



Universidade de Brasília

Instituto de Ciências Exatas
Departamento de Ciência da Computação

Metadados em Internet das Coisas

Jônatas R. S. Pires

Monografia apresentada como requisito parcial
para conclusão do Curso de Engenharia da Computação

Orientadora
Prof.a Dr.a Maristela de Holanda

Brasília
2017



Universidade de Brasília

Instituto de Ciências Exatas
Departamento de Ciência da Computação

Metadados em Internet das Coisas

Jônatas R. S. Pires

Monografia apresentada como requisito parcial
para conclusão do Curso de Engenharia da Computação

Prof.a Dr.a Maristela de Holanda (Orientadora)
CIC/UnB

Prof. Dr. Donald Knuth Dr. Leslie Lamport
Stanford University Microsoft Research

Prof. Dr. Ricardo Pezuol Jacobi
Coordenador do Curso de Engenharia da Computação

Brasília, 15 de novembro de 2017

Dedicatória

Eu dedico este trabalho aos meus pais, Marta e Sérgio. À minha irmã Tata e seu Marido Pedro. Aos meus avós, José e Virgínia.

Agradecimentos

Agradeço aos meus pais, Marta e Sérgio, sem o apoio deles eu não conseguiria chegar ao fim da graduação. À minha irmã Tata e seu marido Pedro que sempre me escutaram, aconselharam e me aguentaram. À minha namorada Amanda que me suportou durante os períodos mais estressantes. À minha orientadora Prof.a Dr.a Maristela de Holanda, uma excelente mentora e amiga. E finalmente à minha cachorrinha Hera, sua inestimável companhia durante as incontáveis horas de trabalho ajudaram tornar este período tolerável.

Resumo

O *resumo* é um texto inaugural para quem quer conhecer o trabalho, deve conter uma breve descrição de todo o trabalho (apenas um parágrafo). Portanto, só deve ser escrito após o texto estar pronto. Não é uma coletânea de frases recortadas do trabalho, mas uma apresentação concisa dos pontos relevantes, de modo que o leitor tenha uma ideia completa do que lhe espera. Uma sugestão é que seja composto por quatro pontos: 1) o que está sendo proposto, 2) qual o mérito da proposta, 3) como a proposta foi avaliada/validada, 4) quais as possibilidades para trabalhos futuros. É seguido de (geralmente) três palavras-chave que devem indicar claramente a que se refere o seu trabalho. Por exemplo: *Este trabalho apresenta informações úteis a produção de trabalhos científicos para descrever e exemplificar como utilizar a classe L^AT_EX do Departamento de Ciência da Computação da Universidade de Brasília para gerar documentos. A classe define um padrão de formato para textos do CIC, facilitando a geração de textos e permitindo que os autores foquem apenas no conteúdo. O formato foi aprovado pelos professores do Departamento e utilizado para gerar este documento. Melhorias futuras incluem manutenção contínua da classe e aprimoramento do texto explicativo.*

Palavras-chave: metadados, IoT, Internet das Coisas

Abstract

O *abstract* é o resumo feito na língua Inglesa. Embora o conteúdo apresentado deva ser o mesmo, este texto não deve ser a tradução literal de cada palavra ou frase do resumo, muito menos feito em um tradutor automático. É uma língua diferente e o texto deveria ser escrito de acordo com suas nuances (aproveite para ler [http://dx.doi.org/10.6061/2Fclinics%2F2014\(03\)01](http://dx.doi.org/10.6061/2Fclinics%2F2014(03)01)). Por exemplo: *This work presents useful information on how to create a scientific text to describe and provide examples of how to use the Computer Science Department's L^AT_EX class. The class defines a standard format for texts, simplifying the process of generating CIC documents and enabling authors to focus only on content. The standard was approved by the Department's professors and used to create this document. Future work includes continued support for the class and improvements on the explanatory text.*

Keywords: metadata, IoT, Internet of Things

Sumário

1	Introdução	1
1.1	Bancos de Dados	1
1.1.1	Definição	1
1.1.2	Esquema de Banco de Dados	1
1.2	Internet das Coisas	1
1.2.1	Definição	2
1.2.2	Tecnologias essenciais	4
1.2.3	Desafios	5
1.3	dados	5
1.3.1	Metadados	6
1.3.2	Metadados de sensores	7
2	Metodologia	8
2.1	Revisão Sistemática	8
2.1.1	Planejamento	8
	Referências	12

Lista de Figuras

1.1 O paradigma Internet das Coisas como um resultado de diferentes visões. . .	2
1.2 Visão geral técnica de IoT.	3

Lista de Abreviaturas e Siglas

IoT Internet das Coisas.

ITU International Telecommunication Union.

RFID Identificação por Radio Frequência.

RSSF Redes de Sensores Sem Fio.

SensorML Linguagem de Modelagem de Sensores.

Capítulo 1

Introdução

O presente trabalho visa a obtenção e gerenciamento de metadados em informações provenientes de dispositivos categorizados em Internet das Coisas (IoT).

1.1 Bancos de Dados

intro

1.1.1 Definição

1.1.2 Esquema de Banco de Dados

1.2 Internet das Coisas

A Internet das Coisas é um novo paradigma tecnológico idealizado como uma conexão global de máquinas e dispositivos capazes de interagir entre si. A proposta de IoT consiste em vários objetos do cotidiano trocando informações mutuamente, através da internet, para serem mais eficientes e realizarem diversas tarefas. Os objetos passam a agir de forma mais inteligente e sensorial, de modo a favorecer diversos setores como: indústria, hospitais, agropecuária, transporte público e muitos outros. A partir desta disponibilidade astronômica de recursos, a IoT é reconhecida com uma das áreas mais importantes em termos de tecnologia do futuro e está recebendo cada vez mais atenção de desenvolvedores, usuários e indústrias.

Um dos objetivos principais da Internet das Coisas é permitir que humanos e máquinas possuam maior consciência de seus arredores. Esse maior entendimento do seu ambiente se torna viável através da utilização de diversos tipos de dispositivos sensíveis (sensores)

e, após a percepção de seu ambiente, é possível realizar ações por meio de dispositivos atuadores ou fazer análises.

A Internet das Coisas surgiu a partir do conjunto de diferentes visões como podemos observar na Figura 1.1, cada qual com seus objetivos específicos.

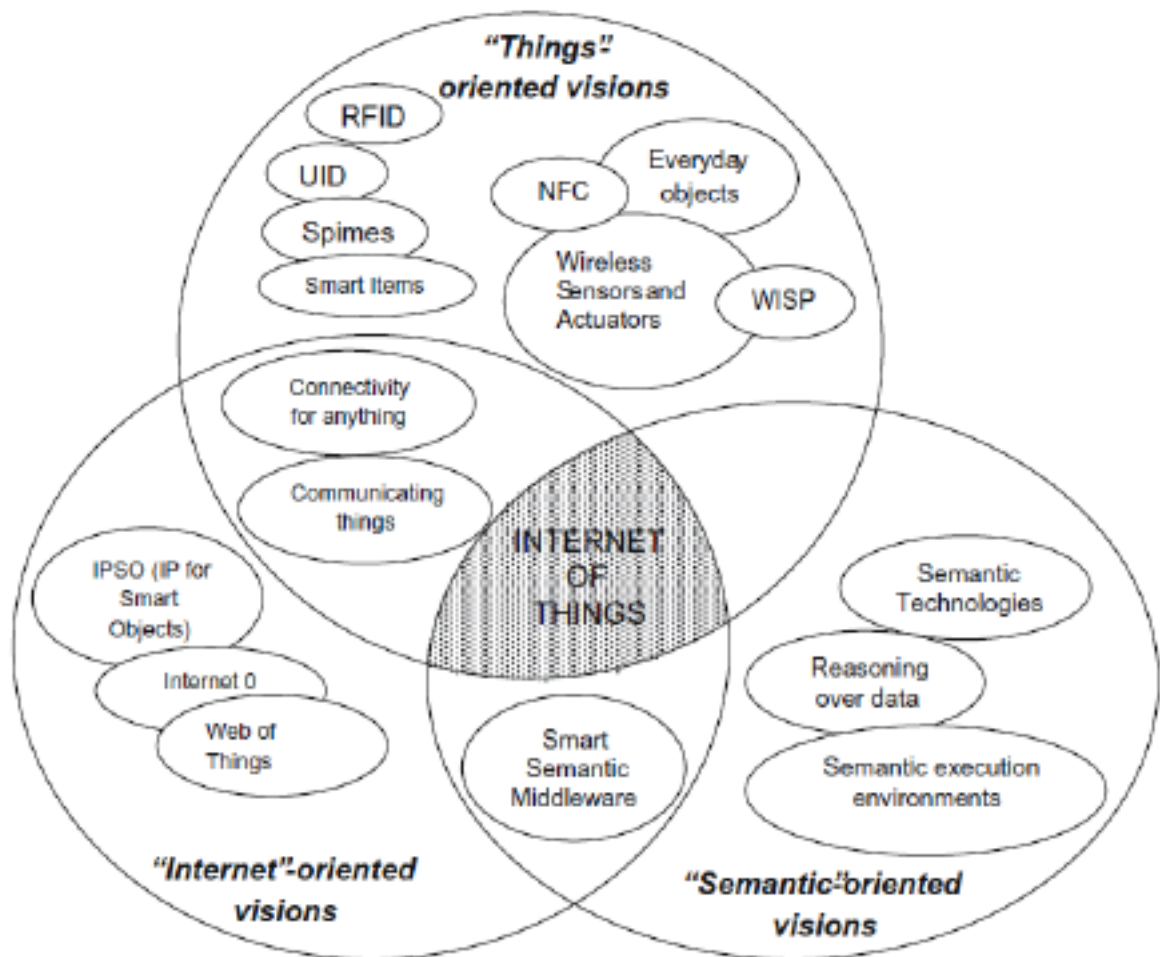


Figura 1.1: O paradigma Internet das Coisas como um resultado de diferentes visões.

1.2.1 Definição

Em 2012, a International Telecommunication Union (ITU) realizou estudos sobre infraestrutura de informação global, aspectos de protocolos de internet e redes da próxima geração. A partir desse estudo foi construída a recomendação ITU-T Y.2060 [1] que trata sobre a Internet das Coisas e possui o intuito de esclarecer o conceito e o escopo de IoT, identificar

as características fundamentais e os requerimentos de alto-nível.

No documento produzido pela ITU, foram consolidadas as definições de:

- Internet das Coisas, "uma infraestrutura global para a Sociedade de Informações, permitindo serviços avançados ao interconectar (fisicamente e virtualmente) coisas devido à existência e evolução da interoperabilidade de tecnologias de comunicação e informação"[1];
- Dispositivo, no contexto de IoT, é um equipamento que, obrigatoriamente, possui a capacidade de comunicação e, opcionalmente, possui capacidade de sensibilidade, atuação, captura de dados, armazenamento de dados e/ou processamento de dados [1];
- Coisas, no contexto de IoT, são "objetos no mundo físico (objetos físicos) ou no mundo das informações (objetos virtuais), os quais são capazes de serem identificados e integrados a uma rede de comunicações". Objetos físicos podem sentir, atuar e conectar. Objetos virtuais podem ser armazenados, processados e acessados.[1]

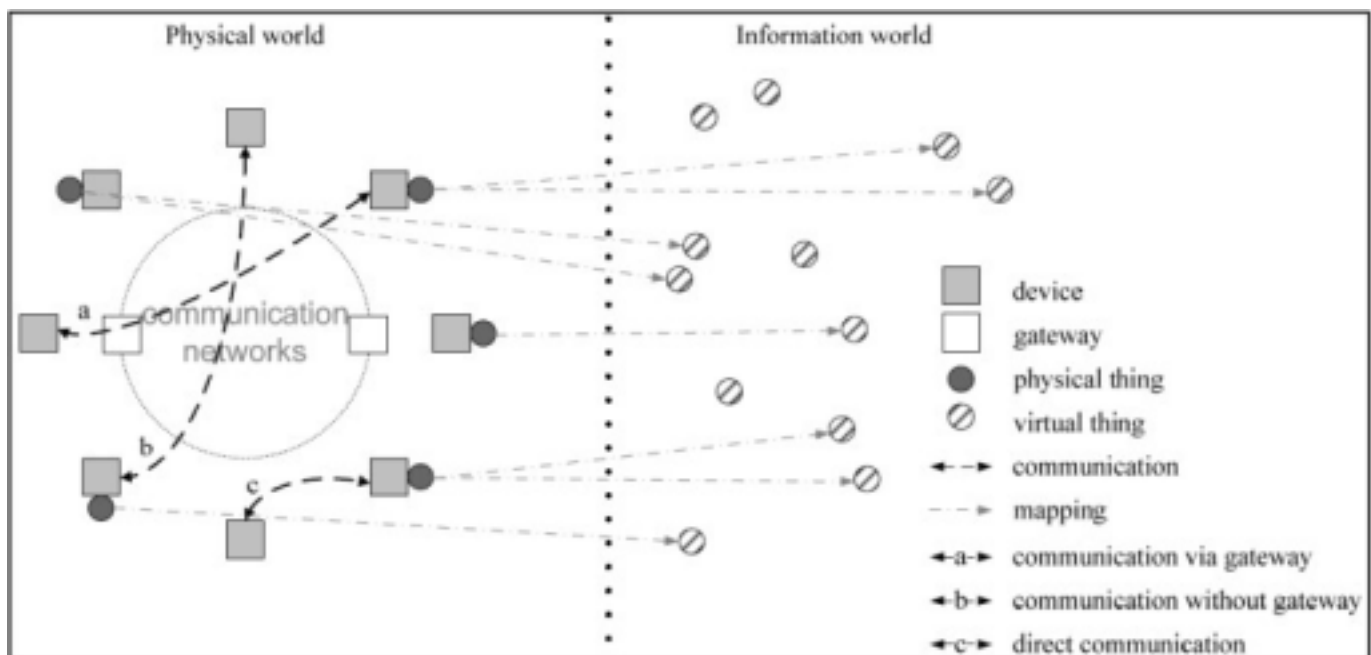


Figura 1.2: Visão geral técnica de IoT.

1.2.2 Tecnologias essenciais

Identificação por Radio Frequência (RFID)

- Esta tecnologia permite identificação automática e captura de informação por meio de rádio frequência. Dividem-se os dispositivos RFID em duas grandes categorias: ativos e passivos. Dispositivos ativos dependem de uma fonte de energia constante para manter ativa e transmitir a informação. Dispositivos passivos não necessitam de energia constante, um campo eletromagnético energiza o dispositivo, o qual se torna apto a transferir a informação contida nele. [2]

Redes de Sensores Sem Fio (RSSF)

- Esta tecnologia consiste na distribuição de dispositivos sensíveis autônomos para monitorar condições físicas ou ambientais e podem cooperar com sistemas RFID para medir de forma mais eficaz localização, temperatura e movimentação, por exemplo. [3]

Middleware

- O middleware é a camada de abstração entre aplicações de software para facilitar para os desenvolvedores realizar a comunicação entre softwares e operações de recebimento e envio de dados. O objetivo do middleware no contexto de IoT é simplificar a integração entre dispositivos heterogêneos.

Computação em nuvem

- Computação em nuvem é um modelo para acesso de recursos compartilhados conforme a necessidade de um serviço. Um dos resultados mais notáveis da IoT é a enorme quantidade de dados gerados por dispositivos conectados à internet [4]. A computação em nuvem é importante para o contexto de Internet das Coisas ao permitir um ambiente com alta escalabilidade.

Aplicações de software

- Aplicações IoT permitem interações dispositivo-dispositivo e humano-dispositivo de uma forma confiável e robusta. As aplicações nos dispositivos devem garantir que as informações são recebidas e processadas de maneira adequada, no momento adequado.

1.2.3 Desafios

A Internet das Coisas possui diversos desafios devido à sua própria concepção, essas dificuldades devem ser ultrapassadas para que a IoT possa ser amplamente e devidamente implantada. Alguns fatores críticos podem ser elencados:

Infraestrutura de rede

O custo para interconectar os dispositivos é alto. Para uma grande rede de sensores é necessário a distribuição de toda infraestrutura, de cabeamento ou infraestrutura sem fio.

Segurança

Uma das principais dificuldades num ambiente de Internet das Coisas é a segurança dos dados que são coletados pela grande quantidade de dispositivos sensíveis da rede. Uma invasão de um sistema crítico pode significar consequências graves em um ambiente automatizado.

Espaço de armazenamento

Um ambiente IoT gera uma grande quantidade de dados, suponhamos um sistema de uma cidade inteligente que possui 10000 sensores de diversas categorias, cada sensor gerando uma mensagem de 10 kB a cada 5 minutos, por dia, temos um total de aproximadamente 29 GB de dados; após um ano teríamos 10,5 TB de dados apenas para esta cidade.

Consumo de energia

É imperativo que o consumo de energia dos dispositivos sensíveis seja o menor possível. Em muitos cenários é improvável a presença de uma rede de energia elétrica e até mesmo de manutenção constante, então os equipamentos devem conseguir se manter funcionais por meio de baterias por uma quantidade de tempo considerável.

1.3 dados

Um ambiente de Internet das Coisas tem como objetivo a compreensão do ambiente em que está situado utilizando as informações geradas por diversos dispositivos sensíveis. Este entendimento é baseado em três tipos de dados [5]:

- Dados gerados pelos dispositivos;
- Dados que descrevem os dispositivos;

- Dados que descrevem o ambiente.

Normalmente, dispositivos IoT tem sua semântica descrita em termos de suas capacidades sensitivas. A semântica do ambiente é determinada de acordo com o domínio da aplicação [6]. Consequentemente, modelos de suporte à decisão são construídos baseados nos metadados que descrevem os dispositivos e seu ambiente.

Os esforços de pesquisa em Internet das Coisas estão principalmente focados nos desafios de interoperabilidade, escalabilidade e integração entre dispositivos heterogêneos [6], entretanto o desafio do dinamismo dos metadados em IoT tem sido inexplorado [7].

1.3.1 Metadados

Metadados são as principais ferramentas para descrever e gerenciar recursos de informações extremamente dinâmicos, como os dados contidos na rede mundial de internet.

Princípios

Os seguintes princípios são considerados as linhas de base para o desenvolvimento de soluções práticas em desafios de semântica e interoperabilidade de dispositivos em qualquer domínio e utilizando qualquer conjunto de metadados [8].

- **Modularidade** de metadados é um princípio chave para a organização de ambientes caracterizados pela diversidade de fontes e estilos de conteúdo e abordagens à descrição de recursos. Permite que projetistas de esquemas de metadados criem novos esquemas baseados em projetos já existentes e se beneficiar de boas práticas já observadas. Em um ambiente com metadados modulares, elementos de informação de diferentes esquemas podem ser combinados de forma interoperável tanto sintaticamente quanto semanticamente.
- **Flexibilidade.** Sistemas de metadados precisam ser flexíveis para acomodar particularidade de uma determinada aplicação. Arquiteturas de metadados devem se adequar facilmente a noção de um esquema base e elementos adicionais necessários para uma aplicação local ou um domínio específico sem comprometer a interoperabilidade proporcionada pelo esquema base.
- **Refinamento.** O nível de detalhes necessário para cada domínio de aplicação pode variar consideravelmente. Para evitar gastos desnecessários com armazenamento e processamento, o processo de desenvolvimento dos padrões de metadados devem permitir que os projetistas escolham o nível de detalhes apropriado para uma dada aplicação.

- **Multilinguismo.** O multilinguismo é essencial ao adotar arquiteturas de metadados que respeitem a diversidade linguística e cultural. Por ter a possibilidade de conectar sistemas de diversas partes do planeta, é importante que a comunicação dos metadados não tenham a linguagem e formatação como desafios a serem ultrapassados.

1.3.2 Metadados de sensores

Metadados de sensor é o modelo que descreve o sensor e suas capacidades como por exemplo:

- Modelo do sensor;
- Localização do sensor;
- Unidade de medida utilizada;
- Grau de confiabilidade.

A Linguagem de Modelagem de Sensores (SensorML) [9] é uma coleção de padrões desenvolvida para representar informações de sensores em formato XML. O propósito da SensorML é:

- Prover descrições de sensores e sistemas de sensores para gerenciamento de inventário;
- Providenciar informação sobre o sensor e sobre processamento;
- Auxiliar o processamento e análise de dados coletados por sensores;
- Suportar informações de geolocalização de valores coletados;
- Fornecer informações de desempenho;
- Providenciar uma descrição explícita sobre o processo em que os dados foram obtidos;
- Prover uma cadeia de processos executável para derivar novos produtos de informação;
- Arquivar propriedades fundamentais e suposições sobre os sistemas de sensores.

Capítulo 2

Metodologia

Este capítulo descreve os métodos utilizados para a realização deste trabalho.

2.1 Revisão Sistemática

Foi definido um processo de revisão que consiste em três fases:

- Planejar;
- Executar;
- Documentar.

A fase de planejamento constitui-se na especificação do protocolo. A fase de execução representa a coleta dos dados de forma a atender as especificações exigidas na fase de planejamento. A fase de documentação implica na consolidação dos dados obtidos.

2.1.1 Planejamento

O objetivo desta revisão sistemática é a identificação de trabalhos acadêmicos que expõem resultados, projeções, explicações ou elucidações sobre o tema de armazenamento de metadados para Internet das Coisas.

Questões de Estudo

Esta revisão tem como objetivo responder às seguintes questões:

- Há estudos sobre o armazenamento de metadados para Internet das Coisas?
- Quais são os métodos mais utilizados para o armazenamento de metadados para IoT?

- Quais são as características dos middlewares utilizados para estes estudos?
- Quais aplicações os metadados estão sendo utilizados nos trabalhos

Estratégia de Busca

É necessário determinar uma estratégia para a realização das buscas nas bases de dados escolhidas.

Definição da String de Busca

- **População:** A população é o gerenciamento de metadados para IoT. Para procurar, foram utilizadas as palavras-chave 'Internet of things metadata' e 'IoT metadata';
- **Intervenção:** A intervenção é verificar middlewares e esquemas. Os termos utilizados para pesquisa foram 'middleware', 'management', 'model' e 'schema';
- **Comparação:** O foco deste trabalho não se limitou a estudos comparativos;
- **Resultado:** Tem-se como objetivo a procura de avaliações, definições, validações e implementações de middlewares e/ou esquemas utilizados em pesquisas científicas. Desta forma, obtemos as seguintes palavras-chave 'validation', 'evaluation' e 'implementation', 'definition'.

A string de busca gerada é a seguinte: (('Internet of things metadata') **OR** ('IoT metadata')) **AND** (('middleware') **OR** ('schema') **OR** ('management') **OR** ('model')) **AND** (('validation') **OR** ('evaluation') **OR** ('implementation') **OR** ('definition'))

Fontes de Busca

Foram escolhidas as seguintes bases digitais para que as buscas sejam realizadas:

- Google Acadêmico (<https://scholar.google.com.br/>)
- JSTOR (<https://www.jstor.org/>)
- Periódicos CAPES (<https://www.periodicos.capes.gov.br/>)

Bases escolhidas devido a sua relevância e sua grande abrangência sobre diversos temas.

Idioma

Para este trabalho, o idioma de preferência para seleção de artigos será a língua inglesa, entretanto, trabalhos científicos escritos em língua portuguesa não serão descartados, desde que atinjam os requisitos para inclusão.

Seleção dos Estudos

Critérios de Inclusão e Exclusão

Critérios de Inclusão

- Os trabalhos devem estar disponíveis nas bases de dados escolhidas previamente;
- Serão considerados apenas publicações posteriores à 2006(ESCOLHER DATA POR UM MOTIVO MELHOR), salvo em casos de fontes relevantes que contenham definições necessárias para a realização deste trabalho;
- O trabalho deve possuir menção a metadados em ambientes Internet das Coisas.

Critérios de Exclusão

- Trabalhos publicados anteriormente à 2006, exceto fontes relevantes que contenham definições pertinentes para este estudo;
- Trabalhos não disponíveis nas bases de dados digitais escolhidas;
- O trabalho não possuir menção a metadados em ambientes Internet das Coisas;
- Publicações que forem interpretadas como fora do escopo, ou seja, fogem do tema metadados em ambientes IoT;
- Trabalhos que não propõe, comparam ou avaliam métodos para o gerenciamento de metadados;
- Trabalhos duplicados, ou seja, publicados em múltiplas bases de dados.

Processo de Seleção dos Estudos

Os artigos obtidos por meio da estratégia descrita acima passarão por um processo de avaliação sistêmico, com base nos critérios especificados anteriormente. Desta forma, os artigos que atingirem os parâmetros serão adicionados à base de estudos da revisão sistemática. A estratégia para a pesquisa e seleção é:

1. Pesquisa de trabalhos científicos nas bases de dados definidas utilizando as strings de busca;
2. Leitura do título, resumo, palavras chave e data de publicação, aplicando os critérios de inclusão e exclusão definidos;
3. Leitura da introdução e conclusão dos trabalhos que foram mantido na fase anterior;

4. Os trabalhos selecionados na fase anterior serão lidos por completo, e as informações pertinentes serão coletadas destes trabalhos.

Referências

- [1] 13, ITU T Study Group: *Recommendation itu-t y.2060. overview of the internet of things*, 2007. <http://handle.itu.int/11.1002/1000/11559>. 2, 3
- [2] Want, Roy: *An introduction to rfid technology*. IEEE pervasive computing, 5(1):25–33, 2006. 4
- [3] Luigi Atzori, Antonio Iera, Giacomo Morabito: *The internet of things: A survey*. Computer networks, 54(15):2787–2805, 2010. 4
- [4] Jayavardhana Gubbi, Rajkumar Buyya, Slaven Marusic Marimuthu Palaniswami: *Internet of things (iot): A vision, architectural elements, and future directions*. Future generation computer systems, 29(7):1645–1660, 2013. 4
- [5] Payam Barnaghi, Wei Wang, Cory Henson Kerry Taylor: *Semantics for the internet of things: early progress and back to the future*. International Journal on Semantic Web and Information Systems, 8(1):1–21, 2012. 5
- [6] Charu C. Aggarwal, Naveen Ashish, Amit Sheth: *The internet of things: a survey from the data-centric perspective*. Em *Managing and mining sensor data*, capítulo 12, páginas 383–428. Springer, Boston, MA, 2013. 6
- [7] Hassan, Umairul et al.: *A collaborative approach for metadata management for internet of things*. Collaborative Computing: Networking, Applications and Worksharing, páginas 593–598, 2013. 6
- [8] Duval, Erick et al.: *Metadata principles and practicalities*, 2002. <http://www.dlib.org/dlib/april02/weibel/04weibel.html>? 6
- [9] Mike Botts, Alexandre Robin: *Opengis sensor model language (sensorml) implementation specification*, 2007. 7