

Murilo Silveira Gomes

**PROPOSTA DE ARQUITETURA PARA ECOSISTEMA
DE INOVAÇÃO EM DADOS ABERTOS**

Dissertação submetido ao Programa de Engenharia e Gestão do Conhecimento da Universidade Federal de Santa Catarina para a obtenção do Grau de Mestre em Engenharia e Gestão do Conhecimento.

Orientador: Prof. Dr. Marcelo Macedo

Florianópolis

2017

Ficha de identificação da obra elaborada pelo autor,
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Gomes, Murilo Silveira

Proposta de Arquitetura para Ecossistema de Inovação em
Dados Abertos / Murilo Silveira Gomes ; orientador,
Marcelo Macedo - Florianópolis, SC, 2017.
104 p.

Dissertação (mestrado) - Universidade Federal de Santa
Catarina, Centro Tecnológico. Programa de Pós-Graduação em
Engenharia e Gestão do Conhecimento.

Inclui referências

1. Engenharia e Gestão do Conhecimento. 2. Ecossistema
de Inovação. 3. Arquitetura. 4. Dados Abertos. I. Macedo,
Marcelo. II. Universidade Federal de Santa Catarina.
Programa de Pós-Graduação em Engenharia e Gestão do
Conhecimento. III. Título.

Murilo Silveira Gomes

PROPOSTA DE ARQUITETURA PARA ECOSSISTEMA DE INOVAÇÃO EM DADOS ABERTOS

Esta Dissertação foi julgada adequada para obtenção do Título de “Mestre em Engenharia e Gestão do Conhecimento”, e aprovado em sua forma final pelo Programa de Pós-graduação em Engenharia e Gestão do Conhecimento da Universidade Federal de Santa Catarina – UFSC.

Florianópolis, 16 de fevereiro de 2017.

Prof. Roberto Carlos do Santos Pacheco, Dr.
Coordenador do Curso

Banca Examinadora:

Prof.^a Marcelo Macedo, Dr.
Orientador
Universidade Federal de Santa Catarina

Prof.^a Fernando A. Ostuni Gauthier, Dr.
Universidade Federal de Santa Catarina

Prof. Alexandre Moraes Ramos, Dr.
Universidade Federal de Santa Catarina

Prof. João Artur de Souza, Dr.
Universidade Federal de Santa Catarina

Este trabalho é dedicado a minha esposa
Lidiane que sempre me apoia e à minha
família.

Deixem que o futuro diga a verdade e avalie cada um de acordo com o seu trabalho e realizações. O presente pertence a eles, mas o futuro pelo qual eu sempre trabalhei pertence a mim.

(Nikola Tesla)

RESUMO

Os ecossistemas de inovação e de software utilizam de tecnologias para apoiar seus objetivos. Estes ecossistemas possuem particularidades mas também muita semelhança em alguns aspectos, como por exemplo, a inovação. As tecnologias semânticas por sua vez vem sendo destaque em pesquisas e no desenvolvendo de algumas tecnologias que auxiliam no vislumbre semântico dos dados. Não foi identificado na literatura nenhum estudo que realizasse um paralelo entre os dois ecossistemas e que utilizasse como base as tecnologias semânticas em um ecossistema. Desta forma, esta dissertação tem como objetivo propor uma modelagem de um ecossistema de inovação baseado em tecnologias semânticas utilizando elementos que compõem os ecossistemas de inovação e de software. Para a coleta dos dados, foi utilizado do método de revisão sistemática realizada em três bases de dados SCOPUS, IEEE e *Web of Science* e priorizando os *papers* dos últimos cinco anos para coleta de dados sobre os ecossistemas de inovação e de software, e uma revisão narrativa para dar embasamento sobre as tecnologias semânticas. Como resultados obtidos, foram elencados características e/ou fatores de sucesso dos ecossistemas de inovação e software, realizando o seu agrupamento, e foram identificados também na literatura as tecnologias e ciclos de vida que auxiliam na publicação de dados. Como resultado final desta dissertação, foi proposto uma arquitetura e recomendações que utilizam os elementos pertencentes de ambos os ecossistemas de inovação e de software, agregando um ciclo de vida e as tecnologias que auxiliem na publicação de dados a este ecossistemas de inovação em tecnologias semânticas. É observado neste estudo uma oportunidade de execução para obtenção de resultados reais, contribuindo com uma evolução para este trabalho, abrindo novas oportunidades de pesquisas e uma nova perspectiva de serviços e produtos que podem ser oferecidos a sociedade

Palavras-chave: Ecossistema de Inovação, Tecnologia Semântica, Publicação de Dados

ABSTRACT

Innovation and software ecosystems use technologies to support their goals. These ecosystems have particularities but also a lot of similarity in some aspects, such as innovation. Semantic technologies in turn have been highlighted in research and in the development of some technologies that help in the semantic glimpse of the data. In this way no study was identified in the literature that realized a parallel between the two ecosystems and that used as base the semantic technologies in an ecosystem. In this way, this dissertation aims to propose a modeling of an ecosystem of innovation based on semantic technologies using elements that make up the ecosystems of innovation and software. In order to obtain the results of this research, a systematic review method was used in three SCOPUS, IEEE and Web of Science databases, prioritizing the papers of the last five years for data collection on innovation and software ecosystems, and a review Narrative to support semantic technologies. As a result, the characteristics and/or factors of success of the innovation and software ecosystems were identified and their grouping was carried out, and the technologies and life cycles that aid in the publication of data were also identified in the literature. As a final result of this dissertation, it was proposed an architecture and recommendations that use the elements belonging to both ecosystems of innovation and software, adding a life cycle and the technologies that help in the publication of data to this ecosystem of innovation in semantic technologies. Finally, with the conclusion of this dissertation, an opportunity of execution is seen to obtain real results, contributing with an evolution to this work, opening new research opportunities and a new perspective of services and products that can be offered to society.

Keywords: Ecosystem of Innovation, Semantic Technology, Data Publication

LISTA DE FIGURAS

| | | |
|-----------|--|----|
| Figura 1 | Interesse ao Longo do tempo (Fonte: Do Autor, 2017) | 16 |
| Figura 2 | Países que mais pesquisaram (Fonte: Do Autor, 2017) | 17 |
| Figura 3 | Interesse ao Longo do tempo (Fonte: Do Autor, 2017) | 18 |
| Figura 4 | Interesse ao Longo do tempo (Fonte: Do Autor, 2017) | 19 |
| Figura 5 | Processo Metodológico de Pesquisa (Fonte: Do Autor, 2016) | 23 |
| Figura 6 | Triplice Hélice (ETZKOWITZ; LEYDESDORFF, 2000) | 26 |
| Figura 7 | Ciclo de vida para ecossistema de inovação da NASA (HARDASH et al., 2014) | 32 |
| Figura 8 | Innovation Union Scoreboard (HUHTAM; RUBENS et al., 2016) | 32 |
| Figura 9 | Framework Saúde do Ecossistema de Software (MANIKAS; HANSEN, 2013b) | 41 |
| Figura 10 | Arquitetura Orientada a Serviço (SOUZA et al., 2006) | 43 |
| Figura 11 | Arquitetura E-gov (PACHECO et al., 2015) | 44 |
| Figura 12 | Histórico da WEB (RAMALHO, 2010) | 45 |
| Figura 13 | Arquitetura da Web Semântica (W3C, 2016c) | 46 |
| Figura 14 | Nuvem LOD 2014 (LOD..., 2016) | 51 |
| Figura 15 | Classificação da publicação LOD (BERNERS-LEE, 2006) | 52 |
| Figura 16 | Ciclo de Vida LOD - 1 (VILLAZÓN-TERRAZAS et al., 2011) | 53 |
| Figura 17 | Ciclo de Vida LOD - 2 (AUER et al., 2012) | 55 |
| Figura 18 | Ciclo de Vida LOD - 3 (BROEK; VEENSTRA; FOLMER, 2014) | 56 |
| Figura 19 | Arquitetura do Ecossistema- (Autor, 2016) | 70 |
| Figura 20 | Arquitetura do Ecossistema proposta pelos entrevistados- (Autor, 2016) | 87 |

LISTA DE TABELAS

| | | |
|----------|--|----|
| Tabela 1 | Trabalhos oriundo do EGC..... | 24 |
| Tabela 2 | Agrupamento das características..... | 28 |
| Tabela 2 | Agrupamento das características..... | 29 |
| Tabela 3 | Agrupamento - Ecossistema de Software..... | 35 |
| Tabela 3 | Agrupamento - Ecossistema de Software..... | 36 |
| Tabela 3 | Agrupamento - Ecossistema de Software..... | 37 |
| Tabela 4 | Representação do conjunto de dados presente na Nuvem LOD(LOD..., 2016)..... | 52 |
| Tabela 5 | Comparativo Ecossistemas (Autor, 2016)..... | 66 |
| Tabela 5 | Comparativo Ecossistemas (Autor, 2016)..... | 67 |
| Tabela 5 | Comparativo Ecossistemas (Autor, 2016)..... | 68 |
| Tabela 6 | Verificação Objetivo Específico (1)..... | 80 |
| Tabela 6 | Verificação Objetivo Específico (1)..... | 81 |
| Tabela 7 | Verificação Objetivo Específico (2)..... | 83 |
| Tabela 8 | Verificação Objetivo Específico (3)..... | 84 |

LISTA DE ABREVIATURAS E SIGLAS

| | |
|-------|--|
| CKAN | Comprehensive Kerbal Archive Network |
| EC | Engenharia do Conhecimento |
| EUA | Estados Unidos da América |
| HTTP | Hypertext Transfer Protocol |
| ICT | Institutos de Ciência e Tecnologia |
| IRI | Internacional Resource Identifier |
| JSON | JavaScript Object Notation |
| LD | Linked Data |
| LED | Linked Enterprise Data |
| LOD | Linked Open Data |
| MS | Microsoft |
| NASA | National Aeronautics and Space Administration |
| OGD | Open Government Data |
| P&D | Pesquisa e Desenvolvimento |
| PPEGC | Programa de Pós-graduação em Engenharia e Gestão do Conhecimento |
| RDF | Resource Description Framework |
| RDFS | Resource Description Framework Schema |
| RIF | Rule Interchange Format |
| SECO | Software Ecosystem |
| SOA | Service Oriented Architecture |
| SQL | Structured Query Language |
| TI | Tecnologia da Informação |
| TIC | Tecnologia da Informação e Comunicação |
| UFSC | Universidade Federal de Santa Catarina |
| URI | Uniform Resource Identifier |
| W3C | World Wide Web Consortium |
| WWW | World Wide Web |
| XML | Extensible Markup Language |

SUMÁRIO

| | | |
|----------|--|----|
| 1 | INTRODUÇÃO | 13 |
| 1.1 | OBJETIVOS | 15 |
| 1.1.1 | Objetivo Geral | 15 |
| 1.1.2 | Objetivos Específicos | 15 |
| 1.2 | JUSTIFICATIVA | 15 |
| 1.3 | METODOLOGIA | 21 |
| 1.4 | ADERÊNCIA AO EGC | 22 |
| 2 | REVISÃO DE LITERATURA | 25 |
| 2.1 | ECOSSISTEMA | 25 |
| 2.1.1 | Ecossistema de Inovação | 25 |
| 2.1.2 | Ecossistema de Software | 33 |
| 2.2 | ARQUITETURA | 42 |
| 2.3 | EVOLUÇÃO DA WEB | 44 |
| 2.3.1 | Web Semântica | 45 |
| 2.4 | LINKED DATA | 48 |
| 2.4.1 | Linked Open Data | 50 |
| 2.4.2 | Ciclo de Vida do Linked Open Data | 53 |
| 2.4.3 | Tecnologias | 59 |
| 3 | ANÁLISE E MODELAGEM | 65 |
| 3.1 | ANÁLISE ENTRE OS ECOSSISTEMAS DE INOVAÇÃO E SOFTWARE | 65 |
| 3.2 | MODELAGEM DO ECOSSISTEMA DE INOVAÇÃO EM TECNOLOGIAS SEMÂNTICAS | 69 |
| 4 | VERIFICAÇÃO | 80 |
| 4.1 | RESULTADO VERIFICAÇÃO | 86 |
| 5 | CONCLUSÃO | 89 |
| 5.1 | ESTUDOS FUTUROS | 90 |
| | REFERÊNCIAS | 92 |

1 INTRODUÇÃO

A palavra ecossistema é oriunda da ecologia e significa um conjunto de diferentes seres vivos em um mesmo meio ambiente, tendo cada ser vivo um papel fundamental para a sobrevivência do ecossistema (CHRISTIAN, 2009). Ao longo do tempo surgiram outras abordagens empregando o termo ecossistema, como por exemplo, o ecossistema de inovação e o ecossistema de software.

O ecossistema de inovação proporciona a integração de uma comunidade econômica apoiada por organizações e indivíduos, com objetivo de promover a inovação através de Pesquisa e Desenvolvimento (P&D) colaborando com as necessidades do mercado, mantendo o ecossistema ativo e em constante crescimento. Dessa forma, o ecossistema possibilita o desenvolvimento novos produtos ou serviços, gerando valores que impactam positivamente na economia (BOSCH-SIJTSEMA; BOSCH, 2014; JACKSON, 2011).

O ecossistema de software tem por objetivo melhorar a reutilização de softwares na indústria e as relações entre empresas e demais interessados, possibilitado a construção de sistemas com o uso de plataformas comuns e a participação de desenvolvedores externos e internos ao ecossistema (SILVA, 2012; LARRUCEA; NANCLARES; SANTAMARIA, 2016). Dessa forma, o ecossistema de software também possibilita a inovação, por meio da competitividade e cooperativismo a fim de desenvolver novos produtos (SANTOS; WERNER, 2011). O ecossistema de software se destaca por ser um meio eficaz no desenvolvimento e no uso de sistemas de software que fazem uso de plataformas tecnológicas (MANIKAS; HANSEN, 2013b).

Muitas das plataformas tecnológicas desenvolvidas atualmente possuem seu foco na web (O'REILLY, 2005; MARTINS, 2013). A web, por sua vez, está em evolução e vem beneficiando a humanidade (RAMALHO, 2010). Desta forma, a web atual atende somente os seres humanos que possuem a capacidade de compreensão, diferentemente das máquinas. Sendo assim a web semântica é apresentada como uma extensão da web atual, vislumbrando a compreensão das informações presentes na Web, por homens e máquinas. A fim de alcançar o vislumbre da web semântica surge o conceito de LD e LOD.

Linked Data (LD) é apresentado como as melhores práticas para se publicar dados na Web Semântica (BERNERS-LEE, 2006). Sendo que, posteriormente, o conceito foi estendido quando o governo dos EUA

apresentou a proposta de abertura dos dados com intuito de deixar o governo mais transparente. Deste modo, iniciou-se o movimento de dados abertos (DEFINITION..., 2016), e o conceito de LD evoluiu para Linked Open Data (LOD) herdando do LD suas características, ascendo também as princípios de dados abertos (BIZER; HEATH; BERNERS-LEE, 2009). Desde então dados vem sendo publicados em várias áreas, como exemplo, governo, geográfica, ciências da vida, meios de comunicação e entre outros (LOD..., 2016). Para auxiliar na publicação dos dados em LOD foram realizadas, pesquisas e desenvolvidos ciclos de vida e tecnologias. O consumo dos dados publicados possibilita o desenvolvimento de novos produtos de softwares, novas tecnologias e ferramentas que podem auxiliar no cotidiano da sociedade ou proporcionar diferentes análises dos dados e informações (BANDEIRA, 2015; MAGALHÃES, 2012).

Ecossistema de inovação e ecossistema de software possuem vários elementos base em comum, ganhando destaque a inovação onde ambos os ecossistemas trabalham a fim de gerar novos produtos e/ou serviços agregando valor ao ecossistema, sendo ele um valor monetário ou apenas um reconhecimento (GOMES et al., 2016; MANIKAS, 2016). Desta forma, a importância deste trabalho está na possibilidade de aliar os benefícios apontados pelo ecossistema de inovação em gerar a inovação através de novos produtos e serviços, proporcionando um impacto positivo na economia com a capacidade em dar o suporte necessário ao desenvolvimento de tecnologia apresentado no ecossistema de software e o benefício da publicação de dados através do LD e LOD. Nesse sentido, é proposto o desenvolvimento de um ecossistema de inovação baseado em tecnologias semânticas a fim de aliar essas três vertentes proporcionando inovações através da P&D.

Considerando os benefícios de convergência de dois ecossistemas em um único e aliando estes a dados abertos ligados, apresenta-se as seguintes perguntas de pesquisa: "Quais elementos do ecossistema de inovação e ecossistema de software podem servir como base a um ecossistema de inovação baseado em tecnologias semânticas?" e "Qual será a arquitetura que dará o suporte à este ecossistema?"

1.1 OBJETIVOS

1.1.1 Objetivo Geral

Modelar um ecossistema de inovação baseado em tecnologias semânticas que envolva dados abertos provenientes de academia, empresas e governo.

1.1.2 Objetivos Específicos

- Apresentar os principais elementos do ecossistema de software e de ecossistema de inovação.
- Abordar as tecnologias semânticas e os ciclos de vida para a publicação de dados em empresas, academia e governo
- Propor um conjunto de recomendações e uma arquitetura para um ecossistema de inovação baseado em tecnologias semânticas para dados abertos.

1.2 JUSTIFICATIVA

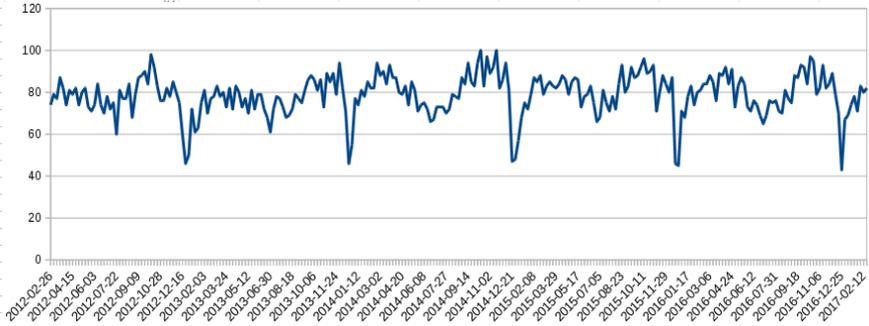
A web está em constante evolução, sendo a web semântica apresentada como uma extensão da web atual (PICKLER, 2007). Com o advento das tecnologias e da difusão da web nos últimos anos, a web semântica vem ganhando espaço e conquistando diversas áreas com sua aplicabilidade (DIAS; SANTOS, 2003; MORAIS; AMBRÓSIO, 2007).

O conceito de LD foi o começo para o desenvolvimento da nuvem de dados. Seu objetivo consiste em não apenas publicar dados na web, mas publicar de acordo com as melhores práticas (BERNERS-LEE, 2006). O uso de LD, além de possibilitar a publicação de dados na web, possibilita uma uniformidade das informações facilitando sua comunicação, a referenciação dos dados, a integrabilidade onde os dados podem ser facilmente integrados entre diferentes conjuntos de dados ou sistemas, entre outros benefícios (AUER; LEHMANN; NGOMO, 2011). Como pode ser observado, os dados retirado do *Google Trends*¹ uma ferramenta de coleta de palavras chaves mais pesquisadas em toda a

¹www.trends.google.com

web. É possível observar na Figura 1 e Figura 2 o interesse ao longo do tempo e os países que mais tiveram buscas pela palavra *Linked Data*.

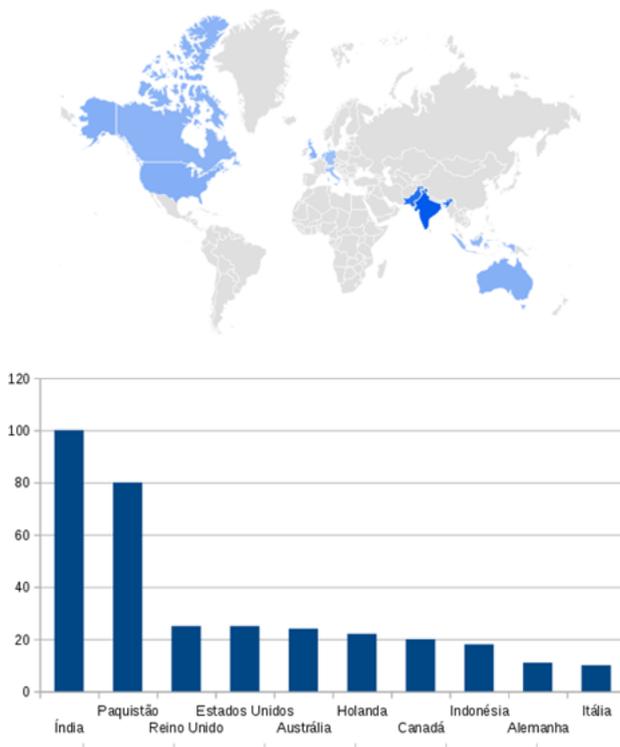
Figura 1 – Interesse ao Longo do tempo (Fonte: Do Autor, 2017)



Como pode ser observado na Figura 1, O valor 120 é o pico de popularidade do *Linked Data*, já o valor 50 corresponde a metade da popularidade e o valor 0 corresponde a menos de 1% das pesquisas realizadas. Pode ser observado também que nos anos de 2012, 2014 e 2016 houve uma maior procura do termo *Linked Data*.

Na Figura 2 é possível observar o pico de popularidade e quais os países que pesquisaram sobre *Linked Data*, dando destaque para a Índia e o Paquistão.

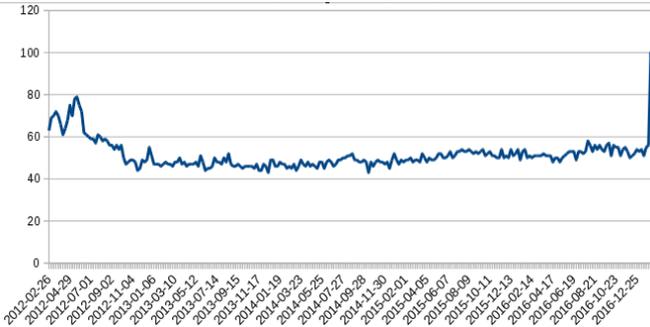
Figura 2 – Países que mais pesquisaram (Fonte: Do Autor, 2017)



Com o início da abertura dos dados pelo governo americano e outros governos de países desenvolvidos, houve a difusão do conceito de LOD com o objetivo de contribuir com o crescimento da web de dados, também conhecida como a nuvem LOD (LOD..., 2016). A nuvem LOD é composta por oito áreas, sendo que o governo foi o primeiro a ganhar foco nas pesquisas para a publicação de dados em LOD, com o intuito de promover maior transparência, responsabilidade democrática e proporcionar à sociedade a possibilidade de consumir os dados para o desenvolvimento de novos produtos ou serviços, entre outros (JANSSEN; CHARALABIDIS; ZUIDERWIJK, 2012). Este movimento de publicação de dados governamentais é denominado de *open government data* (UBALDI, 2013), mas o governo não foi o único a utilizar de LD, as empresas também participam deste meio, obtendo alguns

benefícios como: enriquecimento de dados existentes na organização, redução da complexidade da interoperabilidade entre aplicações, entre outros. Tal movimento é denominado de *Linked Enterprise Data* (SERVANT, 2008; ANTIDOT, 2016). As academias também fazem uso de LD, com foco na aprendizagem, obtendo também alguns benefícios, como a interoperabilidade das bases de dados educacionais e o enriquecimento da informação, viabilizando assim uma aprendizagem mais efetiva. Tal movimento não possui uma denominação definida, mas foi identificado na literatura algumas, como exemplo, *Linked Education*, *Linked University*, *Open University*, entre outras (DIETZE et al., 2012; D'AQUIN; DIETZE, 2014). Também é possível observar os dados retirados do *Google Trends* na Figura 3 e Figura 4 os anos e os países que mais tiveram buscas pela palavra *LOD* e *Linked Open Data*.

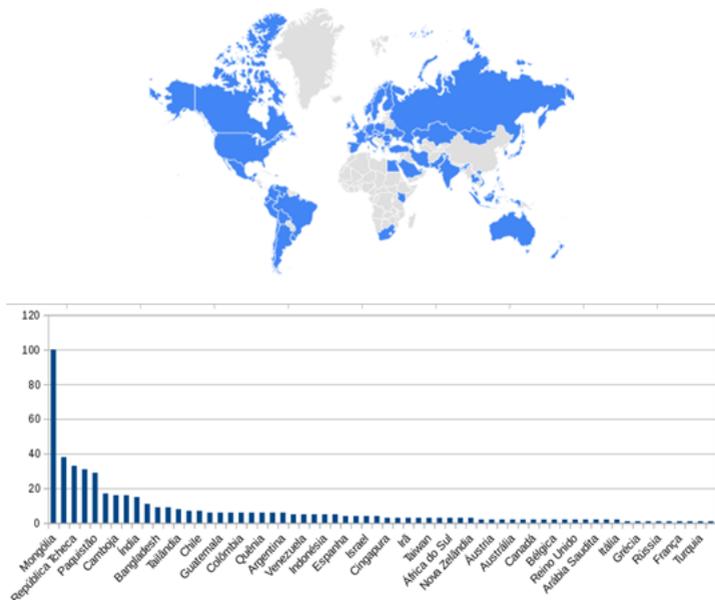
Figura 3 – Interesse ao Longo do tempo (Fonte: Do Autor, 2017)



Como pode ser observado na Figura 3, O valor 120 é o pico de popularidade do *LOD*, já o valor 50 corresponde a metade da popularidade e o valor 0 corresponde a menos de 1% das pesquisas realizadas. Pode ser observado também que nos anos de 2012 e 2017 houve uma maior procura do termo *LOD*.

Na Figura 4 é possível observar o pico de popularidade e quais os países que pesquisaram sobre *LOD*, sendo um total de 67 países que pesquisam sobre o tema, dando destaque para a Mongólia, República Tcheca, Paquistão que ficaram entre os que mais pesquisam e salienta-se que os países como EUA, Alemanha e Brasil com uma média de dois pontos das pesquisas realizadas no globo sobre o tema.

Figura 4 – Interesse ao Longo do tempo (Fonte: Do Autor, 2017)



A inovação é apresentada como fonte de vantagem competitiva, possibilidade de crescimento econômico, crescimento do mercado entre outros benefícios. O ecossistema de inovação possibilita a interação de indivíduos a fim de proporcionar inovações através de P&D, proporcionando o desenvolvimento de tecnologia e inovação (OH et al., 2016; BOSCH-SIJTSEMA; BOSCH, 2014). P&D é viabilizada pelas academias que são compostas por vários grupos de pesquisa, trabalhando para o desenvolvimento de diversas áreas. No entanto, para que haja o desenvolvimento de tecnologia e inovação se faz necessário a presença de mais duas organizações, o governo auxiliando no fomento das atividades econômicas e as empresas que apresentam as demandas e necessidades. A interação dessas três organizações é considerada um modelo de inovação denominada de tríplice hélice, focada em novos conhecimentos, em inovações tecnológicas e no desenvolvimento econômico (LEYDES-DORFF; ETZKOWITZ, 1998; GOMES; PEREIRA, 2015).

A inovação é importante em uma época de crescente demanda por desempenho e orçamentos, sendo eles congelados ou decrescen-

tes (HARDASH et al., 2014). Os ecossistemas de inovação proporcionam alguns benefícios, como exemplo, a interação e valorização entre os atores do ecossistema, a diversidade de atores que contribuem para a inovação de novos produtos e serviços, a geração do conhecimento, entre outros (OH et al., 2016; ETZKOWITZ, 2003). No Brasil foi criada a Lei de Inovação Tecnológica nº 10.973, aprovada em 2 de dezembro de 2004 e regulamentada em 11 de outubro de 2005 pelo Decreto nº 5.563, sendo que a mesma está organizada em torno de cinco eixos: (i) a constituição de ambiente propício a parcerias estratégicas entre universidades, institutos tecnológicos e empresas; (ii) o estímulo à participação de Instituições Científica, Tecnológica e de Inovação (ICT) no processo de inovação; (iii) o estímulo à inovação nas empresas, (iv) estímulo ao inventor independente e (v) fundos de investimento (CONGRESSO, 2016), oportunizando o desenvolvimento de ambientes propícios à inovação, ou seja, um ecossistemas de inovação. A lei de inovação 10.973 sofreu algumas alterações, passando ser a lei nº 13.243, de 11 de janeiro de 2016 possibilitando algumas melhorias ao setor de pesquisa e inovação.

As tecnologias estão presentes em todas as áreas de maneira direta ou indireta (BENAKOUCHE, 2005). A maioria das inovações e demandas de pesquisa, estão ligadas a tecnologias, sendo que a mesma auxilia no desenvolvimento de novos produtos, serviços tecnológicos ou no uso da tecnologia para atender uma necessidade (KON, 2016).

Os ecossistemas de software possibilitam a reutilização dos softwares e/ou tecnologias poupando recursos para serem aplicados em novas pesquisas e na geração de novas tecnologias (SANTOS; WERNER, 2011; LETTNER et al., 2014). O ecossistema de software é um conceito recente apresentado nas pesquisas, onde sua exploração teve início à partir de 2010, mas já são apresentas diversas pesquisas relevantes (MANIKAS, 2016).

Os ecossistemas de software na era tecnológica são comuns, como exemplo, o Google. A Google possui vários envolvidos trabalhado para geração de novas tecnologias, como exemplo: os programadores que podem colaborar com o Google, oportunizando programadores externos a publicarem seus softwares no ecossistema desde que estejam nas normas exigidas pela Google. A Google possui também pesquisadores que trabalham em prol da pesquisa e desenvolvimento. E por fim, os cliente são os consumidores de serviços e das tecnologias fornecidos pela Google. Outro ecossistema de software é da Apple cujo funcionamento é semelhante o da Google (PELLICCIONE, 2014; URLI et al., 2014; CHRISTENSEN et al., 2014; SOUZA et al., 2016).

Observa-se que os ecossistemas de inovação estão em crescimento

e ganhando o apoio do governos com leis de incentivo e que as tecnologias estão presente fortemente neste meio. Os ecossistemas de software também estão ganhando força com a difusão de grandes ecossistemas de software como Google e Apple. A inovação está presente em ambos ecossistemas, sendo fomentada e explorada. Observa-se também que existem movimentos que abrangem e incentivam o uso e a pesquisa do LD e LOD, tais movimentos podem proporcionar a P&D, tanto de novos produtos, como de novos serviços. Desta forma observa-se a oportunidade de unir os dois ecossistemas fazendo uso de LD, oportunizando novas perspectivas de pesquisa, possibilidades de desenvolvimento de novas tecnologias e impactando na economia.

1.3 METODOLOGIA

Para elaboração de qualquer trabalho científico se faz necessário estabelecer alguns procedimentos metodológicos, como exemplo, a definição da sequência das atividades a fim de atingir os objetivos finais e os métodos e procedimentos a serem utilizados. Deste modo possibilita a orientação do desenvolvimento da pesquisa na coleta de dados, a análise e interpretação dos resultados (MARCONI; LAKATOS, 2011; GIL, 2002).

Para a estruturação desta pesquisa a fim de atender os objetivos propostos foram realizadas pesquisas de cunho narrativo e sistemática. Primeiramente foram realizadas as revisões sistemáticas, onde uma estratégia de busca utilizando-se de métodos sistematizados é definida. Para isso, utiliza-se de bases de dados para buscar determinado tema e, após, sintetiza-se um resumo das evidências e faz-se uma análise crítica. Tal método é apresentado como útil, pois pode trazer uma visão macro de um determinado tema escolhido pelo pesquisador, possibilitando vislumbrar trabalhos futuros, a identificação de problemas na área de pesquisa e o embasamento robusto sobre determinado tema (SAMPAIO; MANCINI, 2007; KEELE, 2007).

Deste modo foi realizada a pesquisa em três bases de dados, sendo, SCOPUS, IEEE, Web of Science, utilizando das palavras chaves "*software ecosystem*" e "*innovation ecosystem*", utilizando somente os *papers* publicados nos últimos cinco anos, a fim de identificar na literatura as características e/ou fatores de sucesso para a implementação dos ecossistemas.

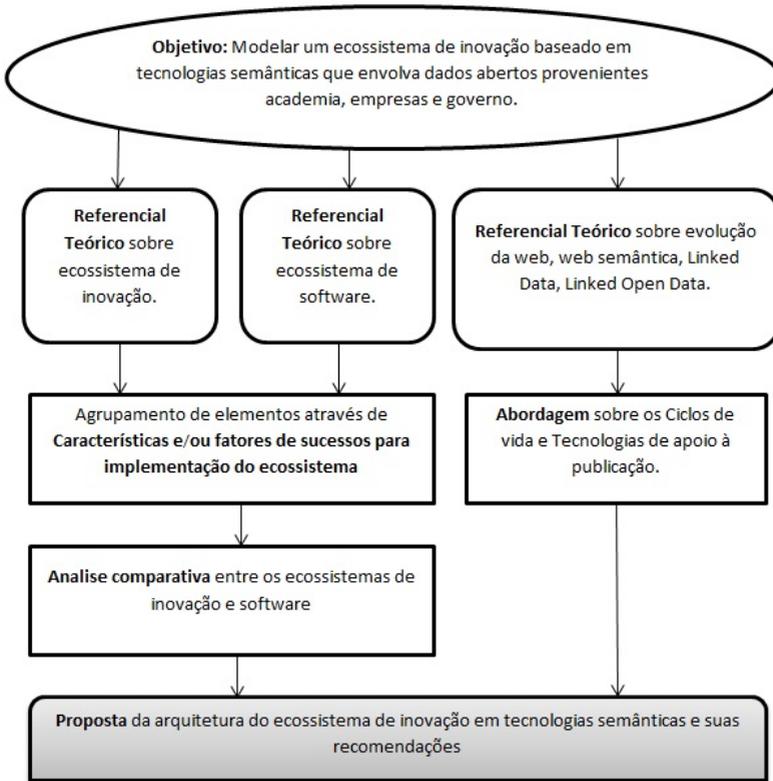
Para a obtenção do referencial teórico sobre a evolução da web, web semântica, Linked Data e Linked Open Data fez-se o uso de revi-

são narrativa, sendo utilizado uma estrutura previamente definida pelo autor, a fim de dar o embasamento necessário para a descrição e compreensão do tema abordado. A revisão narrativa é apropriada para descrever o atual estado da arte de determinado assunto sobre uma ótica teórica ou contextual (CORDEIRO et al., 2007).

Desta forma, com a obtenção de todo referencial teórico, foi realizado uma análise comparativa entre os ecossistemas, e elicitado os ciclos de vida e as tecnologias que apoiam a publicação de dados.

Por fim realizou-se o desenvolvimento da arquitetura do ecossistema de inovação baseado em tecnologias semânticas e suas recomendações com foco na publicação de dados a fim de atender governo, academia e empresas. A seguir é apresentado na Figura 5 os procedimentos metodológicos utilizados nesta pesquisa.

Figura 5 – Processo Metodológico de Pesquisa (Fonte: Do Autor, 2016)



1.4 ADERÊNCIA AO EGC

Esta dissertação está contextualizada na área de concentração de Engenharia do Conhecimento (EC), do Programa de Pós-graduação em Engenharia e Gestão do Conhecimento (PPEGC) da Universidade Federal de Santa Catarina (UFSC), na linha de pesquisa de Engenharia do Conhecimento Aplicada. Esta linha estuda a concepção, desenvolvimento e implantação de soluções da Engenharia do Conhecimento nas organizações (EGC, 2016).

A EC é amparada pela visão epistemológica do cognitivismo, sendo que esta visão surgiu com o início dos computadores e da inteligência artificial. No cognitivismo entende-se que o conhecimento pode ser adquirido por métodos técnico-científicos e pode ser processado por máquina. Portanto, percebe-se o conhecimento como tratável em computadores, armazenado em bancos de conhecimentos e relacionado a arquivos, manuais ou rotinas que promovam o compartilhamento em uma organização (PACHECO, 2016). Ao considerar todo conhecimento e informação que são produzidos dentro de um ecossistema, a EC pode auxiliar no processo de geração de valor através do conhecimento.

A dissertação trata, mais especificamente, da aplicação de tecnologias semânticas – relacionadas à Engenharia do Conhecimento – em campo de aplicação da área de Gestão do Conhecimento do EGC. Mais especificamente, de um lado, trata-se de *Linked Data* - (LD) e *Linked Open Data* - (LOD) - tecnologias semânticas que têm ganho destaque em diversos domínios de publicação e catalogação dos dados (ROCHA, 2014; KUČERA; CHLAPEK; MYNARZ, 2012). Nesta dissertação, estuda-se sua aplicação nos chamados ecossistemas de inovação, que preveem a interação de diferentes atores (individuais e organizacionais) em um mesmo ambiente em prol do desenvolvimento de uma economia baseada em conhecimento (ETZKOWITZ, 2003).

O propósito deste trabalho é guiar o processo de disponibilização de novas fontes de dados a partir de um ecossistema de inovação baseado em tecnologias semânticas, envolvendo os atores da tríplice hélice para o processo de publicação dos dados e propor uma arquitetura e suas recomendações.

Em relação à memória do EGC nos temas da dissertação, buscas no banco de teses e dissertações do EGC revelam uma pequena quantidade de trabalhos na área de LOD e LD, sendo os trabalhos de (SÁ, 2012; PIZZOL, 2014; SPERONI, 2016). Em relação ao ecossistema foi encontrado somente o trabalho (FIATES, 2014). Na tabela 1 é apresentado os trabalhos mencionados com o nome dos pesquisadores, o título

dos trabalhos, o ano em que foram apresentados e se é Tese (T) ou Dissertação (D).

Tabela 1: Trabalhos oriundo do EGC

| Linked Data ou Linked Open Data | | | | |
|--|--|------------|------------|--|
| Autor | Título | Ano | D/T | |
| Sá, Paulo Roberto Oliveira de | Engenharia do Conhecimento Aplicada a Criação Automatizada de Conteúdo Interativo para TV Digital | 2012 | D | |
| Pizzol, Leandro Dal | Uso da Web de Dados como Fonte de Informação no Processo de Inteligência Competitiva Setorial | 2014 | D | |
| Speroni, Rafael de Moura | Modelo de Referência Para Indicadores de Inovação Regional Suportado por Dados Ligados | 2016 | T | |
| Ecosistema | | | | |
| Fiates, José Eduardo Azevedo | Influência dos Ecosistemas de Empreendedorismo Inovador na Indústria de Venture Capital: Estratégias de Apoio às Empresas Inovadoras | 2014 | T | |

2 REVISÃO DE LITERATURA

Este capítulo tem por objetivo elencar através da literatura o entendimento sobre ecossistema abordando sua raiz ecológica. Neste contexto, busca-se apresentar também as definições do ecossistemas de inovação e ecossistema de software assim como os elementos que os compõem e suas características.

Será abordado também um breve histórico evolutivo da Web, contextualizando o surgimento da Web Semântica e de sua arquitetura. Também será contextualizado *Linked Data* e *Linked Open Data* mostrando alguns dos ciclos de vidas que guiam a publicação em *Linked Data* ou *Linked Open Data*, assim como são apresentadas as tecnologias que auxiliam a cada etapa do processo de publicação.

2.1 ECOSSISTEMA

O conceito de ecossistema ao longo dos anos vem sendo utilizado nos mais diversos contextos e popularmente é conhecido como um ambiente que envolve diversos indivíduos (CHRISTIAN, 2009). O estudo do ecossistema relaciona-se basicamente aos processos que ligam diferentes seres vivos em um ambiente mutuo (KLING, 2008). Deste modo, esta denominação vem sendo aplicada em diferentes contextos.

O termo ecossistema é oriundo da ecologia, no entanto com o passar do tempo foi evoluindo para uma compreensão mais profunda refletindo em um conceito geral (CHRISTIAN, 2009). Atualmente o termo ecossistema é aplicado em diversos ambientes, herdando algumas características da sua raiz ecológica.

A seguir será apresentado os conceitos acerca de ecossistema de inovação, bem como os elementos que o compõe.

2.1.1 Ecossistema de Inovação

Um ecossistema é apresentado como uma comunidade econômica apoiada por organizações e indivíduos que interagem com o objetivo de sobreviver e coevoluir juntamente com o ecossistema (BOSCH-SIJTSEMA; BOSCH, 2014). A inovação é a implementação de um novo produto, um bem, um serviço, um processo ou um método, podendo ser também o melhoramento dos mesmos (CUNNINGHAM; CUNNINGHAM;

EKENBERG, 2014). Um ecossistema de inovação parte de duas economias distintas, sendo uma delas a economia do conhecimento que é impulsionada pela P&D - Pesquisa e Desenvolvimento e a economia comercial que é impulsionada pelo mercado (JACKSON, 2011).

A inovação é a uma fonte essencial de vantagem competitiva e crescimento econômico mundial. Para que haja tal inovação é apresentado o ecossistema de inovação, contendo alguns atores envolvidos, como, a universidade, governo e empresa (HUHTAM; RUBENS et al., 2016). O envolvimento destes três atores em um meio, é denominado tríplice-hélice, onde governo, empresas e universidades trabalham no desenvolvimento regional, em processos dinâmicos de ciência, tecnologia, pesquisa e desenvolvimento, resultando em inovações (SILVA; NARCIZO; CARDOSO, 2012) . A Figura 6 apresenta a ilustração da tríplice hélice.

Figura 6 – **Triplíce Hélice** (ETZKOWITZ; LEYDESDORFF, 2000)



Na literatura encontram-se diversas definições para ecossistema de inovação, dentre elas, para Jackson (2011), ecossistemas de inovação "são as complexas relações que se formam entre os atores ou entidades cujo objetivo funcional é para permitir o desenvolvimento de tecnologia e inovação". No entanto, para Bosch-Sijtsema e Bosch (2014) um ecossistema de inovação é "um acordo de colaboração através do qual as empresas combinam suas ofertas individuais em uma solução coerente voltada para o cliente". A pesquisa realizada por Autio e Thomas (2014) define ecossistema de inovação como "Um conjunto de organizações interconectadas, em torno de uma empresa focal ou uma plataforma, incorporando produtores e desenvolvedores, focando em

produzir novos recursos ou valores através da inovação".

Para a identificação dos elementos que compõem o ecossistema de inovação foram determinados alguns critérios com base na literatura (DURST; POUTANEN, 2013; BOSCH-SIJTSEMA; BOSCH, 2014; OH et al., 2016). Deste modo, observou-se termos classificados como características presentes na estrutura de um modelo ou tratando-se de um fator de sucesso para a implementação do ecossistema. Sendo assim é apresentado a baixo os termos identificados na literatura com maior frequência pelos autores estudados.

- | | |
|---------------------------|-----------------------------------|
| a) Recursos | l) Capital Humano |
| b) Governança | m) Talentos |
| c) Estratégia e Liderança | n) Densidade de pesquisadores |
| d) Cultura | o) Empresários e instituições |
| e) Recursos Humanos | p) Acesso ao capital |
| f) Tecnologia | q) Ambiente regulatório favorável |
| g) Parceiros | r) Organizador |
| h) Clustering | s) Relação sinérgica entre atores |
| i) Plataforma | |
| i) Organizações | t) Diversidade de atores |

Com base nessas características foi elaborado um agrupamento para melhor organizar os elementos com base na literatura. Deste modo, abaixo a tabela 2 apresenta os agrupamentos conforme as características identificadas na literatura, as demais pesquisa que foram coletadas através da revisão sistemática e com maior frequência pelos autores estudados.

Tabela 2: Agrupamento das características

| | DURST, POUTANEN, 2013 | BOSCH- SIJTSEMA, BOSCH, 2014 | OH et al., 2016 | Outros autores |
|--------------------|--|---|--|---|
| Arquitetura | Recursos, Tecnologia | Tecnologia | - | Weil,Sabhlok, Cooney, 2014 |
| Atores | Clustering, Recursos Humanos, Parceiros | - | Densidade de pes- quisa- dores; Talentos ; Empre- sários e institu- ções | Weil,Sabhlok, Cooney (2014); Kolosky, Speroni, Gauthier (2015; Silva, Narcizo, Cardoso, 2012; Hardash Et Al., 2014 |
| Cultura | Cultura Organizaci- onal | - | Cultura empresarial, Ambiente regu- latório favorável | Hardash et al., 2014 |

Tabela 2: Agrupamento das características

| | DURST, POUTANEN, 2013 | BOSCH- SIJTSEMA, BOSCH, 2014 | OH et al., 2016 | Outros autores |
|--|---|---|---|---|
| Gestão | Governança, Estratégia e Liderança, Gestão de Recursos Humanos | Estratégia, Organizador | Acesso ao capital | Weil,Sabhlok, Cooney, 2014 |
| Inovação/ Inovação Aberta | Cultura de inovação Foco em P&D | Foco em P&D | Diversidade de atores | Huhtam, Rubens et al., 2016; Hardash et al., 2014; Silva, Nar- cizo, Car- doso, 2012; Weil, Sabh- lok, Cooney, 2014 |
| Indicadores | - | - | - | Huhtam, Rubens et al., 2016 |
| Plataforma | Tecnologia | Tecnologias, Plataforma | Relação Sinérgica entre atores | Huhtam, Rubens Et Al., 2016 Weil, Sabh- lok, Cooney, 2014 Hardash et al. (2014) |

A tabela 2 foi efetuado o agrupamento das características encontradas na literatura, sendo este agrupamento denominado de elementos,

são eles a **arquitetura** visando a estrutura do ecossistema abordando os recursos e as tecnologias. Os **atores** abordando todos os envolvidos no ecossistema de inovação. Utilizou-se do termo atores por ser a que mais aparece na literatura analisada. A **cultura** apresentando a sua importância ao ecossistema, **Gestão** sendo abordado toda a governança, estratégia, liderança e a própria gestão de pessoas que engloba todo o ecossistema. Utilizou-se deste termo por aparecer descrito na literatura pela maioria dos artigos analisados. A **inovação** que é apresentado como o foco de um ecossistema de inovação e a abordagem da P&D que pode proporcionar. Os **indicadores** como meio de mensurar a evolução e o desenvolvimento do ecossistema. Por fim a **plataforma** cujo o objetivo é dar o aporte tecnológico necessário para que haja uma relação sinérgica entre os envolvidos no ecossistema. A seguir é apresentado com maior detalhamento cada elemento identificado na literatura.

Arquitetura é apresentada como sendo parte de modelos ou tipos de estruturas. A mesma esta relacionada também com a organização interna e tecnológica do ecossistema de inovação (BOSCH-SIJTSEMA; BOSCH, 2014). Em ecossistemas de inovação a arquitetura que os compões pode ser diferenciada em dois tipos, a arquitetura modular que visa a flexibilidade da relação entre os atores do ecossistema e a arquitetura integral cujo o foco é a fixação da relação com os atores (WEIL; SABHLOK; COONEY,). A arquitetura é apresentada como um fator de sucesso para a implementação do ecossistema (DURST; POUTANEN, 2013).

Ator refere-se aos diferentes tipos de participantes do ecossistema, se dividindo inicialmente em duas partes: **(i)** recursos materiais (sendo os fundos, equipamentos e instalações, dentre outros) e **(ii)** recursos humanos (sendo os alunos, professores, governos e empresas, entre outros) (JACKSON, 2011). Outros autores encontrados na literatura são: empresas, escritórios, capitais de risco, investidores, universidades, políticos, institutos de pesquisa, bancos, indústrias, fundações, organizações científicas, organizações econômicas e organizações governamentais (WEIL; SABHLOK; COONEY, ; KOLOSKY; SPERONI; GAUTHIER, 2015). Outro tipo de ator presente na literatura é denominado de híbrido, pois o mesmo gradualmente toma o lugar e as funções do outro ator dentro do ecossistema (SILVA; NARCIZO; CARDOSO, 2012). Além da denominação ator encontrada na literatura, encontra-se também a denominação participante, sendo estes: cientistas, tecnólogos, pesquisadores, empresas, governo, indústria, professores, alunos e empreendedores (HARDASH et al., 2014).

Cultura é apontada como fator de sucesso e definida como um

elemento não tecnológico importante para o ecossistema, assim como os elementos tecnológicos (OH et al., 2016; DURST; POUTANEN, 2013). A cultura de inovação deve estar presente no ecossistema sendo vista como um fator necessário, pois a mesma deve estar no cotidiano dos atores do ecossistema, tendo com objetivo o avanço econômico (HARDASH et al., 2014).

Gestão é responsável por guiar o ecossistema de inovação a fim de fortalece-lo, possibilitando o gerenciamento das suas atividades e a criação e desenvolvimento de valores. Uma técnica de gestão propostas pelos autores é a utilização da estratégia de inovação *outsourcing*, que consiste na terceirização da gestão estratégica da inovação no ecossistema (WEIL; SABHLOK; COONEY,).

Inovação é a principal diferença para distinguir um ecossistema de inovação de qualquer outro, cujo o seu foco não seja inovação (OH et al., 2016).

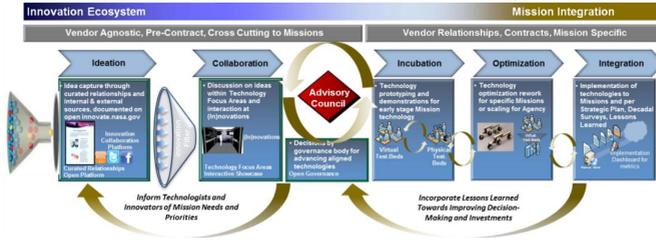
A inovação é apresentada como fonte de vantagem competitiva e crescimento econômico global (HUHTAM; RUBENS et al., 2016). Observa-se na literatura a importância do uso adequado da palavra inovação, pois a mesma não se trata apenas de uma palavra que está na moda, mas inovação é colocada como um fator indispensável em um ecossistema (HARDASH et al., 2014). A inovação é abordada como uma criação de valor para o mercado que possibilita o aumento de competitividade (DURST; POUTANEN, 2013). Sendo vista como um resultado de um processo complexo e dinâmico com relações entre as áreas de ciência, tecnologia, pesquisa e desenvolvimento envolvendo a Rede de universidades, empresas e governo (SILVA; NARCIZO; CARDOSO, 2012).

A inovação Aberta é apresentada como uma diversidade de atores que contribuem para um objetivo comum (OH et al., 2016), sendo um fator importante para obtenção de novas perspectivas de P&D, auxiliando assim, no crescimento do ecossistema (BOSCH-SIJTSEMA; BOSCH, 2014). Através da inovação aberta se possibilita a aceleração da inovação e expansão do mercado (WEIL; SABHLOK; COONEY,), o aumento da participação dos atores externos para o processo de inovação focando em P&D, proporcionando o aceleramento da inovação interna e externa para o mercado (DURST; POUTANEN, 2013). Para a representação da inovação aberta é apresentado o modelo utilizado por Hardash et al. (2014) aonde todas as ideias podem ser utilizadas e retrabalhadas a fim de atender a um objetivo. O modelo utilizado no ecossistema de inovação da NASA¹, contempla várias etapas, desde a captação das ideias, colaboração, aprimoramento, análise das ideias, incubação, otimização

¹<https://www.nasa.gov>

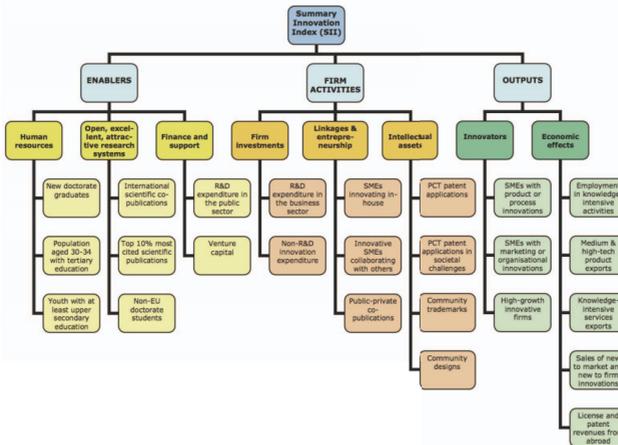
e prototipagem até que a ideia possa ser utilizada em uma missão, como apresentado na Figura 7.

Figura 7 – Ciclo de vida para ecossistema de inovação da NASA (HARDASH et al., 2014)



Indicadores foi identificado na literatura o *Framework Innovation Union Scoreboard* utilizado para mensurar o desempenho da inovação na União Europeia, que consiste em três principais indicadores e oito dimensões de inovação obtendo um total de 25 indicadores (HUHTAM; RUBENS et al., 2016). A seguir é apresentado o *Framework Innovation Union Scoreboard* na Figura 8:

Figura 8 – **Innovation Union Scoreboard** (HUHTAM; RUBENS et al., 2016)



Plataforma é apresentada como meio em que os clientes, usuários e desenvolvedores possam construir uma relação sinérgica e possam potencializar a valorização do ecossistema (OH et al., 2016), também sendo a base para a inovação aberta objetivando ancorar todo o ecossistema (BOSCH-SIJTSEMA; BOSCH, 2014). Uma plataforma é um meio de conexão entre os atores, deste modo deve-se haver uma preocupação para que a plataforma atenda adequadamente as necessidades do ecossistema (HUHTAM; RUBENS et al., 2016; WEIL; SABHLOK; COONEY,). Uma plataforma aberta e publica possibilita que interessados conheçam os projetos, assim como compartilhem suas ideias, podendo assim colaborar com a geração de novas ideias e beneficiar a todos os envolvidos (HARDASH et al., 2014).

2.1.2 Ecossistema de Software

Os ecossistemas de software surgiram com objetivo de melhorar a reutilização de softwares na indústria, considerando as relações entre empresas e demais interessados (SANTOS; WERNER, 2011). O ecossistema de software é oriundo do ecossistema de negócios, de tal maneira que as empresas possam co-evoluir suas capacidades em torno de uma inovação, trabalhando competitivamente e cooperativamente para apoiar novos produtos, satisfazendo as necessidades dos clientes e eventualmente agregando mais à inovação (LARRUCEA; NANCLARES; SANTAMARIA, 2016). Os ecossistemas de software ou SECO como também são apresentados, é um meio eficaz na construção de grandes sistemas de software utilizando como base uma plataforma tecnológica comum, onde desenvolvedores internos e externos possam enriquecer o ecossistema contribuindo com componentes de software, como é o caso do Android ou IOS (MANIKAS; HANSEN, 2013b).

Atualmente os ecossistemas de software *Open Source* são os mais estudados, em decorrência da expansão do software livre tanto para computadores quanto para mobile, como é o caso do sistema Android. Logo, tem-se um crescimento do tema em áreas multidisciplinares (MANIKAS, 2016).

Para que haja sobrevivência do ecossistema e que o mesmo possa funcionar adequadamente é de suma importância o monitoramento da saúde do ecossistema. Na literatura apresentam-se três indicadores para que se possa fazer este monitoramento, sendo: a produtividade onde é verificado o quão produtivo o ecossistema se encontra, a robustez onde é analisado o quão preparado o ecossistema se encontra para

enfrentar níveis de stress ou situação inesperadas, como a saída de um desenvolvedor em pleno projeto e o nicho onde também é avaliada a capacidade em que o ecossistema tem de criar novas oportunidades de crescimento e para novos atores (MANIKAS; KONTOGIORGOS, 2015; SANTOS; WERNER, 2011; MANIKAS; KONTOGIORGOS, 2015; ALAMI; RODRÍGUEZ; JANSEN, 2015). A saúde de um ecossistema de software é uma indicação de quão bem o ecossistema está funcionando e a avaliação da saúde pode apontar questões que precisam ser abordadas e melhoradas no ecossistema (MANIKAS; HANSEN, 2013a)

Na literatura observa-se a existência de alguns conceitos de ecossistema de software, onde o mais atual encontrado é de Manikas (2016) que define o ecossistema de software como "a interação do software e ator em relação a uma infra-estrutura tecnológica comum, que resulta em um conjunto de contribuições que influênciam diretamente ou indiretamente no ecossistema". Foram encontrados outros conceitos de ecossistema de software na literatura como o de Jansen, Finkelstein e Brinkkemper (2009) que define o ecossistema de software como:

"um conjunto de empresas que funcionam como uma unidade e interagindo com um mercado comum para software e serviços, juntamente com as relações entre eles. Essas relações são frequentemente apoiadas por uma plataforma tecnológica comum ou mercado e operam através da troca de informações, recursos e artefatos"pg 187-188.

. Outro conceito encontrado foi de Manikas e Hansen (2013a) onde o ecossistema de software é apresentado como:

"um ecossistema de software é uma interação de um conjunto de atores sobre uma plataforma tecnológica comum, que resulta em um número de soluções ou serviços de software. Cada ator é motivado por um conjunto de interesses ou modelos de negócio e está conectado aos demais atores e ao ecossistema de software como um todo por meio de relacionamentos simbióticos. Por sua vez, a plataforma tecnológica está estruturada para permitir o envolvimento e a contribuição dos diferentes atores"pg 1297-1298.

Para a identificação dos elementos que compõem o ecossistema de software foram determinados alguns critérios com base na literatura (MANIKAS, 2016; LETTNER et al., 2014; PELLICCIONE, 2014). Deste

modo utilizou-se alguns termos classificados como características, apresentados pelos autores, sendo assim apresenta-se abaixo os elementos identificados.

- | | |
|--|--|
| a) Desenvolvedores internos e externos | m) Controle do Ecossistema de Software |
| b) Rede de Atores | n) Plataforma tecnológica comum |
| c) Estrutura Organizacional | o) Keystone |
| d) Estrutura de Negócios | p) Stakeholders |
| e) Estrutura de Software | q) Facilitador da comunicação |
| f) Gerenciamento | r) Técnicas de Qualidade |
| g) Governança | s) Arquitetura de Software |
| h) Saúde | t) Compartilhamento |
| i) Plataforma Restritiva | u) Plataforma Restritiva |
| i) Variabilidade | v) Reutilização |
| l) Controle de Qualidade | |

Essas características foram agrupadas para melhor descrever os elementos dos ecossistemas de software com base na literatura. Deste modo, a tabela 3 apresenta os agrupamentos.

Tabela 3: Agrupamento - Ecossistema de Software

| | Manikas, 2016 | Lettner et al., 2014 | Pelliccione, 2014 | Outros autores |
|--------------------|--|-----------------------------|-------------------------------------|--------------------------|
| Arquitetura | Estrutura Organizacional, Estrutura de Negócios, Estrutura de Software | Foco na Variabilidade | Comunicação Facilitada entre atores | Christensen et al., 2014 |

Tabela 3: Agrupamento - Ecossistema de Software

| | Manikas, 2016 | Lettner et al., 2014 | Pelliccione, 2014 | Outros autores |
|------------------|------------------------------|-------------------------------------|--------------------------|--|
| Atores | Desenvolvedor Rede de Atores | Desenvolvedor interno e/ou externo | Keystone, Stakeholder | Manikas, Hansen, 2013a; Christensen Et Al., 2014; Manikas, Kontogiorgos, 2015; Larrucea, Nanclares, Santamaria, 2016; Axelsson, Skoglund, 2016 |
| Qualidade | - | Controle de Qualidade | Técnicas de Qualidade | Axelsson, Skoglund, 2016; Hansen, Zhang, 2013 |
| Gestão | Gerência, Governança | Controle do Ecossistema de Software | - | Axelsson, Skoglund, 2016; Santos, Werner, 2011; Manikas, Hansen, 2013a |
| Inovação | - | - | Compartilhamento | Axelsson, Skoglund, 2016; Manikas, Hansen, 2013a; Larrucea, Nanclares, Santamaria, 2016 |

Tabela 3: Agrupamento - Ecossistema de Software

| | Manikas, 2016 | Lettner et al., 2014 | Pelliccione, 2014 | Outros autores |
|---------------------|----------------------|------------------------------|--------------------------|---|
| Indicadores | Saúde | - | - | Santos, Werner, 2011; Manikas, Kontogiorgos, 2015; Alami, Rodríguez, Jansen, 2015; Manikas, Hansen, 2013b |
| Plataforma | - | Plataforma Tecnológica Comum | Plataforma Restritiva | Axelsson, Skoglund, 2016; Schmid, 2013 |
| Reutilização | - | - | Replicar e Reutilizar | - |

Na tabela 3 são apresentados os agrupamentos das características encontradas na literatura, sendo cada agrupamento é denominado de elemento, sendo estes: a **arquitetura** visando a estrutura do ecossistema abordando três pilares centrais: a estrutura organizacional cujo o foco é determinar qual será o tipo de orquestração, a estrutura de negócios com foco em proporcionar a geração de valor e a estrutura de software objetivando em quais tecnologias serão utilizadas no ecossistema. Os **atores** abordam os diferentes tipo de participantes no ecossistema. Utilizou-se deste termo pois é o que mais aparece na literatura. A **qualidade** trata de como os participantes externos ao ecossistema podem contribuir com o desenvolvimento de softwares ou no crescimento do ecossistema sem prejudica-lo. A **gestão** é focada na governança e gerenciamento do ecossistema. Utilizou-se deste termo por ser mais utilizado na literatura analisada. A **inovação** é apresentada como reflexo do compartilhamento de produtos e projetos do ecossistema. Os **indicadores** são abordados como um meio de mensurar o crescimento e a evolução do ecossistema, também é colocado o termo "saúde" como meio de monitorar o funcionamento adequando do ecossistema. A **plataforma** apresentado como um meio comum em que os

atores ou participantes do ecossistema possam se comunicar e trabalhar em conjunto. Por fim a **reutilização** que possui foco na documentação de projetos legíveis e compreensíveis para que possa ser reutilizada e replicada. A seguir são apresentados com maior detalhamento cada elemento com base no agrupamento realizado.

Ator é apresentado como sendo os diferentes tipos de participantes do ecossistema, por exemplo: o *keystone*, os dominadores, *niche player*, orquestrador, academias, desenvolvedores, clientes, usuários, fornecedores, comunidades, entidades do governo e grupos de indústria de software (MANIKAS; HANSEN, 2013a; CHRISTENSEN et al., 2014; MANIKAS; KONTOGIORGOS, 2015; LARRUCEA; NANCLARES; SANTAMARIA, 2016; AXELSSON; SKOGLUND, 2016).

Arquitetura deve apoiar a natureza do ecossistema, ou seja, deve ser adaptada conforme a necessidade do ecossistema (MANIKAS, 2016). A modelagem da arquitetura do ecossistema de software faz uso de três estruturas, **(i)** estrutura organizacional abordando os elementos de software e os atores relacionando com a governança e organização do ecossistema, **(ii)** estrutura de negócios visando como os atores e os elementos de softwares criam, entregam e capturam valores para o ecossistema e **(iii)** estrutura software voltado a produção de aplicações de software (CHRISTENSEN et al., 2014). Destaca-se a arquitetura como sendo o pilar central do ecossistema onde toda decisão relacionada a arquitetura deve ser analisada com cuidado, visto que pode comprometer o ecossistema (PELLICCIONE, 2014).

Estrutura organizacional se define através do tipo de orquestração (ou gerencia/gestão) aplicada ao ecossistema. Pode ser classificada em **(a) Monarquia** é a orquestração de um ecossistema por um ator, onde ele é responsável por um ecossistema ou possui uma grande influencia sobre, **(b) Federal** é a orquestração por um conjunto de atores representativos, **(c) Coletivo** é a orquestração por meio de processos que envolvam todos os atores, como exemplo a votação e **(d) Anarquia** é caracterizada pela falta de orquestração geral e cada ator atua por conta própria, com base nas necessidades locais (MANIKAS, 2016).

Estrutura de Negócios responsável por gerar valor no ecossistema. Podendo ser classificada em **(a) Proprietário** onde a criação de valor no ecossistema é baseada em contribuições de propriedade. Ecossistemas deste tipo são normalmente protegidos pela propriedade intelectual processos de gestão e o valor refere-se a uma compensação monetária, **(b) Open Source** quando as contribuições são tipicamente abertas para o resto dos atores ou público. Valor, neste ecossistema

,normalmente, se refere à exceção de compensações monetárias, por exemplo, conhecimento e experiência ou necessidade de satisfação e **(c) Híbrido** o ecossistema que suporta ambas as contribuições fonte proprietário e abertas (MANIKAS, 2016).

Estrutura de software representa as tecnologias comuns presentes no ecossistema de software. Pode ser **(a) Plataforma** quando as contribuições são tipicamente em cima ou estendendo-se uma plataforma de software comum, **(b) Protocolo** toda interação do software é baseada em protocolos. Para a contribuição nesse tipo de ecossistema se utiliza de normas e compromisso com as especificações técnicas, **(c) Standard** onde interação software é alcançado com base no cumprimento de um conjunto padrão de (regras). Isto é semelhante ao protocolo, mas geralmente mais controlado e limitado. Contribuições aqui pode normalmente incluir mais conformidade técnica, enquanto o nível de cumprimento das normas de cada ator pode ser avaliado por terceiros, **(d) Infraestrutura** toda a interação do ator ocorre durante o desenvolvimento, isto é, utiliza-se uma plataforma de comum de desenvolvimento(ex: Github), sendo normalmente independente e muito diferente, implicando em contribuições de atores de domínio diferentes(ex: uma plataforma Github onde esta sendo desenvolvido um sistema de moveis e um ator de medicina pode contribuir) (MANIKAS, 2016).

Qualidade tem por objetivo oferecer qualidade sobre um produto de software que obteve a contribuição de desenvolvedores externos, evitando assim pacotes maliciosos e defeituosos que possam comprometer a saúde do ecossistema (PELLICCIONE, 2014), logo cumprindo também os requisitos definidos pelos *stakeholders* garantindo a qualidade dos conjuntos de atividades realizadas para que o sistema tenha as condições mínimas exigidas (AXELSSON; SKOGLUND, 2016). Outro ponto apresentado pela literatura é que em um ecossistema de software a qualidade não é avaliada apenas pelos softwares ali desenvolvidos, mas também por toda a influencia que o ecossistema exerce em cima de todos envolvidos (HANSEN; ZHANG, 2013).

Gestão são pilares centrais com foco no tipo de estrutura de negócio presente no ecossistema podendo ser uma estrutura proprietária, *Open Source* ou híbrida (MANIKAS, 2016; AXELSSON; SKOGLUND, 2016). O negócio também é apresentado com foco no fluxo de conhecimento, isto é, foco nos artefatos, foco nos recursos e informações desenvolvidas no ecossistema (SANTOS; WERNER, 2011). Outro ponto chave apresentado é a motivação dos atores no ecossistema, sendo que cada ator possui seu próprio interesse e seu modelo de negócio, caso o

ator não esteja alinhado poderá impactar negativamente no desenvolvimento do ecossistema (MANIKAS; HANSEN, 2013a).

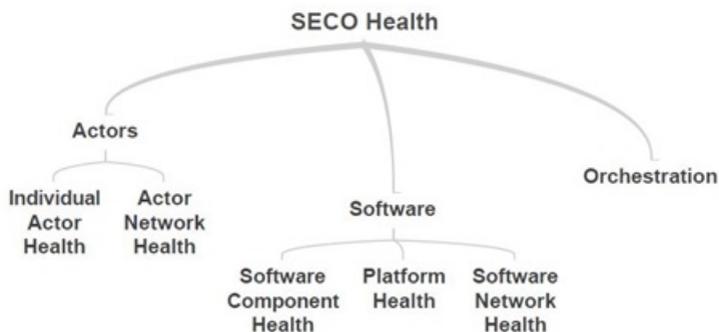
Inovação é apresentada como aliada juntamente ao compartilhamento da plataforma, pois deste modo auxilia a aceleração da inovação de um determinado produto (PELLICCIONE, 2014; AXELSSON; SKOGLUND, 2016). A inovação é apresentada com foco na produtividade do ecossistema e na capacidade do ecossistema em converter matéria prima em inovação com baixo custo ou novos produtos, possibilitando ao ecossistema o aumento da diversidade ao longo do tempo com a inovação (MANIKAS; HANSEN, 2013a). Outro ponto importante é a capacidade das empresas em co-evoluir em torno da inovação gerada através do ecossistema (LARRUCEA; NANCLARES; SANTAMARIA, 2016).

Indicadores esses elementos que compõem o ecossistema são responsáveis por avaliar o seu desenvolvimento, crescimento e sua sustentabilidade. Algumas métricas são apresentadas na literatura como: a saúde, a interação, a performance, as entradas e saídas, a concorrência e métodos de coordenação (SANTOS; WERNER, 2011). Alguns indicadores também são definidos para avaliar o ecossistema como: a produtividade que é responsável por analisar o nível de atividade no ecossistema, a robustez que aborda o nível em que o ecossistema está preparado para lidar com situações inesperadas como a saída repentina de um ator ou um avanço tecnológico que pode afetar toda a estrutura e o indicador nicho que se refere a capacidade do ecossistema tem de criar novas oportunidades para novos atores e proporcionar novas oportunidades de negócios para todos os envolvidos (MANIKAS; KONTOGIORGOS, 2015; ALAMI; RODRÍGUEZ; JANSEN, 2015).

Também é apresentado na literatura um *Framework* de Saúde cujo o objetivo é auxiliar no monitoramento do ecossistema utilizando-se de três pilares centrais como o **(a)** ator sendo subdivididos em (i) saúde do ator individual onde é mesurada a produtividade com foco no desenvolvimento de novos produtos e funções, já a robustez avalia o desgaste e stress sofrido pelo ator ao longo do tempo no ecossistema, (ii) a saúde da rede de atores trata também da produtividade e robustez do conjunto de atores, balanceando a baixa produtividade de um ator individual com a rede de atores equilibrando e protegendo o ecossistema, sendo importante ressaltar que os atores possuem influência direta na saúde do ecossistema. O elemento **(b)** software é subdividido em três sub-elementos (i) saúde do componente do software que pode ser mensurada através de quatro indicadores como: a confiabilidade responsável pela segurança que o software proporciona aos usuários e ao ecossistema; disponibilidade sendo resistente a falhas e mantendo

o software sempre em funcionamento; interoperabilidade que garante a comunicação com outros softwares ou sistemas independente a sua plataforma de desenvolvimento; e a possibilidade de modificação e prevenção do efeito cascata de um software no ecossistema. (ii) saúde da plataforma é definida como a integridade dos quatro indicadores da saúde do software sendo aplicada à plataforma. Também pode haver um acréscimo à plataforma dependendo da organização do ecossistema, possuindo o orquestrador como parte da plataforma, sendo o responsável pela organização e gestão de todo o ecossistema e (iii) saúde da rede de software que é a integração dos softwares com os envolvidos do ecossistema formando assim a rede de software, podendo ser mensurado através da conectividade entre os elementos do ecossistema. O último elemento apresentado no framework é o **(c)** orquestrador que tem um papel importante no ecossistema e é o responsável por todo o gerenciamento, monitoramento e avaliação, guiando o ecossistema para o seu constante crescimento e sobrevivência (MANIKAS; HANSEN, 2013b). A seguir a Figura 9 ilustra o *framework* de saúde do ecossistema.

Figura 9 – **Framework Saúde do Ecossistema de Software** (MANIKAS; HANSEN, 2013b)



Plataforma é uma tecnologia presente em ecossistemas, sendo apresentada como um meio comum para desenvolvedores e usuários trabalharem, um exemplo é o *GitHub*² (MANIKAS, 2016). Outro ponto é a capacidade da plataforma em moldar sua própria arquitetura e impor restrições para sistemas de software, podendo ser uma plataforma restritiva ou uma plataforma semi-aberta, onde é liberado o acesso parcial da plataforma a terceiros (PELLICCIONE, 2014; AXELSSON; SKOGLUND, 2016). Uma plataforma pode ser dividida em três modos: plataforma *sigle* que utiliza do modelo vertical do ecossistema, *multi-plataforma* que utiliza do modelo horizontal do ecossistema e a plataforma especializada, onde somente os stakeholders contribuem na plataforma (SCHMID, 2013). Uma plataforma deve possuir uma estrutura extensível e adaptável contendo um conjunto de componentes reutilizáveis e um conjunto de soluções prontas para utilização, sendo possível o compartilhamento e permitindo contribuições de participantes externos ao ecossistema, assim deve ser monitorado todas as contribuições para que se possa manter a qualidade do ecossistema (LETTNER et al., 2014).

Reutilização é um elemento de grande importância para o ecossistema, pois deve-se manter toda a documentação de software legível e compreensível para que possa haver uma replicação e reutilização. A reutilização também é apresentada como um novo paradigma onde a reutilização de recursos, componentes, códigos e bibliotecas podem dar origem a novos produtos ou a uma nova família de produtos, podendo ser assemelhando a linhas de produtos de software (LETTNER et al., 2014).

2.2 ARQUITETURA

Na história contemporânea a arquitetura nada mais é que a construção criada a fim de ordenar e organizar um espaço para determinada finalidade (COSTA, 1995). Com o passar do tempo, é observado o uso da arquitetura em outros contextos, sendo cada contexto uma aplicabilidade específica, como exemplo, em ecossistemas, softwares, computação, TI e entre outros.

A arquitetura é apresentada em um ecossistema de inovação como meio de organizar papéis, responsabilidades, processos e ferramentas (BOSCH-SIJTSEMA; BOSCH, 2014). Em sistemas baseados em componentes, a arquitetura é responsável pelas especificações de componentes do sistemas e a comunicação entre eles, garantindo o geren-

²<https://github.com/>

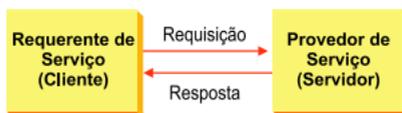
ciamento na construção de um sistema e sua manutenção (LEAVENS; SITARAMAN, 2000). Na arquitetura de informação é apresentada como maneira de organizar, estruturar, gerenciar toda a informação (ROSENFELD; MORVILLE, 2002). Por sua vez a arquitetura de software tem como preocupação a maneira em que o sistema vai ser organizado e como será a estrutura geral do sistema (SOMMERVILLE, 2011).

Como pode ser observado a arquitetura está presente em vários contextos, mas sempre seguindo sua essência independente sua aplicação. A seguir é apresentado dois tipos arquitetura.

- **Arquitetura Orientada a Serviço (SOA)**

- A arquitetura orientada a serviço, é uma arquitetura de suporte a computação orientada à serviço derivada da computação distribuída. O SOA vem sendo muito utilizada no meio empresarial onde é possível a adaptação e reestruturação organizacional e tecnológica, dando uma maior agilidade na implementação de novos serviços e a reutilização de serviços existentes (JR; OLIVEIRA; MEIRA, 2012; SOUZA et al., 2006). A seguir na Figura 10 é apresentado um modelo simplificado da arquitetura do SOA:

Figura 10 – **Arquitetura Orientada a Serviço** (SOUZA et al., 2006)



- **Arquitetura E-Gov**

- Esta arquitetura foi desenvolvida à partir de um estudo abordando e-gov e TIC onde foi possível verificar tendências tecnológicas e modelos de e-gov , como por exemplo, web 2.0, e-gov2.0, governo aberto, dados abertos, web semântica, e-gov semântico, que possibilita a viabilização de sistemas de conhecimento, tal estudo deu origem a arquitetura denominada de e-governança (E-Gov) (PACHECO et al., 2015). A seguir na Figura 11 é apresentado a arquitetura E-Gov

Figura 11 – **Arquitetura E-gov** (PACHECO et al., 2015)



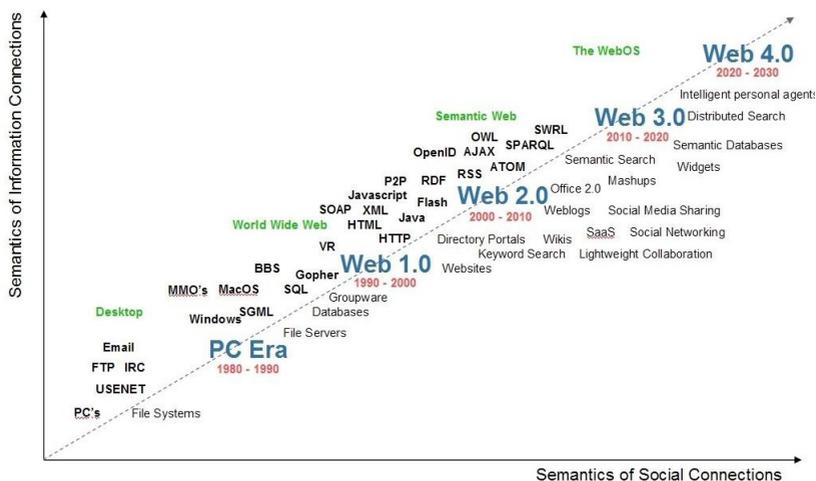
2.3 EVOLUÇÃO DA WEB

A *World Wide Web* ou apenas Web, vem evoluindo com passar do tempo, sendo visto atualmente como um meio de encurtar distâncias, compartilhar informação e expressar opiniões. No ano de 1989 Tim Berners-Lee, propôs um ambiente onde se pudesse apresentar textos, imagens, vídeos e animações de modo que fossem integrados por meio da internet, conhecido atualmente como Web (BERNERS-LEE, 1989).

Com a difusão da Web, Tim Berners-Lee fundou o Consórcio World Wide Web ou simplesmente W3C, oriundo do laboratório de ciência de computação do Massachusetts Institute of Technology - MIT, objetivando em regulamentar e desenvolver protocolos que ajudem a WEB crescer ao longo do tempo, para obter o seu rendimento máximo (W3C, 2016d).

A evolução da WEB vem sendo apresentada como etapas, sendo a WEB 1.0 onde o objetivo era apenas divulgar conteúdo para os usuários, com a evolução para WEB 2.0 os usuários tomaram uma posição mais ativa, podendo interagir com a WEB, compartilhando conteúdo, com o uso de blogs e wikis (RAMALHO, 2010). A seguir na Figura 12 apresenta um histórico evolutivo da WEB.

Na Figura 12 observa-se a presença de mais duas eras da Web, sendo a Web 3.0, denominada por alguns autores como a terceira onda, que tem por objetivo interligar dados e dar significados a eles possibilitando a interpretação por meio de máquinas (SANTOS; NICOLAU, 2015).

Figura 12 – **Histórico da WEB**(RAMALHO, 2010)

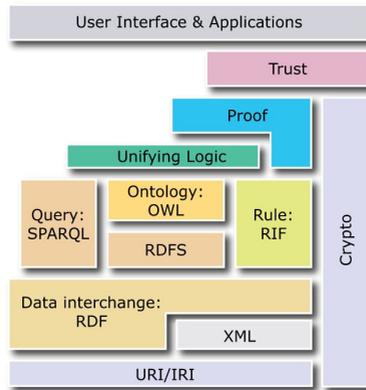
Na pesquisa realizada por Ramalho e Ouchi (2011) eles apontam que em 2006 o jornalista John Markoff fez uso pela primeira vez do termo Web 3.0 a denominando como a terceira geração de serviços Web baseados em sistemas inteligentes, sendo capaz de interpretar conteúdos e proporcionar respostas mais precisas ao usuário. Referente a Web 4.0, Aghaei, Nematbakhsh e Farsani (2012) relatam em sua pesquisa, como sendo um tema muito subjetivo ainda, mas acreditam que o seu objetivo será a interação de máquina e humanos em simbiose, sendo possível a construção de interfaces mais poderosas.

2.3.1 Web Semântica

Os seres humanos são capazes de interpretar e distinguir o significado das palavras conforme determinado contexto, isso não acontece com as máquinas, elas não conseguem compreender o significado das palavras (ALVES et al., 2005). A Web Semântica tem por objetivo tornar possível a interpretação de conteúdos de forma automática. Para isso ocorrer os documentos e dados precisam ser representados de modo que as máquinas possam compreendê-los (BERNERS-LEE, 2001). Os

padrões e tecnologias fazem que as máquinas compreendam o significado das informações presentes na WEB, possibilitando o compartilhamento e reutilização da informação (GOMES, 2013). O conjunto dos padrões e tecnologias presente na Web Semântica formam a arquitetura, sendo proposta originalmente pela W3C, que pode ser visualizada na Figura 13.

Figura 13 – **Arquitetura da Web Semântica** (w3c, 2016c)



Observando a representação da arquitetura da Web Semântica fica evidente a utilização de camadas e a interoperabilidade. Os padrões são amplamente utilizados por serem consolidados e recomendados da W3C. A seguir apresenta-se um breve detalhamento sobre as camadas presentes na Web Semântica.

- **URI / IRI:** *Uniform Resource Identifier* (URI) de forma simples e extensível à identificação de recursos, usando de uma sentença compacta de caracteres, assegurando a sua unicidade (BERNERS-LEE; FIELDING; MASINTER, 1998; GOÑI; FERNANDES; LUCENA, 2002). A *Internacional Resource Identifier* (IRI) é um novo protocolo que complementa a URI, utilizando de caracteres do *Universal Character*³(w3c, 2016e).
- **XML:** *Extensible Markup Language* (XML) é classificada como linguagem de marcação, é baseada na representação de informações estruturadas utilizando apenas textos simples (w3c, 2016f). O XML possui padrões relacionados, como os *namespace* (NS),

³Mais detalhes em <http://www.unicode.org/>

usando de esquemas para estruturar dados na Web sem nenhuma comunicação e nenhum significado. Sendo a base para outras tecnologias semânticas (AGHAEI; NEMATBAKHSH; FARSANI, 2012).

- **RDF: *Resource Description Framework* (RDF)** é um modelo de dados que objetiva representar informações, utilizando URIs para identificar recursos baseados na Web e descrever relações entre os recursos. Esta estrutura pode melhorar o processamento das informações por máquina, suportando normalmente a interoperabilidade a nível semântico. Quando se faz uso de RDF, consequentemente há uma semântica, deste modo os agentes de software são capazes de fazer inferências lógicas a fim de executar funções baseadas em metadados (LIMA; CARVALHO, 2005; AGHAEI; NEMATBAKHSH; FARSANI, 2012). A estrutura das informações declaradas são em forma de triplas, consistindo de Sujeito, Predicado e Objeto. Tanto sujeito como objeto denotam um recurso na Web, enquanto o predicado representará a relação entre os dois, sendo identificada por URIs (ANTONIOU, 2004; HARTH; JANIK; STAAB, 2011).
- **RDFS: *Resource Description Framework Schema* (RDFS)** é responsável por definir vocabulários no modelo RDF, descrevendo as propriedades e classes de recursos específicos que serão aplicadas em determinado domínio. Possibilitando assim uma estrutura simples de raciocínio para realizar inferências nos tipos de recursos (AGHAEI; NEMATBAKHSH; FARSANI, 2012; GOMES, 2013; HARTH; JANIK; STAAB, 2011).
- **Ontologia:** Tem por objetivo especificar formalmente um domínio ou discurso, tendo como principal característica uma lista finita de termos e seus relacionamento (ANTONIOU, 2004). A ontologia é representada computacionalmente utilizando o *Web Ontology Language* (OWL), permitindo descrever os aspectos semânticos e relacionamentos entre os conceitos de um domínio (RAMALHO; OUCHI, 2011; GOMES, 2013). Apresentando também como complemento para documentos RDF e RDFS, possibilitando aumentar ainda mais o nível de expressividade e a realizações de inferências (FILHO; LÓSCIO, 2016).
- **RIF: O *Rule Interchange Format* (RIF)**, é uma linguagem derivada do XML e um complemento para o OWL, sendo um padrão para o intercâmbio de regras orientadas a lógica de primeira ordem, possibilitando uma interoperabilidade entre as regras lógicas

utilizadas em outros sistemas (DOMINGUE; FENSEL; HENDLER, 2011; GOMES, 2013; RAMALHO; OUCHI, 2011).

- SPARQL: é um protocolo utilizado para consulta declarativa semelhante ao SQL, utilizado sobre dados RDF. Não possui meios para realizar inferências, sendo assim somente possibilita a manipulação das informações armazenadas (DOMINGUE; FENSEL; HENDLER, 2011; GOMES, 2013; RAMALHO; OUCHI, 2011).
- Logic: tem por objetivo facilitar a realização de inferências por agentes de software. Esta camada contém as regras que poderão ser utilizadas na execução de serviços inteligentes, podendo auxiliar no relacionamento e processamento das informações contidas em estrutura de dados (CARLAN, 2006; AGHAEI; NEMATBAKHS; FARSANI, 2012; GOMES, 2013).
- Proof: se refere a prova, sendo responsável pela verificação da consistência dos dados que serão acessados pela Web Semântica. Esta verificação se dá através da execução das regras de inferências definidas na camada Lógica, visando garantir a segurança e a veracidade ao conteúdo publicado (CARLAN, 2006; GOMES, 2013; AGHAEI; NEMATBAKHS; FARSANI, 2012).
- Trust: é responsável por analisar a confiabilidade da prova dada (KOIVUNEN; MILLER, 2001). Esta análise se dá avaliando com auxílio da assinatura digital, a autenticidade das fontes e a confiabilidade das informações consultadas pelos agentes (CARLAN, 2006; GOMES, 2013; AGHAEI; NEMATBAKHS; FARSANI, 2012).

2.4 LINKED DATA

O termo Linked Data foi cunhado por Tim Berners-Lee em 2006, sendo que no contexto de Web Semântica Linked Data não se objetiva apenas em inserir dados na Web, mas sim em como realizar as ligações de modo que uma máquina ou um humano possa explorar a Web. Deste modo Linked Data é definido como "*um conjunto de boas práticas para conectar e publicar dados estruturados na Web de diferentes fontes*" (BERNERS-LEE, 2006).

A publicação de dados na Web já vem sendo realizada a algum tempo, mas tais abordagens tradicionalmente são realizadas com a utilização de dados brutos em formato CSV, XML ou HTML, sem preocupação com a integração ou com sua semântica (SHETH, 2011). A Web

tradicional é formada pela composição de *links hypertext*, tal composição não é suficiente para ter uma maior expressividade e significância nos dados presentes na web atual. Deste modo o Linked Data utiliza de *links hyperdata*, com informações expressas em RDF onde não apenas os documentos podem ser ligados, mas também as informações contidas neles (SHETH, 2011; BERNERS-LEE, 2006). O uso de Linked Data traz alguns benefícios (AUER; LEHMANN; NGOMO, 2011), sendo estes apresentados a seguir:

- **Uniformidade:** Com a utilização do modelo de dados RDF toda a informação publicada é representada em fatos expressos por termos consistindo por sujeito, predicado e objeto. Deste modo todos os dados são baseados em um modelo uniforme.
- **Capacidade de ser referenciado:** Utiliza-se de URIs para identificar entidades e permite a localização, recuperação de recursos descritos e representados na web por estas entidades assim como uma URL.
- **Coerência:** As triplas RDF possuem URIs sendo estruturadas com sujeito, predicado e objeto. Estas triplas estabelecem uma ligação entre as entidades identificadas pelo sujeito e o objeto sendo interligadas por um predicado que representa a ligação entre eles, agregando maior qualidade no dado.
- **Integrabilidade:** Todas as fontes de dados em Linked Data são baseado em um único mecanismo para representar a informação, sendo de fácil integração semântica por diferentes conjuntos de dados.
- **Pontualidade:** Todas publicações e atualização do Linked Data são simples, facilitando a rápida disponibilização.

Para a utilização do Linked Data Tim Berners-Lee criou quatro recomendações básicas para que possa publicar dados em Linked data (BERNERS-LEE, 2006), sendo:

- I - Use URIs como nomes para as coisas;
- II - Use HTTP URIs para que as pessoas possam procurar esses nomes;
- III - Quando alguém procura um URI, fornecer informações úteis, usando os padrões (RDF ou SPARQL);

IV - Incluir links para outros URIs. Para que se possa descobrir mais informações.

Através do Linked Data se originaram outros projetos como o Linked Open Data visando a publicação de dados abertos, o Linked Government Data visando a publicação dos dados abertos governamentais e Enterprise Linked Data visando a publicação de dados de empresas (PIZZOL et al., 2015), a seguir será abordado um destes projetos.

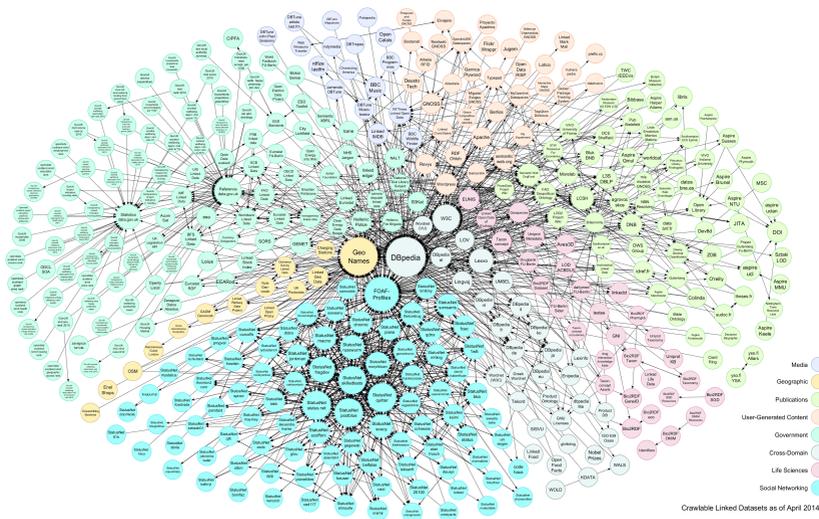
2.4.1 Linked Open Data

Em 2007 Tim Berners-Lee iniciou o projeto Linked Open Data (LOD), com objetivo de construir a Web de dados utilizando de *datasets* com licença aberta e transformando-os em RDF de acordo com as recomendações de Linked Data para sua publicação (BIZER; HEATH; BERNERS-LEE, 2009).

Em 2009 dados abertos começaram a ganhar visibilidade, quando governos de países desenvolvidos como Estados Unidos, Reino Unido, Canadá e Nova Zelândia iniciaram o incentivo da abertura de suas informações governamentais ao público (HANDBOOK..., 2016). Desde então iniciaram-se estudos e abordagens sobre dados abertos, sendo considerado o contexto que "os Dados Abertos e conteúdos podem ser livremente usado, modificado, e compartilhadas por qualquer um para qualquer finalidade" (DEFINITION..., 2016).

A tecnologia Linked Data permite que haja o compartilhamento dos dados auxiliando as iniciativas de publicação de dados governamentais, de tal modo que todos os dados conectados ajudaram a incrementar a nuvem de *Linked Open Data* (BIZER; HEATH; BERNERS-LEE, 2009). Esta nuvem é baseada em *datasets* conectados e publicados por indivíduos ou organizações que contribuem com o projeto. A seguir é apresentado a Figura 14, mostrando a nuvem LOD de 2014, sendo sua última atualização disponível.

Figura 14 – Nuvem LOD 2014(LOD..., 2016)



Observado a nuvem LOD, é dividida em oito domínio representados por uma coloração diferenciada, totalizando 1014 *datasets*. A seguir na Tabela 4 apresenta os domínios, quantidade de *dataset* por domínio, o percentual de predominância de cada domínio em relação a nuvem e sua coloração. Salienta-se que a nuvem possui dois domínios com maior destaque, sendo o governo com 183 *datasets* com 18.05% e a Web Social com 520 *datasets* com 51.28% de predominância nas publicações, juntas totalizam 69,3%, sendo mais da metade dos *datasets* presentes na nuvem.

dados abertos, sendo que a quarta estrelas refere-se a utilização de dados com uso de RDFs para que possam ser apontados e a quinta estrelas refere-se a vincular os dados publicados com outros dados também publicados a fim de fornecer um contexto, deste modo os dois últimos níveis referem-se a dados apresentados no formato LOD.

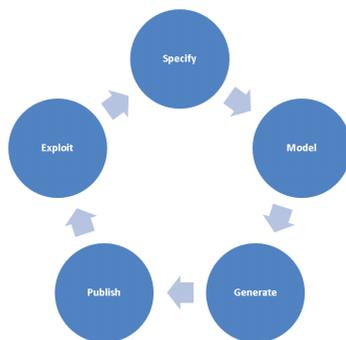
Neste contexto, para o auxílio na publicação de dados em formato LOD, foram desenvolvidos ciclos de vida para a publicação de dados, sendo que estes serão discutidos a seguir.

2.4.2 Ciclo de Vida do Linked Open Data

O ciclo de vida do LOD foi desenvolvido para auxiliar, de maneira detalhada, como efetuar a publicação dos dados no formato Linked Data e Linked Open Data. O ciclo de vida é definido como um sistema proposto para abordar todas suas fases de existência, sendo muitas vezes associada com o desenvolvimento de software (BROEK; VEENSTRA; FOLMER, 2014).

Na sequência são apresentados três tipos de ciclo de vida para publicação de dados. O primeiro, foi proposto como um guia metodológico com objetivo de auxiliar governos a publicarem dados para que possam ser re-utilizados posteriormente pelo público (VILLAZÓN-TERRAZAS et al., 2011), como é apresentado na Figura 16.

Figura 16 – Ciclo de Vida LOD - 1 (VILLAZÓN-TERRAZAS et al., 2011)

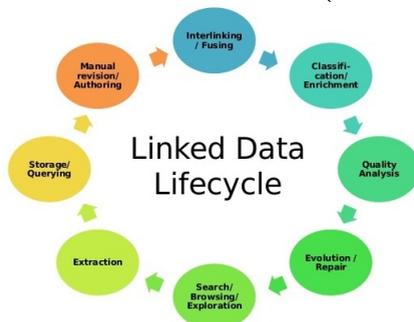


Como pode ser observado este ciclo de vida é composto por cinco etapas, **(1)** Especificação, **(2)** Modelagem, **(3)** Geração, **(4)** Publicação e **(5)** Exploração. A seguir é detalhado cada etapa:

- **Especificação:** Esta etapa trata do levantamento dos requisitos mais importantes referente a quais dados podem ser publicados. Para a realização desta etapa um conjunto de tarefas são necessárias, dentre estas estão as tarefas de **a** analisar e identificar as fontes de dados, **b** design da URI e **c** a definição da licença.
- **Modelagem:** Esta etapa trata das necessidades de utilizar a ontologia para modelagem de domínio das fontes de dados que foram identificadas, selecionadas e analisadas na etapa anterior. Mas uma parte importante para o desenvolvimento da ontologia é a utilização dos vocabulários possibilitando a reutilização.
- **Geração:** Esta etapa aborda a geração das triplas em formato RDF conforme os princípios de Linked Data. Esta transformação para RDF é realizado com base nos vocabulários criados na atividade de modelagem da etapa anterior. Para esta etapa são determinadas algumas tarefas para a conclusão da mesma, sendo estas: **a** transformação dos dados em RDF, **b** limpeza dos dados e **c** a ligação dos dados.
- **Publicação:** Esta etapa trata simplesmente da publicação dos dados em RDF. Para realização desta etapa são determinadas as seguintes tarefas: **a** publicar os *datasets*, **b** publicar os metadados e **c** disponibilizar a descoberta dos dados na Web.
- **Exploração:** Esta etapa aborda o incentivo para o público e comércio utilizar e reutilizar os dados publicados em Linked Open Data, a fim de desenvolver aplicativos que forneçam interfaces gráficas aos cidadãos.

Outro ciclo de vida proposto é o criado pela Universidade de Leipzig na Alemanha que tem por objetivo auxiliar a publicação em Linked Data. Este ciclo de vida possui três pilares centrais, que compreendem um quadro metodológico e tecnológico para integrar os componentes heterogêneos, com o intuito de gerar um framework consistente (AUER; LEHMANN; NGOMO, 2011; AUER et al., 2012). A seguir na Figura 17 é apresentado o ciclo de vida.

Figura 17 – Ciclo de Vida LOD - 2(AUER et al., 2012)



Este modelo é composto por oito etapas para a realização da publicação de dados em Linked Data. Na sequência serão brevemente detalhadas estas etapas.

- **Extração:** esta etapa aborda o mapeamento de informações em formato estruturado, semi-estruturados ou não estruturados para o modelo de dado em RDF.
- **Armazenamento e Consulta:** trata do banco de dados em RDF, onde é armazenado, indexado e realizado as consultados dos dados armazenados.
- **Autoria:** trata da etapa onde o usuário tem oportunidade de criar ou atualizar as informações.
- **Link:** etapa onde diferentes editores de dados fornecem informações sobre os mesmos ou entidades relacionadas são interligadas.
- **Enriquecimento:** esta etapa possui como foco o enriquecimento dos dados com estruturas de nível superior, a fim de agregar e consultar dados de forma mais eficiente. Deste modo o enriquecimento visa suprir a falta de classificação, estrutura e informações presentes nos esquemas.
- **Qualidade:** esta etapa aborda a avaliação da qualidade dos dados que serão publicados na Web.
- **Evolução:** esta etapa se refere a parte após a identificação de problemas existentes nos dados, onde é necessário o reparo dos mesmos, para que se possa haver uma evolução do dado ligados.

- **Exploração:** esta etapa final trata da informação publicada e disponível na Web de dados, de forma fácil e rápida onde os usuários possam navegar e explorar os dados.

O último ciclo de vida do LOD a ser abordado, foi desenvolvido a fim de atender não só as etapas técnicas abordados em outros ciclos de vida, mas também considerando a pós publicação dos dados levando em conta a re-utilização deles. Este modelo foi desenvolvido com base em um estudo literário sobre os modelos de ciclo de vida com a aplicação de um estudo de caso em uma empresa semi-pública da Holanda, tendo como objetivo a abertura dos dados com estratégia organizacional (BROEK; VEENSTRA; FOLMER, 2014).

Figura 18 – Ciclo de Vida LOD - 3 (BROEK; VEENSTRA; FOLMER, 2014)



Neste ciclo de vida observa-se que há diferenças entre os ciclos dos autores Boris Villazón-Terrazas, Luis. M. Vilches-Blázquez, Oscar Corcho, Asunción Gómez-Pérez, Sören Auer, Lorenz Bühmann, Christian Dirschl, Orri Erling, Michael Hausenblas, Robert Isele, Jens Lehmann, Michael Martin, Pablo N. Mendes, Bert Van Nuffelen, Claus Stadler, Sebastian Tramp, Hugh Williams, sendo que possui cinco etapas principais: (i) Identificação, (ii) Preparação, (iii) Publicação, (iv) Reuso e (v) Avaliação, sendo dividida em sub etapas com os responsáveis por cada uma delas, sendo que os responsáveis podem ser: (a) Alta gerencia, (b) Gestor da Informação, (c) Consultor Jurídico, (d) Gestor da Comunidade e (e) Proprietário do dado. A seguir será detalhada cada

uma das etapas do ciclo de vida.

- **Identificação:** nesta primeira etapa do ciclo, se tem a responsabilidade dos processos de definição e identificação dos dados que serão publicados.
 - Definir Estratégia: esta sub-etapa se refere ao desenvolvimento de uma estratégia para abertura dos dados que contribua com a missão organizacional, sendo responsabilidade da alta gerência. A estratégia criada deve além de identificar quais dados serão publicados deve verificar também quais podem se reutilizados. Outro ponto é a promoção de atividades econômicas, visando futuros usuários e suas necessidades e demandas.
 - Selecionar Dados: esta segunda sub-etapa se refere a seleção dos dados que serão abertos, deste modo o gestor de informação e os proprietários dos dados identificam os dados que poderão ser publicados com base na estratégia definida na etapa anterior.
- **Preparação:** esta segunda etapa refere-se ao preparo dos *dataset* para publicação, salienta-se a importância do consultor jurídico nesta fase.
 - Definir Requisitos: nesta sub-etapa o gerente de informação e o consultor jurídico elaboram as exigências dos dados levando em consideração os requisitos técnicos como o nível de qualidade dos dados, padrões, metadados, as exigências econômicas como a proposta de valor e modelo de negócio e os requisitos legais como a autorização aberta e licenças. O gerente de projeto tem como papel nesta etapa envolver as partes interessadas na definição dos requisitos dos dados para que não haja eventuais problemas relacionado ao projeto.
 - Preparar os Dados: esta etapa refere-se a preparação técnica dos dados, sendo de responsabilidade do gestor de informação e do proprietário dos dados para que tenha um gerenciamento de um conjunto de dados específico. Conforme os requisitos definidos na etapa anterior pode-se incluir a modelagem, descrição, conversão e o armazenamento dos dados. Outro ponto importante é a clareza das propriedades dos dados para que possam ser publicados.

- **Publicação:** esta etapa refere-se a efetiva publicação dos dados, sendo salientado que para que haja a reutilização dos dados, se faz necessário além da preocupação da organização na utilização técnica dos dados o alinhamento com a comunidade, pois este é um fator importante para garantir a sua reutilização.
 - Garantir a Usabilidade: para garantir a reutilização dos dados, a primeira sub-etapa considera certificar-se de que os dados publicados serão encontrados pelos utilizadores, sendo possível registrando os dados e metadados em um catalogo de dados já existentes como *dados.gov.br* ou *datahub.io*. Para isso, é importante encontrar uma plataforma adequada para a publicação, que atraia a atenção dos usuários, mas também deve-se considerar a criação de um portal de dados próprio para a publicação. Isso deve ser realizado pelo gerente do projeto juntamente com o gerente de informação.
 - Publicidade dos dados: esta etapa é de responsabilidade do gerente de comunidade, que possui a tarefa de realizar o incentivo e a promoção dos dados através de atividades como comunicados de imprensa, uso de blogs, competições de desenvolvimento de aplicativos, *hackatons*, jornais de informação ou uma premiação de aplicativos.
- **Reuso:** esta etapa trata da reutilização dos dados, tendo como responsável o gerente da comunidade que orienta os proprietários dos dados à publicarem dados menos complexos quando forem utilizados para incentivar o uso em eventos.
 - Construção da Comunidade: esta sub-etapa refere-se a construção de uma comunidade ativa que reutilize os dados publicados pela organização, visto que uma comunidade ativa ajuda no processo de *feedback* a fim de melhorar a qualidade dos dados publicados e refletindo em uma maior reutilização. O Gerente da comunidade tem um papel importante, pois ele é responsável pelos contatos com parceiros externos como organizações, empresários, incubadoras e instituídos de pesquisa interessados em construir uma rede ativa de reutilização dos seus dados, promovendo novos produtos, soluções ou inovação.
 - Gestão dos Dados: esta sub-etapa é de responsabilidade do gestor da informação e do proprietário dos dados, que estes

possuem um papel ativo após a publicação, onde a construção de um plano para a gerência e a certificação dos dados é importante para manter a qualidade dos dados permanecendo em no nível desejado.

- **Avaliação:** esta etapa refere-se ao processo de avaliação dos dados publicados.
 - Avaliar a proposição dos dados: esta sub-etapa refere-se a fase de avaliação da proposta de valor dos dados publicados, ou seja, é a avaliação dos dados publicados em relação ao caso de negócio construído juntamente com a alta gerência. Além desse detalhe o gerente de projeto deve analisar o impacto dos dados publicados utilizando de indicadores como, **(a)**número de download, **(b)**combinação com outros *datasets*, **(c)**usuários, **(d)**aplicações e **(e)**usuários finais das aplicações. Toda a avaliação deve ser compartilhada e avaliada junto a alta gerência.
 - Incorporar estratégia nos processos de organização e de trabalho: esta sub-etapa aborda a avaliação da incorporação de LOD na estratégia e no processo organizacional, tendo em vista que a alta gerência deve acompanhar os *feedbacks* da implementação do LOD na estratégia organizacional e observar qualquer mudança cultural dentro da organização.

2.4.3 Tecnologias

Para a realização da publicação dos dados em LOD, é necessário fazer uso de tecnologias que auxiliem neste processo. As tecnologias apresentadas na maioria são *open-source* e utilizadas em sistema operacional Linux, sendo grande parte funcional somente na distribuição Debian pois é uma distribuição robusta e com uma grande comunidade ativa. A seguir serão apresentadas algumas ferramentas com base no ciclo de vida do LOD e com base na literatura (AUER; LEHMANN; NGOMO, 2011; AUER et al., 2012; VILLAZÓN-TERRAZAS et al., 2011; BIZER, 2016).

- **Triplify:**⁴ é responsável por realizar o mapeamento de HTTP-URI nos bancos de dados relacionados para que possa publicar em

⁴AKSW (2016g). **Triplify**. Disponível em: <<http://aksw.org/Projects/Triplify.html>>. Acesso em: 07 Setembro 2016.

Linked Data através da serialização dos dados. Esta ferramenta é utilizada na fase de Extração.

- **D2R Server:**⁵ usado para realizar a serialização dos dados, realizando o mapeamento dos banco de dados relacionais para RDF, especificando os recursos que são identificados e quais propriedades são usadas para descrever os RDFs. Esta ferramenta também é utilizada na etapa de Extração.
- **DBpedia Extraction:**⁶ para auxiliar na extração de triplas RDF presentes no Dbpedia. Esta ferramenta é utilizada também na fase de Extração.
- **Open Link Virtuoso:**⁷ é um banco de dados moderno e híbrido, comportando alguns tipos de dados, como: SQL, XML, RDF, TURTLE, RDF/XML, JSON, JSON-LD e muito utilizado em aplicações semânticas. Esta ferramenta é utilizada na etapa Storage/Query.
- **OntoWiki:**⁸ utilizado com base de conhecimento apresentado como uma mapa de informações, permite a criação de conteúdos semânticos com objetivo de auxiliar na gestão do conhecimento e proporcionando uma melhor interação do usuário. Esta ferramenta é comumente utilizada na fase de Aatoria.
- **RDFaCE Text Annotation:**⁹ é uma ferramenta utilizada na fase de Aatoria, que visa criar conteúdos e utilizar APIs semânticos para facilitar a anotação e edição de RDFa.
- **Silk:**¹⁰ é um *framework open source* que integra as fontes de dados heterogêneos. Sendo utilizado na fase de Ligação.

⁵Eisenberg (2016).**D2R Server**. Disponível em: <<http://d2rq.org/getting-started>>. Acesso em: 07 Setembro 2016.

⁶Lyu (2016).**DBpedia Extractor**. Disponível em: <<https://github.com/dbpedia/extraction-framework/wiki>>. Acesso em: 07 Setembro 2016.

⁷Web (2016). **Open Link Virtuoso**. Disponível em:<<http://virtuoso.openlinksw.com>>. Acesso em: 07 Setembro 2016.

⁸AKSW (2016d). **OntoWiki**. Disponível em:<<http://aksw.org/Projects/OntoWiki.html>>. Acesso em: 07 Setembro 2016.

⁹AKSW (2016e). **RDFaCE Text Annotation**. Disponível em:<<http://aksw.org/Projects/RDFaCE.html>>. Acesso em: 07 Setembro 2016.

¹⁰Isele et al. (2016). **Silk**. Disponível em:<<http://silkframework.org/>>. Acesso em: 07 Setembro 2016.

- **LIMES:**¹¹ é um *framework* para descoberta de links na web de dados, sendo utilizado também na fase de Ligação.
- **DL-Learner:**¹² é uma ferramenta baseada em lógica descritiva, onde ela aprende os conceitos a partir de exemplos, sendo apresentado como uma aprendizagem de máquina baseado em lógica descritiva. Sendo utilizada na fase de Enriquecimento.
- **ORE:**¹³ ferramenta responsável por designar e agregar recursos na Web de dados. Esta ferramenta é utilizada na fase de Evolução.
- **Cubviz:**¹⁴ é uma ferramenta de dados estatísticos que faz uso de RDF em formato de cubo de dados, possibilitando vários *facets* de forma interativa, possibilitando a visualização em gráficos. Esta ferramenta é utilizada na fase de Exploração.
- **TopBraid Composer:**¹⁵ é uma ferramenta proprietária, deste modo para utilizá-la deve-se comprar uma licença mas há a possibilidade de utilizar para teste em um período de 30 dias. Esta ferramenta foca no auxílio da modelagem, construção e teste de configurações de ontologias e gráficos em RDF.
- **Protégé:**¹⁶ é uma ferramenta para o criação e edição de ontologias. Sendo disponível para versão desktop: Windows, Linux, MAC e também disponível em uma versão WEB.
- **Neon Tool Kit:**¹⁷ é um suíte para o desenvolvimento de engenharia de ontologias, sendo disponibilizado para sistemas operacionais Windows, Linux e MAC.

¹¹AKSW (2016c). **LIMES**. Disponível em:<<http://aksw.org/Projects/LIMES.html>>. Acesso em: 07 Setembro 2016.

¹²Lehmann (2016). **DL-Learner**. Disponível em:<<http://dl-learner.org/one-page-introduction/>>. Acesso em: 07 Setembro 2016.

¹³Archives (2016). **Open Archives - Initiative Object Reuse and Exchange**. Disponível em:<<http://www.openarchives.org/ore/1.0/datamodel>>. Acesso em: 07 Setembro 2016.

¹⁴AKSW (2016a). **Cubviz**. Disponível em:<<http://aksw.org/Projects/CubeViz.html>>. Acesso em: 07 Setembro 2016.

¹⁵Quadrant (2016). **TopBraid Composer**. Disponível em:<<http://www.topquadrant.com/tools/modeling-topbraid-composer-standard-edition/>>. Acesso em: 07 Setembro 2016.

¹⁶Research (2016). **Protégé**. Disponível em:<<http://protege.stanford.edu/>>. Acesso em: 07 Setembro 2016.

¹⁷Project (2016). **Neon ToolKit**. Disponível em:<<http://neon-toolkit.org/>>. Acesso em: 07 Setembro 2016.

- **Ultrawrap**:¹⁸ é uma ferramenta que integra os dados dinamicamente, independentemente da plataforma de banco de dados, sem a necessidade do armazenamento dos dados.
- **CKAN**:¹⁹ é um catálogo de dados online, todos os dados publicados em vários formatos como PDF, CSV, RDF, HTML e outros, são publicados nessa plataforma a fim de deixar os dados à fácil acesso, para que possam ser reutilizados posteriormente por outros usuários.
- **Tabulator**:²⁰ é um navegador de dados da Web de dados, sendo um plug-in que pode ser instalado no navegador *Firefox*.
- **Apache Stanbol**:²¹ é um suíte que auxilia na gerência dos conteúdos semânticos.
- **DBpedia Spotlight**:²² é uma ferramenta que identifica entidades automaticamente no texto em vários idiomas conforme a o grau de confiança determinado pelo usuário e realiza o vínculo com a DBpedia.
- **Jena**:²³ é um *framework* em java *open source* para construir aplicações em *Linked Data* e Web Semântica.
- **SemMap**:²⁴ é uma ferramenta para apresentar dados geoespaciais em dados semânticos em RDF através de consultas SPARQL.
- **Sieve**:²⁵ é uma ferramenta que avalia a qualidade dos dados e possui também um módulo de fusão, onde esta ferramenta analisa os dados de modo que se forem semelhantes só diferindo por tipo

¹⁸Capsenta (2016). **Ultrawrap**. Disponível em:<<https://capsenta.com/ultrawrap/>>. Acesso em: 07 Setembro 2016.

¹⁹Foundation (2016). **CKAN**. Disponível em:<<http://ckan.org/>>. Acesso em: 07 Setembro 2016.

²⁰W3C (2016a). **Tabulator**. Disponível em:<<https://www.w3.org/2005/ajar/tab.html>>. Acesso em: 07 Setembro 2016.

²¹Apache (2016b). **Apache Stanbol**. Disponível em:<<https://stanbol.apache.org/>>. Acesso em: 07 Setembro 2016.

²²Team (2016). **Bdpedia Spotlight**. Disponível em:<<https://github.com/dbpedia-spotlight/dbpedia-spotlight/wiki>>. Acesso em: 07 Setembro 2016.

²³Apache (2016a). **Jena**. Disponível em:<<https://jena.apache.org/>>. Acesso em: 07 Setembro 2016.

²⁴AKSW (2016f). **SemMap**. Disponível em:<<http://aksw.org/Projects/SemMap>>. Acesso em: 07 Setembro 2016.

²⁵<<http://sieve.wbsg.de/>>

de expressão por localidade ou nacionalidade é realizado a fusão dos dados, isso garante a qualidade dos mesmos.

- **Vilant**:²⁶ é uma ferramenta de extração e transformação que utiliza do *eXtensible Stylesheet Language for Transformation* (XSLT) para documentos em XML ou RDF.
- **LDIF**:²⁷ é um *framework open source* que objetiva-se na integração de dados podendo ser utilizado por aplicativos Linked Data para a conversão dos dados da Web, normalizar as URIs e mantendo o controle de proveniência de dados.
- **NG4J**:²⁸ é uma extensão da para o Jena a fim de analisar, manipular e serializar conjuntos de gráficos nomeados da web semântica.
- **Marbles**:²⁹ é um navegador semântico, onde possibilita usuário a explorar da web de dados.
- **WIQA**:³⁰ é uma *framework* que aborda a qualidade da informação na web de dados através de um conjunto de componente que abrange algumas políticas da qualidade da informação a fim de filtrar as informações na web.
- **Pubby**:³¹ é uma interface para o uso SPARQL Endpoint a fim de deixar a interação com o usuário mais amigável através do uso de uma aplicação web em JAVA.
- **RAP**:³² é um pacote de software que proporciona a análise, manipulação, serialização de RDP para a linguagem de programação PHP.

²⁶W3C (2016b). **Vilant**. Disponível em: <<https://www.w3.org/XML/2000/04rdf-parse/>>. Acesso em: 07 Setembro 2016.

²⁷Schultz et al. (2016). **LDIF**. Disponível em: <<http://ldif.wbsg.de>>. Acesso em: 03 Dezembro 2016.

²⁸Bizer, Cyganiak e Hartig (2016). **NG4J**. Disponível em: <<http://wifo5-03.informatik.uni-mannheim.de/bizer/ng4j/>>. Acesso em: 03 Dezembro 2016.

²⁹Bizer e Becker (2016). **Marbles**. Disponível em: <<http://mes.github.io/marbles/>>. Acesso em: 03 Dezembro 2016.

³⁰Bizer, Cyganiak e Gauss (2016). **WIQA**. Disponível em: <<http://wifo5-03.informatik.uni-mannheim.de/bizer/wiqa/>>. Acesso em: 03 Dezembro 2016.

³¹Bizer e Cyganiak (2016). **Pubby**. Disponível em: <<http://wifo5-03.informatik.uni-mannheim.de/pubby/>>. Acesso em: 03 Dezembro 2016.

³²Bizer e Project (2016). **RAP**. Disponível em: <<http://wifo5-03.informatik.uni-mannheim.de/bizer/rdfapi/>>. Acesso em: 03 Dezembro 2016.

Também observa-se que por meio destas ferramentas e utilizando o REST é possível convergir em um novo software ou plataforma para a publicação e consumo de dados em LOD. Deste modo, foram construídas algumas aplicações para o uso de *Linked Open Data* como o desenvolvido pela Universidade de Leipzig na Alemanha o *GeoKnow*³³ que apresenta dados geoespaciais ou informações geográficas identificando recursos naturais, construções e limites sobre a Terra.

Considerações finais do capítulo

Neste capítulo de revisão da literatura observa-se a origem dos ecossistemas de software e inovação e sua semelhança com o ecossistema da ecologia. Observa-se também que os ecossistemas de inovação e ecossistemas software, apresentam um grande volume de pesquisa e ambos os ecossistemas possuem elementos que são particularidades de cada um, mas identificam-se elementos que são comuns a ambos os ecossistemas, como por exemplo, os atores, a inovação, a arquitetura, entre outros.

No âmbito da Web pode-se notar a evolução nestes últimos anos e as contribuições que a Web Semântica vem proporcionando. Observa-se que desde a apresentação do conceito de web semântica muito estudos veem sendo realizados, dando origem a novos conceitos, como o de *Linked Data*(LD) e *Linked Open Data*(LOD).

Tais conceitos de LD e LOD possibilitaram o desenvolvimento de tecnologias e ciclo de vidas que auxiliam na publicação de dados, sendo uma abordagem para que qualquer indivíduo possa publicar dados possibilitando aos consumidores destes dados a utilização dos mesmos a fim de desenvolver novos produtos e proporcionar novos serviços. Os ciclos de vida são apresentados como guias no processo de publicação dos dados, sendo destacado o ciclo de vida Figura 18 (BROEK; VEENSTRA; FOLMER, 2014) identificado como o mais atual presente na literatura. Este ciclo de vida foi aplicado em uma organização semi-pública na Holanda e seu desenvolvimento foi focado na pós-publicação, sendo que o mesmo apresenta uma abordagem mais estratégica apontando os *stakeholders* envolvidos em cada etapa da publicação dos dados.

³³AKSW (2016b).

3 ANÁLISE E MODELAGEM

Neste capítulo será apresentada a análise realizada referente aos ecossistemas de inovação e de softwares a fim de mostrar suas diferenças e similaridades. Para a realização desta análise foi feito um agrupamento das características e/ou fatores de sucesso de ambos os ecossistemas, a fim elaborar um comparativo entre eles. Referente a abordagem das tecnologias e ciclos de vida que auxiliam na publicação dos dados, foi utilizado o último ciclo de vida identificado na literatura e as tecnologias mais citadas na literatura. Tal comparativo e a abordagem das tecnologias e ciclos de vida foram necessárias para a proposição da arquitetura e as recomendações para o ecossistema de inovação baseados em tecnologias semânticas.

3.1 ANÁLISE ENTRE OS ECOSSISTEMAS DE INOVAÇÃO E SOFTWARE

Como pode ser constatado na literatura os ecossistemas de inovação e ecossistemas software são semelhantes ao ecossistema da ecologia. Assim, como na ecologia, os ecossistemas de inovação e software possuem uma base conceitual semelhante, onde diferentes indivíduos trabalham em um mesmo ambiente, e cada indivíduo é responsável por uma parte do ecossistema, sendo que todos trabalham para a sobrevivência do mesmo. O ecossistema de inovação e ecossistema software possuem alguns elementos que os compõem, porém pode ser observado que alguns desses elementos são comuns entre eles. Deste modo, a seguir na Tabela 5 é apresentado um comparativo entre os elementos dos ecossistemas e, após, são abordados os elementos que são comum entre ambos os ecossistemas.

Tabela 5: Comparativo Ecosystemas (Autor, 2016)

| Elementos | Ecosystema de Inovação | Ecosystema de Software |
|-------------|--|---|
| Atores | São apresentados os recursos humanos: os alunos, professores, empresas, escritórios, capitais de risco, investidores, universidades, políticos, institutos de pesquisa, bancos, indústrias, governos e empreendedores. | O keystone, os dominadores, nicho player, orquestrador, academias, clientes, usuários, fornecedores, comunidades, entidades do governo e grupos de indústria de software. |
| Arquitetura | Possui foco na organização interna e tecnológica do ecossistema, podendo ser diferenciada em dois tipos a arquitetura: modular e integral. | Possui foco no apoio as necessidades do ecossistema e o uso de três estruturas: estrutura organizacional, estrutura de negócios e estrutura software. |
| Plataforma | Um meio em que os clientes, usuários e desenvolvedores possam construir uma relação sinérgica e possam potencializar valorização do ecossistema. | Um meio comum para desenvolvedores e usuários trabalharem, podendo ser do tipo single, multi-plataforma ou especialista. |
| Gestão | Apresentado como o responsável por fortalecer o ecossistema, gerando e agregando valor. | Foco no tipo de estrutura utilizada no ecossistema para a sua gestão e geração de valor, podendo ser do tipo proprietária, open source ou híbrida. |

Tabela 5: Comparativo Ecossistemas (Autor, 2016)

| Elementos | Ecossistema de Inovação | Ecossistema de Software |
|------------------------|---|---|
| Inovação | Fonte de vantagem competitiva, crescimento econômico global, a diversidade de atores que contribuem para o crescimento da inovação, aceleração da inovação e possibilidade de novas perspectivas P&D. | Foco no aceleração da produtividade do ecossistema e na capacidade do ecossistema em inovar com baixo custo e na co-evolução dos atores. |
| Qualidade | - | Foco na qualidade de produtos de software desenvolvidos por desenvolvedores externos ao ecossistema, mas também apresentado a qualidade em que o ecossistema se encontra e influencia sobre os todos os envolvidos. |
| Métricas e Indicadores | Apresentado o modelo <i>Innovation Union Scoreboard</i> , sendo utilizado para mensurar a qualidade de um ecossistema de inovação. | Foco na avaliação do desenvolvimento, crescimento e sustentabilidade do ecossistema, é também salientado a possibilidade da utilização do Framework de saúde (pg 41) para efetuar um monitoramento do ecossistema. |
| Cultura | Um elemento não tecnológico importante a fim de instigar os atores a inovar. | - |

Tabela 5: Comparativo Ecosystemas (Autor, 2016)

| Elementos | Ecosistema de Inovação | Ecosistema de Software |
|--------------|---|--|
| Reutilização | É aprestando na literatura com o termo de garagem tecnológica, que objetiva a reutilização de tecnologias ou produtos já desenvolvidos a fim de poupar recursos para serem aplicados à novas pesquisas. | Foco em manter a documentação, os códigos, componentes e bibliotecas de software legível e compreensível para que possa ser reutilizada e replicada. |

O elemento **arquitetura** é comum em ambos os ecossistemas, destacando-se como crucial para o seu funcionamento, e está relacionado com toda a organização interna e tecnológica do ecossistema. Deste modo, a arquitetura abrange toda a estrutura organizacional do ecossistema, como exemplo, quem participa do ecossistema, o papel dos indivíduos, quais são os elementos de base para o seu funcionamento entre outros.

O elemento **ator** também é comum em ambos os ecossistemas sendo observado a presença de vários indivíduos, como exemplo, as instituições, governo, indústrias entre outros. Utilizou-se o termo ator por ser o que mais aparece na literatura.

O elemento **inovação** também está presente em ambos, sendo o elemento em destaque, pois mesmo que os ecossistemas possuam objetivos distintos, também possuem o foco na inovação, em busca pela criação de novos produtos, serviços e/ou crescimento do ecossistema.

O elemento **plataforma** está presente em ambos ecossistemas, sendo importante observar que é a base tecnológica em que os atores interagem a fim de atingir um objetivo mutuo. Outro ponto que é destacado é a importância da plataforma com a possibilidade do compartilhamento e a interação de atores externos ao ecossistema. Esta interação de diferentes atores possibilita uma integração dos atores proporcionando diferentes perspectivas de P&D e consequentemente o crescimento do ecossistema.

O elemento **gestão** abrange o gerenciamento do ecossistema, sendo responsável por mantê-lo fortalecido e produzindo valor. A gestão se encontra evidente na arquitetura do ecossistema de software quando

apresentada suas estruturas, como exemplo, a estrutura de negócios, onde é classificada em três tipos de negócios (Proprietária, *Open Source* e híbrida). Cada tipo de negócios gera um valor diferente para o ecossistema. No ecossistema de software o elemento gestão também aparece com a terminologia de orquestração ou como em um papel de orquestrador, com foco na gestão do ecossistema de software.

O elemento **métricas e indicadores** dão o suporte para a avaliação e monitoramento do ecossistema, possibilitando que os gestores guiem o ecossistema da melhor forma possível a fim de mantê-lo funcional. Existem alguns meios para que possa ser realizada esta avaliação e monitoramento, nos ecossistemas de inovação existe o *Innovation Union Scoreboard* e o *framework* de saúde para os ecossistemas de software.

O elemento **reutilização** objetiva poupar recursos e tempo no ecossistema, ou seja, reutilizar recursos e tecnologias já desenvolvidos a fim de poupar os recursos para que possam ser utilizados em novas tecnologias e investidos em novas pesquisas. O foco da reutilização está também em manter toda a documentação, todos os códigos e componentes criados no ecossistema legíveis e compreensíveis possibilitando, assim, o seu reaproveitamento.

3.2 MODELAGEM DO ECOSISTEMA DE INOVAÇÃO EM TECNOLOGIAS SEMÂNTICAS

A arquitetura proposta tem como propósito apresentar um ecossistema de inovação baseado em tecnologias semânticas em uma perspectiva macro, sendo inteiramente baseado na literatura. Será apresentado um detalhamento da camada de publicação de dados, a fim de explicar o processo de publicação sendo que o mesmo foi baseado integralmente no ciclo de vida aplicado em uma organização semi-pública na Holanda, apresentado na Figura 18 (pg 56).

Para a estruturação desta arquitetura utilizou-se das três estruturas básicas provenientes da arquitetura do ecossistema de software (Estrutura organizacional, Estrutura de negócios e Estrutura de software). Utilizou-se como base a arquitetura do ecossistema de software por dar o suporte tecnológico ao ecossistema de inovação baseado em tecnologia semântica.

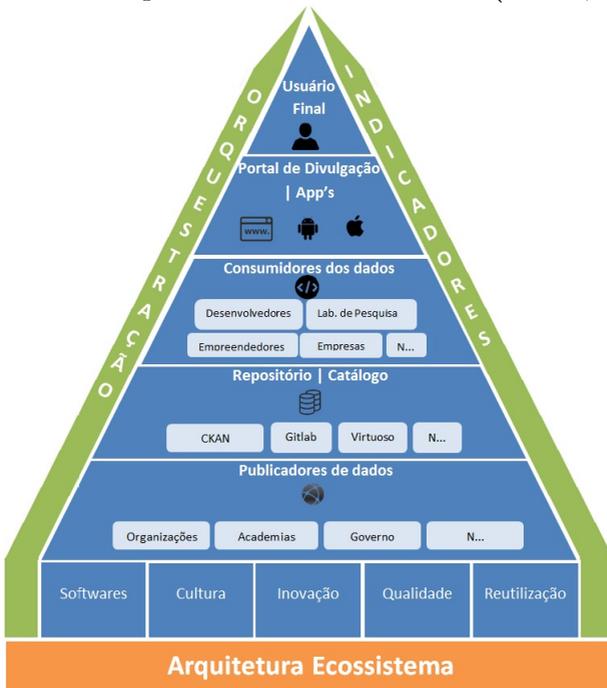
Referente à estrutura organizacional recomenda-se ao ecossistema o uso do tipo federativo, a fim de possuir um integrante de cada organização pertencente ao ecossistema. Esta prática possibilita melhor interação entre organizações a fim de melhorar a orquestração do

ecossistema.

Referente a estrutura de negócios recomenda-se o uso do tipo híbrido, onde partes dos produtos desenvolvidos são classificadas como *Open Source*, ou seja, todos os softwares desenvolvidos devem deixar o código, componentes, bibliotecas e a documentação disponível para que possibilite a reutilização. Contudo é possível que alguns softwares ou serviços disponibilizados pelo ecossistema possam ser cobrados, através de treinamentos ou demandas por projetos específicos. Esta prática promove valores do tipo monetário ou valores não-monetários.

Referente a estrutura de software é recomendado o uso de plataformas como um meio tecnológico comum a todos os participantes do ecossistema, possibilitando o cooperativismo, a integração e facilitando a comunicação dos indivíduos e das organizações. A Figura 19 (pg 70) apresenta a arquitetura proposta.

Figura 19 – Arquitetura do Ecossistema- (Autor, 2016)



A arquitetura apresentada é composta por uma base constituída

por cinco pilares (Softwares, Cultura, Inovação, Qualidade e Reutilização) que servirão como suporte a toda a arquitetura. Os softwares são conjuntos de programas com intuito de auxiliar nas tarefas existentes no ecossistema. Os softwares abrangem desde softwares tradicionais que auxiliam nas tarefas diárias dos usuários como o pacote *Microsoft Office* (Word, Excel...), visualizadores de PDF, navegadores WEB e abrangendo até softwares mais complexos e semânticos como o OntoWiki, D2R entre outros. A Cultura se refere a todo aquele complexo que inclui costumes, hábitos, crenças e atitudes adquiridas pelo ser humano, não somente a nível organizacional como também a nível familiar e social. A inovação é colocada como um ato ou ação de inovar ou criar algo novo, podendo mudar velhos hábitos ou processos. A qualidade prima pelo grau de excelência para os produtos e serviços ofertados e desenvolvidos no ecossistema. A reutilização refere-se a reutilizar os recursos quando possível a fim de poupar despesas e tempo para que se possa investir em novos projetos.

Estes cinco pilares se relacionam a fim de dar suporte a toda a arquitetura. Os softwares podem dar suporte à cultura, por exemplo, no auxílio com ambientes virtuais proporcionando maior interação social entre os participantes e as organizações do ecossistema sem se preocupar com as limitações geográficas ou podendo auxiliar propiciando ambientes para o compartilhamento e discussão de ideias. Na inovação os softwares podem contribuir com ambientes virtuais, por exemplo, fóruns, *web pages* e *wiki's* onde os participantes do ecossistema ou interessados externos podem acessar os projetos e/ou aplicativos, a fim de conhecer e utilizar os conhecimentos, produtos ou serviços oferecidos no ecossistema. Deste modo possibilita aos indivíduos o incentivo à contribuir com novas ideias ou até mesmo em os estimular fazendo parte do ecossistema. Na qualidade os softwares podem auxiliar no monitoramento dos produtos e serviços realizados e ofertados no ecossistema, mantendo assim um nível de excelência e padronização. Na reutilização os softwares dão apoio através de repositórios onde todos os recursos, programas, aplicativos, códigos, componentes, bibliotecas e a documentação criada no ecossistema sejam armazenados e possibilite a reutilização sempre que possível.

A cultura também esta relacionada com os demais pilares, apoiando a inovação tendo em vista o incentivo e proporcionando que os participantes e as organizações criem o habito, espaços ou momentos onde haja o compartilhamento das ideias, possibilitando a criação de novos produtos, serviços, projetos, entre outros ou o aprimoramento dos mesmos proporcionando a inovação no ecossistema e a geração de

novos valores monetários e não-monetários. Na qualidade a cultura ajuda no incentivo dos participantes a manter o grau de excelência nos produtos desenvolvidos ou serviços ofertados. Na reutilização a cultura auxilia no incentivo aos participantes se habituarem a reutilizar sempre que possível, poupando tempo e recursos.

A qualidade esta relacionada com os outros pilares, dando o apoio à reutilização tendo em vista que todos os projetos, aplicativos desenvolvidos no ecossistema devem ser documentados com clareza para que possa ser reutilizado ou melhorado quando necessário.

A reutilização esta relacionada com todos os outros pilares como foi apresentado anteriormente, além de objetivar poupar recursos e tempo, a reutilização é importante, pois toda a documentação criada no ecossistema não fica vinculada somente aos participantes, assim o ecossistema não corre o risco de um membro sair do ecossistema e o conhecimento por ele desenvolvido ser totalmente perdido.

Com a apresentação destes cinco pilares que suportam toda a arquitetura, são apresentadas as próximas camadas do ecossistema. Na primeira camada da arquitetura são apresentados os publicadores de dados, fazendo parte deste conjunto às organizações que participam do ecossistema. Nesta arquitetura são apresentados três conjuntos organizacionais (Academia, Governo e Indústria). Salienta-se que nesta camada pode haver outras organizações. Nesta camada as organizações poderão fazer uso do ciclo de vida apresentado por Broek, Veenstra e Folmer (2014) para a publicação dos seus dados, dados que serão consumidos posteriormente.

Para a publicação dos dados cada organização terá cinco stakeholders (Alta Gerência, Gestor da Informação, Gerente Jurídico, Gerente da Comunidade e o Proprietário dos Dados) responsáveis por participar de cada uma das etapas de publicação, sendo divididas em cinco etapas principais (Identificação, Preparação, Publicação, Reuso e Avaliação), como são apresentadas a seguir.

Identificação

Os envolvidos nesta primeira etapa têm o desafio de identificar quais dados poderão ser publicados criando assim uma estratégia de publicação, tendo como foco também a promoção de novas atividades econômicas e visando novos consumidores para estes dados. Para que seja realizada a construção desta estratégia é necessário o envolvimento de alguns stakeholders, sendo apontado a Alta Gerência para identificação de quais dados poderão ser publicados e assegurar que os dados

disponibilizados não prejudicarão as organizações e o ecossistema. O Gerente de Informação verifica quais dados poderão ou não ser de fato utilizados para a publicação, o Gestor da Comunidade analisa quais dados serão úteis para serem consumidos pela comunidade após sua publicação. Desta forma estes três stakeholders poderão criar uma estratégia de publicação de dados adequada a missão organizacional.

Após a estratégia criada, o Gestor da Informação e o Proprietário dos Dados devem selecionar quais dados serão de fato publicados, conforme a estratégia definida anteriormente. Para esta seleção pode-se utilizar de alguns softwares que auxiliam neste processo, podendo ser utilizados DBpedia extraction, Apache Stabot e Dbpedia Spotlight.

Recomenda-se para que um dataset seja publicado, deve-se observar a propriedade dos dados, ou seja, se os dados pertencem à organização ou se foi adquirida uma autorização para que os mesmos possam ser publicados. Também deve-se tomar o cuidado de não publicar dados que: contêm informações confidenciais, ou seja, que possam vir a prejudicar a organização ou o ecossistema; e que os dados não contêm informações que possam estar ligadas a indivíduos.

Preparação

Nesta segunda etapa é realizada a preparação dos dados para publicação. Se faz necessário à presença do Consultor Jurídico, pois ele é o responsável por verificar se os dados podem ser publicados sem acarretar nenhum problema legal para a organização ou ao ecossistema. Também se faz necessário nesta etapa a presença do Gestor da Informação, o responsável pela parte técnica e o Proprietário dos Dados, e responsável por optar por dados menos complexos, pois isso possibilitará um maior consumo, mas salienta-se que esta etapa é o reflexo da estratégia já definida anteriormente pelos stakeholders. Esta etapa objetiva definir os requisitos (I) técnicos para que possa ser efetuada a publicação dos dados, sendo observados a qualidade, padrões e os metadados, (II) as exigências econômicas como a preposição de valores e modelo de negócios e (III) os requisitos legais, tais como as autorizações para publicação.

No que refere-se a preparação dos dados a qualidade é analisada em consideração aos níveis de publicação determinado por Berners-Lee (2006), apresentado na sessão 2.4.1, porém a qualidade é dependente da estratégia adotada. Deste modo é recomendado a análise dos datasets e a verificação de sua classificação atual para publicação. Após isso é necessário determinar o nível desejado de publicação a fim de construir

um plano de publicação ate atingir o nível desejado, que dever ser o nível 4 ou 5 apresentado na Figura 15 (pg 52).

Após definido os requisitos para a publicação dos dados, deve ser realizada a preparação técnica dos dados, sendo observado o anonimato de qualquer informação, também deve ser efetuado a modelagem dos conceitos com base nos três requisitos levantados pelos stakeholders e as ligações dos dados.

Tecnologias que podem auxiliar no processo de preparação e modelagem dos dados, são: R2R, D2R, Protégé, Virtuoso, RDFaCE text Annotation, Neon Toolkit, Sieve, ORE, Silk, LIMES, Ultrawrap, Top-Braid Composer e Vilant.

Publicação

Nesta terceira etapa ocorre a publicação dos dados após seu tratamento, neste momento envolve-se somente o Gestor da Informação onde o mesmo tem a responsabilidade de garantir sua usabilidade considerando o registro dos dados em catálogo de dados já existente como o *datahub.io*¹. Destaca-se a possibilidade do uso de um catálogo de dados próprio utilizando a plataforma CKAN, mas para isso deve haver um conhecimento prévio e publicitário do catálogo em relação a comunidade que irá consumir estes dados. Outro ponto importante é a presença do Gestor da Comunidade onde ele será o responsável por incentivar o consumo dos dados publicados, organizar eventos como *hackttons*, notas de imprensa para a comunidade, destaque em blogs ou até mesmo a organização de uma competição com prêmios para incentivar o consumo dos dados. Dessa forma, com o passar do tempo acredita-se que pode ser criado o habito cultural de utilizar os dados e tornar difundido o uso e consumo de LOD no ecossistema.

Existem tecnologias que podem ser utilizadas para o consumo dos dados, mesmo que as organizações que participam do ecossistema não tenham o foco da publicação externa dos dados. As organizações pode utilizar do LD para obter uma perspectiva diferenciada de análise dos dados, com o uso de ferramentas, como por exemplo, o Cubviz ou uso de navegadores semânticos como o Tabulator, mas caso a organização tenha uma equipe de desenvolvedores pode utilizar do LD ou LOD para criar uma nova aplicação própria que atenda suas necessidades.

¹Um catálogo de dados onde indivíduos podem publicar seus dados, podendo criar grupos ou comunidades. <https://datahub.io/>

Reuso

Nesta etapa aborda-se a preocupação com o consumo dos dados publicados, é importante salientar que comunidades de dados abertos são importantes nesta etapa. Deste modo, o Gestor da Comunidade juntamente com todos os demais *stakeholders* podem manter a comunidade ativa no consumo de dados, podendo convidar para participar da comunidade os empresários, universidades, incubadoras. Esta integração com a comunidade possibilita *feedbacks* sobre os dados publicados melhorando a qualidade dos próximos dados e conseqüentemente aumentando o consumo. Outro ponto a ser considerado é a construção de descritivos, ou seja, pautas a fim de manter organizadas todas as melhorias e sugestões dadas pela comunidade e mantendo as melhorias alinhado com a estratégia definida anteriormente pela Alta Gerência na primeira etapa do ciclo de vida.

Outro ponto importante nesta etapa é a preocupação com a gestão dos dados, pois o Gestor da Informação e o Proprietário dos Dados tem um papel ativo na pós-publicação visando manter a qualidade dos dados publicados. Para que haja a qualidade é necessário manter os dados atualizados e garantir sua previsibilidade e usabilidade, também é importante obter sempre os *feedback's* dos usuários sobre os dados publicados e o monitoramento dos consumidores.

Algumas tecnologias podem ser utilizadas, como exemplo, blogs, web pages com espaços para divulgação dos dados publicados e para possibilitar aos usuários darem *feedback's* ou podendo também ser usado o OntoWiki visando uma abordagem semântica.

Avaliação

Na última etapa visa à avaliação do processo de publicação dos dados. Desta forma é necessário o envolvimento de todos os *stakeholders* a fim de avaliar as publicações, utilizando-se de indicadores como: o número de downloads, a combinação com outros *datasets*, verificar as aplicações desenvolvidas a partir dos dados publicados, analisar a popularidade dos aplicativos em relação aos usuários finais, entre outros. Por fim, a Alta Gerencia deve analisar toda a informação coletada a fim de melhorar sua estratégia possibilitando uma melhora na publicação e consumo de dados, assim como verificar se houve alguma mudança cultural em relação à publicação de dados no ecossistema.

Próximas camadas da arquitetura

Na camada **repositório | catálogo**, ficarão armazenados todos os documentos, códigos, software e bases de dados, sendo transversal a todo o ecossistema. O uso desta camada viabiliza a possibilidade de oferecer a estrutura necessária para a publicação e consumo de dados no ecossistema. Para os publicadores ou consumidores de dados que não possuem a estrutura própria e necessitem fazer uso desta camada ou uso de outras tecnologias de armazenamento é necessário verificar junto ao orquestrador do ecossistema as condições para uso do mesmo.

Recomenda-se o uso de plataformas web *Gitlab* que disponibiliza espaços para o controle de versão, compartilhamento de ideias, fóruns e outros. Esta plataforma proporcionam aos consumidores de dados um ambiente que apoia o desenvolvimento de aplicativos e de software, podendo também ser disponibilizado todos os documentos, componentes e bibliotecas de software desenvolvidos. Esta plataforma também possibilita que os desenvolvedores recebam *feedback's* de outros desenvolvedores assim como o compartilhamento dos software para que haja contribuições e melhorias em seus projetos, possibilitando criar versões alternativas e deixando o software mais refinado e usual. Para o armazenamento dos dados utilizando por software é recomendado o uso do banco de dados Virtuoso, visto que é um banco de dados híbrido e atende tanto a aplicações convencionais como semânticas.

O uso do *Comprehensive Kerbal Archive Network* (CKAN) é recomendado pois é uma plataforma de catalogação de dados, onde todos os dados publicados podem ser organizados, possibilitando o monitoramento com intuito de analisar a catalogação dos dados, por exemplo, quem mais cataloga, a frequência de catalogação, qual o tipo de licença mais utilizada entre outros. No processo de catalogação pode ser realizados por indivíduos, podendo também ser criados grupos ou organizações que pretendam publicar os dados e queiram manter uma organização e controle dos dados publicados. Outro ponto que pode ser explorado desta plataforma é o uso da opção *harvest* onde o CKAN irá procurar e coletar metadados de outros servidores de *Catalog Service for the Web* (CSW), de *Web Accessible Folder* (WAF) ou em documentos de metadados acessíveis na web, deste modo o CKAN se transforma em uma especie de "HUB" podendo centralizar vários metadados de várias fontes.

A camada **consumidores de dados**, refere-se a todos os interessados que irão consumir os dados, como por exemplo, os programadores que podem desenvolver software oriundos dos dados publicados

do ecossistema. Os consumidores de dados podem ser laboratórios de pesquisa, desenvolvedores, empreendedores, empresas e/ou outros interessados que participam ou pretendem participar do ecossistema.

Todos os software desenvolvidos pelos consumidores de dados devem possuir uma qualidade determinada pelo ecossistema a fim de possibilitar o uso, a organização dos dados e mantendo o ecossistema seguro de qualquer mal intencionado. Alguns critérios presentes no ecossistema de software que podem ser aplicados, como exemplo, a qualidade da documentação de software desenvolvidos pelos consumidores, deixando a documentação clara e objetiva possibilitando sua reutilização.

Outro ponto é a preferência por uso da política *open source* de software quando possível, ou seja, o código dos software desenvolvidos além de documentados devem ser abertos para que outros indivíduos possam reutiliza-los a fim de melhorar ou aprimorar o software. Em casos que softwares são derivados de projetos proprietários, ou seja, softwares proprietários ou patrocinados por organizações, pode-se manter o seu código fechado sendo disponibilizado apenas a documentação de uso para ao público, sendo que sua documentação detalhada abordando o seu desenvolvimento seja publicado em um espaço restrito com acesso somente a pessoas autorizadas de modo que não prejudique o ecossistema e as organizações envolvidas. Todos os softwares desenvolvidos devem ser armazenados com toda a documentação e o código nesta camada, sendo observados a política de acesso adotada para cada projeto e software.

A camada **portal de comunicação** | App's diz respeito a todos os softwares desenvolvidos, projetos, editais disponíveis e novidades sobre o ecossistema, proporcionando a divulgação do ecossistema, compartilhando do conhecimento e instigando a participação de novas organizações e consumidores ao ecossistema.

Toda a informação disponibilizada e os software desenvolvidos, são importantes a fim de obter *feedback's* para que haja um amadurecimento do ecossistema com o passar do tempo, assim como a divulgação de produtos e serviços gerados. Este portal contará com espaços onde indivíduos poderão conhecer os projetos do ecossistema, os softwares e aplicativos e conhecer os parceiros que atuam no ecossistema.

Os indivíduos também poderão contribuir com ideias no portal. As novas organizações ou consumidores de dados que queiram fazer parte do ecossistema podem entrar em contato com o ecossistema através de um canal de comunicação presente no portal. O portal terá um espaço onde serão divulgados os App's desenvolvidos na plataforma

Android, sendo disponibilizados do Google Play para download, como para plataforma Apple disponibilizados na Apple Store para download. Alguns dos aplicativos terão versões *free*, ou seja, gratuitas com funções limitadas e versões mais completas que será cobrado um valor para acesso do mesmo.

Na camada **usuário final** abrange os indivíduos que irão utilizar das informações, do conhecimento ou dos produtos e serviços oferecidos pelo ecossistema, como exemplos, software, consultorias, cursos, entre outros. Os usuários também podem participar indiretamente do ecossistema acessando os conteúdos informativos, contribuindo com suas ideias deixando seu *feedback*, participando de *hacktons* e demais eventos promovidos pelo ecossistema.

Há possibilidade também dos indivíduos que se interessem em participar do ecossistema tenham esta oportunidade através de grupos de pesquisa, oportunidade de emprego nas organizações ou até mesmo em cursos nas universidades que participam do ecossistema.

Os **indicadores** são o meio que a orquestração irá utilizar na avaliação e análise do ecossistema, como exemplo, quais informações os usuários mais acessam, quais são os softwares mais baixados, qual a área de pesquisa que possui mais interatividade com os usuários, quais são os *feedback's* apresentados em relação aos serviços e produtos ofertados pelo ecossistema, quem são os novos interessados a participar do ecossistema (Organizações, desenvolvedores, laboratórios de pesquisa), o quão inovador o ecossistema e as organizações estão, a quantidade de dados publicados e consumidos, entre outros. Todos estes indicadores auxiliam a orquestração a gerenciar todo o ecossistema e mantê-lo mais atrativo e vivo.

A **orquestração** é uma expressão originada do ecossistema de software, onde o orquestrador realiza a gestão do ecossistema. Deste modo foi adotado o termo orquestração para este ecossistema e foi utilizado do modelo federação (pg 38) pertencente à estrutura organizacional do ecossistema de software, ou seja, a orquestração será realizada com um conjunto de atores representativos que são responsáveis por guiar o ecossistema. Para fazer parte desse comitê de orquestração cada organização irá eleger um responsável para participar deste grupo, mantendo alinhados os interesses das organizações e do ecossistema. O comitê orquestrador tem o papel de gerenciar, regulamentar, monitorar e avaliar todo o ecossistema, mantendo a ordem e harmonia conforme os interesses das organizações que fazem parta do ecossistema.

Considerações finais sobre o capítulo

Neste capítulo de análise e modelagem observa-se que os elementos agrupados na revisão de literatura possuem semelhanças e particularidades entre os ecossistemas de inovação e de software. Deste modo, foi possível a realização da proposição da arquitetura para o ecossistema de inovação baseados em tecnologias semânticas. A arquitetura do ecossistema de inovação baseados em tecnologias semânticas proposto utilizou-se de todos os agrupamentos apresentados na sessão 3.1, a fim de deixar o ecossistema mais robusto unindo as particularidades do ecossistema de inovação e de software. Na sessão seguinte é realizado a verificação da arquitetura proposta por meio de entrevistas semi-estruturadas com três indivíduos: que representam a organização, o desenvolvedor e a uma empresa.

4 VERIFICAÇÃO

Na verificação deste estudo utilizou-se de uma entrevista semi-estruturada. Este método consiste na combinação de questões abertas e fechadas dando maior flexibilidade ao pesquisador permitindo-o explorar mais profundamente determinados assuntos, assim é possível delimitar um número de informações e direcionando-o para o tema possibilitando atingir o objetivo proposto na pesquisa (BONI; QUARESMA, 2005).

A entrevista foi realizada com quatro indivíduos, sendo eles, o especialista com uma visão geral, desenvolvedor de uma empresa ativa no mercado tecnológico com uma visão de consumidores e uma organização com uma visão de publicadora de dados. Esta entrevista utilizou os objetivos específicos para verificação juntamente com os quatro indivíduos entrevistados. Deste modo foi utilizada de três perguntas chaves baseadas nos objetivos específicos para guiar a entrevista, sendo **(a)** O que você percebe sobre a integração entre os ecossistemas de inovação e ecossistema de software?, **(b)** O que você percebe sobre os ciclos de vida e as tecnologias semânticas que auxiliam na publicação dos dados?, **(c)** Qual sua percepção da arquitetura proposta por este trabalho? A seguir serão apresentados os resumos das entrevistas, sendo que cada tabela irá abordar um objetivo.

Tabela 6: Verificação Objetivo Específico (1)

| Verificação do Ecossistema em Tecnologias Semânticas |
|---|
| Objetivo Específico (1): Apresentar os principais elementos do ecossistema de software e ecossistema de inovação. |
| Resumo do Objetivo: Este objetivo tem como proposito identificar os principais elementos que compõem os ecossistemas de inovação e software através de características e/ou fatores de sucesso do ecossistema com base na literatura. Desta forma possibilita realizar um agrupamento que será à base da arquitetura do ecossistema de inovação em tecnologias semânticas. |

Tabela 6: Verificação Objetivo Específico (1)

| Verificação do Ecossistema em Tecnologias Semânticas | |
|--|---|
| Objetivo Específico (1): Apresentar os principais elementos do ecossistema de software e ecossistema de inovação. | |
| Especialista | Cada ecossistema tem sua particularidade mas existe similaridades entre eles, os dois ecossistemas atedem a comunidades próprias e como tal o ecossistema deve ter mecanismos que façam que o ecossistema seja preservado. Ambos os ecossistemas são dinâmicos, sendo assim esse comparativo é valido para servir de base para sustentar essa arquitetura que está sendo proposta |
| Organização | Na visão do entrevistado o ecossistema de software é o meio para um fim e o ecossistema de inovação é este fim. A inovação tem que ter aplicação e os software proporcionam isso. É observado um alinhamento entre os elementos dos ecossistemas e foi um bom levantamento pois serviu como base para o estudo e a arquitetura |
| Desenvolvedor | Acha interessante a abordagem, mas salienta que em um ecossistema os indivíduos não visam somente o lucro "não querem somente ganhar dinheiro" mas querem "uma motivação maior" para fazer parte do ecossistema como "reconhecimento ou querem fazer diferente". |
| Empresa | Não existe dicotomia entre os ecossistemas. O software esta intrínseco em qualquer situação. Mesmo em um ecossistema de inovação o software esta presente em auxiliar na criação de um projeto ou até mesmo em uma simples plataforma de envio de <i>e-mail</i> ou de venda de produtos. |

Na tabela 6 é possível observar uma visão geral da entrevista realizada com os quatro indivíduos. Referente ao primeiro objetivo o o especialista deixa claro que este estudo realizado na literatura ajudou a dar o arcabouço necessário para embasar a proposta da arquitetura, sendo que ambos os ecossistemas são "dinâmicos" e tem elementos que

são semelhantes em ambos.

A organização concorda com o especialista acrescentando que existe um alinhamento entre os ecossistemas, sendo clara a presença de elementos em comum, e completa dizendo: "o ecossistema de software é o meio para o fim, sendo que este fim é o ecossistema de inovação". A empresa já discorda da organização, pois em sua perspectiva "o software e a inovação não existe uma dicotomia entre eles" deixando claro que "o software está intrínseco em tudo atualmente".

Na perspectiva do desenvolvedor o estudo foi interessante, pois deu uma base para entender como é a estrutura dos ecossistemas, mas o grande problema de qualquer ecossistema são as pessoas, abordando muito o fator cultural que impacta em qualquer situação, dizendo ainda "se não há motivação ou interesse por parte dos envolvidos, as coisas não funcionam". Deste modo, o entrevistado explica à partir da experiência vivida que em um ecossistema "as pessoas querem fazer a diferença.

Tabela 7: Verificação Objetivo Específico (2)

| Verificação do Ecossistema em Tecnologias Semânticas | |
|---|--|
| Objetivo Específico (2): Abordar as tecnologias semânticas de apoio para a publicação de dados em empresas, academia e governo. | |
| Resumo do Objetivo: Este objetivo tem como propósito identificar na literatura as tecnologias e métodos que auxiliem na publicação de dados em Linked Data e Linked Open Data. | |
| Especialista | Não resta dúvida que, um ecossistema de inovação, com o advento de dados abertos vai favorecer enormemente o uso dos dados com o propósito de um ecossistema para fazer inovação é o uso de métodos e tecnologias semânticas. |
| Organização | Observa-se a necessidade das tecnologias semânticas e de métodos como o ciclo de vida para guiar no uso dessas tecnologias. Salienta-se que as tecnologias atuais (Abertas como "Google" ou privadas onde há aquisição do software) "não dão conta das demandas" e "acredita-se que as tecnologias semânticas podem atender estas necessidades". |
| Desenvolvedor | Não tem familiaridade com as tecnologias semânticas, mas pelo que observa as tecnologias e os ciclos de vida que auxiliam na publicação de dados semânticos apresentados são "as que todo mundo usa" e "devem ser as melhores para essa finalidade". |
| Empresa | Segundo sua experiência, as tecnologias semânticas devem amadurecer mais, pois não geram valor para o mercado ainda. |

Na segunda tabela 7 é apresentada uma visão geral sobre abordagem referente ao segundo objetivo. O especialista deixa claro que as as tecnologias semânticas podem e vão "fortemente beneficiar o uso de dados no ecossistema", e concorda que utilizar de ciclos de vidas e tecnologias semânticas é o caminho a ser percorrido, e completa: "as tecnologias semânticas podem ajudar na inovação". A organização também compartilha dessa expectativa, deixando evidente que "as tecnologias atuais, muitas vezes não dão conta das demandas", e fecha

dizendo que "acredita que as tecnologias semânticas podem atender as demandas que estão aparecendo".

O desenvolvedor deixa claro que "já viu alguma coisa sobre as tecnologias semânticas" e não tem muito conhecimento sobre as mesmas, mas pelo que analisou diz "as tecnologias que você me apresentou são as que todos usam, pelo que andei pesquisando, e por isso acredito que são as mais adequadas para o uso".

O entrevistado da empresa tecnológica deixa claro com base na sua experiência com tecnologias semânticas dizendo "eu trabalhei com essas tecnologias, gostei e tem potencial, mas o mercado não tem absorção e essas tecnologias não vendem", ele comenta que já viu essas tecnologias sendo usadas em grandes empresas, mas este uso é interno. O entrevistado fecha dizendo que "as tecnologias semânticas são um meio que precisa de mais maturidade e a acadêmica precisa mostrar para o mercado, por exemplo, que tipo de valor (monetário) pode ser gerado com uso dessas tecnologias? Entendo que a academia tem uma visão de longo prazo e por isso digo que está no caminho".

Tabela 8: Verificação Objetivo Específico (3)

| Verificação do Ecossistema em Tecnologias Semânticas | |
|---|--|
| Objetivo Específico (3): Propor um conjunto de recomendações e uma arquitetura para um ecossistema de inovação baseado em tecnologias semânticas para dados abertos | |
| Resumo do Objetivo: Com base nos elementos identificados na literatura e as tecnologias que apoiam a publicação de dados é proposto uma arquitetura a fim de estruturar todo o ecossistema e as recomendações para a publicação e o consumo dos dados. | |
| Especialista | Vista como promissora, como toda a arquitetura tem uma versão onde essa é a versão 1.0 e esta no caminho pois o estudo elencou os elementos que são necessário para a arquitetura no processo de inovação. |
| Organização | A arquitetura proposta é "muito próxima a realidade vivida", visto que na prática o modelo "pode funcionar". A cultura é um ponto onde deve-se trabalhar pois envolve pessoas e pode "dificultar as coisas". |

| | |
|----------------------|--|
| Desenvolvedor | A arquitetura é "interessante" pois apresenta uma visão geral do ecossistema e sua organização. Coloca também que o maior desafio será a questão "cultural", onde os indivíduos devem ser "educados ou terem o hábito" de publicar e consumir os dados gerados pelo ecossistema. |
| Empresa | A ideia da arquitetura é bem interessante, mas é colocado algumas pequenas mudanças, principalmente no quesito orquestração. |

Na terceira tabela 8, pode-se observar que a arquitetura é bem vista e destacada como "promissora" coloca o Especialista. Na visão do Desenvolvedor é uma arquitetura interessante e apresenta uma "perspectiva geral", mas que deve-se investir uma atenção especial na cultura, sendo ela um "desafio" para o funcionamento do ecossistema, pois os dados publicados pelo ecossistema tem que ser atrativos e possibilitam agregar valor para que possa ser consumido. Por fim a organização coloca que esta arquitetura proposta vem de encontro com a "nossa vivência", salientando a possibilidade de colocar em prática, mas deve-se ter uma atenção no aspecto cultura pois envolve pessoas e hábitos que podem dificultar o funcionamento do ecossistema.

A empresa coloca que esta arquitetura de modo geral é boa, mas peca em dois quesitos. O primeiro quesito é a falta de um sexto pilar denominado de *funding* ou dinheiro, "abordando quem paga a conta, que sem isso o ecossistema não vai funcionar". Outro ponto colocado pelo entrevistado é que em um ecossistema o uso de orquestração ou governança não funciona, pois o ecossistema utiliza da teoria de sistemas abertos, e diz: "essa ideia de orquestração e governança só funciona em organizações ou governo, em um ecossistema não funciona, pois ninguém quer receber ordens". O entrevistado ainda complementa que o ecossistema deve usar como base o modelo da W3C, onde você usa de recomendações ", os indivíduos seguem suas recomendações se querem, não são obrigados a fazer algo por determinação de alguém", desta forma ele aconselha mudar o nome para conselheiros ou recomendadores.

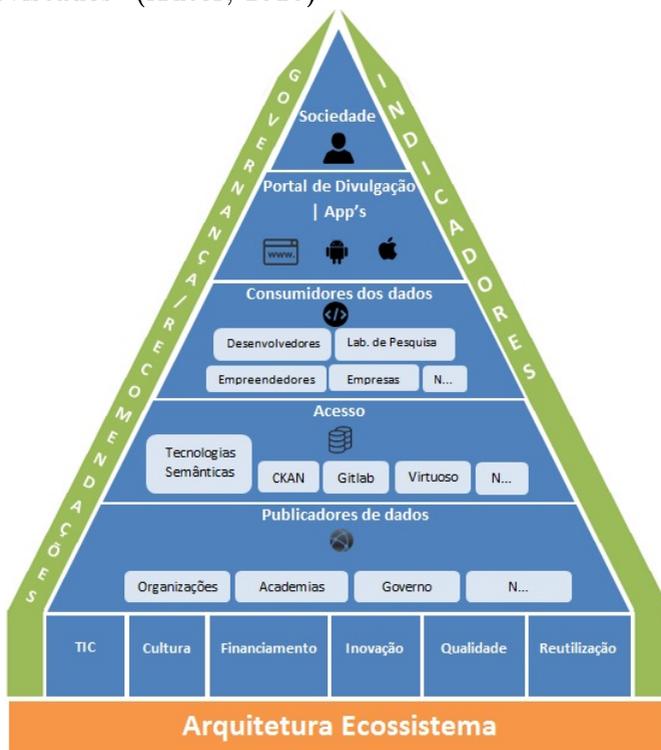
4.1 RESULTADO VERIFICAÇÃO

Com a aplicação da entrevista semi-estruturada foi possível identificar junto aos entrevistados o interesse da temática. A organização e o especialista tem visões e expectativas semelhantes, por outro lado o desenvolvedor e a empresa possuem uma visão diferente do cenário apresentado, deixando evidente a importância da aproximação desses indivíduos em um ecossistema. Este estudo apresenta uma proposta da primeira versão de um ecossistema em tecnologias semânticas e observa-se também uma gama de possibilidades de estudos futuros a fim de dar continuidade e um refinamento desta arquitetura.

Foram identificados através das entrevistas novas perspectivas de melhoras, sendo abordados alguns pontos que merecem uma atenção nesta pesquisa, destacando a cultura, pois ela impacta diretamente no ecossistema sendo vista como o um "desafio", pois pessoas podem atrapalhar o processo de publicação e consumo dos dados. outro ponto destacado pelo especialista é que em sua percepção esta arquitetura proporciona um grande "HUB" onde coleta os dados das organizações (academia, governo e empresa) e disponibiliza como um "grande repositório de dados" proporcionando um grande volume de insumo para um ecossistema de inovação. Outro ponto a ser destacado é a mudança da orquestração por recomendadores ou conselheiros a fim de seguir a teoria de sistema aberto e a adição do pilar *Funding*, como salientado pela empresa.

Como resultado da verificação foi construída a segunda arquitetura baseada na visão dos entrevistados, sendo que foram alterados cinco pontos que segundo a visão deles deixará a arquitetura mais dinâmica e robusta. A seguir é apresentada a Figura 20 e apresentado os cinco pontos de mudança sugeridos pelos entrevistados.

Figura 20 – Arquitetura do Ecosistema proposta pelos entrevistados- (Autor, 2016)



No que se refere os pilares, foi alterado o software para Tecnologia da Informação e Comunicação (TIC), que segundo a visão do especialista esta abordagem é mais ampla abrangendo todas as tecnologias envolvidas no ecossistema, desde hardware ao software e todo o aparato que se refere a comunicação. Sendo assim a TIC abrange todo o conjunto de recursos tecnológicos que se integraram a fim de automatizar e auxiliar na comunicação dos processos de negócio, pesquisas científicas, no ensino e aprendizagem entre outros. Foi também realizado a adição do pilar financiamento, que segundo a visão da empresa entrevistada é um fator chave a ser abordado no ecossistema, sendo que os financiamentos podem ser originados através de subsídios do governo, assim como por investidores anjos, capitais de risco, empresas privadas, entre outros. Deste modo esse pilar tem o objetivo de abor-

dar como será a aquisição dos financiamentos e quem irá pagar por isso.

Outro ponto modificado após as entrevista foi a camada repositório | catálogo, onde o especialista sugere a mudança para acesso, pois esta camada seria responsável por concentrar todo tipo de acesso a informação do ecossistema, ou seja, nesta camada se concentraria os repositórios de dados para o acessos dos consumidores de dados, seria o local onde se concentraria as tecnologias semânticas tendo a possibilidade que trabalhar com a indexação e recuperação de dados semânticos. O entrevistado usa da expressão "um grande *HUB*" onde se concentraria todas os dados e informações das organizações e dos consumidores de dados.

Uma outra modificação foi na camada usuário final, onde foi sugerido pelo especialista a mudança para sociedade, visto que os produtos e serviços oferecidos pelo ecossistema de inovação baseados em tecnologias semânticas tendem atender a sociedade como um todo e não apenas a um único usuário.

Por fim a última mudança sugerida, sendo que foi identificado uma discordância entre dois entrevistados. Na percepção do especialista seria adequada a mudança de orquestração por governança, sendo que governança possui uma abrangência maior focando em premissas, conceitos, valores e práticas relacionado a estrutura organizacional, ao relacionamento entre os indivíduos e nas demais atividades do ecossistema. De modo que anteriormente a orquestração é muito específica e simples em relação ao modelo proposto. No entanto a empresa entrevistada coloca que seria ideal a mudança de orquestração por um comitê coselheiro, onde seria apresentadas as melhores práticas e recomendações a fim de guiar o ecossistema.

De todas as sugestões apresentadas pelos entrevistados, todas foram adicionadas no modelo de arquitetura, sendo que as sugestões dadas pelos entrevistados foram pertinentes pois os mesmos possuem experiência na área.

Referente a discordância entre o especialista e a empresa, foram adotada as duas sugestões, sendo observado a literatura e as perspectivas dos entrevistados. Deste modo optou-se em usar do termo governança por ser mais abrangente como explicitado pelo especialista e a adição do termo recomendação sugerido pela empresa a fim de manter um equilíbrio no ecossistema, onde haja um monitoramento, regulamentação e uma gerência mais macro de modo que não interfira na autonomia dos indivíduos que participam do ecossistema.

5 CONCLUSÃO

Este estudo propôs a modelagem de um ecossistema de inovação baseado em tecnologias semânticas, utilizando de elementos que compõem os ecossistemas de inovação e ecossistema de software. Foram também abordados as tecnologias e ciclos de vida que auxiliam na publicação de dados. Para a identificação dos ciclos de vida e dos elementos embasou-se na literatura para propor a arquitetura e as recomendações a fim de atender o objetivo geral.

Para alcançar o objetivo geral esta dissertação foi segmentada em três outros objetivos específicos que deram o embasamento necessário para a realização da modelagem do ecossistema proposto. Sendo os três objetivos específicos descritos a seguir:

O primeiro objetivo específico foi alcançado através de uma pesquisa sistemática, em três bases de dados, sendo elas, SCOPUS, IEEE e Web Of Science, com intuito de identificar na literatura as características e/ou fatores de sucesso para implementação de um ecossistema possibilitando assim a realização do agrupamento desse resultado a fim de propor elementos que serão utilizados na arquitetura proposta.

Foi possível verificar com este estudo que o termo "Ecossistema" é utilizado em vários contextos, mas sempre seguindo sua essência derivada da biologia. Os ecossistemas pesquisados são semelhantes em vários aspectos, mas possuem pequenas particularidades, como exemplo, a cultura presente no ecossistema de inovação e a qualidade presente no ecossistema de software, ambas essas particularidades definem com mais clareza o objetivo dos ecossistemas. Desta forma foi possível perceber que o uso do termo ecossistema pode fazer uso dos elementos apresentados por esta pesquisa, pois os mesmos são comuns entre os ecossistemas pesquisados, sendo que um ecossistema pode ser específico apenas por suas particularidades.

O segundo objetivo específico foi alcançado através de uma pesquisa narrativa a fim de dar o embasamento literário necessário sobre as tecnologias e ciclo de vida encontrados na literatura. Desta forma foram elicitados e abordados três ciclos de vida e tecnologias existentes na literatura que auxiliam na publicação de dados que podem ser utilizadas pelas três organizações indicadas (empresa, academia e governo).

Por fim o terceiro objetivo foi alcançado através da análise e compilação dos dados e informações coletadas nos objetivos específicos anteriores. Através da análise da literatura identificou-se a relevância de cada um dos elementos e suas respectivas recomendações sendo

proposta a primeira versão da arquitetura do ecossistema de inovação baseado em tecnologias semânticas em uma perspectiva macro e teórica.

Para verificação desta pesquisa foi utilizado de entrevista semi-estruturada com quatro indivíduos a fim de obter uma perspectiva menos teórica e identificar possíveis melhorias. Desta forma foram sugeridas algumas melhorias referente a arquitetura dando origem a uma segunda versão mais dinâmica e robusta. De modo geral todos os objetivos propostos foram atendidos neste trabalho, o comparativo realizado e a arquitetura proposta traz uma contribuição ao meio acadêmico proporcionando uma nova perspectiva de pesquisa e trabalhos futuros.

Espera-se que no futuro possa ser colocado em prática este estudo, a fim de obter resultados reais, comprovando os reais benefícios dos uso das tecnologias semânticas e as inovações que as tecnologias semânticas podem proporcionar a sociedade e possibilitando novas perspectivas de pesquisas a academia. Outro ponto a ser destacado é a importância em que as pesquisas realizadas na academia devem se preocupar com a aplicação no mercado de trabalho gerando *funding*, lembrando o que foi dito pela empresa entrevistada "a academia deve desenvolver dissertações e teses que possam gerar lucro e não apenas mais um trabalho que ao final fica engavetado!".

5.1 ESTUDOS FUTUROS

Como sugestões de trabalhos futuros observa-se uma vasta possibilidade de pesquisa, sendo possível atender as três áreas de concentração do PPEGC. Deste modo a seguir são apresentadas algumas das possibilidades de trabalhos futuros que podem ser derivados desta dissertação.

- Aplicação da arquitetura a fim de obter uma análise e propor uma melhoria através dos resultados coletados.
- Propor um estudo de governança para este ecossistema;(um modelo de governança de ecossistema de inovação baseados em tecnologia semântica para apoiar atividades de publicação e consumo de dados)
- Propor o desenvolvimento de uma plataforma de auxilie a publicação, consumo e governança do ecossistema;
- Propor o desenvolvimento de indicadores de maturidade do ecossistema;

- Propor o desenvolvimento de um ontologia para este ecossistema;
- Propor o aprimoramento de um ciclo de vida que auxilie na publicação de dados no ecossistema;
- Propor um estudo de como pode ser propagada a ideia de consumo de dados e de como a sociedade pode usufruir do ecossistema.
- Realizar uma atualização da arquitetura com foco na sustentabilidade do ecossistema;
- Realizar uma exploração de novas arquiteturas, a fim de propor uma nova arquitetura mais detalhada e/ou completa, envolvendo processos;

Deste modo este estudo prova que há possibilidades de estudos em várias pontas deste ecossistema e como é sua primeira versão, ainda há a possibilidade de se trabalhar com outras vertentes, bem como o aprofundamento dos estudos em cada um dos pilares apontados por estes estudo.

REFERÊNCIAS

AGHAEI, S.; NEMATBAKSHI, M. A.; FARSAANI, H. K. Evolution of the world wide web: From web 1.0 to web 4.0. **International Journal of Web & Semantic Technology**, Academy & Industry Research Collaboration Center (AIRCC), v. 3, n. 1, p. 1, 2012.

AKSW. **CubeViz**. 2016. Disponível em:
<<http://aksw.org/Projects/CubeViz.html>>. Acesso em: 07 Setembro 2016.

AKSW. **GeoKnow Project**. 2016. Disponível em:
<<http://geoknow.eu/Welcome.html>>. Acesso em: 07 Setembro 2016.

AKSW. **LIMES**. 2016. Disponível em:
<<http://aksw.org/Projects/LIMES.html>>. Acesso em: 07 Setembro 2016.

AKSW. **OntoWiki**. 2016. Disponível em:
<<http://aksw.org/Projects/OntoWiki.html>>. Acesso em: 07 Setembro 2016.

AKSW. **RDFaCE**. 2016. Disponível em:
<<http://aksw.org/Projects/RDFaCE.html>>. Acesso em: 07 Setembro 2016.

AKSW. **SemMap**. 2016. Disponível em:
<<http://aksw.org/Projects/SemMap>>. Acesso em: 07 Setembro 2016.

AKSW. **Trypify**. 2016. Disponível em:
<<http://aksw.org/Projects/Tripify.html>>. Acesso em: 07 Setembro 2016.

ALAMI, D.; RODRÍGUEZ, M.; JANSEN, S. Relating health to platform success: Exploring three e-commerce ecosystems. In: **ACM. Proceedings of the 2015 European Conference on Software Architecture Workshops**. [S.l.], 2015. p. 43.

ALVES, R. C. V. et al. Web semântica: uma análise focada no uso de metadados. Universidade Estadual Paulista (UNESP), 2005.

ANTIDOT. **Linked Enterprise Data Principles, uses and benefits**. 2016. Disponível em:

<<https://www.inria.fr/content/download/18171/514403>>. Acesso em: 07 Nov. 2016.

ANTONIOU, G. H. F. v. **A Semantic Web Primer**. Cambridge: MIT Press, 2004. 238p p.

APACHE. **Jena**. 2016. Disponível em: <<https://jena.apache.org/>>. Acesso em: 07 Setembro 2016.

APACHE. **Stanbol**. 2016. Disponível em: <<https://stanbol.apache.org/>>. Acesso em: 07 Setembro 2016.

ARCHIVES, O. **Open Archives - Initiative Object Reuse and Exchange**. 2016. Disponível em:

<<http://www.openarchives.org/ore/1.0/datamodel>>. Acesso em: 07 Setembro 2016.

AUER, S. et al. Managing the life-cycle of linked data with the lod2 stack. In: SPRINGER. **International semantic Web conference**. [S.l.], 2012. p. 1–16.

AUER, S.; LEHMANN, J.; NGOMO, A.-C. N. Introduction to linked data and its lifecycle on the web. In: **Reasoning Web. Semantic Technologies for the Web of Data**. [S.l.]: Springer, 2011. p. 1–75.

AUTIO, E.; THOMAS, L. **Innovation Ecosystems: Implications for Innovation Management**. [S.l.]: Oxford University Press, 2014.

AXELSSON, J.; SKOGLUND, M. Quality assurance in software ecosystems: A systematic literature mapping and research agenda. **Journal of Systems and Software**, Elsevier, v. 114, p. 69–81, 2016.

BANDEIRA, J. M. **Dados abertos conectados**. Tese (Doutorado) — Universidade de São Paulo, 2015.

BENAKOUCHE, T. Tecnologia é sociedade: contra a noção de impacto tecnológico. **Cadernos de Pesquisa**, v. 17, 2005.

BERNERS-LEE, T. Information management: A proposal. 1989.

BERNERS-LEE, T. Linked data-design issues (2006). **URL** <http://www.w3.org/DesignIssues/LinkedData.html>, 2006.

BERNERS-LEE, T.; FIELDING, R.; MASINTER, L. **Uniform resource identifiers (URI): generic syntax**. [S.l.]: RFC 2396, August, 1998. 4-8 p.

BERNERS-LEE, T. E. a. The semantic web-a new form of web content that is meaningful to computers will unleash a revolution of new possibilities. **Scientific American**, 2001.

BIZER, C. **Open Source Software Projects**. 2016. Disponível em: <<http://www.wiwiss.fu-berlin.de/en/fachbereich/bwl/pwo/bizer/team/BizerChristian.html>>. Acesso em: 07 Setembro 2016.

BIZER, C.; BECKER, C. **Marbles**. 2016. Disponível em: <<http://mes.github.io/marbles/>>. Acesso em: 07 Setembro 2016.

BIZER, C.; CYGANIAK, R. **Pubby - A Linked Data Frontend for SPARQL Endpoints**. 2016. Disponível em: <<http://wifo5-03.informatik.uni-mannheim.de/pubby/>>. Acesso em: 07 Setembro 2016.

BIZER, C.; CYGANIAK, R.; GAUSS, T. **WIQA - Web Information Quality Assessment Framework**. 2016. Disponível em: <<http://wifo5-03.informatik.uni-mannheim.de/bizer/wiqa/>>. Acesso em: 07 Setembro 2016.

BIZER, C.; CYGANIAK, R.; HARTIG, O. **NG4J - Named Graphs API for Jena**. 2016. Disponível em: <<http://wifo5-03.informatik.uni-mannheim.de/bizer/ng4j/>>. Acesso em: 07 Setembro 2016.

BIZER, C.; HEATH, T.; BERNERS-LEE, T. Linked data-the story so far. **Semantic Services, Interoperability and Web Applications: Emerging Concepts**, p. 205–227, 2009.

BIZER, C.; PROJECT, T. **RAP - RDF API for PHP**. 2016. Disponível em: <<http://wifo5-03.informatik.uni-mannheim.de/bizer/rdfapi/>>. Acesso em: 07 Setembro 2016.

BONI, V.; QUARESMA, S. J. Aprendendo a entrevistar: como fazer entrevistas em ciências sociais. **Em Tese**, v. 2, n. 1, p. 68–80, 2005.

BOSCH-SIJTSEMA, P. M.; BOSCH, J. Aligning innovation ecosystem strategies with internal r&d. In: IEEE. **Management of**

Innovation and Technology (ICMIT), 2014 IEEE International Conference on. [S.l.], 2014. p. 424–430.

BROEK, T.; VEENSTRA, A. F.; FOLMER, E. Walking the extra byte: a lifecycle model for linked open data. *Linked Open Data*, 2014.

CAPSENTA. **Ultrawrap**. 2016. Disponível em: <<https://capsenta.com/ultrawrap/>>. Acesso em: 07 Setembro 2016.

CARLAN, E. *Ontologia e web semântica*. 2006.

CHRISTENSEN, H. B. et al. Analysis and design of software ecosystem architectures—towards the 4s telemedicine ecosystem. **Information and Software Technology**, Elsevier, v. 56, n. 11, p. 1476–1492, 2014.

CHRISTIAN, R. Concepts of ecosystem, level and scale. **Ecology-Volume I**, Encyclopedia of Life Support Systems (EOLSS), p. 34, 2009.

CONGRESSO, N. **Lei da Inovação N° 10.973 DE 2 DE DEZEMBRO DE 2004**. 2016. Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2004-2006/2004/lei/110.973.htm> Acesso em: 07 Nov. 2016.

CORDEIRO, A. M. et al. Revisão sistemática: uma revisão narrativa. **Rev. Col. Bras. Cir**, v. 34, n. 6, p. 428–431, 2007.

COSTA, L. **Lucio Costa: registro de uma vivência**. [S.l.]: Empresa das Artes, 1995.

CUNNINGHAM, P. M.; CUNNINGHAM, M.; EKENBERG, L. Baseline analysis of 3 innovation ecosystems in east africa. In: **IEEE. Advances in ICT for Emerging Regions (ICTer), 2014 International Conference on.** [S.l.], 2014. p. 156–162.

D'AQUIN, M.; DIETZE, S. Open education: A growing, high impact area for linked open data. **ERCIM News**,(96), 2014.

DEFINITION, Open. 2016. 05/07/2016. Disponível em: <<http://opendefinition.org/>>.

DIAS, T. D.; SANTOS, N. Web semântica: Conceitos básicos e tecnologias associadas. **Cadernos do IME**, v. 14, 2003.

DIETZE, S. et al. Linked education: interlinking educational resources and the web of data. In: ACM. **Proceedings of the 27th annual ACM symposium on applied computing**. [S.l.], 2012. p. 366–371.

DOMINGUE, J.; FENSEL, D.; HENDLER, J. A. **Handbook of semantic web technologies**. [S.l.]: Springer Science & Business Media, 2011.

DURST, S.; POUTANEN, P. Success factors of innovation ecosystems: Initial insights from a literature review. **Pros. Co-Create**, 2013.

EGC. **EGC - Áreas de Concentração**. 2016. Disponível em: <<http://www.egc.ufsc.br/pos-graduacao/programa/areas-de-concentracao>>. Acesso em: 25 Fev. 2016.

EISENBERG, V. **D2R Server**. 2016. Disponível em: <<http://d2rq.org/getting-started>>. Acesso em: 07 Setembro 2016.

ETZKOWITZ, H. Innovation in innovation: The triple helix of university-industry-government relations. **Social science information**, Sage Publications, v. 42, n. 3, p. 293–337, 2003.

ETZKOWITZ, H.; LEYDESDORFF, L. The dynamics of innovation: from national systems and 'mode 2' to a triple helix of university–industry–government relations. **Research policy**, Elsevier, v. 29, n. 2, p. 109–123, 2000.

FAZENDA, I. C. A. **Didática e interdisciplinaridade**. [S.l.]: Papyrus Editora, 2008.

FIATES, J. E. A. **Influência dos Ecossistemas de Empreendedorismo Inovador na Indústria de Venture Capital: Estratégias de Apoio às Empresas Inovadoras**. Tese (Doutorado), 2014.

FILHO, F. W.; LÓSCIO, B. F. Web semântica: Conceitos e tecnologias. 2016. Acessado em: 04/07/2016.

FOUNDATION, O. K. **Comprehensive Kerbal Archive Network (CKAN)**. 2016. Disponível em: <<http://ckan.org/>>. Acesso em: 07 Setembro 2016.

GIL, A. C. Como elaborar projetos de pesquisa. **São Paulo**, v. 5, p. 61, 2002.

GOMES, F. C. d. R. Arquitetura de repositório semântico de dados para organização de pesquisa agropecuária. Universidade Federal do Rio de Janeiro/Instituto Brasileiro de Informação em Ciência e Tecnologia, 2013.

GOMES, M. A. S.; PEREIRA, F. E. C. Hélice tríplice: Um ensaio teórico sobre a relação universidade-empresa-governo em busca da inovação. **International Journal of Knowledge Engineering and Management (IJKEM)**, v. 4, n. 8, p. 136–155, 2015.

EGC/UFSC (Ed.). **Estudo comparativo dos elementos estruturais do ecossistema de inovação e ecossistema de software**. [S.l.: s.n.], 2016.

GOÑI, J. L.; FERNANDES, M. C. P.; LUCENA, C. J. P. de. **E-learning e a Web Semântica**. [S.l.]: PUC, 2002. 4,5 p.

HANDBOOK, Open Data. 2016. 05/07/2016. Disponível em: <http://opendatahandbook.org/guide/pt_BR/introduction/>.

HANSEN, K. M.; ZHANG, W. Towards structure-based quality awareness in software ecosystem use. In: SPRINGER. **International Conference on Service-Oriented Computing**. [S.l.], 2013. p. 469–479.

HARDASH, J. et al. Nasa innovation ecosystem: Host to a government technology innovation network. In: IEEE. **Aerospace Conference, 2014 IEEE**. [S.l.], 2014. p. 1–10.

HARTH, A.; JANIK, M.; STAAB, S. Semantic web architecture. In: **Handbook of Semantic Web Technologies**. [S.l.]: Springer, 2011. p. 43–75.

HUHTAM, J.; RUBENS, N. et al. Exploring innovation ecosystems as networks: Four european cases. In: IEEE. **2016 49th Hawaii International Conference on System Sciences (HICSS)**. [S.l.], 2016. p. 4505–4514.

ISELE, R. et al. **SILK**. 2016. Disponível em: <<http://silkframework.org/>>. Acesso em: 07 Setembro 2016.

JACKSON, D. J. What is an innovation ecosystem. **research paper, Engineering Research Centers, National Science Foundation**), http://erc-assoc.org/sites/default/files/topics/policy_studies/DJackson_Innovation%20Ecosystem_03-15-11.pdf, 2011.

JANSEN, S.; FINKELSTEIN, A.; BRINKKEMPER, S. A sense of community: A research agenda for software ecosystems. In: **IEEE. Software Engineering-Companion Volume, 2009. ICSE-Companion 2009. 31st International Conference on** [S.l.], 2009. p. 187–190.

JANSSEN, M.; CHARALABIDIS, Y.; ZUIDERWIJK, A. Benefits, adoption barriers and myths of open data and open government. **Information Systems Management**, Taylor & Francis, v. 29, n. 4, p. 258–268, 2012.

JR, J. J. L. D.; OLIVEIRA, J.; MEIRA, S. d. L. Pontos chaves para adoção de uma arquitetura orientada a serviços: Uma análise comparativa de modelos de maturidade soa da indústria. **VIII Simpósio Brasileiro de Sistemas de Informação, SBSI. São Paulo**, 2012.

KEELE, S. Guidelines for performing systematic literature reviews in software engineering. In: **Technical report, Ver. 2.3 EBSE Technical Report**. EBSE. [S.l.: s.n.], 2007.

KLING, G. W. **The Concept of the Ecosystem**. 2008. Disponível em:

<<http://www.globalchange.umich.edu/globalchange1/current/lectures/kling/ecosystem>>
Acesso em: 15 Jul. 2016.

KOIVUNEN, M.-R.; MILLER, E. W3c semantic web activity. **Semantic Web Kick-Off in Finland**, p. 27–44, 2001.

KOLOSKY, M. A. N.; SPERONI, R. d. M.; GAUTHIER, F. O. Ecosistemas de inovação ? uma revisão sistemática da literatura. **Espacios**, v. 36, p. 13, 2015.

KON, A. Ecosistemas de inovação: a natureza da inovação em serviços. **Revista de Administração, Contabilidade e Economia da Fundace**, v. 7, n. 1, 2016.

KUČERA, J.; CHLAPEK, D.; MYNARZ, J. Czech ckan repository as case study in public sector data cataloging. **Systémová Integrace**, v. 19, n. 2, p. 95–107, 2012.

LARRUCEA, X.; NANCLARES, F.; SANTAMARIA, I. A method for defining a regional software ecosystem strategy: Colombia as a case study. **Technological Forecasting and Social Change**, Elsevier, v. 104, p. 247–258, 2016.

LEAVENS, G. T.; SITARAMAN, M. **Foundations of component-based systems**. [S.l.]: Cambridge University Press, 2000.

LEHMANN, J. **DL-Learner**. 2016. Disponível em: <<http://dl-learner.org/one-page-introduction/>>. Acesso em: 07 Setembro 2016.

LETTNER, D. et al. A case study on software ecosystem characteristics in industrial automation software. In: **ACM. Proceedings of the 2014 International Conference on Software and System Process**. [S.l.], 2014. p. 40–49.

LEYDESDORFF, L.; ETZKOWITZ, H. The triple helix as a model for innovation studies. **Science and public policy**, Oxford University Press, v. 25, n. 3, p. 195–203, 1998.

LIMA, J. C. de; CARVALHO, C. L. de. **Resource description framework (rdf)**. [S.l.], 2005.

LOD Cloud 2014, State. 2016. <http://linkeddatacatalog.dws.informatik.uni-mannheim.de/state/>. Acessado em: 06/07/2016.

LYU, C. **DBpedia Extractor**. 2016. Disponível em: <<https://github.com/dbpedia/extraction-framework/wiki>>. Acesso em: 07 Setembro 2016.

MAGALHÃES, R. P. **Um ambiente para processamento de consultas federadas em linked data Mashups**. Tese (Doutorado), 2012.

MANIKAS, K. Revisiting software ecosystems research: a longitudinal literature study. **Journal of Systems and Software**, Elsevier, v. 117, p. 84–103, 2016.

MANIKAS, K.; HANSEN, K. M. Reviewing the health of software ecosystems—a conceptual framework proposal. In: **IWSECO@ICSOB**. [S.l.: s.n.], 2013. p. 33–44.

MANIKAS, K.; HANSEN, K. M. Software ecosystems—a systematic literature review. **Journal of Systems and Software**, Elsevier, v. 86, n. 5, p. 1294–1306, 2013.

MANIKAS, K.; KONTOGIORGOS, D. Characterizing software activity: The influence of software to ecosystem health. In: **ACM. Proceedings of the 2015 European Conference on Software Architecture Workshops**. [S.l.], 2015. p. 46.

MARCONI, M. d. A.; LAKATOS, E. M. **Técnicas de pesquisa: planejamento e execução de pesquisas, amostragens e técnicas de pesquisa, elaboração, análise e interpretação de dados**. [S.l.]: Atlas, 2011.

MARTINS, P. M. O. Back office para plataformas web de gestão de conferências científicas: Desenvolvimento e avaliação. **Porto: Universidade do Porto**, 2013.

MORAIS, E. A. M.; AMBRÓSIO, A. P. L. Ontologias: conceitos, usos, tipos, metodologias, ferramentas e linguagens. **Universidade Federal de Goiás**, 2007.

OH, D.-S. et al. Innovation ecosystems: A critical examination. **Technovation**, Elsevier, v. 54, p. 1–6, 2016.

O'REILLY, T. O que é web 2.0: padrões de design e modelos de negócios para a nova geração de software. **Disponível in <http://www.cipedia.com/doc/102010>**. Acesso em, v. 5, 2005.

PACHECO, R. C. d. S. et al. Governo na web: reflexões teóricas e práticas. In: _____. [S.l.]: Campinas: Alínea, 2015. cap. 3, p. 57–110.

PACHECO, R. C. S. Em interdisciplinaridade: Universidade e inovação social e tecnológica. In: _____. [S.l.]: Organizado por Joana Maria Pedro e Patrícia de Sá Freire, Curitiba, CRV, 2016. cap. Coprodução em Ciência, Tecnologia e Inovação: Fundamentos e Visões, p. 21–62.

- PELLICCIONE, P. Open architectures and software evolution: the case of software ecosystems. In: IEEE. **2014 23rd Australian Software Engineering Conference**. [S.l.], 2014. p. 66–69.
- PICKLER, M. E. V. Web semântica: ontologias como ferramentas de representação do conhecimento. **Perspectivas em Ciência da Informação**, SciELO Brasil, v. 12, n. 1, p. 65–83, 2007.
- PIZZOL, L. D. **Uso da Web de Dados como Fonte de Informação no Processo de Inteligência Competitiva Setorial**. 2014.
- PIZZOL, L. D. et al. Análise bibliométrica da produção científica sobre linked data. **Informação & Informação**, v. 20, n. 3, p. 77–112, 2015.
- PROJECT, N. T. **Neon ToolKit**. 2016. Disponível em: <<http://neon-toolkit.org/>>. Acesso em: 07 Setembro 2016.
- QUADRANT, T. **Top Braid Composer**. 2016. Disponível em: <<http://www.topquadrant.com/tools/modeling-topbraid-composer-standard-edition/>>. Acesso em: 07 Setembro 2016.
- RAMALHO, R. A. S. Desenvolvimento e utilização de ontologias em bibliotecas digitais: uma proposta de aplicação. Universidade Estadual Paulista (UNESP), 2010.
- RAMALHO, R. A. S.; OUCHI, M. T. Tecnologias semânticas: novas perspectivas para a representação de recursos informacionais. **Informação & Informação**, v. 16, n. 3, p. 60–75, 2011.
- RESEARCH, S. C. for B. I. **Protégé**. 2016. Disponível em: <<http://protege.stanford.edu/>>. Acesso em: 07 Setembro 2016.
- ROCHA, R. P. d. Web semântica, dados ligados e web 2.0: explorando novas fronteiras para os arquivos abertos. 2014.
- ROSENFELD, L.; MORVILLE, P. **Information architecture for the world wide web**. [S.l.]: "O'Reilly Media, Inc.", 2002.
- SÁ, P. R. O. d. **Engenharia do Conhecimento Aplicada a Criação Automatizada de Conteúdo Interativo para TV Digital**. 2012.

SAMPAIO, R. F.; MANCINI, M. C. Estudos de revisão sistemática: um guia para síntese criteriosa da evidência científica. **Braz. J. Phys. Ther.(Impr.)**, v. 11, n. 1, p. 83–89, 2007.

SANTOS, E.; NICOLAU, M. Web do futuro: a cibercultura e os caminhos trilhados rumo a uma web semântica ou web 3.0. **Temática**, v. 8, n. 10, 2015.

SANTOS, R. P. dos; WERNER, C. M. L. A proposal for software ecosystems engineering. In: **IWSECO@ ICSOB**. [S.l.: s.n.], 2011. p. 40–51.

SCHMID, K. Variability support for variability-rich software ecosystems. In: IEEE. **Product Line Approaches in Software Engineering (PLEASE), 2013 4th International Workshop on**. [S.l.], 2013. p. 5–8.

SCHULTZ, A. et al. **LDIF ? Linked Data Integration Framework**. 2016. Disponível em: <<http://ldif.wbsg.de/>>. Acesso em: 07 Setembro 2016.

SERVANT, F.-P. Linking enterprise data. In: **LDOW**. [S.l.: s.n.], 2008.

SHETH, A. **Semantic Services, Interoperability and Web Applications: Emerging Concepts: Emerging Concepts**. [S.l.]: IGI Global, 2011.

SILVA, C. E. L. da; NARCIZO, R. B.; CARDOSO, R. The proximity between academy, industry and government: Towards a more sustainable development of a brazilian oil region. **Procedia-Social and Behavioral Sciences**, Elsevier, v. 52, p. 100–109, 2012.

SILVA, T. S. da. **RECONHECIMENTO DE ENTIDADES NOMEADAS EM NOTÍCIAS DE GOVERNO**. Tese (Doutorado) — Universidade Federal do Rio de Janeiro, 2012.

SOMMERVILLE, I. **Engenharia de Software 9.ed.** [S.l.]: Pearson View, 2011.

SOUZA, C. R. de et al. The social side of software platform ecosystems. In: ACM. **Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems**. [S.l.], 2016. p. 3204–3214.

SOUZA, V. A. S. M. de et al. Uma arquitetura orientada a serviços para desenvolvimento, gerenciamento e instalação de serviços de rede. Campinas, SP, 2006.

SPERONI, R. S. **Modelo de Referência Para Indicadores de Inovação Regional Suportado por Dados Ligados**. Tese (Doutorado), 2016.

TEAM, D. S. **DBpedia Spotlight**. 2016. Disponível em: <<https://github.com/dbpedia-spotlight/dbpedia-spotlight/wiki>>. Acesso em: 07 Setembro 2016.

UBALDI, B. Open government data. OECD Publishing, 2013.

URLI, S. et al. Managing a software ecosystem using a multiple software product line: a case study on digital signage systems. In: **IEEE. 2014 40th EUROMICRO Conference on Software Engineering and Advanced Applications**. [S.l.], 2014. p. 344–351.

VILLAZÓN-TERRAZAS, B. et al. Methodological guidelines for publishing government linked data. In: **Linking government data**. [S.l.]: Springer, 2011. p. 27–49.

W3C. **Tabulator**. 2016. Disponível em: <<https://www.w3.org/2005/ajar/tab.html>>. Acesso em: 07 Setembro 2016.

W3C. **Transforming RDF with XSLT**. 2016. Disponível em: <<https://www.w3.org/XML/2000/04rdf-parse/>>. Acesso em: 07 Setembro 2016.

W3C. **W3C Arquitetura Web Semântica**. 2016. <https://www.w3.org/2007/03/layerCake.svg>. Acessado em: 28/06/2016.

W3C. **W3C Fatos sobre a W3C**. 2016. Acessado em: 28/06/2016. Disponível em: <<https://www.w3.org/Consortium/facts#history>>.

W3C. **W3C Internationalized Resource Identifiers (IRIs)**. 2016. <https://www.w3.org/International/0-URL-and-ident.html>. Acessado em: 01/07/2016.

W3C, C. **XML Technology**. 2016. <http://www.w3.org/standards/xml/>. Acessado em: 29/06/2016.

WEB, O. L. S. **Open Link Virtuoso**. 2016. Disponível em:
<<http://virtuoso.openlinksw.com>>. Acesso em: 07 Setembro 2016.

WEIL, H. B.; SABHLOK, V. P.; COONEY, C. L. The dynamics of innovation ecosystems: A case study of the us biofuel market. **Energy Strategy Reviews**, v. 3, p. 88–99.