

Um *Survey* sobre Plataformas de Mediação de Dados para Internet das Coisas

Emanoel Carlos G. F Silva¹, Marcelo Iury S. Oliveira¹,
Eduardo Oliveira², Kiev Santos da Gama¹, Bernadette Farias Lóscio¹

¹Centro de Informática -- Universidade Federal de Pernambuco (UFPE)
Caixa Postal 50.740-560 -- Recife -- PE -- Brasil

²Department of Computing and Information Systems, The University of Melbourne
Parkville 3010 VIC, Australia

{ecgfs, miso, kiev, bfl}@cin.ufpe.br, eduardo.oliveira@unimelb.edu.au

Abstract. *The internet of things area is emerging rapidly. Because of this, there are currently numerous platforms developed with the purpose of data mediation, and each of them presents different characteristics. Some specialist who plans to develop applications to Internet of Things has the difficult task to gather and consider all these features and, finally, adopt only one of these platforms. The objective of this study is to map the data mediation platforms for Internet of Things in the market, presenting a comparative point of view of data models, communication protocols and features offered.*

Resumo. *Apesar dos avanços no desenvolvimento de novas plataformas para coleta, integração e análise de dados na Internet das Coisas (IoT), poucos trabalhos apresentam um mapeamento e análise quanto as soluções existentes e disponíveis no mercado. Selecionar qual a plataforma mais apropriada para uma necessidade específica da indústria não é uma tarefa trivial. Esta seleção é complexa porque grande parte das plataformas existentes e disponíveis em mercado faz uso de soluções e tecnologias que usualmente não estão disponíveis, detalhadas ou avaliadas em trabalhos científicos. Neste artigo, realizamos um mapeamento das plataformas de mediação de dados disponíveis para internet das coisas existentes, apresentando um comparativo do ponto de vista de modelos de dados e funcionalidades de gerenciamento e processamento de dados.*

1. Introdução

Avanços recentes nas áreas de comunicação móvel sem fio, microeletrônica, sensoria-mento e sistemas embarcados permitiram o advento de um novo modelo de sistemas e comunicação conhecido como Internet das Coisas (do inglês, *Internet of Things* - IoT), que consiste na presença distribuída e conectada de uma grande variedade de coisas ou objetos ao nosso redor que se comunicam com o objetivo de trocar informações e interagir entre si [Atzori et al. 2010]. IoT é um conceito emergente e em franca evolução que permitirá novas oportunidades para o desenvolvimento, tanto econômico como tecnológico.

De acordo com previsão do NIC (*US National Intelligence Council*), indivíduos, em-presas e governos não estão preparados para este possível futuro da Internet [Council 2008]. Segundo o NIC, em 2025 os objetos e coisas que trafegarão na Internet residirão em ativida-des cotidianas e diárias – pacotes de alimentos, móveis, documentos em papel e muito mais.

Enquanto novos tipos de dispositivos conectados são trazidos ao mercado, investimentos significativos já foram feitos e implementados em ambientes industriais existentes. Segundo a Intel, estima-se que “mais de 85% dos dispositivos conectados que poderiam se beneficiar mais da Internet das Coisas já existem dentro de infraestruturas instaladas” [INTEL 2014].

Esta “nova” Internet contará com um número incalculável de objetos heterogêneos gerando uma massiva e crescente quantidade de dados. Além disso, esta massa de dispositivos ainda traz consigo diversos desafios acerca da heterogeneidade (diferentes objetos, sensores, protocolos e aplicações), da resiliência, do armazenamento dos dados oriundos de redes de sensores e da Web e, da análise inteligente destes dados.

Plataformas de mediação de dados podem ser utilizadas para facilitar o desenvolvimento de aplicações e a integração das diferentes tecnologias no âmbito da IoT. Entre as funcionalidades fornecidas, há o provimento de serviços para gerenciar diversos dispositivos, auxiliar no armazenamento e na recuperação de dados, e processar dados usados para disparar cálculos, estatísticas, alertas e muito mais.

Diversas plataformas de mediação de dados foram propostas ou desenvolvidas para conectar pessoas, sensores e *softwares* entre si [Yan et al. 2008, Oliveira et al. 2015]. A proposição de arquiteturas e plataformas para IoT tem sido alvo de vários estudos reportados na literatura, entretanto muitos trabalhos ainda estão em fase conceitual com poucas validações práticas. Em contraposição, existem também diversas plataformas prontas para uso e de propósito genérico, tais como [Axeda 2015, Xively 2015, Thingspeak 2015], que são disseminadas na internet e estão sendo aplicadas em diversas iniciativas de cidades inteligentes e automação residencial. Em particular, uma parcela majoritária dessas plataformas é proprietária e oferecida, em grande parte, por grandes fornecedores de *software* e serviços, tais como IBM, Siemens e NEC.

Como há uma abundância de plataformas de mediação de dados, selecionar qual a solução mais apropriada para um cenário específico não é uma tarefa trivial. Esse trabalho se torna ainda mais complexo pelo fato de que uma grande quantidade de plataformas existentes envolve o uso de soluções mercadológicas que usualmente não estão especificadas ou avaliadas em trabalhos científicos. Isso faz com que métodos de mapeamento e revisões sistemáticas excluam estes tipos de soluções.

Diante deste cenário, neste artigo apresentamos um *survey* de plataformas de mediação de dados para Internet das Coisas, motivado pela seguinte questão: *Quais as plataformas de IoT para mediação de dados utilizadas hoje em dia e quais suas características, levando em consideração modelos de dados, padrões de dados adotados, e funcionalidades de gerenciamento e processamento da informação?* O restante deste trabalho está organizado da seguinte forma: Seção 2 apresenta e discute acerca de diversas plataformas de software existentes para mediação de dados para IoT. Seção 3 discute trabalhos correlatos. A Seção 4 apresenta a metodologia utilizada para a realização do *survey*. A Seção 5 faz uma análise sobre as plataformas mapeadas. E, por fim, a Seção 6 conclui o trabalho com um breve sumário das contribuições do *survey* e trabalhos futuros.

2. Plataformas de Mediação de Dados

Plataformas de mediação de dados para IoT tem como principal objetivo conectar e unificar inúmeros objetos e sistemas heterogêneos, permitindo a coleta e o processamento de

informações em larga escala [Atzori et al. 2010]. Todos os dados coletados devem ser compartilhados para que aplicações e serviços, que não precisam necessariamente ter conhecimento ou domínio das tecnologias envolvidas, possam utilizá-los.

As plataformas de mediação de dados em IoT usualmente fornecem serviços de registro, monitoramento e descoberta para dispositivos, *gateways* e aplicações de controle [Guinard et al. 2010]. Estes serviços de registro devem permitir que dispositivos sejam capazes de publicar detalhes sobre seus serviços para que possam ser encontrados por outros dispositivos ou usuários. E, alternadamente, o serviço de descoberta deve permitir que usuários sejam capazes de descobrir os dispositivos mais adequados às suas necessidades. O serviço de descoberta deve retornar dados suficientes para que usuários possam avaliar os dispositivos de forma independentemente, antes de contatá-los ou integrá-los às suas aplicações.

A heterogeneidade é um dos principais desafios a serem tratados por uma plataforma de mediação de dados [Lenzerini 2002, Wiederhold 1992]. Em IoT, não há padronização entre os dispositivos, seja em relação aos protocolos de comunicação, ao modo de operação ou aos modelos de dados usados. Em especial, os dados coletados podem estar em diferentes formatos (*e.g.* texto, tabelas, objetos, etc) e apresentarem sintaxes e semânticas distintas. As plataformas de mediação devem permitir a integração e interoperabilidade dos diferentes modelos tecnológicos, semânticos e lógicos (*i.e.* dados) existentes.

Sobre modelo de dados, entende-se pela estrutura de representação dos sensores, atuadores, processos e as observações trafegadas como *stream* de dados. A importância de adotar um modelo de dados reflete questões como capacidade de descrição, indexação e busca de sensores e atuadores, integração entre diferentes sistemas, busca de dados (*i.e.* histórico ou em tempo real) entre outros. A interoperabilidade entre modelos heterogêneos é usualmente atendida através da adaptação dos modelos de origem heterogênea para modelos padronizados suportados pela plataforma [Wiederhold 1992]. No mais, é importante que as plataformas suportem diferentes protocolos de comunicação para viabilizar a transmissão de dados e o acesso a informação para diversos tipos de dispositivos.

O modelo de dados ainda pode ter associado tecnologias semânticas que fornecem informações sobre o significado dos dados disponíveis e seus relacionamentos semânticos. A associação de semântica abre a possibilidade de alguns aprimoramentos tais como integração automática de dados e geração de novos conhecimentos através de inferência, juntamente com outras técnicas inteligentes.

Devido ao grande volume de dispositivos e dados é importante que as plataformas utilizem mecanismos para garantir que os dados sejam coletados de forma adequada, integrando-os e armazenando-os, em alguma forma escalável de armazenamento com a ajuda de, por exemplo, infraestruturas de *Cloud Computing* e bases de dados NoSQL, permitindo a análise por diversos mecanismos (*e.g.*, mineração de dados, contexto computacional) que possam estabelecer correlações e gerar informação de maior valor agregado [Li et al. 2012].

Levando tudo isso em consideração, as plataformas de mediação de dados em IoT, embora possam variar de acordo com os seus objetivos, possuem um conjunto de requisitos essenciais para viabilizar a integração e processamento de dados [Wiederhold and Genesereth 1997, Wiederhold 1992]. Esses requisitos podem ser resumidos de acordo com a Tabela 1.

Tabela 1. Principais Requisitos

1. Modelo de Dados Padrão:
I. Conjunto mínimo de atributos obrigatórios;
II. Capacidade de extensão;
III. Aderência a especificações de modelo de dados existentes;
IV. Semântica;
2. Funcionalidades de Gerenciamento de Dados:
I. Serviços de registro e descoberta dos dispositivos;
II. Serviços armazenamento e processamento de dados;
III. Suporte a protocolos de acesso aos dados;
IV. Provedimento de segurança e privacidade dos dados;
V. Suporte a Contexto;
VI. Suporte a <i>Complex Event Processing</i> ;
VII. Escalabilidade

3. Trabalhos Relacionados

Avaliação de soluções para o ambiente de IoT é um campo de pesquisa que tem sido alvo de vários estudos reportados na literatura. No que diz respeito à avaliação e *surveys* dessas plataformas de mediação de dados, há trabalhos similares ao nosso, entretanto nenhum deles as avalia sob a ótica de modelo de dados e funcionalidades de gerenciamento de dados. Além disso, a grande maioria desses trabalhos cobre uma pequena quantidade de plataformas analisando apenas aquelas mais citadas por outros trabalhos. Isso faz com que plataformas emergentes não sejam mapeadas pela literatura.

Os trabalhos [Balamuralidhara et al. 2013] e [Castro et al. 2012] elaboraram *surveys* de algumas plataformas de mediação de dados no âmbito da IoT e *Machine-to-Machine* (M2M). Ambos apresentam um panorama geral da evolução de tecnologias e soluções emergentes para, posteriormente, definir as tendências e requisitos essenciais destas áreas. Castro e demais [Castro et al. 2012] argumentam que versatilidade, escalabilidade, segurança e padronização são funcionalidades essenciais para as plataformas de mediação de dados. Balamuralidhar e colegas classificaram os requisitos fundamentais em quatro grande grupos, sendo eles: conectividade, segurança e privacidade, análise, plataformas de sensoriamento [Balamuralidhara et al. 2013]. Ambos os trabalhos apesar de conter também uma análise de plataformas de mediação de dados, possuem um escopo restrito a cerca de 10 soluções. Além disso, os trabalhos apresentam um breve resumo sobre cada uma das plataformas. Inclusive, o trabalho [Balamuralidhara et al. 2013] não apresenta uma comparação entre suas ferramentas.

[Mazhelis and Tyrvaïnen 2014] seguiram a mesma abordagem dos trabalhos anteriores. Os autores avaliaram 12 plataformas de mediação de dados de acordo com funcionalidades consideradas por eles essenciais para as plataformas de mediação. Os requisitos elencados foram suporte a múltiplos elementos de aplicação (*i.e.* dispositivos, *gateways*, interface de usuário e interface web), registro e gerenciamento dos elementos, suporte a monitoramento e monetização. Diferente do nosso trabalho, não há neles um levantamento quanto ao tipo de dados suportado e nem sobre funcionalidades de suporte a semântica e contexto computacional.

O trabalho realizado por Köhler e demais [Köhler et al. 2014] argumenta que vários *surveys* não apresentam critérios transparentes de seleção das ferramentas, selecionando apenas um conjunto pequeno de ferramentas. Em função disso, eles realizaram a seleção através de um grupo focal com sete especialistas. Por sua vez, os especialistas escolheram 6 plataformas com base em fatia de mercado, posicionamento estratégico e grau de inovação. Além disso, eles propuseram um modelo de referência arquitetural para análise das plataformas.

Apesar de relevante, o modelo proposto engloba critérios já analisados por outros trabalhos, tais como [Castro et al. 2012, Mazhelis and Tyrvaiven 2014].

O estudo mais exaustivo foi realizado por Mineraud e colegas em [Mineraud et al. 2015] que analisaram 37 plataformas. Apesar deste trabalho ser aquele mais similar ao nosso em termos de representatividade de amostra e em relação a interseção de alguns critérios (*e.g.* tipo de arquitetura, licença de uso, protocolos de acesso e funcionalidades), ele difere quanto ao foco. Enquanto o nosso trabalho tem como objetivo avaliar as plataformas no que se refere a mediação de dados, eles buscam realizar o levantamento das lacunas a serem preenchidas pelas plataformas.

4. Método de Pesquisa

O objetivo deste trabalho é realizar um *survey* comparativo das plataformas de mediação de dados para IoT. Em geral, *surveys* permitem o levantamento de conceitos, modelos, características e tendências emergentes que aparecem de forma desconexa e esparsa na literatura. Visto que a temática deste trabalho envolve o uso de soluções mercadológicas, estas podem não participar de mapeamentos e revisões sistemáticas da literatura por não terem sido publicados nos veículos convencionais científicos. Então, as revisões exploratórias são essenciais porque possibilitam a descoberta de trabalhos oriundos de diversos veículos de publicação “cinza”, tais como: sites, manuais, relatórios técnicos e *white papers*.

Segundo [Pollitt 2007], revisões exploratórias não envolvem a definição de processo de pesquisa. Usualmente, consistem no levantamento de trabalhos com a informação desejada para, então, classificá-los a partir de métodos específicos. Assim como na revisão sistemática [Kitchenham et al. 2009], os trabalhos devem ser analisados para a extração dos conceitos e respostas às perguntas de pesquisa.

O presente *survey* foi realizado em três etapas bem definidas: (i) definição dos critérios de avaliação, (ii) seleção das plataformas, (iii) extração dos dados e análise. A análise das plataformas de mediação selecionadas após as etapas é apresentada na Seção 5.

4.1. Critérios de Avaliação

Os critérios de avaliação das plataformas foram construídos com base nas funcionalidades essenciais para mediação de dados elencadas na Seção 2. Além desses, foram adicionados um grupo de critérios de cunho geral com o objetivo de ajudar a classificar as plataformas. Em particular, o critério Q1 (*i.e.* Tipo de Arquitetura) substitui o requisito escalabilidade definido como uma funcionalidade essencial. Essa substituição foi realizada em virtude da arquitetura ser a forma mais utilizada pelas plataformas para alcançar escalabilidade. Os critérios de avaliação definidos estão apresentados na Tabela 2.

4.2. Seleção das Plataformas

A seleção das plataformas mostrou-se uma tarefa complexa, pois o presente *survey* teve como desafio cobrir o maior número de plataformas de mercado possíveis. Grande parte das plataformas utilizadas em mercado não estão mapeadas em bases de dados de artigos, como: Scopus, IEEE Xplore, Science Direct e ACM Digital Library. Neste sentido, inicialmente, foi realizada a busca de plataformas através do *site* de busca GoogleTM, usando a seguinte *string*:

Tabela 2. Critérios de Avaliação

Grupo I. Características Gerais:
Q1: Tipo de arquitetura: <i>Cloud</i> , Centralizada ou Descentralizada;
Q2: Tipo de licença de uso: <i>OpenSource</i> ou Comercial;
Q3: Qualidade da documentação: <i>inexistente, ad-hoc</i> , completa, informal;
Q4: Ano de lançamento;
Grupo II. Modelo de Dados
Q5: Padronização de modelo de dados: Sim ou Não;
Q6: Define tipos obrigatórios: Sim ou Não;
Q7: Formatos de dados: JSON, XML, CVS e outros;
Q8: Permite extensão: Sim ou Não;
Q9: Aderência a especificações de modelo de dados: Não ou, por exemplo, SensorML, NGSII;
Q10: Suporte a semântica: Sim ou Não;
Grupo III. Funcionalidades de Gerenciamento de Dados
Q11: Provimento de serviços de registro e descoberta dos dispositivos: Sim ou Não;
Q12: Provimento de serviços armazenamento e processamento de dados: Sim ou Não;
Q13: Suporte a protocolos de acesso aos dados: Sim ou Não;
Q14: Provimento de segurança e privacidade dos dados: Sim ou Não;
Q15: Suporte a Contexto: Sim ou Não;
Q16: Suporte a <i>Complex Event Processing</i> : Sim ou Não;

(“Internet of Things” OR IoT OR “Internet das coisas” OR “Machine-to-Machine” OR “M2M”) AND (architecture OR middleware OR platform OR solution)

Os resultados mais relevantes dessa estratégia vieram de sites especializados em acompanhar tendências de mercado em IoT. No entanto, esta abordagem retornou muitas páginas redundantes ou páginas que não apresentavam conteúdo relacionado com a busca. Desta forma, foi aplicado um método alternativo que consistia na procura de artigos relacionados para a partir deles extrair as plataformas a serem analisadas. As bases escolhidas para esta nova abordagem de pesquisa foram SCOPUS e Science Direct.

Como critério de seleção dos artigos, foi definido a seleção de trabalhos que tenham como foco a realização de *survey*, levantamentos ou análises de plataformas ou soluções de mediação de dados para IoT. Para esta estratégia foi definida a seguinte *string* de busca:

(“Internet of Things” OR IoT OR “Internet das coisas” OR “Machine-to-Machine” OR “M2M”) AND (survey OR levantamento OR analysis OR análise) AND (architecture OR middleware OR platform OR solution)

Como critério de exclusão dos artigos, foram utilizados os seguintes filtros:

- O trabalho não apresenta um levantamento de plataformas de mediação de dados para IoT;
- A plataforma foi desenvolvida para uso específico;
- O trabalho não está escrito em inglês ou português;
- O trabalho não está disponível na Internet.

Aplicando a estratégia de busca e os critérios de inclusão, foram mapeadas ao todo 41 plataformas. No entanto, 6 delas eram de propósito específico e, portanto, pelo critério de exclusão não participaram da análise do mapeamento. O objetivo deste critério de exclusão é manter apenas plataformas extensíveis a qualquer domínio de aplicação. Foram consideradas como uso específico aquelas cujo escopo de aplicação é bem definido e delimitado, como por exemplo: monitoramento médico, automação residencial, acompanhamento de cadeia de suprimentos entre outros. Outro resultado dessa estratégia de busca foi a seleção dos trabalhos relacionados, cujo artigos principais estão relatados na Seção 5.

4.3. Extração dos Dados

Nesta fase, buscou-se obter informações das plataformas para responder os critérios de avaliação previamente definidos. Os dados foram extraídos a partir de documentação oficial de cada plataforma, podendo incluir artefatos como manuais, diagramas arquiteturais, *white papers*, fóruns, vídeos, apresentações e páginas de divulgação da plataforma.

Infelizmente, não foi possível extrair dados para responder todos os critérios de avaliação. Isso acontecia essencialmente devido a dois fatores: (1) documentação oficial incompleta ou inexistente e (2) documentação oficial privada. Neste último caso, a documentação somente era acessível a partir da aquisição de um *development kit* ou através de um login corporativo.

A extração dos dados também envolveu uma interpretação subjetiva da documentação, exigindo a experiência dos envolvidos em identificar as respostas semelhantes que usaram termos sinônimos, o que gerou muitas discussões internas. Neste mapeamento, cada plataforma recebeu um código que foi utilizado para facilitar as referências durante as fases da análise. As plataformas mapeadas e seus respectivos códigos de referência são encontrados na Tabela 3.

5. Análise

Esta seção faz uma análise das plataformas de mediação de dados para Internet das Coisas mapeadas, orientada pelas questões de pesquisa elencadas na Seção 4.

5.1. Características Gerais

A seguir serão respondidas as questões de pesquisa *Q1*, *Q2*, *Q3* e *Q4*, definidas na Seção 4.

Pela Figura 1, pode-se acompanhar o crescimento da quantidade de plataformas lançadas ao longo dos anos, bem como suas respectivas licenças de uso. Percebe-se que a maioria das plataformas disponíveis atualmente são comerciais, tendo sido lançadas principalmente nos anos de 2013 e 2014. Já as iniciativas *opensource* tiveram uma participação representativa apenas no ano de 2014, com quatro plataformas. Foram elas: [P08], [P19], [P22] e [P35]. As plataformas *opensource* lançadas nos anos 2010, 2011 e 2013 foram, respectivamente: [P12], [P21] e [P31] (ver Tabela 3).

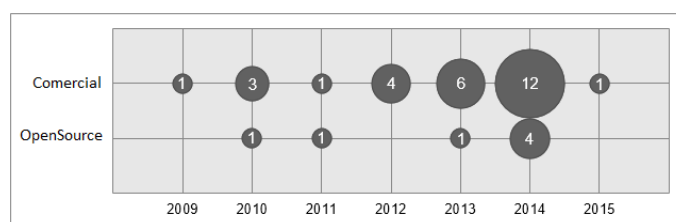


Figura 1. Ano de Lançamento x Licença das plataformas mapeadas

Como visto na Seção 4, as questões de pesquisa foram extraídas diretamente da documentação das plataformas. No entanto, devido aos diferentes níveis de profundidade de conteúdo apresentado, definimos quatro categorias para representar a qualidade da documentação disponível. Foram elas: inexistente, informal, *ad hoc* e completa.

Documentação *inexistente* é aquela restrita apenas a clientes ou não é existente. *Informal* é aquela que não obedece uma sequência linear de conteúdo e é organizada de forma

Tabela 3. Plataformas mapeadas e Critérios gerais de avaliação

Código	Plataforma	URL	Ano	Licença	Documentação	Arquitetura
[P01]	2lemetry	http://2lemetry.com/iot-platform/	2012	Comercial	adhoc	Cloud
[P02]	Ameee	http://blog.ameee.com/products/ameeerealttime/	2012	Comercial	Informal	-
[P03]	Arrayent	http://www.arrayent.com/platform/iot-platform-overview	2014	Comercial	Inexistente	Cloud
[P04]	Axeda	http://www.axeda.com	2010	Comercial	Adhoc	Cloud
[P05]	Aylandnetworks	https://www.aylandnetworks.com	2013	Comercial	Inexistente	Cloud
[P06]	Bosch IoT Suite	https://www.bosch-si.com/products/bosch-iot-suite/benefits.html	2015	Comercial	Inexistente	Cloud
[P07]	Carriots	https://www.carriots.com	2012	Comercial	Adhoc	Cloud
[P08]	DeviceHive	http://devicehive.com/	2014	Open Source	Adhoc	Cloud
[P09]	Digi	http://www.digi.com/cloud-overview	2013	Comercial	Adhoc	Cloud
[P10]	Evrythng	https://evrythng.com	2012	Comercial	Adhoc	Cloud
[P11]	Exosite	http://exosite.com/products/onep	2009	Comercial	Adhoc	Cloud
[P12]	Fi-ware	http://www.fi-ware.org	2011	Open Source	Adhoc	Cloud
[P13]	Glassbeam	http://www.glassbeam.com/spl-technology	2013	Comercial	Inexistente	Cloud
[P14]	Greenwave Systems	http://www.greenwavesystems.com	2014	Comercial	Inexistente	Cloud
[P15]	IBM Bluemix	https://internetofthings.ibmcloud.com	2014	Comercial	Adhoc	Cloud
[P16]	Intel IoT	http://www.intel.com/content/www/us/en/internet-of-things	2014	Comercia	Adhoc	Cloud
[P17]	Iobridge Connect	http://connect.iobridge.com	2014	Comercial	Adhoc	Cloud
[P18]	Jasper	https://www.jasper.com	2014	Comercial	Inexistente	-
[P19]	Kaaproject	http://www.kaaproject.org	2014	Open Source	Completa	Híbrida
[P20]	Kii	http://en.kii.com/platform/internet-of-things	2014	Comercial	Informal	Cloud
[P21]	Nimbits	http://www.nimbits.com	2013	Open Source	Informal	Cloud
[P22]	Octoblu	http://www.octoblu.com	2014	Open Source	Informal	Cloud
[P23]	Parstream	https://www.parstream.com	2014	Comercial	Inexistente	Cloud
[P24]	Prismtech	http://www.prismtech.com/vortex	2014	Comercial	Informal	Cloud
[P25]	Samsung IoT	http://developer.samsung.com/iot	2014	Comercial	Inexistente	Cloud
[P26]	m2mAIR Cloud	http://www.telit.com/trym2maircloud/	2014	Comercial	Inexistente	Cloud
[P27]	Seecontrol	http://www.seecontrol.com	2010	Comercial	Inexistente	Cloud
[P28]	Smarthings	http://www.smarthings.com/developers	2013	Comercial	Completa	-
[P29]	Sofia2	http://sofia2.com/sofia2incloud_en.html	2014	Comercial	Completa	Cloud
[P30]	TempoIQ	https://tempoIQ.com	2011	Comercial	Adhoc	Cloud
[P31]	Thingspeak	https://thingspeak.com	2010	Open Source	Inexistente	Cloud
[P32]	Thingworx	http://www.thingworx.com	2010	Comercial	Inexistente	Cloud
[P33]	Xively	https://xively.com	2013	Comercial	Adhoc	Cloud
[P34]	Zatar	http://www.zatar.com/features	2013	Comercial	Adhoc	Cloud
[P35]	Zettajs	http://www.zettajs.org	2014	Open Source	Adhoc	Cloud

Sendo: - não informado

distribuída. *Adhoc* é aquela que tem uma estrutura já organizada em tópicos e aborda os principais assuntos de interesse da plataforma, mas ainda de forma pontual (por exemplo, tutoriais para registro de sensores, busca, gerenciamento de *feed*, alertas entre outros). Por fim, *completa* é aquela documentação com todos os insumos necessários para que de usuários comuns a desenvolvedores consigam achar informações de interesse.

De acordo com a Tabela 3, um total de 3 plataformas possui documentação completa, 15 *adhoc*, 12 são inexistentes e 5 são informais.

5.2. Modelo de dados

A seguir serão respondidas as questões de pesquisa *Q5*, *Q6*, *Q7*, *Q8*, *Q9* e *Q10*, definidas na Seção 4.

A Tabela 4 apresenta o posicionamento das plataformas mapeadas quanto às questões relacionadas ao modelo de dados adotado. Percebe-se que 20 plataformas afirmam possuir um modelo de dados, 5 não adotam e 10 não apresentam esta informação em sua documentação. Destas 20 plataformas, 7 adotaram algum padrão de modelo de dados proposto por algum consórcio, são elas: [P12] que adota os padrões NGSI e SensorML, [P33] que adota o EEML, [P19] que adota o Apache Avro, [P02] que adota o AMON, [P35] com o Siren, [P29] com o AMON e o *Common Alerting Protocol* (CAP) e, por fim, [P24] com o *Object Management Group* (OMG) e o *Data Distribution Service* (DDS).

Sobre os formatos de dados, 22 das plataformas dão suporte a dados em JSON, 9 suportam dados em XML, 4 em CSV, 3 em RDF, 1 em formato textual e 11 não explicitam em sua documentação qual o formato de dados suportado. O somatório ultrapassa a quantidade de plataformas, uma vez que algumas delas suportam mais de um tipo de dados. Estes números evidenciam que JSON vem se consolidando como um padrão de formato de dados ampla-

Tabela 4. Questões relativas ao modelo e gerenciamento de dados

Código	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q14	Q15	Q16
[P01]	Sim	Sim	Qualquer um, mas armazenado apenas em JSON	Sim	Não	-	Sim	Sim	Sim	-	Sim
[P02]	Sim	Sim	JSON	Sim	Sim	-	Sim	Sim	-	-	-
[P03]	-	-	-	-	-	-	Sim	Sim	Sim	-	-
[P04]	Sim	Sim	JSON	Sim	Não	-	Sim	Sim	Sim	-	Sim
[P05]	-	-	-	-	-	-	Sim	Sim	-	-	-
[P06]	Sim	-	-	Sim	Não	-	Sim	Sim	Sim	-	-
[P07]	Não	-	JSON e XML	-	Não	-	Sim	Sim	-	-	-
[P08]	Não	Não	JSON	Não	Não	-	Sim	Sim	-	-	-
[P09]	Sim	Não	JSON e XML	Não	Não	-	Sim	Sim	-	-	-
[P10]	Sim	Sim	JSON	Sim	Não	-	Sim	Sim	-	-	-
[P11]	Sim	Sim	JSON, XML, CSV	Sim	Não	-	Sim	Sim	Sim	-	-
[P12]	Sim	Sim	JSON, XML, RDF	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Sim
[P13]	-	-	TEXT, XML, JSON, CSV	-	-	Sim	Sim	Sim	Sim	Sim	-
[P14]	-	-	-	-	-	-	Sim	Sim	-	-	-
[P15]	Sim	Sim	JSON	Sim	Não	-	Sim	Sim	Sim	-	-
[P16]	Sim	Sim	-	Sim	-	-	Sim	Sim	Sim	-	-
[P17]	Sim	Sim	JSON, XML	Não	Não	-	Sim	Sim	-	-	-
[P18]	-	-	-	-	-	-	-	-	-	-	-
[P19]	Sim	Não	JSON	Sim	Sim	-	Sim	Sim	Sim	-	-
[P20]	-	-	JSON	Sim	Não	-	Sim	Sim	-	-	-
[P21]	Não	-	JSON	Não	Não	-	Sim	Sim	-	-	-
[P22]	Não	Não	JSON	Sim	Não	-	Sim	Sim	-	-	Sim
[P23]	-	-	-	-	-	-	-	-	-	-	-
[P24]	Sim	Sim	XML	Sim	Sim	-	Sim	Sim	-	-	-
[P25]	-	-	-	-	Não	-	Sim	Sim	Sim	-	-
[P26]	-	-	-	-	-	-	Sim	Sim	Sim	-	-
[P27]	Sim	Sim	-	Sim	-	-	-	-	-	-	-
[P28]	Sim	-	-	-	-	-	Sim	Sim	Sim	-	-
[P29]	Sim	Sim	JSON	Sim	Sim	-	Sim	Sim	-	-	-
[P30]	Sim	Sim	JSON	Sim	Não	-	Sim	Sim	-	-	-
[P31]	Sim	Não	JSON, XML, CSV	Não	Não	-	Sim	Sim	-	-	-
[P32]	-	-	-	-	-	-	Sim	Sim	-	-	-
[P33]	Sim	Não	JSON, XML, CSV	Sim	Sim	-	Sim	Sim	-	-	-
[P34]	Não	Sim	JSON	Sim	Não	-	Sim	Sim	-	-	-
[P35]	Sim	Sim	JSON	Não	Sim	-	Sim	Sim	-	-	-

Sendo: - não informado

mente aceito para plataformas de Internet das Coisas. De fato, ele não é apenas o formato mais adotado por essas plataformas, mas é aproximadamente 2, 4 vezes mais utilizado que XML, o segundo mais utilizado.

Ao escolher uma plataforma de Internet das Coisas sob a perspectiva de dados, além de saber o formato de dados suportado, e eventualmente os padrões de formato de dados utilizados, é interessante também levar em consideração a capacidade que o modelo de dados tem de ser extensível. Sendo assim, através da Tabela 4, percebe-se que 18 plataformas permitem ter seu modelo de dados estendido, 6 não permitem e 11 não informam ter essa capacidade.

Como um limitador da observação feita anteriormente sobre a capacidade de estender o modelo de dados, há a exigência que alguns modelos de dados tem sobre alguns elementos ou atributos em instâncias do seu esquema. Por exemplo, [P30] exige que cada sensor armazene seus dados como *datapoints*, sendo cada *datapoint* uma díade *timestamp/valor*. Não necessariamente o tipo obrigatório, quando é exigido, atue somente sobre a estrutura das observações. [P19], por exemplo, não exige um tipo obrigatório em seus dados, mas por fazer uso da especificação completa do Apache Avro, exige apenas que o elemento root do JSON seja do tipo *record*.

Muito embora as plataformas tenham algum nível de semelhança entre si, levando em consideração as funcionalidades da Tabela 4, apenas [P12] oferece algum nível maduro de interpretação semântica de seus dados.

5.3. Funcionalidades de Gerenciamento de Dados

A seguir serão respondidas as questões de pesquisa Q11, Q12, Q13, Q14, Q15 e Q16, definidas na Seção 4.

De acordo com a pergunta Q13: *Protocolos de Comunicação*, este trabalho buscou quais protocolos são suportados pelas plataformas mapeadas. A Tabela 5 elenca os protocolos utilizados, sejam eles utilizados para a comunicação entre os dispositivos e os *gateways*, ou entre os *gateways* e a plataforma propriamente dita.

Tabela 5. Protocolos utilizados

Protocolo	Plataformas
802.15.4	[P10]
AMMP	[P04]
AMQP	[P32], [P24]
AllJoyn	[P22]
BLE	[P22], [P10]
Bluetooth	[P25], [P04]
Bluetooth low energy	[P25]
CFX	[P24]
Cellular	[P10], [P16]
CoAP	[P22], [P34], [P19], [P10], [P32], [P11], [P01]
DDS	[P32]
Ethernet	[P17], [P19], [P28]
HTTP Ultralight	[P12]
JMS	[P24]
MQTT	[P04], [P12], [P33], [P22], [P19], [P10], [P32], [P25], [P16], [P01], [P07], [P29], [P24], [P15], [P26]
ModBus	[P04]
MometID	[P24]
OMA Lightweight M2M	[P12]
REST	[P04], [P33], [P22], [P19], [P31], [P10], [P02], [P21], [P11], [P17], [P30], [P35], [P20], [P25], [P06], [P16], [P01], [P27], [P07], [P29], [P24], [P13], [P15], [P08], [P26]
RF	[P16]
SOAP	[P04]
Serial	[P22]
TCP	[P04], [P34], [P19], [P24]
UDP	[P04], [P34], [P11], [P27]
Websockets	[P24]
WiFi	[P22], [P19], [P10], [P25], [P03]
XMPP	[P19], [P32], [P24]
Yo	[P22]
ZigBee	[P04], [P19], [P10], [P17], [P25], [P03], [P14], [P28]
JenNet	[P14]
NFC	[P10], [P25]
Z-Wave	[P03], [P14], [P28]
DeviceHive Binary Protocol	[P08]

O protocolo mais utilizado foi o REST, sendo suportado por 25 das 35 plataformas. Seguindo do REST, tem-se MQTT (*Message Queuing Telemetry Transport*) e CoAP (*Constrained Application Protocol*) como as três mais utilizadas. Cinco plataformas não informaram quais protocolos são utilizados.

Uma característica desejável em plataformas de mediação de dados para IoT é o uso de contexto computacional. De acordo com [Dey et al. 2001], contexto é qualquer informação que pode ser usada para caracterizar a situação de qualquer entidade. Uma entidade é uma pessoa, lugar ou objeto que é considerado relevante para a interação entre um usuário ou uma aplicação, incluindo os próprios usuários e as aplicações. Dessa forma, sistemas cientes de contexto para IoT tem um grande potencial de aplicação, por exemplo executando buscas mais precisas dos sensores mais adequados à execução de uma tarefa em específico, ou trazendo informações de monitoramento de uma cidade adequadas ao perfil do usuários responsável pelo monitoramento. Das plataformas mapeadas, apenas [P12] e [P13] dão suporte a contexto.

Por fim, espera-se cada vez mais que plataformas para IoT possuam módulos que tratem do Processamento de Eventos Complexos (CEP - *Complex Event Processing*). De acordo com [Zang and Fan 2007], CEP vem se mostrando como uma poderosa ferramenta para identificar relacionamento entre eventos distintos, levando em consideração fatores como temporalidade, casualidade similaridade em *streams* de dados (ou de outros eventos) em tempo real. Como benefícios, estão a automação do sistema, suporte a tomada de decisão por especialistas, baixo acoplamento entre provedores e consumidores de dados, suporte à detecção de relação entre eventos aparentemente distintos, agregação e composição de novos

eventos complexos em novos eventos complexos abstratos [Buchmann and Koldehofe 2009] e, por fim, CEP se mostra como uma estratégia de implementação de contexto computacional baseado em regras. Considerando as plataformas mapeadas, verificou-se que [P01], [P04], [P12] e [P22] suportam o processamento de eventos complexos, não sendo possível extrair essa informação para as outras plataformas a partir de sua documentação.

6. Conclusão

O mercado de plataformas para mediação de dados em Internet das Coisas tem crescido no Brasil e no mundo. Todavia, por apresentarem diversos requisitos e características distintas, encontrar a plataforma que melhor se adapte as diferentes necessidades da indústria é um grande desafio.

Com o objetivo de ajudar o mercado a se beneficiar da escolha mais adequada quanto a plataforma de mediação de dados para IoT que melhor se adapte ao seu perfil, este trabalho fez um mapeamento de 35 plataformas *opensource* ou comerciais mais utilizadas atualmente. Como resultado, neste artigo foram analisadas questões relacionadas ao modelo de dados utilizado, os padrões de dados adotados, restrições sobre a estrutura desses dados, protocolos de comunicação suportados e funcionalidades ofertadas. Além disso, foram observadas características desejáveis, como contexto computacional, semântica e processamento de eventos complexos.

Como trabalho futuro, pretendemos aumentar o número de bibliotecas pesquisadas, aumentando consequentemente a cobertura deste trabalho. Desta forma, novas pesquisas, com foco maior nas necessidades da indústria, poderão ser direcionadas a partir deste trabalho.

7. Agradecimentos

Este trabalho teve suporte dado pelo INES e pela FACEPE sob o processo IBPG-0773-1.03/13. Eduardo A. Oliveira gostaria ainda de agradecer ao CNPq pelo apoio com seu pós-doutorado (provido sob processo nº BEX 9213/13-9). E por fim, Marcelo Iury de Sousa Oliveira que é bolsista do CNPq–Brasil na modalidade doutorado.

Referências

- Atzori, L., Iera, A., and Morabito, G. (2010). The internet of things: A survey. *Computer networks*, 54(15):2787–2805.
- Axeda (2015). Axeda. <http://www.axeda.com>. Acessado em 1 de abril de 2015.
- Balamuralidhara, P., Misra, P., and Pal, A. (2013). Software platforms for internet of things and m2m. *Journal of the Indian Institute of Science*, 93(3):487–498.
- Buchmann, A. and Koldehofe, B. (2009). Complex event processing. *it-Information Technology Methoden und innovative Anwendungen der Informatik und Informationstechnik*, 51(5):241–242.
- Castro, M., Jara, A. J., and Skarmeta, A. F. (2012). An analysis of m2m platforms: challenges and opportunities for the internet of things. In *Innovative Mobile and Internet Services in Ubiquitous Computing (IMIS), 2012 Sixth International Conference on*, pages 757–762. IEEE.

- Council, C. (2008). Date/dept/type info. *City*, 2:19.
- Dey, A. K., Abowd, G. D., and Salber, D. (2001). A conceptual framework and a toolkit for supporting the rapid prototyping of context-aware applications. *Human-computer interaction*, 16(2):97–166.
- Guinard, D., Trifa, V., Karnouskos, S., Spiess, P., and Savio, D. (2010). Interacting with the soa-based internet of things: Discovery, query, selection, and on-demand provisioning of web services. *Services Computing, IEEE Transactions on*, 3(3):223–235.
- INTEL (2014). Intel iot platforms. <http://www.intel.com.br/content/www/br/pt/internet-of-things/iot-platform-solution-brief.html>. Acessado em 1 de abril de 2015.
- Kitchenham, B., Brereton, O. P., Budgen, D., Turner, M., Bailey, J., and Linkman, S. (2009). Systematic literature reviews in software engineering—a systematic literature review. *Information and software technology*, 51(1):7–15.
- Köhler, M., Wörner, D., and Wortmann, F. (2014). Platforms for the internet of things – an analysis of existing solutions. In *5th Bosch Conference on Systems and Software Engineering (BoCSE)*.
- Lenzerini, M. (2002). Data integration: A theoretical perspective. In *Proceedings of the twenty-first ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems*, pages 233–246. ACM.
- Li, T., Liu, Y., Tian, Y., Shen, S., and Mao, W. (2012). A storage solution for massive iot data based on nosql. In *Green Computing and Communications (GreenCom), 2012 IEEE International Conference on*, pages 50–57. IEEE.
- Mazhelis, O. and Tyrvaïnen, P. (2014). A framework for evaluating internet-of-things platforms: Application provider viewpoint. In *Internet of Things (WF-IoT), 2014 IEEE World Forum on*, pages 147–152. IEEE.
- Mineraud, J., Mazhelis, O., Su, X., and Tarkoma, S. (2015). A gap analysis of internet-of-things platforms. *arXiv preprint arXiv:1502.01181*.
- Oliveira, E. A., Kirley, M., and Fonseca, J. C. (2015). Device nimbus: An intelligent middleware for smarter services for health and fitness.
- Pollitt, M. M. (2007). An ad hoc review of digital forensic models. In *Systematic Approaches to Digital Forensic Engineering, 2007. SADFE 2007. Second International Workshop on*, pages 43–54. IEEE.
- Thingspeak (2015). Thingspeak. <https://thingspeak.com/>. Acessado em 1 de abril de 2015.
- Wiederhold, G. (1992). Mediators in the architecture of future information systems. *Computer*, 25(3):38–49.
- Wiederhold, G. and Genesereth, M. (1997). The conceptual basis for mediation services. *IEEE Expert*, 12(5):38–47.
- Xively (2015). Xively. <http://www.xively.com>. Acessado em 1 de abril de 2015.
- Yan, L., Zhang, Y., Yang, L. T., and Ning, H. (2008). *The Internet of things: from RFID to the next-generation pervasive networked systems*. CRC Press.
- Zang, C. and Fan, Y. (2007). Complex event processing in enterprise information systems based on rfid. *Enterprise Information Systems*, 1(1):3–23.