de dados para as *smart cities* e descrever suas características, bem como modelar seus principais atores, mecanismos e necessidades de informação. Com o uso deste modelo, espera-se apoiar projetos que usem estas fontes de dados para gerar conhecimento, propor soluções e estudos para *smart cities*.

2 FUNDAMENTAÇÃO: SMART CITIES, BIG DATA E OPEN DATA

Primeiramente são apresentados os conceitos de *big data*, *smart cities*, *open data*, e modelos conceituais, conforme segue.

2.1 DEFININDO SMART CITIES

O termo *smart cities* tem estado em evidência nos últimos anos. Segundo Townsend (2013, tradução nossa), *smart cities* são cidades que conseguem capturar dados, transformar em informação e adaptar suas ações em tempo real. Caragliu, Bo e Nijkamp (2011) descrevem uma *smart city* como uma cidade onde se investe em capital humano e social, utilizando-se de infraestrutura de tecnologia de informação para promover o crescimento econômico sustentável e a qualidade de vida. Destacando o papel do conhecimento nestas cidades, Zygiaris (2012) define uma *smart city* como sendo uma cidade que possui uma infraestrutura urbana voltada à proteção ambiental e capacidade de se basear no conhecimento e no capital humano e criativo para gerar inovação, aumentando assim sua competitividade econômica.

É comum encontrar as *smart cities* definidas em domínios de atuação considerados chaves para seu sucesso. Neirotti et al. (2014), corroborando os conceitos de Giffinger et al. (2007), definem uma *smart city* como cidades que possuem alto desempenho e ações em seis domínios de atuação: *Smart Environment* (Meio Ambiente), *Smart Mobility* (Mobilidade), *Smart People* (Capital Humano e Social), *Smart Economy* (Economia), *Smart Government* (Governo Transparente) e *Smart Living* (Qualidade de Vida). Estes ainda podem ser agrupados em duas categorias: *hard* e *soft*, onde a primeira categoria aponta os domínios onde o uso de TICs e *big data* tem maior impacto, como no gerenciamento de recursos naturais e no controle de trânsito por exemplo. O segundo (*soft*), agrupa os domínios que dependem mais de ações governamentais do que de tecnologias diretamente, como empreendedorismo e capital humano.

2.2 BIG DATA: CONCEITO

Segundo relatórios da IBM (2013), 90% dos dados existentes no mundo atual foram criados nos últimos dois anos. Em 2020, a tendência é que aproximadamente 26 bilhões de aparelhos sejam conectados a IoT (internet das coisas, do inglês, *internet of things*) (RIVERA; MEULEN, 2013), produzindo cada vez mais dados. Este alto crescimento na quantidade de dados produzida pela nossa sociedade é associado ao fenômeno conhecido como *big data*. Este tema tem sido comumente relacionado a volumes de dados extremamente grandes e heterogêneos, atributos que dificultam seu processamento e análise (CHEN; MAO; LIU, 2014). Fan e Bifet (2012 tradução nossa), utilizando os conceitos de Laney (2001), explicam *big data* em três atributos, com chamado modelo 3Vs:

- *Volume*: há hoje mais dados do que jamais houve (quantidade), porém, não há tantas ferramentas que possam processá-los;

- *Variety*: existem muitos tipos diferentes de dados, como textos, dados de sensores, áudios, vídeos, imagens etc;
- *Velocity*: os dados chegam em fluxos contínuos e é necessário analisá-los em tempo real.

Fan e Bifet (2012, tradução nossa), afirmam que hoje em dia também são consideradas duas dimensões importantes em *big data*:

- *Variability*: corresponde aos diferentes esquemas presentes em *big data* e como estes são interpretados;
- *Value*: valor de negócio, que possibilita às organizações tomarem decisões que antes (de *big data*) não era possível.

Ainda definindo este conceito, alguns autores adicionam novas características, como *veracity*, que diz respeito ao fato de que estes dados devem ser confiáveis para que as organizações possam tomar suas decisões (IBM, 2013):

- *Veracidade*: uma porcentagem considerável em grandes volumes de dados pode não ser confiável para a tomada de decisão (ex.: *posts* em redes sociais) (IBM, 2013).

Como causas deste fenômeno pode-se apontar o avanço do uso de TICs em diversos setores da sociedade (GROBELNIK, 2012). O uso de *cloud computing* e da web 2.0, assim como a alta inserção de *smartphones*, sensores e outros dispositivos produtores de dados na rotina das pessoas, são em termos gerais, características do avanço tecnológico que vivemos nos dias atuais.

2.3 OPEN DATA

Os dados do fenômeno *big data* produzidos pelas *smart cities* precisam ter formatos e licenças abertas para que seu uso pela sociedade seja potencializado. Nesta lógica, governos de diversos países, como dos EUA (U.S. GOVERNMENT, 2015), do Reino Unido (REINO UNIDO, 2015), Brasil (BRASIL, 2015) e diversos outros, além de setores da iniciativa privada e do meio acadêmico tem investido na abordagem *Open Data* (dados abertos) (MANYKA et al., 2011). Conforme relatório da empresa McKinsey, existem atualmente mais de 1 milhão de *datasets* (bases de dados) abertas (*open data*) em governos no mundo inteiro (levantamento no ano de 2013), e que representam uma estimativa de geração de mais de 3 trilhões de dólares na economia mundial (MANYKA et al., 2011). Este montante estimado se baseia no uso de dados abertos governamentais pelo governo e sociedade, que devem melhorar a eficiência de processos organizacionais destes setores (públicos e iniciativa privada), criação de novos produtos, serviços, e conhecimento, além de benefícios sociais não quantificados (MANYKA et al., 2011, tradução nossa).

Tornar um dado aberto significa deixá-lo sob uma licença que não restrinja o seu uso e compartilhamento por terceiros, e ainda disponibilizá-lo via internet se possível, conforme segue sua definição:

- Reutilização e Redistribuição: os dados (abertos) devem ser fornecidos sob termos que permitam a reutilização e a redistribuição, inclusive a combinação com outros conjuntos de dados;
- Disponibilidade e Acesso: os dados devem estar disponíveis como um todo e sob custo não maior que um custo razoável de reprodução, preferencialmente possíveis de serem baixados pela internet. Os dados devem também estar disponíveis de uma forma conveniente e modificável;
- Participação Universal: todos devem ser capazes de usar, reutilizar e redistribuir - não deve haver discriminação contra áreas de atuação ou contra pessoas ou grupos. Por exemplo, restrições de uso não-comercial que impediriam o uso comercial, ou restrições de uso para certos fins (ex.: somente educativos) excluem determinados dados do conceito de abertos (OPENDATAHANDBOOK, 2014).

Conforme o Laboratório Brasileiro de Cultura Digital (2011), os principais benefícios da abertura de dados governamentais são o aumento da transparência e controle democrático; participação popular; empoderamento dos cidadãos; melhores ou novos produtos e serviços privados; incentivo à inovação e empreendedorismo; melhora na eficiência de serviços governamentais; melhora na efetividade de serviços governamentais; medição do impacto das políticas; geração de novos conhecimentos a partir da combinação de fontes de dados abertas.

2.4 MODELOS CONCEITUAIS

Modelos conceituais, ou modelagem de conceitos, são utilizados em diversas pesquisas e disciplinas, com seu foco variando conforme a finalidade de seu uso. Em geral, pesquisadores utilizam modelos conceituais para estruturar e guiar suas pesquisas (PREEZ; MEYER, 2011). Diversos trabalhos são encontrados na literatura que utilizam modelos conceituais aplicados nas mais distintas áreas, servindo de base para o entendimento e investigação de problemas de pesquisa em geral.

A Engenharia de Conhecimento auxilia na modelagem de conceitos. Esta disciplina possui teorias, métodos e ferramentas para desenvolver aplicações intensivas em conhecimento, bem com modelar este conhecimento (SCHREIBER, 2008). Deste arcabouço, o método OntoKEM (RAUTENBERG et al., 2010) auxilia na modelagem dos principais conceitos (termos) e suas relações em um determinado domínio de estudo. Bunge (2003) também contribui para modelar conceitos e relações na medida em que fornece um método que permite visualizar de forma sistêmica um determinado domínio de estudo, tratando-o como um sistema. O autor fornece a quádrupla CESM (Componente, Ambiente, Estrutura e Mecanismos), que permite modelar

sistemas de diversas naturezas, como sistemas de mobilidade urbana, sistemas biológicos, sistemas de aviação, dentre outros.

2.5 PROCEDIMENTOS METODOLÓGICOS

Conforme Gil (2002), uma pesquisa é um procedimento racional e sistemático que tem como objetivo encontrar respostas aos problemas que são investigados. Ainda segundo o autor, é possível classificar uma pesquisa realizada mediante a diversos critérios. Essa tarefa auxilia na construção dos procedimentos metodológicos a serem implementados. Sendo assim, a primeira classificação deste trabalho se faz conforme sua essência, contexto e principalmente finalidade. Este trabalho caracteriza-se como uma pesquisa tecnológica, pois está preocupado em produzir um artefato funcional como resultado final (CUPANI, 2006; FREITAS JUNIOR et al., 2012).

Em relação a seus procedimentos, esta pesquisa classifica-se como exploratória/bibliográfica (GIL, 2002). Seu delineamento busca implementar esta classificação, conforme pode ser visto no Quadro 1.

Quadro 1 - Delineamento da Pesquisa.

PASSOS
1 - Pesquisa bibliográfica: leitura analítica e leitura interpretativa
2 - Identificação dos atores (<i>stakeholders</i>) e seus papéis
3 - Definição de Requisitos
4 - Desenho e descrição do Modelo
5- Desenvolvimentos dos artefatos
6 - Aplicação do Modelo em um cenário de uso

Fonte: Autores (2016).

Esta pesquisa utiliza-se de diversas fontes bibliográficas como consulta para gerar uma melhor familiaridade com o problema. Foi utilizado como base as coleções virtuais do Portal Capes, Biblioteca Universitária da UFSC, livros, relatórios e *white papers* de grandes empresas do setor (com destaque para os relatórios produzidos pela IBM e McKinsey). Após a revisão bibliográfica, foi desenvolvido o Modelo. Foram utilizadas como ferramentas o processo de desenvolvimento de ontologias OntoKEM, o framework de Zachman, e o modelo CESM para modelagem de sistemas. Como base teórica para ser usada por estas ferramentas foram utilizados os diversos autores e trabalhos relevantes encontrados na pesquisa bibliográfica.

O OntoKEM é um processo de desenvolvimento de ontologias que combina algumas práticas de metodologias para construção de ontologias, como o *Ontology Development 101* (processo iterativo para desenvolvimento de ontologias), *On-to-Knowledge* (definição do domínio e o escopo da ontologia; utilização de questões de competência), e *Methontology* (rico

conjunto de artefatos de documentação) (RAUTENBERG et al., 2010). Este processo prevê quatro fases de desenvolvimento: especificação (identificar propósito, identificar escopo, considerar reuso, identificar fontes), aquisição de conhecimento (gerar questões de competência, listar termos, definir termos, definir propriedades, agregar elementos, definir relações, definir restrições, criar instâncias), implementação (valorar propriedades, valorar relações, valorar restrições) e avaliação (avaliar segundo fontes, avaliar segundo requisitos, avaliar com usuários) (RAUTENBERG et al., 2010). A escolha pelo OntoKEM se justifica pelo fato de ontologias serem comumente usadas para representar o conhecimento e suas estruturas conceituais (RAUTENBERG; TODESCO; GAUTHIER, 2009). Dentre as práticas previstas pelo OntoKEM, foram utilizadas principalmente estas: “Questões de Competência”, “Listar Termos”, “Definir Termos”, “Agregar Elementos” e “Definir Relações”. As Questões de Competência foram combinadas às interrogativas do framework de Zachmann (2003), que é um modelo que realiza seis palavras interrogativas às quais, se respondidas, auxiliam a entender ideias complexas e objetos de estudo de diversos domínios. Os resultados são descritos a seguir.

2.6 O MODELO DESENVOLVIDO

As questões de competência desenvolvidas para construir o modelo conceitual foram agrupadas em três camadas:

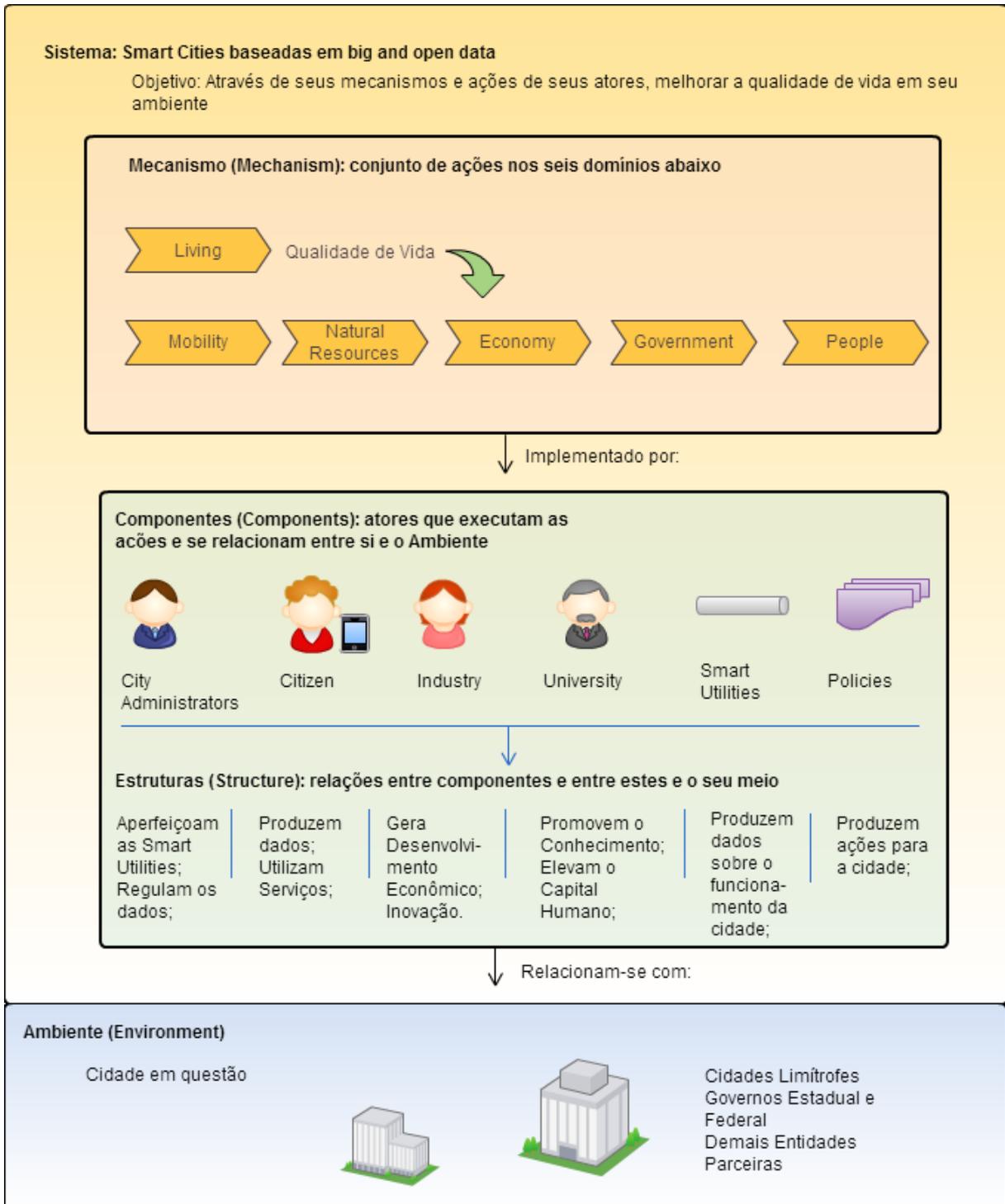
- **Camada Conceitual** (*Conceptual Layer*):
 - O que é uma *smart city* (*What is a smart city*)? Quais são os termos associados (*What are the related terms*)? Como estes termos diferenciam-se entre si (*How are the many terms different from each other*)?
- **Camada de Domínios e Indicadores** (*Domain Layer*):
 - Quais dados coletar (*What data to collect*)? Qual o escopo a ser abordado (*What is the scope or subjects to analyse*)? Como extrair informações sobre o mecanismo (os seis domínios) das *smart cities*, para analisar seu funcionamento (*How to extract information about the mechanism/ six domains of smart cities, to analyze its operation*)? Quem são os principais atores nas *smart cities* (*Who are the stakeholders*)? Como eles influenciam as cidades inteligentes (*How do they influence the smart cities*)?
- **Camada de Dados** (*Data Layer*):
 - Quais são as principais fontes de dados (*What are the main data sources*)? Dados sobre quais domínios pode-se encontrar nestas fontes (*What smart cities domains data can be found*)? Em quais formatos são encontrados estes dados (*What are the main content format of the data sources*)? Quais suas principais características (*What are its main*

characteristics)? Qual a periodicidade em que eles são produzidos (*What is the data production velocity*)?

As questões acima visam elicitar e identificar os principais termos do problema de estudo, definindo quais perguntas o modelo proposto precisa responder. Para responder a estas perguntas, foi aplicada a quádrupla CESM à revisão bibliográfica. Bunge (2003) afirma que qualquer sistema pode ser modelado conforme a quádrupla de elementos “*Composition*” (**composição**), “*Environment*” (**ambiente**), “*Structure*”, (**estrutura**) e “*Mechanism*” (**mecanismo**). Abaixo seguem listados os resultados deste mapeamento. Os termos foram mantidos em inglês para manter seu significado original, mas traduzidos ao seu lado, e ilustrados na Figura 1.

- **CESM: Composição:** os principais atores e elementos do processo principal das *smart cities*: elevar a qualidade de vida de seus habitantes. Seguem abaixo enumerados:
 - *City administrators*: o governo ou os administradores da cidade; 2- *Civil Society*: cidadãos ou sociedade civil; 3- *Industry*: indústria em geral (*industry*), e três subclasses de atores de destaque desta categoria: indústria de TI (*IT industry*), fornecedores de tecnologia para *IoT* (*IoT Providers*) e desenvolvedores (*developers*); 4- *University*: instituições de ensino e pesquisa, como universidades, escolas, faculdades e institutos de pesquisa por exemplo;
 - 5 e 6 - *Utilities* e *Policies*: serviços públicos inteligentes (*utilities* ou *smart utilities*) e políticas públicas (*policies*), respectivamente. Devido a forte importância dada por diversos autores ao papel destes elementos chaves, estes componentes seguem destacados como atores (DOBRE; XHAFA, 2013; VILAJOSANA et al., 2013). Estes dois atores são criados pelo ator *City Administrator* (administradores de cidades) e se relacionam com os demais atores deste sistema, produzindo dados e criando ações, respectivamente. Exemplo: dados produzidos pelos sensores de uma rede de tratamento de esgoto de uma cidade (*utilities*, domínio *hard*) e ações para melhorar a consciência ambiental da população nestas cidades, ou para dispor os seus dados governamentais em formato aberto (*policies*, domínio *soft*);

Figura 1 – CESM aplicado às *smart cities*.



Fonte: Autores (2016).

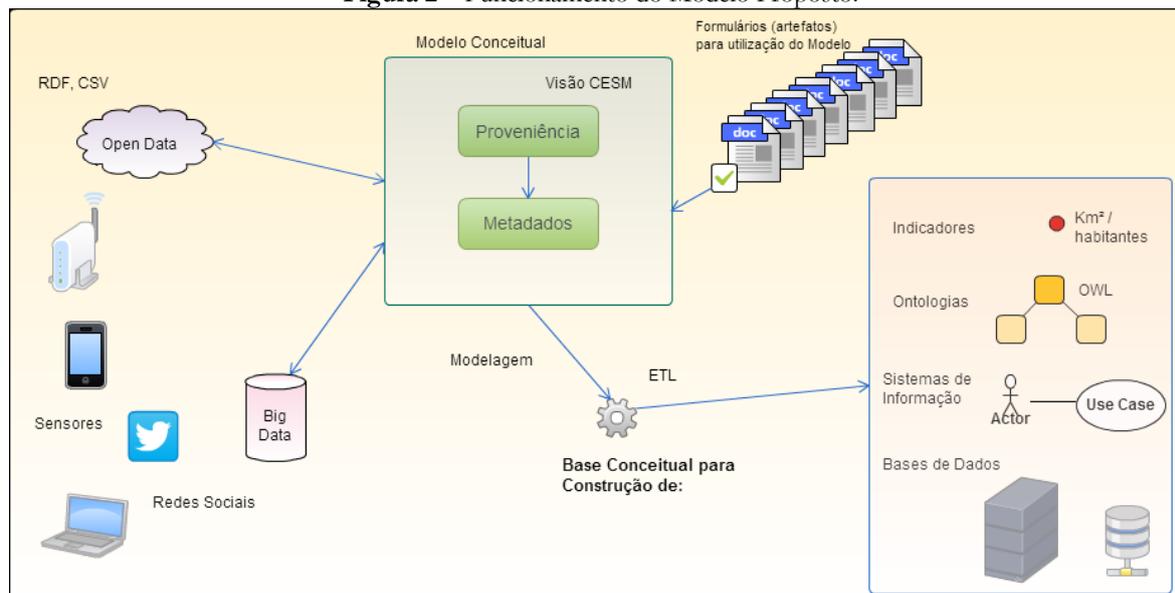
- 7 e 8: *Data* e *Services*: os dados (*big data* e *open data*) e os serviços também são considerados componentes chaves deste sistema. De maneira geral, os dados são produzidos pelas *smart utilities*, mas também são produzidos por atores como *City Administrators*, que produzem os dados governamentais, e pela sociedade civil. O componente *Services* geralmente é realizado em parcerias PPP (Parceria Público-Privada), onde a iniciativa privada e a sociedade civil utilizam os dados públicos para gerar inovação e desenvolvimento econômico usando estes dados (VILAJOSANA et al., 2013);
- **CESM-Ambiente**: a Cidade em Questão, Governos Federal, Estadual, demais Governos e instituições. Governos Municipais de cidades limítrofes também entram neste grupo, pois interagem com os serviços com da Cidade em Questão (exemplo: um sistema de trânsito intermunicipal);
- **CESM-Mecanismo**: são as principais ações que levam as *smart cities* a seu objetivo principal, que é elevar a qualidade de vida de seus habitantes. Estas ações são agrupadas em seis domínios (*domains*), conforme Giffinger et al. (2007). São eles: *Smart Environment* (Meio Ambiente), *Smart Mobility* (Mobilidade), *Smart People* (Capital Humano e Social), *Smart Economy* (Economia), *Smart Government* (Governo Transparente) e *Smart Living* (Qualidade de Vida).
- **CESM-Estrutura**: neste grupo foram identificadas as relações entre os atores e seu ambiente (exosestrutura), e dos componentes entre si (endo estrutura). Exemplo de relação entre atores: os administradores da cidade (*City Administrators*) equipam com sensores os seus serviços públicos (*Utilities*), que são fornecidos pela indústria de *IoT* (*IoT Industry*). Os cidadãos (*Civil Society*) usam estes serviços e produzem dados. As instituições de ensino e pesquisa auxiliam a promover o conhecimento baseado nestes dados (elemento *Data*). Abaixo todas as relações são identificadas e relacionadas a seus atores, escritas em negrito e em inglês para destaque:
 - Relações do ator *City Administrators* (Administradores das Cidades): ***Create Policies, Regulate Data e Improve Utilities***. Este ator se relaciona com a cidade na medida em que criam as políticas públicas (***Create Policies***, exemplo: políticas de proteção ambiental), regulam o uso dos dados (***Regulate Data***, exemplo: publicação de dados governamentais em formato aberto), e aperfeiçoam os serviços públicos (***Improve Utilities***), equipando com sensores os serviços públicos;
 - Relações do ator *Civil Society* (sociedade civil): ***Use Utilities/Produce Data, e Use Services***. A relação ***Use Utilities/Produce Data***, destaca que a sociedade precisa ter acesso aos serviços públicos inteligentes para conseqüentemente gerar dados

sobre estes. Já a relação consumo de serviços (*Use Services*) refere-se ao uso dos serviços desenvolvidos pela *IT Industry* e *Developers*, e que utilizam os dados municipais abertos;

- Relações do ator *Industry: Generate Economic Growth, Provide IoT Infrastructure, Produce Services/Innovation* De maneira geral, a relação deste ator com seu meio é gerar o desenvolvimento econômico das cidades (relação *Generate Economic Growth*). Além da indústria em geral, destacam-se os: 1- os fornecedores de tecnologia para *IoT (IoT Providers)*, que fornecem as tecnologias necessárias para as *smart utilities* (relação *Provide IoT Infrastructure*), como sensores; 2- a Indústria de TI (*IT Industry*) e Desenvolvedores (*Developers*) em geral fornecem demais soluções de software e serviços para estas cidades (relação *Produce Services/Innovation*), gerando inovação;
- Relações do ator *University: Promote Knowledge*. A principal relação deste ator acontece ao promover o conhecimento nestas cidades, elevando seu capital humano e social. Além disso, podem também relacionar-se com estes componentes ao fornecer conhecimento sobre as cidades, através de seus especialistas e pesquisadores. Ou seja, ao analisar os dados das cidades, pesquisadores e/ou especialistas de diversos domínios podem auxiliar as cidades em seus problemas urbanos, e assim inferir e gerar conhecimento sobre as cidades;
- Relações do ator *Policies: Implement Actions*. Este componente chave implementa as ações chaves criadas principalmente pelo ator *City Administrator*. Suas ações estão mais relacionadas com o ambiente, agindo principalmente (mas não exclusivamente) em domínios considerados *soft (smart people, smart economy, smart government)*. Exemplos: a criação de transparência governamental através da abertura de dados, ações de inclusão social dentre outros;
- Relação do ator *Utilities (Serviços públicos inteligentes): Produce (City) Data*. Esta relação interage com todos os demais atores do sistema *smart cities*, produzindo dados sobre o funcionamento da Cidade em Questão.

Após o mapeamento dos principais elementos explicados acima, o Modelo proposto foi construído. Ele implementa duas tarefas: 1 - identificar a necessidades de informações por domínio para *smart cities*; 2- apontar fontes de dados, utilizando metadados de proveniência adaptados do padrão Dublin Core para identificar estas fontes. A Figura 2 apresenta o funcionamento do modelo e seus formulários implementados.

Figura 2 – Funcionamento do Modelo Proposto.



Fonte: Autores (2016).

Para as duas tarefas do modelo foram disponibilizados formulários a serem utilizados pelos seus usuários, que devem ser especialistas de domínios distintos. Um conjunto de seis formulários auxilia a identificar as necessidades de informação para as *smart cities* (um formulário para cada domínio: *smart environment*, *smart economy* e demais domínios), e outro formulário auxilia no mapeamento de proveniência de fontes de dados para estas informações. Os usuários devem ser especialistas em cada área de atuação das informações tratadas em cada formulário (ex.: um engenheiro ambiental para utilizar o formulário de informações do domínio *smart environment* e assim por diante). Um especialista em banco de dados é recomendado para utilizar o formulário de proveniência de fontes de dados. A seguir, é apresentado um cenário de uso deste modelo.

2.7 CENÁRIO DE USO

Com o propósito de demonstrar a utilização do Modelo proposto foi definido um cenário de uso, descrito a seguir. Foi escolhido o domínio *Smart Environment* (Recursos Naturais) e aplicado à cidade de São Paulo. A verificação é feita em dois passos, conforme segue.

Passo 1: primeiramente, o usuário define qual domínio ele quer analisar. Após isso, o Modelo sugere um conjunto de necessidades de informação para o domínio escolhido. Estas devem ser coletadas e/ou construídos, caso a cidade não disponha destes dados. Estas informações são agrupadas em três subgrupos: fatores, indicadores e ações. Os fatores servem como base para a construção de novos indicadores e orientar ações, conforme a necessidade do seu usuário/cidade (Quadro 2). Para cada domínio, existem de 1 até N fatores; e cada fator possui de 0 (zero) a N indicadores e ações.

Quadro 2- Formulário de informações para o domínio *Smart Environment*

Formulário de metadados do domínio, fatores, indicadores e ações	
Cidade:	São Paulo
Domínio:	<i>Environment</i> (Recursos Naturais)
Identificador:	D01-E
Fatores Indicados:	
Fator: 1	<i>Attractivity of natural conditions</i>
Indicadores:	<i>Sunshine hours, Green space share;</i>
Ações:	--
Fator: 2	<i>Pollution Controll</i>
Indicadores:	<i>Summer smog (Ozon), Particulate matter, Fatal chronic lower respiratory diseases per inhabitant</i>
Ações:	<i>Controlling emissions and effluents by using different kinds of devices. Stimulating decisions to improve the quality of air, water, and the environment in general;</i>
Fator: 3	<i>Environmental protection</i>
Indicadores:	<i>Individual efforts on protecting nature, Opinion on nature protection;</i>
Ações:	<i>Treat the natural environment as a strategic component for the future</i>
Fator: 4	<i>Sustainable Resource Management</i>
Indicadores:	<i>Efficient use of water (e. g. per capita), Efficient use of electricity (e. g. per capita)</i>
Ações:	<p>Smart Grid: <i>Implement Electricity networks able to take into account the behaviours of all the connected users in order to efficiently deliver sustainable, economic, and secure electricity supplies. Smart grids should be self-healing and resilient to system anomalies;</i></p> <p>Smart Building: <i>Various systems existing in a building such as electric networks, elevators, fire safety, telecommunication, data processing, and water supply systems. Computer-based systems to control the electrical and mechanical equipment of a building; Home and office automation systems (HOS) interconnect electric devices such as heaters, lights, air conditioners, TVs, computers, alarms, and cameras through a communication network, allowing them to be remotely controlled, monitored or accessed from any room in the building, as well as from any location in the world by Internet. They help people to optimize their living style, arrange the day-to-day schedule, secure a high living quality and reduce the energy consumption bill;</i></p> <p>Facilities Management: <i>Cleaning, maintenance, property, leasing, technology, and operating modes associated with facilities in urban areas;</i></p> <p>Public lighting: <i>Illumination of public spaces with street lamps that offer different functions, such as air pollution control and Wi-Fi connectivity. Centralised management systems that directly communicate with the lampposts can allow reducing maintenance and operating costs, analysing real-time information about weather conditions, and consequently regulating the intensity of light by means of LED technology;</i></p> <p>Green/renewable Energies: <i>Exploiting natural resources that are regenerative or inexhaustible, such as heat, water, and wind power;</i></p> <p>Water Management: <i>Analysing and managing the quantity and quality of water throughout the phases of the hydrological cycle and in particular when water is used for agricultural, municipal, and industrial purposes;</i></p> <p>Waste Management: <i>Collecting, recycling, and disposing waste in ways that prevent the negative effects of an incorrect waste management on both people and the environment;</i></p> <p>Food and Agriculture: <i>Wireless sensor networks to manage crop cultivation and know the conditions in which plants are growing. By combining humidity, temperature, and light sensors the risk of frost can be reduced and possible plant diseases or watering requirements based on soil humidity can be detected.</i></p>

Fonte: Autores (2016).

Os fatores, indicadores e ações indicados pelo formulário funcionam como metadados genéricos, que sugerem quais dados devem ser analisados pelo seu utilizador. Com o auxílio de especialistas, novas informações também podem ser criadas. No caso específico das ações, os metadados estão no gerúndio pois apontam iniciativas utilizadas em *smart cities* ao redor do mundo, as quais se implementadas podem gerar dados. Com estes metadados, um gestor de cidades pode procurar por fontes, ou basear-se neste apontamento e construir novos fatores, indicadores e ações.

Considerando o retorno do Passo 1, escolheu-se um subconjunto de metadados a serem analisados pelo utilizador. Para cada metadado buscou-se por fontes de dados e aplicou-se o Passo 2, a seguir. Os metadados escolhidos foram seguintes:

- Fator *Attractivity of Natural Conditions* (Atratividade das Condições Naturais):
 - Indicador: *Green space share* (Área Verde Por Habitante);
- Fator *Sustainable Resource Management/Water e Waste Management* (Gereciamento Sustentável de Recursos Naturais, foco em Água e Resíduos Sólidos):
 - Indicadores:
 - *Efficient use of water use per GDP* (Uso eficiente de água);
 - *Efficient use of electricity use per GDP* (Uso eficiente de eletricidade);
- Fator *Environmental Protection* (Proteção Ambiental):
 - Indicadores:
 - *Individual efforts on protecting nature protection* (Iniciativas dos cidadãos para preservação do meio ambiente);
 - *Opinion on nature protection* (opinião sobre a preservação ambiental).

Passo 2: mapear fontes de dados para os metadados indicados no passo anterior. Para isso, foi aplicado o formulário de características e proveniência de fontes de dados para cada fonte de dado identificada. Este formulário foi construído com atributos de proveniência adaptados do padrão Dublin Core (Quan, 2000), em conjunto com questões genéricas de mapeamento de dados citadas por Marins (2008). Eles auxiliam a descrever as principais características de um determinado conjunto de dados encontrado, o que futuramente facilita sua análise, dentre outros benefícios, como rastreamento do publicador, relacionamento com outras fontes e outros.

Para os dados do indicador *Efficient use of electricity use per GDP* (Uso eficiente de eletricidade), foi encontrado o dado “consumo total de eletricidade per capita”, da fonte ELETROPAULO (responsável pela distribuição de energia na cidade de São Paulo), demonstrado no Quadro 3. Esta fonte traz dados de 2009 a 2012.

Como fonte para o indicador de “Uso Eficiente de Água” foram encontrados dados sobre perda de água tratada (percentagem do total de água tratada), cuja fonte é a SABESP (Companhia de Saneamento Básico do Estado de São Paulo). Para o fator *Attractivity of Natural Conditions* (indicador *Green space share*-Área Verde Por Habitante) foram utilizados dados da fonte SVMA – Secretaria do Verde e Meio Ambiente da cidade de São Paulo. Outra fonte de dados foi utilizada para compor este fator, focando em resíduos sólidos (*Waste Management*) foi a SES Secretaria Municipal de Serviços. Para isso, foi adicionado o novo indicador “Reciclagem de Resíduos Sólidos (percentagem)”, pois este dado foi encontrado para uso no Programa Cidades Sustentáveis, portal de agrupamento de dados governamentais do programa Cidades Sustentáveis (EOEKE, 2015), de onde os dados das fontes usadas até agora (SVMA, SES, SABESP e ELETROPAULO) foram obtidas.

Para compor o fator *Environmental Protection*, mais precisamente os indicadores: *Individual efforts on protecting nature protection* (Iniciativas dos cidadãos para preservação do meio ambiente) e *Opinion on nature protection* (opinião sobre a preservação ambiental) não foram encontrados dados para a cidade de São Paulo. Logo, foi utilizado um relatório em PDF que apresenta os resultados de uma pesquisa de opinião realizada com a população brasileira sobre o tema “desenvolvimento sustentável”, elaborado pelo Ministério do Meio Ambiente do Governo Federal.

Quadro 3– Mapeamento de Proveniência de fonte de dados para o fator “Consumo de Energia *per capita*”.

Formulário de Mapeamento de Proveniência de Dados para <i>smart cities</i> baseadas em <i>big data</i> e <i>open data</i> .	
Cidade:	São Paulo
Domínio	(x) <i>Environment</i> () <i>Mobility</i> () <i>Government</i> () <i>Economy</i> () <i>People</i> () <i>Living</i>
Identificador:	DO1-E <Ex.: DO2-M, se o domínio escolhido for <i>Mobility</i> >
Fontes de Dados:	
Fonte 1:	ELETROPAULO
Descrição da Fonte:	Dados sobre consumo de energia per capita na cidade de São Paulo
Entidade Responsável pelos dados:	ELETROPAULO Programa Cidades Sustentáveis (Publicador dos dados).
Contato	https://www.aeseletropaulo.com.br/Paginas/aes-eletropaulo.aspx
Autor dos dados	--
Colaboradores	--
Data de disponibilização dos dados:	--
Acesso ao dado:	http://www.cidadessustentaveis.org.br
Identificador:	http://indicadores.cidadessustentaveis.org.br/br/SP/sao-paulo/indicadores.xls

Formulário de Mapeamento de Proveniência de Dados para *smart cities* baseadas em *big data* e *open data*.

Fonte:	--
Idioma:	Português
Relação:	<relação específica com outros recursos ou fontes de dados >
	Identificador do Recurso: Nome do recurso/fonte: URL: Relação: <input type="checkbox"/> IsPartOf HasPart <input type="checkbox"/> IsVersionOf HasVersion <input type="checkbox"/> IsFormatOf HasFormat <input type="checkbox"/> References IsReferencedBy <input type="checkbox"/> IsBasedOn IsBasisFor <input type="checkbox"/> Requires IsRequiredBy
Cobertura Temporal do dado:	2009-2012
Dado bruto ou transformado ?	<o dado sofreu processo de tratamento (ETL) ou foi coletado diretamente na fonte> Dado Tratado (coletado no site)
Quais softwares são necessários?	<ferramentas de software que foram utilizadas para produzir e ou publicar os dados> Editor de planilha eletrônica, como LibreOffice Calc, Microsoft Excel.
Dados do domínio da <i>smart city</i>	
Principais fatores atendidos pela fonte:	<i>Sustainable Resource Management/Water e Waste Management</i>
Principais Indicadores e ações atendidos pela fonte:	<i>Efficient use of electricity use per GDP (Uso eficiente de eletricidade)</i>
Características <i>big data</i> e <i>open data</i> (<i>Volume, Variety, Velocity, Variability</i> , Licença <i>Open Data</i> , formato ...):	
Tipo do dado:	<categoria do recurso, como texto, imagem, som, dados exportados em listas ou tabelas ...> Tabular
Natureza dos dados:	<input type="checkbox"/> Produzido por Máquinas (ex.: dados de sensores, Internet das Coisas) <input checked="" type="checkbox"/> Produzido por Humanos (ex.: <i>posts</i> em redes sociais, textos extraídos de sites, dados informados por usuários de sistemas de informação) <input type="checkbox"/> Não Informado
Plataforma da Fonte de dados (meio onde foi coletado):	<input type="checkbox"/> Dados coletados de Sensores <input checked="" type="checkbox"/> Web <input type="checkbox"/> Outro. Especificar: _____
Detalhamento da plataforma (meio onde foi produzido):	<input type="checkbox"/> Serviços Públicos <input type="checkbox"/> Dispositivos Móveis <input checked="" type="checkbox"/> Sites <input type="checkbox"/> Outro. Especificar: _____
Forma de disponibilização:	<input type="checkbox"/> Dados Abertos <input type="checkbox"/> <i>Web Services</i> <input type="checkbox"/> Postagens em Redes Sociais <input checked="" type="checkbox"/> Outra. Especificar: Planilha do Microsoft Office Excel
Dados Contextuais (localização do dado)	Possui informações sobre a localização (ex.: coordenadas geográficas) da fonte de dados? <input type="checkbox"/> Sim <input checked="" type="checkbox"/> Não
DGA	Caso for Dado Governamental, descrever de qual agência (ex.: Departamento Municipal de Trânsito): Secretaria do Meio Ambiente da Prefeitura de São Paulo
Licença <i>Open Data</i>	Possui licença Aberta (Dado Aberto) ? <input type="checkbox"/> Sim <input type="checkbox"/> Não <input checked="" type="checkbox"/> Não Informado Caso não for Dado Aberto, possui algum tipo de restrição legal de uso? <input type="checkbox"/> Sim <input type="checkbox"/> Não <input checked="" type="checkbox"/> Não Informado

Formulário de Mapeamento de Proveniência de Dados para *smart cities* baseadas em *big data* e *open data*.

Especificar: _____

Estrutura do Dado:	(x) Dados Estruturados (ex.: dados extraídos de sistemas de informação, com esquema de dados separado do dado) () Semiestruturados (ex.: RDF, XML, onde o esquema de dados está juntamente com o dado) () Não-estruturados (sem esquema de dados, ex.: textos, sites em HTML)
Periodicidade de Produção:	<Ex.: tempo-real, diária, mensal, anual> Anual
Tamanho	Tamanho Aproximado da amostra coletada (ex.: 100 MB, 1 GB, 1 TB): 1,3 MB

Fonte: Autores (2016).

Para somar ao indicador *Opinion on nature protection* (opinião sobre a preservação ambiental), foram também coletados dados da rede social Twitter. Estes *posts* podem ser minerados com técnicas de extração de conhecimento em bases textuais e análise de sentimento por exemplo, para avaliar a opinião emitida pelos usuários, como é visto em muitas pesquisas atuais e aplicativos comerciais. No caso do Twitter, o filtro geográfico por cidade torna pequeno o resultado, apenas 8 postagens, devido à falta desta informação na base de dados. Segundo Reino (2015), mesmo sendo um percentual de 77% o número de *tweets* postados na rede através de dispositivos móveis, apenas uma pequena parte é georreferenciada. O Twitter fornece acesso via sua API aos dados georreferenciados, porém o serviço fornece apenas os *tweets* postados nos últimos sete dias (TWITTER, 2015). Para fim de demonstração de utilização deste formulário em dados não estruturados, foi utilizado o protocolo HTTP GET para filtrar os *tweets* por região e data (desde 2009 até 2016), postados próximo a São Paulo, através dos parâmetros *since*, *until* e *within* respectivamente, oferecidos pelo portal de buscas avançadas do Twitter. Estes *posts* podem ser agora cruzados aos dados das outras fontes para obtenção de novas informações e até mineração de padrões e conhecimento.

Após a aplicação destes formulários, pode-se construir e catalogar estas fontes heterogêneas de dados, iniciando a construção de bases de dados para a cidade escolhida. Pode-se cruzar dados sobre a opinião pública em relação a ações governamentais, com dados sobre o consumo de energia ou de água por habitante por exemplo. Com estes dados como apoio, um gestor de uma *smart city* possui maiores subsídios para os processos de tomada de decisões, realizar processos de *benchmarking* entre cidades e avaliar o desempenho de sua cidade em diversas áreas importantes para seu desenvolvimento. É possível também registrar a proveniência das fontes de dados, o que é um passo importante para a construção de bases de dados, buscar por novas fontes de dados e construir novos indicadores. Estes indicadores podem auxiliar a geração de conhecimento nestas

idades, uma vez que as informações podem ser analisados por seus gestores e especialistas de domínio. A seguir, são apresentadas as conclusões deste trabalho.

3 CONSIDERAÇÕES FINAIS

O problema de pesquisa inicialmente proposto questiona como gerar conhecimento para apoiar as *smart cities*, utilizando *big data* e *open data*. Para isto, construiu-se um modelo conceitual que buscou essencialmente responder às seguintes perguntas: quais dados devem ser coletados para estas cidades, quais informações devem ser produzidas, quais fontes de dados utilizar e quais são os principais atores neste processo. Com este Modelo, administradores de cidades, pesquisadores e desenvolvedores podem iniciar projetos para *smart cities*, como sistemas de informação, sistemas de conhecimento e projetos de pesquisas, auxiliando a gerar conhecimento a partir dos dados destas cidades. Em outras palavras, a contribuição maior do Modelo proposto foi em auxiliar o mapeamento de fontes de dados *big data* e *open data* nos seis domínios identificados nas *smart cities*, sugerindo metadados para isso. Com esta resposta em mãos, é possível realizar a construção de indicadores para a análise do funcionamento destas cidades, identificar ações comumente necessárias em *smart cities*, ou ainda identificar lacunas de bases de dados a serem criadas pelos administradores de cidades.

Como trabalhos futuros propõem-se o desenvolvimento de uma ontologia, com a adição de novos elementos e atributos dos conjuntos de metadados de cada um dos domínios das *smart cities*, a ser feito junto a especialistas de cada domínio (ex.: sociólogos, engenheiros ambientais etc), para somar às informações trazidas pelo Modelo, bem como a identificação, em forma de modelagem, da relação entre os elementos destes domínios (ex.: meio ambiente, transportes, e taxas de emissão de carbono). Nesta linha, sugere-se aplicar este Modelo como um guia em projetos de diversas naturezas, estudos de caso de iniciativas de *smart cities* em determinadas cidades, criação de bases de dados, sistemas de informação e construção de indicadores baseados em *big data* e *open data* para as *smart cities*, dentre outras aplicações a serem exploradas por futuros pesquisadores. Uma aplicação em maior escala deste modelo, com mais fontes de dados também é sugerido como trabalho futuro.

REFERÊNCIAS

- BEYER, Mark A.; LANEY, Douglas. **The Importance of 'Big Data':** A Definition. 2012. Disponível em: <<https://www.gartner.com/doc/2057415/importance-big-data-definition>>. Acesso em: 27 set. 2014.
- BRASIL – Governo Federal. O que são dados abertos? **Portal Brasileiro de Dados Abertos**. 2015. Disponível em: <<http://dados.gov.br/dados-abertos/>>. Acesso em: 17 ago. 2015.

- BUNGE, Mario. **Emergence and Convergence: Qualitative Novelty and the Unity of Knowledge**. Toronto: National Library Of Canada, 2003.
- CARAGLIU, Andrea; BO, Chiara del; NIJKAMP, Peter. Smart Cities in Europe. **Journal Of Urban Technology**. Londres, p. 65-82. 10 ago. 2011. Disponível em: <http://www.tandfonline.com/doi/abs/10.1080/.VBZNMx_Hk8o#.VBZPMZ_7G00>. Acesso em: 11 jun. 2014.
- CHEN, Min; MAO, Shiwen; LIU, Yunhao. Big Data: A Survey. **Springer**, Estados Unidos, v. 2, n.19, p.171-209, Abr.2014. Disponível em: <<http://link.springer.com/article/10.1007%2Fs11036-013-0489-0>>. Acesso em: 13 set. 2014.
- CUPANI, Alberto. La peculiaridad del conocimiento tecnológico. **ScientiaeStudia**, São Paulo, v. 4, n. 3, p. 353-71, 2006.
- DAVENPORT, Thomas. **Big Data at Work: Dispelling the Myths, Uncovering the Opportunities**. Boston: Harvard Business Publishing Corporation, 2014. 229 p.
- DOBRE, Ciprian; XHAFIA, Fatos. Intelligent services for Big Data science. **Future Generation Computer Systems**. Amsterdam, p. 267-281. jul. 2013. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0167739X13001593>>. Acesso em: 05 set. 2014.
- FAN, Wei; BIFET, Albert. Mining Big Data: Current Status, and Forecast to the Future. **SIGKDD Explorations**, China, v. 2, n. 14, p.1-5, mar. 2012. Disponível em: <<http://www.kdd.org/sites/default/files/issues/14-2-2012-12/V14-02-01-Fan.pdf>>. Acesso em: 15 set. 2014.
- FREITAS JUNIOR, Vanderlei et al. A pesquisa científica e tecnológica. **Espacios**, [S.I.], v. 35, n. 9, p.12-12, jul. 2014. Disponível em: <<http://www.revistaespacios.com/a14v35n09/14350913.html>>. Acesso em: 16 out. 2014.
- GANTZ, John; REINSEL, David. Extracting Value from Chaos. **IDC Iview**, Framingham, v. 01, n. 17, p.1-12, jun. 2011. Disponível em: <<http://www.emc.com/collateral/analyst-reports/idc-extracting-value-from-chaos-ar.pdf>>. Acesso em: 13 set. 2014.
- GIFFINGER, Rudolf et al. **Smart cities Ranking of European medium-sized cities**. Vienna: Vienna University Of Technology, 2007. 28 p. Disponível em: <http://www.smart-cities.eu/download/smart_cities_final_report.pdf>. Acesso em: 06 jun. 2014.
- GIL, Antonio Carlos. **Como elaborar projetos de pesquisa**. 5. ed. São Paulo: Atlas, 2010.
- GROBELNIK, Marko. **Big Data Tutorial**. Kalamaki: Jožef Stefan Institute, 2012. Color. Disponível em: <http://videlectures.net/eswc2012_grobelnik_big_data/>. Acesso em: 22 jul. 2014.
- IBM, Software. **The IBM big data platform**. Somers: IBM Software, 2013. 4 p. Disponível em: <<http://public.dhe.ibm.com/common/ssi/ecm/en/imb14135usen/IMB14135USEN.PDF>>. Acesso em: 15 ago. 2014.
- LABORATÓRIO BRASILEIRO DE CULTURA DIGITAL (Org.). **Manual dos dados abertos: governo**. S.i: W3c, 2011. 58 p. Disponível em: <http://www.w3c.br/pub/Materiais/PublicacoesW3C/Manual_Dados_Abertos_WEB.pdf>. Acesso em: 07 ago. 2015.
- LANEY, Doug. 3D Data Management: Controlling Data Volume, Velocity, and Variety. **Metadata Group**. 2001. Disponível em: <<https://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf>>. Acesso em: 13 dez. 2016.

MANYKA, James et al. **Big Data**: The next frontier for innovation, competition, and productivity. Nova York: Mckinsey Global Institute, 2011. 20 p. Disponível em:

<http://www.mckinsey.com/insights/business_technology/big_data_the_next_frontier_for_innovation>. Acesso em: 18 jun. 2014.

MARINS, André. **Modelos Conceituais para Proveniência**. Dissertação de Mestrado. Programa de PósGraduação em Informática da PUC-Rio.2008.

NAM Taewoo; PARDO Theresa A. Conceptualizing Smart City with Dimensions of Technology, People, and Institutions. In: ANNUAL INTERNATIONAL CONFERENCE ON DIGITAL GOVERNMENT RESEARCH, 12th, 2011, New York. **Proceedings**. New York: Acm Magazines And Online Publications, 2011. 282-291.

NEIROTTI, Paolo et al. Current trends in Smart City initiatives: Some stylised facts. **Cities**: the international journal of urban policy and planning. Amsterdã, p. 25-36. jun. 2014. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0264275113001935>>. Acesso em: 15 set. 2014.

O'REILLY, Media. **Big Data Now**. Sebastopol: O'reilly Media, 2012. 119 p.

OPEN DATA HANDBOOK. **Por que Abrir Dados?** Disponível em:

<http://opendatahandbook.org/guide/pt_BR/why-open-data/>. Acesso em: 09 out. 2014.

QUAN, Eilen. **Minnesota metadata guidelines for dublin core metadata**. Minnesota Department of Natural Resources. 2000.

RAUTENBERG, Sandro et al. Ferramenta ontoKEM: uma contribuição à Ciência da informação para desenvolvimento de ontologias. **Perspectivas em Ciência da Informação**, [s.i.], v. 15, n. 1, p.239-258, jan. 2010. Disponível em: <<http://www.scielo.br/pdf/pci/v15n1/14.pdf>>. Acesso em: 15 ago. 2015.

RAUTENBERG, Sandro; TODESCO, José L.; GAUTHIER, Fernando A. O.. Processo de desenvolvimento de ontologias: uma proposta e uma ferramenta. **Revista Tecnológica**, Fortaleza, v. 30, n. 1, p.133-144, jun. 2009.

REINO UNIDO. UNITED KINGDOM (UK) GOVERNMENT. **Opening Up Government**. 2015. Disponível em: <<http://data.gov.uk/about>>. Acesso em: 15 jul. 2015.

RIVERA, Janessa; MEULEN, Rob van Der. **Gartner Says the Internet of Things Installed Base Will Grow to 26 Billion Units By 2020**. 2013. Disponível em:

<<http://www.gartner.com/newsroom/id/2636073>>. Acesso em: 02 jul. 2014.

SCHREIBER, Gus. Handbook of Knowledge Representation. **Elsevier**. 2008. DOI: 10.1016/S1574-6526(07)03025-8.

TOWNSEND, Anthony M.. **Smart Cities**: Big Data, Civic Hackers, and the Quest for a New Utopia. [S.I.]: W. W. Norton & Company, 2013. 320 p.

TWITTER, Engineering. **200 million Tweets per day**. 2011. Disponível em:

<<https://blog.twitter.com/2011/200-million-tweets-day>>. Acesso em: 18 ago. 14.

US GOVERNMENT. **U.S. Government's Open Data**. 2015. Disponível em:

<<http://www.data.gov/>>. Acesso em: 15 jul. 2015.

VILAJOSANA, Ignasi et al. Bootstrapping smart cities through a self-sustainable model based on big data flows. **IEEE Communications Magazine**, Estados Unidos, v. 51, n. 6, p.128-134, jun. 2013. Disponível em: <<http://ieeexplore.ieee.org/xpl/articleDetails.jsp?reload=true&tp;=&arnumber=6525605&queryText;=vilajosana+big+data+bootstraping>>. Acesso em: 12 jul. 2014.

ZACHMAN, John A.. **The Zachman Framework For Enterprise Architecture:** Primer for Enterprise Engineering and Manufacturing. [s.i]: Zachman International, 2003. 15 p.

ZYGIARIS, Sotiris. Smart City Reference Model: Assisting Planners to Conceptualize the Building of Smart City Innovation Ecosystems. **Journal Of The Knowledge Economy**, [S. I.], v. 4, n. 2, p.217-231, 8 mar. 2012. Springer Science + Business Media. DOI: 10.1007/s13132-012-0089-4.