

**UNIVERSIDADE FEDERAL DE SANTA CATARINA
DEPARTAMENTO DE ENGENHARIA DO CONHECIMENTO**

Rafael de Moura Speroni

**MODELO DE REFERÊNCIA PARA INDICADORES DE
INOVAÇÃO REGIONAL SUPOSTADO POR DADOS LIGADOS**

Tese submetida ao Programa de Pós-Graduação em Engenharia e Gestão do Conhecimento da Universidade Federal de Santa Catarina (UFSC) para a obtenção do Grau de Doutor em Engenharia e Gestão do Conhecimento.
Orientador: Prof. Dr. Fernando A. Ostuni Gauthier
Coorientador: Prof. Dr. Marcelo Macedo

Florianópolis
2016

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

Speroni, Rafael de Moura
MODELO DE REFERÊNCIA PARA INDICADORES DE INOVAÇÃO
REGIONAL SUPORTADO POR DADOS LIGADOS / Rafael de Moura
Speroni ; orientador, Fernando A. Ostuni Gauthier ;
coorientador, Marcelo Macedo. - Florianópolis, SC, 2016.
233 p.

Tese (doutorado) - Universidade Federal de Santa
Catarina, Centro Tecnológico. Programa de Pós-Graduação
Multidisciplinar em Saúde.

Inclui referências

1. Saúde. 2. Inovação. 3. Indicadores de Inovação. 4.
Dados Abertos. 5. Dados Ligados. I. Gauthier, Fernando A.
Ostuni. II. Macedo, Marcelo. III. Universidade Federal de
Santa Catarina. Programa de Pós-Graduação Multidisciplinar
em Saúde. IV. Título.

Rafael de Moura Speroni

MODELO DE REFERÊNCIA PARA INDICADORES DE INOVAÇÃO REGIONAL SUPORTADO POR DADOS LIGADOS

Esta Tese foi julgada adequada para obtenção do Título de “Doutor em Engenharia e Gestão do Conhecimento”, na Área de Concentração em Engenharia do Conhecimento, e aprovada em sua forma final pelo Programa de Pós-Graduação em Engenharia e Gestão do Conhecimento da Universidade Federal de Santa Catarina.

Florianópolis, 29 de Abril de 2016.

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

Banca Examinadora:

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

Prof.^a. Branca Regina Cantisano dos
Santos e Silva Riscado Terra, Dr.^a.
Universidade Estadual do Rio de
Janeiro (Videoconferência)

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

Prof. Silvestre Labiak Junior, Dr.
Universidade Tecnológica Federal do
Paraná

Prof. José Leomar Todesco, Dr.
Universidade Federal de Santa
Catarina

Prof. Rogério Cid Bastos
Universidade Federal de Santa Catarina

Dedico este trabalho à minha esposa
Lilia e aos meus pais Theresinha e
Valdemar (*in memoriam*).

AGRADECIMENTOS

A realização deste trabalho teria sido muito mais difícil se não fosse todo o apoio que recebi. Desta forma, gostaria de agradecer a algumas das pessoas e instituições que me deram este suporte.

Registro aqui meu agradecimento ao Instituto Federal Catarinense por ter proporcionado as condições para a minha dedicação a esta empreitada. A todos os colegas desta instituição e, em especial, ao pessoal dos Câmpus Araquari e São Francisco do Sul, equipes das quais me ausentei temporariamente por este nobre motivo.

Agradeço também à Universidade Federal de Santa Catarina, ao Departamento de Engenharia do Conhecimento e toda a sua equipe técnica, docente e discente, família à qual tenho orgulho de pertencer, sendo mais um “Egeciano”.

Meu especial agradecimento aos professores Fernando Gauthier, pela confiança em me acolher mais uma vez entre seus orientandos, e Marcelo Macedo, que passou de colega a coorientador, contribuindo significativamente para a conclusão do trabalho.

Não poderia deixar de agradecer aos amigos que fiz no Departamento de Ciências da Administração da UFSC. Foi uma grande honra trabalhar com vocês durante boa parte desse período. Um forte abraço aos professores Alexandre Ramos, Claudelino Junior, Eduardo Lobo, André Leite e Marcos Dalmau.

Aos colegas de pesquisa do Laboratório de Engenharia do Conhecimento, meu muito obrigado por todos os papos no LED, pelas coletas e tratamentos de dados, pelos artigos que escrevemos e pelos churrascos, que também contribuem para o processo. Contem comigo para os próximos.

Aos amigos do IFSC, Antônio Cândido, Cleverson Tabajara e pela parceria. Ao Marco Neiva Koslosky pelas conversas e direcionamentos nos momentos de dúvidas no alinhamento do tema.

Finalmente, agradeço a toda a minha família. Lilia, meu amor, sempre de malas prontas para ir comigo para onde for. Meus pais, Theresinha e Valdemar (*in memoriam*), vocês foram, são e sempre serão o meu exemplo e a minha inspiração. Mônica, Mateus, e as lindas Laura e Alice, contem sempre com o “Dindo Rafa”. À família Medeiros Cardoso, que me acolheu como um filho, e que acompanhou mais de perto esse período, meu muito obrigado.

“É difícil imaginar o poder que você terá quando tantos diferentes tipos de dados estão disponíveis.”

(BERNERS-LEE, 2007)

RESUMO

A inovação está associada à implementação de um novo, ou significativamente melhorado, produto ou processo, método de *marketing*, ou método organizacional, e é reconhecida como um fator-chave para o desenvolvimento econômico. Em função de suas características, a inovação passou a ser vista como um processo complexo que acontece em um ambiente onde diferentes tipos de atores interagem em um sistema, que, quando considerado sob o contexto regional, é chamado de Sistema Regional de Inovação, e que demanda de estratégias e políticas que incentivem e potencializem o desenvolvimento das atividades de inovação, e de ferramentas para avaliação das ações. Diferentes índices são propostos para a mensuração da inovação ao nível das empresas ou em nível nacional, mas que são de difícil aplicação em nível regional, em função subjetividade de escolha das variáveis e da falta de disponibilidade de dados. Esta tese propõe um modelo de referência concebido a partir da análise dos modelos de indicadores compostos para mensuração da inovação regional apresentados na literatura, propondo uma classificação hierárquica para os indicadores. Valendo-se de tecnologias semânticas, o modelo é suportado por Dados Ligados, objetivando a exploração do potencial de dados regionais disponibilizados na *Web* por iniciativas de dados abertos e transparência pública na definição de índices específicos para a inovação regional, em uma forma tal que possibilitem o seu processamento automatizado. Uma prova de conceito é apresentada com a finalidade de demonstrar a viabilidade de utilização do modelo em aplicações reais. Dados sobre municípios de duas mesorregiões do estado de Santa Catarina foram coletados e publicados na forma de dados ligados e foi desenvolvido um protótipo de aplicação *Web* para visualização e análise comparativa entre diferentes níveis regionais e de agregação de indicadores.

Palavras-chave: Indicadores de Inovação. Inovação Regional. Dados Ligados.

ABSTRACT

Innovation is related to the implementation of new, or significantly improved, product or process, marketing method, or organizational method, and is recognized as a key-factor for economic development. Due to its characteristics, innovation is seen as a complex process that occurs in an environment where different actors interact in a system which, when considered in the regional context, is called Regional Innovation System, and demands strategies and policies that encourage and leverage innovation activities and tools for evaluation. There are many different index systems proposed to measure innovation at firm level or at country level, but it is pointed that the choice of variables is subjective, and that is difficult to apply them in the regional level, due to the lack of data availability. This thesis proposes a reference model for classification of innovation indicators. The model is based on the analysis of the literature on composite indicators models for measuring regional innovation. Through the use of semantic technologies, this model is supported by Linked Data and aims to explore the potential of regional open data, available due to transparency and open data initiatives. A proof of concept is presented for the purpose of demonstrating the feasibility of the model real world applications. Data on municipalities from two regions of the Santa Catarina state were collected, semantically annotated to the model, and published as linked data. A Web application prototype was developed for visualization and comparative analysis of different regional levels and aggregate indicators.

Keywords: Innovation Indicators. Regional Innovation. Linked Data.

LISTA DE FIGURAS

Figura 1 - Procedimentos metodológicos do trabalho.....	23
Figura 2 – Uma taxonomia de Inovações.....	33
Figura 3 - Modelo de Inovação fechada - Funil da Inovação.....	36
Figura 4 - Modelo de Inovação Aberta - Funil da Inovação	37
Figura 5 - O modelo da Triplíce Hélice	42
Figura 6 - Atores de um Sistema de Inovação	43
Figura 7 - Componentes de um Sistema de Inovação	44
Figura 8 – <i>Framework</i> do Manual de Oslo	53
Figura 9 - Fluxo de blocos temáticos da PINTEC	59
Figura 10 - Evolução da Engenharia do Conhecimento.....	64
Figura 11 - Pirâmide metodológica de Schreiber.....	65
Figura 12 - Classificação de ontologias quanto ao grau de dependência em uma tarefa.....	70
Figura 13: OntoCube - um sistema de classificação integrado	71
Figura 14 - Fase de competências no ontoKEM	75
Figura 15 - Fase de Vocabulário no ontoKEM	76
Figura 16 - Fase de Hierarquia de classes.....	76
Figura 17 - Fase de Dicionário de classes no ontoKEM	77
Figura 18 - Relação entre identificador, recurso e representação na <i>Web</i>	88
Figura 19 - Trecho de documento XML contendo informações sobre artigo científico	89
Figura 20 - Grafo RDF com dois nós (Sujeito e Objeto) e conexão entre eles (Predicado).....	91
Figura 21 - Exemplo de triplas RDF escritas em pseudocódigo	91
Figura 22 - Exemplo de grafo RDF	92
Figura 23 – Diagrama da nuvem <i>Linking Open Data</i>	97
Figura 24 - Arquitetura de publicação de dados RDF em um Triple Store.....	101
Figura 25 - Exemplo de arquitetura de publicação com mapeamento entre relacional e RDF.....	102
Figura 26 - Exemplo de tripla RDF	103
Figura 27 - Exemplo de consulta SPARQL	103
Figura 28 - Resultado da consulta SPARQL	103
Figura 29 - Exemplo de multicubo	106
Figura 30 - Representação das operações OLAP	107
Figura 31 - Abordagem usual para definição de índices de inovação.....	110
Figura 32 - Modelo de Referência para Indicadores de Inovação Regional ...	111
Figura 33 – Modelo de Índice de Inovação Regional: Níveis zero e um	113
Figura 34 – Classificação dos Indicadores de Inovação Regional	118
Figura 35 - Questões de competência registradas no ontoKEM	124
Figura 36 – Registro e categorização de termos no ontoKEM.....	125
Figura 37 - Definição da hierarquia de classes no ontoKEM.....	126
Figura 38 - Representação parcial das classes da ontologia.....	127
Figura 39 - Consulta SPARQL: lista de subclasses	128

Figura 40 - Resultado da consulta por subclasses.....	128
Figura 41 - Instância de modelo específico	129
Figura 42 - Classes instanciadas com os indicadores do RIS	130
Figura 43 - Passos do processo de publicação de dados ligados.....	131
Figura 44 - Visualização da URI de uma classe do modelo no navegador	132
Figura 45 - Exemplo de consulta no <i>endpoint</i> SPARQL.....	133
Figura 46 - Resultados de consulta em diferentes formatos	134
Figura 47 - Publicação de Dados Ligados sobre a Inovação Regional.....	138
Figura 48 - Hierarquia de classes da ontologia.....	141
Figura 49 - Atribuição de peso a um indicador em um modelo.....	142
Figura 50 - Trecho de arquivo em formato CSV	143
Figura 51 - Ilustração do mapeamento de municípios	144
Figura 52 - Arquivo CSV após o mapeamento.....	145
Figura 53 - Carga dos dados e parametrização do Cubo RDF.....	145
Figura 54 - Trecho de dados transformados em <i>RDF Data Cube</i>	146
Figura 55 - Interface para upload dos dados ligados	147
Figura 56 - Página <i>Web</i> de apresentação do modelo	148
Figura 57 - Apresentação de dados ligados no navegador.....	148
Figura 58 - Consulta SPARQL por um conjunto de dados, agregado por microrregiões	149
Figura 59 - Visualização de tabular e gráfica de conjunto de dados, agregado por microrregiões.....	150
Figura 60 - Visualização comparativa entre diferentes regiões	151
Figura 61 - Visualização comparativa entre regiões utilizando maiores níveis de agregação.....	152

LISTA DE QUADROS

Quadro 1 - Publicações científicas do autor durante o desenvolvimento da tese	24
Quadro 2 – Prós e Contras dos indicadores compostos.	52
Quadro 3 - Comparativo entre indicadores do IUS e RIS.....	55
Quadro 4- Comparação das metodologias da União Européia.....	57
Quadro 5 - Descrição dos eixos do <i>OntoCube</i>	72
Quadro 6 – Tipos de licenças <i>Creative Commons</i>	86
Quadro 7 - As cinco estrelas dos Dados Abertos.....	95
Quadro 8- Classificação dos Indicadores de Entrada utilizados na literatura .	115
Quadro 9- Classificação dos Indicadores de Saída utilizados na literatura.....	117
Quadro 10 – Assuntos abordados pelos indicadores das dimensões de entrada	119
Quadro 11 – Assuntos abordados pelos indicadores das dimensões de saída.	122
Quadro 12 - Lista de indicadores utilizados no <i>Regional Innovation Scoreboard</i>	129
Quadro 13 - Lista de fontes e conjuntos de dados utilizados no protótipo.....	139

LISTA DE ABREVIATURAS E SIGLAS

IBGE

SUMÁRIO

1	INTRODUÇÃO.....	11
1.1	OBJETIVOS.....	14
1.1.1	<i>Objetivo Geral.....</i>	<i>14</i>
1.1.2	<i>Objetivos Específicos</i>	<i>15</i>
1.2	JUSTIFICATIVA.....	15
1.3	ADERÊNCIA AO EGC.....	20
1.4	1PROCEDIMENTOS METODOLÓGICOS.....	22
1.5	ESTRUTURA DO TRABALHO.....	26
2	INOVAÇÃO REGIONAL.....	29
2.1	CONSIDERAÇÕES INICIAIS	29
2.2	O CONCEITO DE INOVAÇÃO	29
2.3	CLASSIFICAÇÕES DE INOVAÇÃO	32
2.4	SISTEMAS DE INOVAÇÃO.....	38
2.4.1	<i>O Sistema Nacional de Inovação Brasileiro</i>	<i>Erro! Indicador não definido.</i>
2.4.2	<i>Sistemas Regionais de Inovação.....</i>	<i>44</i>
2.5	INDICADORES DE INOVAÇÃO	46
2.5.1	<i>Estatísticas de Pesquisa & Desenvolvimento.....</i>	<i>47</i>
2.5.2	<i>Estatísticas de Patentes.....</i>	<i>49</i>
2.5.3	<i>Surveys de Inovação</i>	<i>50</i>
2.5.4	<i>Indicadores de Inovação Compostos</i>	<i>50</i>
2.6	CONSIDERAÇÕES FINAIS.....	61
3	ONTOLOGIAS	63
3.1	CONSIDERAÇÕES INICIAIS	63
3.2	ENGENHARIA DO CONHECIMENTO.....	63
3.3	DEFINIÇÃO DE ONTOLOGIA.....	66
3.4	TIPOS DE ONTOLOGIA	69
3.5	METODOLOGIAS DE DESENVOLVIMENTO DE ONTOLOGIAS	73
3.6	CONSIDERAÇÕES FINAIS.....	77
4	DADOS LIGADOS	79
4.1	CONSIDERAÇÕES INICIAIS	79
4.2	DADOS ABERTOS	79
4.2.1	<i>Formato e Licenças de Dados Abertos.....</i>	<i>83</i>
4.3	WEB SEMÂNTICA.....	87
4.4	DADOS LIGADOS (<i>LINKED DATA</i>)	93

4.4.1	<i>Dados Abertos Ligados (Linked Open Data)</i>	95
4.4.2	<i>Publicação de Dados Ligados</i>	98
4.5	DADOS ESTATÍSTICOS LIGADOS	104
4.6	CONSIDERAÇÕES FINAIS	107
5	MODELO DE REFERÊNCIA PARA INDICADORES DE INOVAÇÃO REGIONAL SUPOTADOS POR DADOS LIGADOS	109
5.1	CONSIDERAÇÕES INICIAIS	109
5.2	ABORDAGEM DO MODELO PROPOSTO	109
5.3	MODELO CONCEITUAL DE INDICADORES DE INOVAÇÃO REGIONAL	112
5.4	MODELO DE REFERÊNCIA SUPOSTADO POR DADOS LIGADOS ..	123
5.4.1	<i>Construção da Ontologia</i>	123
5.4.2	<i>Publicação e uso da Ontologia</i>	127
5.5	INFRAESTRUTURA PARA A PUBLICAÇÃO DE DADOS ESTATÍSTICOS LIGADOS SOBRE A INOVAÇÃO REGIONAL	130
5.6	CONSIDERAÇÕES FINAIS.....	134
6	PROVA DE CONCEITO: DADOS LIGADOS SOBRE A INOVAÇÃO NOS MUNICÍPIOS DE SANTA CATARINA	137
6.1	CONSIDERAÇÕES INICIAIS	137
6.2	MODELO ESPECÍFICO DE INDICADORES DE INOVAÇÃO	137
6.3	DADOS LIGADOS SOBRE A INOVAÇÃO REGIONAL.....	138
6.3.1	<i>Identificação das Fontes de Dados</i>	139
6.3.2	<i>Criação da Ontologia</i>	141
6.3.3	<i>Mapeamento dos Dados</i>	142
6.3.4	<i>Publicação dos Dados</i>	146
6.3.5	<i>Consumo e visualização dos Dados</i>	147
6.4	CONSIDERAÇÕES FINAIS.....	152
7	CONCLUSÕES	155
7.1	SUGESTÕES PARA TRABALHOS FUTUROS	159
	REFERÊNCIAS	161
	APÊNDICE A – QUESTÕES DE COMPETÊNCIA DA ONTOLOGIA (ONTOKEM)	181
	APÊNDICE B – TERMOS CATEGORIZADOS COMO CLASSES NA ONTOLOGIA (ONTOKEM)	183

APÊNDICE C – LEVANTAMENTO DOS MODELOS E INDICADORES UTILIZADOS NA LITERATURA DE REFERÊNCIA	189
---	------------

1 INTRODUÇÃO

O cenário dinâmico observado na sociedade e na economia atuais é uma consequência das novas características impostas pela chamada Economia do Conhecimento. O desenvolvimento econômico de uma empresa, região ou país está associado à capacidade que demonstram para se manterem competitivos frente ao ambiente em que estão inseridos. Os países em desenvolvimento, em especial, têm que realizar um enorme esforço para avançar na geração e utilização do conhecimento técnico-científico, criando capacidades e competências em áreas estratégicas.

O termo inovação é bastante amplo, e possui várias definições, mas está normalmente associado ao desenvolvimento de novos produtos, serviços ou processos, além dos resultados em termos de comercialização destas novidades (HANSEN; WAKONEN, 1997; CROSSAN; APAYDIN, 2010, p. 12). Devido à sua complexidade, a inovação passa a ser vista cada vez menos como um processo interno e isolado, dando lugar ao entendimento de que tão ou mais importantes do que os fatores internos, são os fatores externos, as redes de relacionamento e a pesquisa e desenvolvimento colaborativos, com ênfase na inovação aberta (CHESBROUGH, 2003; LORENZO, 2008; CARAYANNIS; CAMPBELL, 2011). Ganham importância, portanto, as configurações que consideram que a inovação acontece em um ambiente onde diferentes atores interagem de forma a coevolurem em torno de objetivos comuns.

Com base em teorias econômicas, passou-se a analisar a inovação por seu papel na economia de regiões como forma de explicar a organização das atividades econômicas no espaço, bem como o crescimento ou estagnação destas ao longo do tempo. O argumento é de que o conjunto de atores da região produz efeitos sistêmicos que encorajam as empresas a desenvolver formas específicas de capital que derivam das relações sociais, normas, valores e interações com a comunidade, reforçando a competitividade e a capacidade de inovação regional (DOLOREUX; PARTO, 2004).

As formas de organização do ambiente de inovação ganham diferentes nomenclaturas, como Sistemas Nacionais de Inovação (LUNDVALL, 1992), Sistemas Regionais de Inovação (DOLOREUX; PARTO, 2004; ZABALA-ITURRIAGAITIA *et al.*, 2007), Ecossistemas de Inovação (NAMBISAN; BARON, 2013; KOSLOSKY; SPERONI; GAUTHIER, 2015). Devido às características esperadas para estes atores dentro do ambiente de inovação, sua interação é apresentada na literatura sob

diferentes pontos de vista. Para representá-la, são citadas estruturas como o triângulo de Sábato, Hélice Tríplice (ETZKOWITZ; LEYDESDORFF, 1995), Hélice Sêxtupla (LABIAK JUNIOR, 2012), que representam as interações entre os diferentes atores em um sistema de inovação.

Diante do entendimento da importância destas características para o desenvolvimento econômico, diversos esforços governamentais vêm sendo empreendidos pelo mundo todo, no sentido de fomentar a inovação. A estrutura das redes de colaboração em Ciência e Tecnologia, incluindo o suporte governamental em áreas estratégicas para o desenvolvimento do país, representa uma estratégia para a promoção de uma maior integração dos atores da inovação, com a construção de pontes para aumentar substancialmente a transferência e aplicação do conhecimento, investimento em pesquisa e desenvolvimento (P&D) e absorção de pesquisadores altamente qualificados nas companhias (VARRICHIO *et al.*, 2012).

No Brasil, buscou-se o aumento na interlocução entre Governo Federal, Estados Municípios, agências de fomento, institutos de pesquisa, universidades e empresas, buscando-se a consolidação institucional de um sistema de Ciência, Tecnologia e Inovação (CT&I). A criação da Estratégia Nacional de Ciência, Tecnologia e Inovação (ENCTI) demonstra o esforço empreendido para a articulação e coordenação de políticas públicas relacionadas com o desenvolvimento científico e tecnológico do país. A prioridade é traduzir este desenvolvimento em progresso material e bem estar social para o conjunto da população brasileira, por meio da revolução do sistema educacional e da incorporação da inovação ao sistema produtivo como mecanismo de reprodução e ampliação do potencial social e econômico (MCTI, 2012).

Mais especificamente, no estado de Santa Catarina, a Política Catarinense de Ciência, Tecnologia e Inovação (FAPESC, 2010) visa promover o avanço científico, tecnológico e de inovações no ambiente produtivo, nas instituições de ensino, pesquisa e extensão, nos agentes econômicos e sociais e nos órgãos de governo, visando à qualidade de vida dos habitantes e ao desenvolvimento social e econômico do estado, com sustentabilidade ambiental e equilíbrio regional. Além disso, em decorrência da Lei Catarinense de Inovação (SANTA CATARINA, 2008), o estado destina dois por cento de sua receita líquida a órgãos de fomento, pesquisa, desenvolvimento e inovação.

Os esforços governamentais empreendidos no fomento da inovação podem ter seus resultados medidos por uma série de

indicadores. Comumente, são verificadas as estatísticas de patentes, bem como os dispêndios em P&D, mas estes são indicativos parciais do esforço total de inovação realizado (MAKKONEN; HAVE, 2013).

Em outras abordagens, entretanto, a inovação é medida utilizando-se índices compostos, onde fatores de entrada e saída são utilizados simultaneamente, sintetizando as informações providas por vários indicadores em um só (EUROPEAN UNION, 2012; CORNELL UNIVERSITY; INSEAD; WIPO, 2013; IBGE, 2013). Makkonen e Have (2013) observam, contudo, que a grande maioria dos trabalhos aborda o uso de índices compostos de inovação ao nível das empresas ou em nível nacional, enquanto poucos são os trabalhos que tratam de indicadores para o nível regional. Além disso, muitos deles, não podem ser empiricamente verificados devido à falta de dados estatísticos (ZABALA-ITURRIAGAGOITIA *et al.*, 2007).

Um exemplo representativo do que acontece em relação a essa falta de disponibilidade de dados estatísticos é verificável no *Regional Innovation Scoreboard*, uma publicação que apresenta a análise comparativa de regiões na Europa. Para isto aplica o modelo composto de 25 de indicadores proposto no *Union Innovation Scoreboard*, cujo foco é a análise comparativa dos países. Na aplicação em nível regional, entretanto, somente 11 indicadores são utilizados, em função da disponibilidade de dados (EUROPEAN UNION, 2012; EUROPEAN UNION, 2014a).

Se, de um lado, a falta de disponibilidade de dados é um problema recorrente nas iniciativas que visam à mensuração da inovação, há outras iniciativas que tratam justamente da disponibilização de grandes conjuntos de dados por meio de sua publicação na *Web*. Nos últimos anos, os governos de vários países, incluindo o Brasil, vêm demonstrando crescente preocupação com a necessidade de participação da sociedade no governo. Dentre as iniciativas neste sentido, salientam-se as que incentivam, e até mesmo obrigam, a publicação de dados abertos por órgãos governamentais. A motivação pela disponibilização de dados é justificada pela esperança do aumento do envolvimento dos cidadãos no governo, do aumento da transparência e da melhoria na tomada de decisão, alinhadas com a ambição de governos mais inteligentes (BERTOT; CHOI, 2013; CONRADIE; CHOENNI, 2014).

Em 2009, em seu primeiro dia de mandato, o presidente dos EUA, Barack Obama, emitiu um memorando sobre transparência e governo aberto, vindo a resultar na criação do portal *data.gov*. Muitos outros países seguiram esta tendência, criando seus próprios portais de

dados abertos governamentais. No Brasil, a Lei nº 12.527/2011, conhecida como Lei de Acesso à Informação, começou a vigorar em maio de 2012 e estabelece a obrigatoriedade de que os órgãos da administração pública disponibilizem dados abertos na Web, bem como de que respondam a solicitações dos cidadãos por novos conjuntos de dados que ainda não estejam disponíveis. Como resultado, o governo brasileiro tornou-se membro da Parceria de Governo Aberto (*Open Government Partnership - OGP*) e criou o Portal Brasileiro de Dados Abertos.

Embora tais iniciativas tenham possibilitado a disponibilização de grandes quantidades de dados na *Web*, a falta de padronização e estrutura destes dados impõe limitações quanto ao seu uso. As tecnologias da *Web Semântica* (BERNERS-LEE; HENDLER; LASSILA, 2001) surgem como uma forma de acrescentar significado e estrutura bem definidos aos dados na *Web*. Seguindo princípios de publicação de dados ligados (*Linked Data*), os dados são interligados em um espaço global, conhecido como *Web de Dados*, de forma que seja possível “navegar” nos dados de forma semelhante ao que se faz nos documentos da *Web* tradicional, além de torna-los aptos para o processamento automatizado (BIZER; HEATH; BERNERS-LEE, 2009).

Uma das características fundamentais da *Web Semântica*, e que a diferencia da abordagem tradicional, é justamente a existência de uma camada adicional de semântica, na qual os dados são contextualizados, ou seja, há uma representação explícita de seus significados. Isto é possível graças à utilização de ontologias, que provêm um vocabulário que define formalmente os termos de um determinado domínio e os relacionamentos que existem entre eles.

Caracterizadas a importância da inovação como indutor do crescimento econômico, bem como a disponibilidade de dados abertos decorrente de iniciativas públicas e privadas de publicação de dados na *web*, apresenta-se a seguinte pergunta de pesquisa:

De que forma o uso de dados ligados pode contribuir para representação de índices específicos de inovação regional?

1.1 OBJETIVOS

1.1.1 Objetivo Geral

Desenvolver um modelo de referência suportado por dados ligados para a representação de indicadores de inovação regional.

1.1.2 Objetivos Específicos

Para o cumprimento do Objetivo Geral do trabalho são definidos os seguintes Objetivos Específicos:

- Caracterizar e definir a inovação no contexto regional;
- Identificar os principais constructos dos indicadores de inovação regional;
- Discutir o papel dos dados abertos como possível fonte para a composição de indicadores da inovação regional;
- Propor um modelo de referência para a representação de indicadores de inovação regional;
- Desenvolver um protótipo baseado em dados ligados que possibilite a representação de múltiplos índices de inovação a partir de dados abertos.

1.2 JUSTIFICATIVA

A inovação tornou-se um fator chave para o crescimento e competitividade (PINTO; GUERREIRO, 2010), e serve, atualmente, como um tipo de atrator intelectual que leva em conta os desafios impostos por uma sociedade baseada em conhecimento e pela competitividade econômica e global entre empresas, regiões, países e até continentes (JALONEN, 2013). Como afirmam Bessant e Tidd (2009), se uma empresa não muda os bens ou serviços oferecidos, bem como a forma como os cria e oferta no mercado, corre o risco de ser superada por outras que o façam. Da mesma forma, municípios, regiões, estados e países buscam manterem-se competitivos e em crescimento sustentável e, para isto têm empregado esforços no desenvolvimento de políticas de incentivo à inovação.

Incentivar a inovação envolve, costumeiramente, ações que promovam a formação de recursos humanos, empreendedorismo, atração de investimentos, e uma aproximação entre universidades e empresas, visando o aumento da transferência tecnológica. Neste contexto, as políticas públicas desempenham importante papel, devido às necessidades de reconfiguração dos sistemas nacionais de inovação, tendo em vista a criação de um ambiente mais propício ao desenvolvimento das capacidades para a inovação (VAN KLEEF; ROOME, 2007).

No Brasil, a Estratégia Nacional de Ciência, Tecnologia e Inovação (ENCTI) (MCTI, 2012) foi criada como um marco para a

articulação e coordenação intertemporal das políticas públicas setoriais e transversais e das iniciativas do setor privado relacionadas com o desenvolvimento científico e tecnológico do País.

“A prioridade agora, é principalmente traduzir o desenvolvimento científico e tecnológico em progresso material e bem estar social para o conjunto da população brasileira, o que passa pela convergência de dois macro-movimentos estruturais: a revolução do sistema educacional e a incorporação sistemática ao processo produtivo, em seu sentido amplo, da inovação como mecanismo de reprodução e ampliação do potencial social e econômico do País. Esse é o caminho para transformar a ciência, a tecnologia e a inovação em eixo estruturante do desenvolvimento brasileiro” (MCTI, 2012, p.12).

Há que se considerar, entretanto, a grande extensão territorial e as diversidades regionais do Brasil, que dificultam o estabelecimento de políticas adequadas às necessidades específicas das regiões. Dessa forma, muitas iniciativas de inovação ocorrem no âmbito dos estados, por meio de alianças entre as esferas federal e estadual de governo. O estado de Santa Catarina, por exemplo, conta com a Lei Catarinense de Inovação (SANTA CATARINA, 2008), que destina dois por cento da receita líquida do orçamento do estado à Fundação de Amparo à Pesquisa e Inovação do Estado de Santa Catarina (FAPESC) e à Empresa de Pesquisa Agropecuária e Extensão Rural de Santa Catarina (EPAGRI).

O art. 1º da Lei Catarinense de Inovação define:

“Esta Lei estabelece medidas de incentivo à pesquisa científica e tecnológica e à inovação no ambiente produtivo, visando à capacitação em ciência, tecnologia e inovação, o equilíbrio regional e o desenvolvimento econômico e social sustentável do Estado, em conformidade com os arts. 176 e 177 da Constituição do Estado de Santa Catarina” (SANTA CATARINA, 2008, p.1).

Em decorrência da atenção do estado à importância da inovação, foi desenvolvida a Política Catarinense Ciência, Tecnologia e Inovação (FAPESC, 2010), cujo objetivo é promover o avanço do conhecimento científico, tecnológico e de inovações no ambiente produtivo, nas instituições de ensino, pesquisa e extensão, nos agentes

econômicos e sociais e nos órgãos de governo, visando à qualidade de vida dos habitantes e ao desenvolvimento social e econômico do Estado de Santa Catarina, com sustentabilidade ambiental e equilíbrio regional (FAPESC, 2010).

Em função dos esforços já empreendidos, a produção científica brasileira apresentou um importante crescimento nos últimos anos. O número de doutores formados também cresceu, demonstrando o potencial do país em CT&I (CGEE, 2010). Contudo, embora a produção científica tenha saído de um patamar de fraco desempenho para números que colocam o Brasil em estágio intermediário de desenvolvimento científico e tecnológico, persiste um hiato em relação aos países desenvolvidos. O relativo atraso do país no registro de patentes nas instituições internacionais especializadas resulta da dissociação entre o avanço científico e a incorporação do progresso tecnológico à base produtiva (MCTI, 2012).

Uma das formas de orientar as políticas públicas é o desenvolvimento de indicadores que possam medir os resultados do processo de inovação. Zabala-Iturriagagoitia *et al.* (2007) afirmam que, embora muitos estudos tenham proposto metodologias para medir a capacidade de inovação, ainda havia muito trabalho a ser feito, uma vez que os estudos existentes, bem como as estatísticas disponíveis não levavam em conta os aspectos institucionais, interações, acordos de cooperação e etc. que são considerados elementos cruciais de um Sistema de Inovação (LUNDVALL, 1992; ETZKOWITZ; LEYDESDORFF, 2000).

Medidas comumente aplicadas, por serem consideradas *proxy* para a inovação, são as estatísticas de Pesquisa e Desenvolvimento (P&D) e de patentes, sob o argumento de que o aumento nos esforços de P&D em uma região resultam em uma taxa maior de invenção, assim como o de que um número maior de patentes leva a mais inovações em uma região.

Embora P&D e Patentes sejam considerados indicadores válidos para medir a inovação, tais relações lineares podem não ser necessariamente verdadeiras, visto que estes indicadores apresentam fragilidades (NELSON, 2009; SCHMOCH; GAUCH, 2009; MAKKONEN; HAVE, 2013). As estatísticas de P&D apresentam limitações na mensuração de inovação, uma vez que nem todos os esforços de P&D estão necessariamente relacionados a resultados de inovação comercializados ou de sucesso (NELSON, 2009; MAKKONEN; HAVE, 2013). Há, ainda, o entendimento de que P&D pode ser considerado

apenas um fator de entrada (*input*) da inovação (KLEINKNECHT; VAN MONTFORT; BROUWER, 2002; MAKKONEN; HAVE, 2013).

De forma semelhante, as estatísticas de patentes têm vulnerabilidades, uma vez que nem todas as invenções patenteadas tornam-se inovações, e o conteúdo e valor econômico das patentes pode diferir largamente entre elas. Além disso, alguns setores tecnológicos geram mais patentes do que outros, e o tamanho das empresas também tem influência em sua propensão de patentear (KLEINKNECHT; VAN MONTFORT; BROUWER, 2002; COAD; RAO, 2008; NELSON, 2009; GRUPP; SCHUBERT, 2010)..

Outra importante forma de mensurar a inovação é a quantificação direta dos resultados da inovação (MAKKONEN; HAVE, 2013). Tais informações são comumente obtidas por meio de questionários aplicados a empresas, por meio dos quais são explorados os fatores que influenciam o comportamento inovador da empresa, suas estratégias, incentivos e barreiras para a inovação, o escopo das várias atividades de inovação e, sobretudo, os resultados e efeitos da inovação. Esta abordagem é conhecida como *subject-approach*, uma vez que as próprias empresas respondentes são o alvo da pesquisa (OECD, 2005).

Embora a abordagem *subject-approach* seja uma prática largamente aplicada, algumas deficiências do método são apontadas na literatura. Dentre elas citam-se: a falta de objetividade nas respostas; a falta de entendimento do respondente em relação ao que é inovação; a existência de respostas estimadas; as baixas taxas de resposta; a heterogeneidade de valor econômico das inovações reportadas; o custo de produção das medidas de inovação; e a limitada disponibilidade de dados (ACS, 2002; KLEINKNECHT; VAN MONTFORT; BROUWER, 2002; RATANAWARAH; POLENSKE, 2007; EUROPEAN UNION, 2012).

A complexidade da inovação reflete-se, portanto, na dificuldade de mensurá-la. Como o entendimento de Schumpeter sobre a inovação não é dado em termos de formalismos matemáticos, tornou-se uma prática comum a combinação de diversos indicadores de ciência, tecnologia e inovação para formar números agregados ou compostos, normalmente apresentados na forma de um *ranking* ou normalizado em um intervalo fechado (GRUPP; SCHUBERT, 2010).

Diversas iniciativas tratam da definição de indicadores compostos para a inovação (OECD, 2005; EUROPEAN UNION, 2010; 2012; CORNELL UNIVERSITY; INSEAD; WIPO, 2013; IBGE, 2013), possibilitando estudos comparativos entre diferentes regiões. A dimensão regional revela-se importante devido às diferenças socioculturais, de atividades de inovação e de recursos humanos existentes entre diferentes regiões, e

estão relacionadas às características dos chamados Sistemas de Inovação (SI). Makkonen e Have (2013) observam, contudo, que a grande maioria dos trabalhos tratam do uso de indicadores compostos de inovação em nível das empresas ou em nível nacional, enquanto poucos são os trabalhos que tratam de indicadores para o nível regional. Além disso, muitos deles, não podem ser empiricamente verificados devido à falta de dados estatísticos em nível regional (ZABALA-ITURRIAGAGOITIA *et al.*, 2007).

No Brasil, uma das mais importantes iniciativas em termos de indicadores de inovação é a Pesquisa de inovação tecnológica (PINTEC) (IBGE, 2010), que tem o objetivo de fornecer informações para a construção de indicadores das atividades de inovação das empresas brasileiras. Sua referência conceitual e metodológica é baseada no Manual de OSLO (OECD, 2005), e no modelo *Community Innovation Survey* (CIS) (EUROSTAT, 2010), permitindo, assim, comparações com os resultados de estudos realizados em outros países.

Na visão de Makkonen e Have (2013), portanto, um dos maiores obstáculos envolvidos na avaliação do desempenho dos Indicadores de Inovação Compostos é a escassa disponibilidade de medidas de referência confiáveis de inovação. A falta de dados é uma referência encontrada em publicações que apresentam índices compostos, tais como o *Global Innovation Index* (CORNELL UNIVERSITY; INSEAD; WIPO, 2013), o *Regional Innovation Scoreboard* (EUROPEAN UNION, 2012), o *Community Innovation Survey*, o *European Competitiveness Index*, a metodologia japonesa NISTEP, o *European Innovation Survey* (ZABALA-ITURRIAGAGOITIA *et al.*, 2007).

Mesmo no caso do Manual de Oslo, que serve como base metodológica para coleta de dados sobre inovação em muitos outros trabalhos, inclusive há a referência sobre a dificuldade das empresas em informar um conjunto completo de dados, independente da abordagem de coleta utilizada (OECD, 2005).

Se, por um lado, a dificuldade de obtenção de dados é apontada como um empecilho na construção de índices de inovação compostos, há uma área de estudos que se preocupa justamente com a disponibilização de dados por parte de governos e organizações privadas, para livre consulta e utilização da sociedade. Tais iniciativas tratam do conceito de Dados Abertos, aqueles que qualquer um é livre para usar, reusar e redistribuir, estando sujeito a, no máximo, a exigência de crédito de sua autoria e de compartilhamento pela mesma licença (OPEN DEFINITION, 2009).

A abertura de dados é uma prática que vem sendo adotada pelos governos de diversos países, que passam a disponibilizar seus dados para livre consulta e utilização da sociedade. No Brasil, em função de leis específicas sobre o tema, os órgãos da administração pública passaram a ser obrigados a abrir seus dados na *Web*. Tal ação tem por objetivo a publicidade como preceito geral, a divulgação de interesse público, independente de solicitações, o desenvolvimento da cultura da transparência e do controle social da administração pública (BRASIL, 2011).

Ainda que algumas diretivas de publicação sejam especificadas em cartilhas governamentais e em páginas *Web* de órgãos que tratam de dados abertos, há pouca padronização no formato de publicação dos dados disponibilizados, o que dificulta o seu processamento de forma automatizada. Iniciativas como as de Berners-Lee (2006) e Bizer; Heath e Berners-Lee (2009), entretanto, apresentam uma alternativa na forma de publicação de dados na *Web*, com o objetivo de transformá-la de uma rede de dados interligados, por meio um conjunto de práticas caracterizadas como *Linked Data*.

Conforme observa Berners-Lee (2006), *linked data* é essencial para conectar a *web* semântica. A utilização das tecnologias semânticas faz com que seja possível navegação pelos dados, possibilitando a descoberta de outros conjuntos de dados, bem como o seu significado. Conseqüentemente, novos usos para os dados podem surgir, de acordo com novas necessidades dos usuários, à medida que a *Web* de Dados se desenvolve.

ADERÊNCIA AO EGC

Esta tese apresenta o desenvolvimento de um modelo de referência para indicadores da inovação regional suportado por dados ligados. Trata-se de um trabalho interdisciplinar, onde são abordadas as áreas de pesquisa relacionadas à gestão da inovação regional e seus indicadores, bem como as tecnologias computacionais associadas à *web* semântica e à utilização de dados abertos.

O trabalho enquadra-se na área de concentração de Engenharia do Conhecimento (EC) do Programa de Pós-Graduação em Engenharia e Gestão do Conhecimento. Mais especificamente, está contextualizado na linha de pesquisa denominada Teoria e Prática em Engenharia do Conhecimento, cujo objetivo é estudar metodologias e técnicas da Engenharia do Conhecimento e da Inteligência Computacional para a produção de plataformas de governo eletrônico.

A inovação ocupa papel importante no contexto da sociedade atual. Em uma economia onde as mudanças são constantes, a capacidade de adaptação é um fator crucial para a sobrevivência das organizações. Neste sentido, o conhecimento forma a base para o sucesso econômico e a inovação, e a Gestão do Conhecimento é a estratégia para a criação, o acesso e o apoio a este recurso vital (MEYER; SUGIYAMA, 2007).

Conforme afirmam (SCHREIBER *et al.*, 1999), a Engenharia do Conhecimento trata do desenvolvimento de sistemas de informação nos quais o conhecimento e o raciocínio têm papel fundamental. A contextualização desta tese na área de concentração da Engenharia do Conhecimento justifica-se, uma vez que propõe a construção de um modelo que, por meio de técnicas oriundas da Ciência da Computação e Sistemas de Informação, combinadas com as de gestão da inovação, permita a construção de painéis de indicadores de inovação a partir de dados abertos.

No âmbito do Programa de Pós-Graduação em Engenharia e Gestão do Conhecimento (PPEGC) o termo inovação aparece como objeto de estudo em vários outros trabalhos defendidos, sejam dissertações de mestrado ou teses de doutorado. O termo “Dados Ligados”, entretanto, não é encontrado como palavra-chave ou parte do título de nenhuma dissertação ou tese. Ainda, foram encontradas quatro teses que propõem modelos de referência ou ontologias de referência.

No contexto do processo de inovação, os trabalhos abordam a geração de ideias (PRADA, 2009; MIGUEZ, 2012; DOROW, 2013), a criação do conhecimento (MANHÃES, 2010; OROFINO, 2011), a aquisição do conhecimento (SCHMITT, 2013) e a proteção do conhecimento (MOTA, 2011).

O entendimento de que a inovação acontece em uma configuração de ambiente complexo é representado por trabalhos que tratam das instituições de ensino e pesquisa e sua importância para a inovação (DALL’AGNOL, 2010; GUBIANI, 2011; SILVA, 2011; TOSTA, 2012; CADORI, 2013), bem como das diferentes configurações, tais como Parques Científicos e Tecnológicos (GIUGLIANI, 2011), Sistemas de Inovação (SARTORI, 2011; LABIAK JUNIOR, 2012; MALDONADO, 2012) e Ecossistemas de Inovação (FIATES, 2014) abordam as instituições de ensino e pesquisa e sua potencial de contribuição no processo de inovação.

A utilização de técnicas semânticas associadas à inovação aparece no trabalho de Kinceler (2013), que apresenta um *framework* baseado em ontologia para apoio à Gestão Estratégica da Inovação.

Também são encontrados no EGC trabalhos que visaram o desenvolvimento de modelos de referência (CASTRO, 2010; BRITO, 2012; GUEMBAROVSKI, 2014) e de ontologia de referência (FACHIN, 2011), este último voltado à representação de periódico científico digital.

As próximas seções do trabalho apresentam os procedimentos metodológicos seguidos no desenvolvimento desta tese, bem como a estrutura deste documento.

1.3 PROCEDIMENTOS METODOLÓGICOS

Esta seção apresenta a classificação e os procedimentos metodológicos utilizados no desenvolvimento desta tese.

Quanto à sua natureza, esta tese caracteriza-se como uma pesquisa aplicada (tecnológica), definida por Gil (2008) como aquela que, embora dependa das descobertas e se enriqueça com o desenvolvimento da pesquisa básica, tem como característica fundamental o interesse na aplicação, utilização e consequências práticas dos conhecimentos e no desenvolvimento de artefatos. (CUPANI, 2006). Tais características aplicam-se a esta tese, uma vez que o trabalho nela descrito visa a proposição de um modelo que aplica de forma prática os conceitos abordados na literatura, fornecendo um artefato para a representação de indicadores de inovação regional, suportado pelas tecnologias semânticas que embasam os dados ligados.

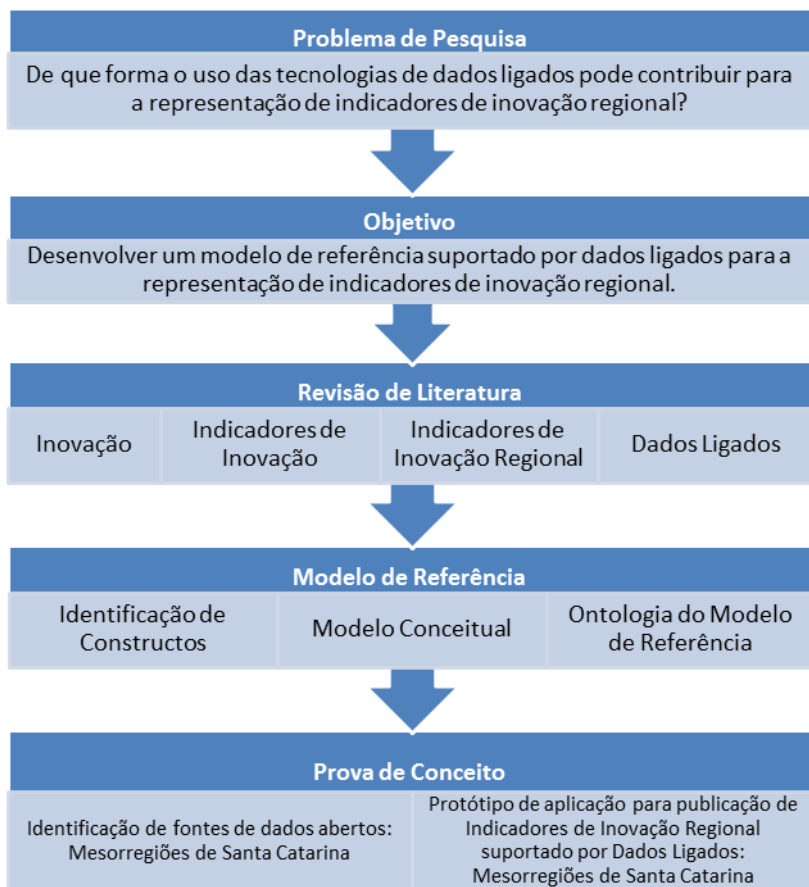
Quanto ao seu objetivo, caracteriza-se como uma pesquisa exploratória, aquela que tem como principal finalidade desenvolver, esclarecer e modificar conceitos e ideias, tendo em vista a formulação de problemas mais precisos ou hipóteses pesquisáveis para estudos posteriores (GIL, 2008). A tese explora a complexidade da inovação regional pelo viés da multiplicidade de indicadores e a falta de disponibilidade de dados sobre a inovação regional, conforme relatado na literatura. Ao trazer à discussão as iniciativas de dados abertos e as tecnologias semânticas, o modelo lança um olhar sobre a possibilidade de exploração destes dados enquanto subsídio para a avaliação da inovação regional.

Em relação à abordagem do problema, a pesquisa classifica-se como qualitativa, que é uma categoria caracterizada por considerar a relação dinâmica entre o mundo real e o sujeito, isto é, um vínculo indissociável entre o mundo objetivo e a subjetividade do sujeito, que não pode ser traduzido em números (MINAYO, 2007). A tese propõe um modelo para que se possa explorar o potencial de utilização dos dados

abertos em aplicações voltadas à mensuração da inovação regional. Embora trate da representação de indicadores de inovação, sua aplicação será realizada por meio do desenvolvimento de um ferramental de suporte a ações de gestão e avaliação da inovação regional.

A Figura 1 representa as etapas cumpridas no desenvolvimento deste trabalho.

Figura 1 - Procedimentos metodológicos do trabalho



Fonte: O autor

Para a seleção de bibliografia utilizada nesta tese foram realizadas buscas sistemáticas de literatura, que tiveram por objetivo a identificação dos principais autores e grupos de pesquisa nos temas, bem como verificação das referências mais significativas em termos de

aderência aos termos de busca e ao número de citações no corpo de literatura. Tais buscas, bem como outros trabalhos de pesquisa desenvolvidos no decorrer do desenvolvimento desta tese foram publicados, e estão listados no Quadro 1.

A partir do levantamento de literatura, foram identificados os trabalhos que apresentavam modelos de indicadores compostos aplicados à inovação regional e catalogados os principais constructos utilizados na classificação dos indicadores. Uma classificação em três níveis é proposta como modelo de referência para a representação de indicadores da inovação regional.

Quadro 1 - Publicações científicas do autor durante o desenvolvimento da tese

- FERREIRA, Odilon A.; RODRIGUES, Felipe M.; SPERONI, Rafael de Moura; GAUTHIER, Fernando A. Ostuni; TODESCO, José Leomar. Aplicando *Linked Open Data* na visualização de informações socio-econômicas e geográficas sobre os municípios brasileiros. VI Foro Científico Estudantil do Instituto Superior Tupy – FOCIEST. Joinville, SC, Brasil, 2012.
- ZANCANARO, Airton; DAL PIZZOL, Leandro; SPERONI, Rafael de Moura; TODESCO, José Leomar; GAUTHIER, Fernando A. Ostuni. Publishing Multidimensional Statistical Linked Data. eKNOW 2013, The Fifth International Conference on Information, Process and Knowledge Management. Nice, France, 2013.
- SILVA, Gesiel da; FRANZEN, Greici Baretta; TODESCO, José Leomar; GAUTHIER, Fernando A. Ostuni; SPERONI, Rafael de Moura. Processo de Publicação de Linked Data para a Distribuição de Informação e Conhecimento. III Congresso Internacional do Conhecimento e Inovação. Porto Alegre, RS, Brasil, 2013.
- VIDOTTO, Juarez D. F. ; SPERONI, Rafael de Moura ; NERES, Wudson A. ; BASTOS, Rogério Cid . Entrevista de desligamento como ferramenta de Gestão de Conhecimento. In: III Congresso Internacional do Conhecimento e Inovação, 2013, Porto Alegre. ANAIS DO III CONGRESSO INTERNACIONAL DE CONHECIMENTO E INVOAÇÃO (ciKi). Florianópolis: Universidade Federal de Santa Catarina, 2013. v. v. 3.
- VIDOTTO, Juarez D. F.; SPERONI, Rafael de Moura; NERES, Wudson A.; BASTOS, Rogério Cid . Entrevista de desligamento como ferramenta de Gestão de Conhecimento. Navus Revista de Gestão e Tecnologia, v. 3, p. 52-58, 2013.
- SPERONI, Rafael de Moura; TRINDADE, Evelin Priscila; KOSLOSKY, Marco A. Neiva; MACEDO, Marcelo; GAUTHIER, Fernando A. Ostuni. Avaliação da Produção Científica sobre *Enterprise Linked Data*. Linked Open

Data Brasil 2014. Florianópolis, SC, Brasil, 2014.

- SPERONI, Rafael M.; RAMOS, Alexandre Moraes; GAUTHIER, Fernando A. Ostuni; DA LUZ, Rafael Ramos. DIAS JÚNIOR, Claudelino Martins. Avaliação do ensino superior público no Brasil: protótipo de aplicação *Linked Data*. Linked Open Data Brasil 2014. Florianópolis, SC, Brasil, 2014.
- FREITAS FILHO, Fernando Luiz; SPERONI, Rafael de Moura; DANDOLINI, Gertrudes Aparecida; SOUZA, João Artur. Mensuração da Inovação: um estudo de caso em uma empresa no setor de eletrodomésticos. 3º Congresso de Engenharia de Produção da Região Sul. CONEPROSUL. Joinville, SC, Brasil, 2014.
- FREITAS JUNIOR, Vanderlei; WOSZEZENKY, Cristiane; ANDERLE, Daniel Fernando; SPERONI, Rafael de Moura; NAKAYAMA, Marina Keiko. A pesquisa científica e tecnológica. *Espacios (Caracas)* v.35 N°09, p. 52-58, 2013.
- KOSLOSKY, Marco Antônio Neiva; SPERONI, Rafael de Moura; GAUTHIER, Fernando A. Ostuni. Ecossistemas de Inovação – Uma revisão de literatura. *Espacios*. Vol 36 N°03, Caracas, 2015.
- TRINDADE, Evelin P.; SPERONI, Rafael de Moura; MACEDO, Marcelo; GAUTHIER, Fernando A. Ostuni; BOTELHO, Louise de Lira Roedel. Framework 4I's: a Bibliometric Review. Congresso Internacional de Conhecimento e Inovação – CIKI. Joinville, 2015.
- SPERONI, Rafael de Moura; DANDOLINI, Gertrudes; SOUZA, João Artur; GAUTHIER, Fernando A. Ostuni. ESTADO DA ARTE DA PRODUÇÃO CIENTÍFICA SOBRE INDICADORES E ÍNDICES DE INOVAÇÃO. RAI: revista de administração e inovação, Brasil, v. 12, n. 4, p. 49-75, dez. 2015.
- DAL PIZZOL, Leandro; SPERONI, Rafael de Moura; ZANCANARO, Airton; TODESCO, José Leomar; GAUTHIER, Fernando A. Ostuni. Análise bibliométrica da produção científica sobre Linked Data. *Informação & Informação*, [S.l.], v. 20, n. 3, p. 77 - 112, dez. 2015.
- SPERONI, Rafael de Moura; MACEDO, Marcelo; GAUTHIER, Fernando A. Ostuni. Revisão Sistemática de Literatura sobre Indicadores da Inovação Regional. *Espacios*. Vol 37, N° 05. Caracas, 2016.
- SPERONI, Rafael; MACEDO, Marcelo; GAUTHIER, Fernando A. Ostuni. Modelo de Referência para Indicadores de Inovação Regional. *Espacios*. Vol. 37, N° 10. Caracas, 2016.

Fonte: o autor

Visando ao atendimento do objetivo de que o modelo seja suportado por dados ligados, fez-se necessária a representação

semântica do modelo. Para tal, uma ontologia foi construída, cuja primeira versão foi desenvolvida segundo as etapas propostas no ontoKEM (TODESCO *et al.*, 2009), que aplica conceitos e etapas de diferentes metodologias de desenvolvimento de ontologias.

Finalmente, um protótipo foi desenvolvido como prova de conceito, visando à demonstração do potencial de aplicação do modelo proposto. Para isto, foram coletados dados sobre municípios de duas mesorregiões do estado de Santa Catarina, que foram mapeados e anotados semanticamente, e transformados, para um vocabulário de cubo multidimensional em RDF. A publicação destes dados seguiu os passos descritos em Zancanaro *et al.* (2013), onde é apresentado um processo de publicação de dados ligados estatísticos, fazendo uso do vocabulário *RDF Data Cube* (CYGANIAK; REYNOLDS; TENNISON, 2014).

O protótipo possibilita a visualização dos dados em formato tabular ou gráfico, possibilitando a análise comparativa entre diferentes regiões. A visualização gráfica pode ser feita segundo diferentes níveis de agregação regional e segundo diferentes níveis de agregação dos indicadores, conforme o modelo de referência. Um gráfico de radar é construído dinamicamente, sendo que cada eixo representa um indicador específico ou seu agrupador de classificação, e cada conjunto de dados exibido corresponde a uma região.

1.4 ESTRUTURA DO TRABALHO

Esta seção apresenta a forma como o trabalho está estruturado, ao longo de seus sete capítulos. O primeiro capítulo foi dedicado à introdução, tendo por objetivo a contextualização do tema, caracterização do problema, definição de objetivos, justificativa, aderência do assunto da tese ao programa de Pós-Graduação em Engenharia e Gestão do Conhecimento, e os procedimentos metodológicos.

O segundo capítulo aborda a Inovação Regional, apresentando conceitos e características da inovação e sua importância como peça chave para o desenvolvimento social e econômico. A seguir é trazido o conceito de Sistema de Inovação, um *framework* conceitual para o entendimento dos processos de inovação ao nível da economia, que quando aplicado em nível regional, é chamado de Sistema Regional de Inovação. São apresentados, ainda, os indicadores de inovação, e as considerações gerais acerca das principais abordagens.

O capítulo três é dedicado à apresentação das ontologias. É feita uma breve descrição da Engenharia do Conhecimento, a fim de contextualizar as ontologias como ferramentas para a modelagem de conhecimento. São abordadas, ainda, as classificações de ontologia, bem como algumas metodologias de desenvolvimento.

O capítulo quatro é apresenta aos Dados Ligados (*Linked Data*). A revisão de literatura passa pela conceituação e caracterização de Dados Abertos, apresentando iniciativas de publicação de dados abertos governamentais, responsáveis pela disponibilização de grandes volumes de conjuntos de dados na *Web* para livre acesso. São abordados a *Web Semântica* e suas tecnologias, bem como aquelas relacionadas aos Dados Ligados, que possibilitam a publicação de dados na *Web* utilizando padrões semânticos, com o objetivo de construção de um repositório global de dados.

O Capítulo cinco traz a contribuição mais significativa desta tese. Trata-se do desenvolvimento de um modelo de referência para indicadores de inovação regional suportado por dados ligados. Baseado nas teorias da inovação e Sistemas Regionais de Inovação, assim como nos modelos específicos de indicadores para a inovação regional, o modelo proposto baseia-se em um conjunto de constructos segundo os quais os indicadores estão organizados, e serve de base para a especificação de índices específicos de inovação regional que, fazendo uso da tecnologia de Dados Ligados, visa explorar a disponibilidade de dados regionais publicados na *Web*.

No capítulo seis, uma prova de conceito é apresentada. Um protótipo de aplicação *Web* é descrito, tendo por finalidade demonstrar a viabilidade de utilização do modelo. São apresentados os processos de coleta, transformação e publicação dos dados ligados, que, pelo vínculo estabelecido com o modelo proposto, são utilizados como insumos para a representação de indicadores da inovação regional. O protótipo conta com visualização gráfica que permite a análise comparativa entre regiões, segundo diferentes níveis de agregação regional e de indicadores.

Por fim, o sétimo capítulo, traz as considerações finais sobre o trabalho. São apresentadas, ainda, possibilidades e sugestões para trabalhos futuros que possam ser desenvolvidos a partir desta tese.

2 INOVAÇÃO REGIONAL

2.1 CONSIDERAÇÕES INICIAIS

O termo inovação desperta interesse nas mais variadas áreas de conhecimento, tendo se tornado uma palavra cuja ocorrência é bastante comum em discursos políticos ou mesmo em campanhas publicitárias. Isso ocorre em função da percepção de que a inovação tornou-se um fator chave para o crescimento e competitividade. Seja na esfera privada, onde a necessidade de superar a concorrência faz com que empresas busquem desenvolver novos produtos, ou fazê-los de forma diferente, ou na esfera pública, onde políticas públicas são propostas para criar um ambiente que potencialize as condições para a inovação.

A complexidade do fenômeno da inovação conduziu os estudiosos do assunto a uma visão sistêmica, que engloba os diferentes tipos de organizações, instituições e estruturas socioeconômicas que, juntos, direcionam e ditam o ritmo da inovação. Quando analisados sob o contexto de uma determinada região, caracteriza-se a estrutura do *framework* conceitual chamado Sistema Regional de Inovação (SRI) (DOLOREUX; PARTO, 2004).

Para criar as condições desejadas, a implantação de políticas públicas de inovação necessita de instrumentos que permitam a avaliação das condições estruturais do SRI em análise, de modo que seja possível direcionar os esforços e avaliar sua efetividade. Assim, diferentes modelos de indicadores são utilizados a fim de retratar a complexidade do processo de inovação.

Este capítulo apresenta as definições para a inovação, sua importância enquanto fator de desenvolvimento econômico regional. Traz, ainda, uma busca sistemática em bases científicas sobre o tema “Indicadores de inovação”, com o objetivo de analisar a literatura disponível sobre o assunto, bem como apresentar diferentes abordagens sobre as formas de medir a inovação.

2.2 O CONCEITO DE INOVAÇÃO

Embora receba muito destaque na literatura, o termo “inovação” é notoriamente ambíguo e falta uma definição ou medida única (ADAMS; BESSANT; PHELPS, 2006). A primeira conceituação para inovação é atribuída ao economista Joseph Alois Schumpeter, em seu livro “Teoria do Desenvolvimento Econômico”, publicado em 1934.

Em sua teoria sobre o desenvolvimento econômico, Schumpeter apresenta a ideia de “fluxo circular”, que considera uma economia perfeitamente competitiva, em equilíbrio estacionário, onde há equilíbrio competitivo, e que continua repetindo-se, onde os mesmos produtos são produzidos ao longo dos anos, da mesma maneira.

Para Schumpeter, o desenvolvimento econômico é a mudança espontânea e descontínua nos canais do fluxo circular, perturbação do equilíbrio, o que sempre altera e desloca o estado de equilíbrio anteriormente existente. (SCHUMPETER, 2012, p.64) Consiste na realização de novas combinações para as quais existam possibilidades no estado estacionário, e que acontecem na forma de inovações.

As definições posteriores costumam também estar associadas ao aspecto da novidade, salientando as saídas: um novo bem ou uma nova qualidade do bem, um novo método de produção, um novo mercado, uma nova fonte de abastecimento, ou uma nova estrutura organizacional, que podem ser sumarizadas como ‘fazer as coisas diferentes’ (HANSEN; WAKONEN, 1997).

A relação da inovação com o desenvolvimento econômico é apresentada por Schumpeter, quando este afirma que a inovação é responsável pela “destruição criativa”. Segundo o autor, é através deste fenômeno que são destruídos os modos como se fazia determinada atividade, revolucionando constantemente a estrutura econômica do sistema capitalista, de forma que novas firmas, novas tecnologias e novos produtos substituam constantemente os antigos (SCHUMPETER, 1961).

Para Schumpeter (2012), a inovação pode assumir várias formas, não sendo necessário que se invente algo novo, podendo mesmo submeter uma ideia já existente a uma nova forma de realizá-la ou uma nova situação. Mais especificamente, Crossan e Apaydin (2010) definem:

“Inovação é: produção ou adoção, assimilação e exploração de novidades de valor-agregado em esferas econômicas e sociais; renovação e ampliação de produtos, serviços e mercados; desenvolvimento de novos métodos de produção; e estabelecimento de novos sistemas de gestão. É tanto o processo quanto o resultado” (CROSSAN; APAYDIN, 2010, p. 02, tradução nossa).

O aspecto compartilhado na literatura, portanto, é de que inovação não requer necessariamente a criação de um produto novo. No Manual de Oslo (OECD, 2005), referência para a padronização de

conceitos, metodologias e construção de estatísticas e indicadores de Pesquisa e Desenvolvimento (P&D), define-se:

"Uma inovação é a implementação de um novo, ou significativamente melhorado, produto (bem ou serviço), ou processo, um novo método de marketing, ou novo método organizacional em práticas de negócio, organização do ambiente de trabalho, ou relações externas" (OECD, 2005, p. 46, tradução nossa).

Inovação pode significar uma nova e única aplicação de antigas tecnologias, usando projetos para o desenvolvimento de novos produtos e serviços, novos processos e estruturas para melhorar o desempenho em diversas áreas, criatividade organizacional e iniciativas de setor público para melhoria da entrega dos serviços. Em função deste entendimento, a inovação hoje está, cada vez mais, indo além dos limites da P&D formal para redefinir tudo (NATIONAL INOVATION COUNCIL, 2010).

Guo (2009) apresenta quatro características atribuídas à inovação:

- 1) O ponto fundamental da inovação é aumentar o valor por meio da mudança na função de produção e combinação dos meios de produção. Inovação tecnológica é um processo que envolve uma série de atividades de ciência, tecnologia e negócios que vão da geração de ideias inovadoras até a industrialização e comercialização dos resultados das pesquisas, enquanto que uma simples invenção técnica não se classifica como inovação.
- 2) Inovação é um projeto sistemático que está relacionado à completa interação entre a combinação de ciência, tecnologia e economia, inovação institucional, inovação de mercado e inovação de política e organização.
- 3) A inovação é caracterizada por sua fase, nível e campo específicos. O grau de inovação demonstra hierarquia e evolução com a mudança do contexto histórico e das condições.
- 4) O núcleo da geração de valor da inovação é satisfazer o cliente, perceber a colaboração da vantagem competitiva entre o baixo custo e diferenciação, e melhorar a atração competitiva.

A inovação, portanto, não é uma ação isolada, mas sim um processo total de subprocessos inter-relacionados (BAREGHEH; ROWLEY; SAMBROOK, 2009; QUINTANE *et al.*, 2011). Neste mesmo sentido, Trott (2012) afirma que a inovação é a gestão de todas as atividades envolvidas no processo de geração de ideias, desenvolvimento de tecnologia, fabricação e marketing de um novo

produto (ou aperfeiçoamento) ou de um processo de fabricação ou equipamento. O conhecimento, portanto, está fortemente ligado à inovação, pois a criação de novas possibilidades ocorre por meio da combinação de diferentes conjuntos de conhecimentos (DU PREEZ; LOUW; ESSMANN, 2006).

Nonaka e Takeuchi (1997) observam, ainda, que o conhecimento alimenta inovação, mas por si só não gera a inovação e nem vantagem competitiva. Para os autores, o conhecimento é um fator competitivo, enquanto que a capacidade de inovação de uma organização é uma consequência da aplicação desse conhecimento. A vantagem competitiva muda gradativamente em favor daquelas organizações que conseguem mobilizar conhecimento e avanços tecnológicos e conceber a criação de novidades. O conhecimento em questão pode ser explícito, de forma que outros possam acessá-lo, discuti-lo, transferi-lo, etc. - ou pode ser tácito: conhecido, mas sem formulação (TIDD; BESSANT; PAVIT, 2008).

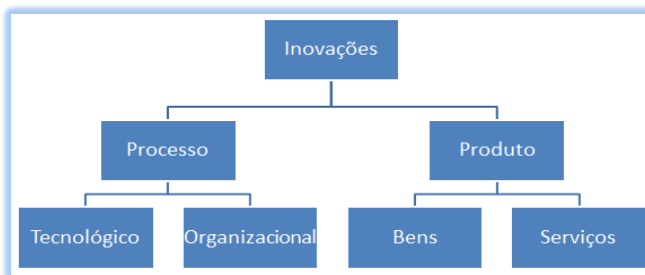
Em sua forma mais simples, o processo de inovação pode ser visto como gerador de novos conhecimentos (educação e treinamento) e tecnologias (desenvolvimento e comercialização), que vai da pesquisa básica até o mercado (WANG, 2010). Um processo de inovação é caracterizado pela criação do conhecimento necessário para entender como a inovação foi gerada, e que este conhecimento deve ser duplicável, considerado novo no contexto em que é introduzido, e útil na prática (Quintane et al., 2011).

2.3 CLASSIFICAÇÕES DE INOVAÇÃO

As classificações da inovação variam dependendo do contexto e dos autores. A seguir são apresentadas as classificações quanto à área de negócio impactada, quanto ao grau de impacto provocado, e quanto ao grau de controle que a empresa exerce sobre o processo.

Quanto à área de negócio impactada, Edquist (2001) apresenta uma taxonomia, representada na Figura 2, onde divide as inovações entre “Processo” e “Produto”, e considera que as inovações de Produto podem ser “Bens” ou “Serviços”, e diz respeito a “o que” está sendo produzido. As inovações de Processo, por sua vez, podem ser “Tecnológicas” ou “Organizacionais”, e dizem respeito a “como” os bens e serviços estão sendo produzidos. Nesta taxonomia, apenas “bens” e “processos tecnológicos” são inovações do tipo material. Os “processos organizacionais” e os “serviços” são intangíveis.

Figura 2 – Uma taxonomia de Inovações



Fonte: Edquist (2001, p.7)

O Manual de Oslo, por sua vez, apresenta quatro tipos de inovação, envolvendo uma ampla gama de mudanças nas atividades das organizações: Inovação de produto; Inovação de processo; Inovação em marketing; e, Inovação organizacional (OECD, 2005).

- **Inovação de Produto:** introdução de um bem ou serviço que seja novo ou significativamente melhorado com respeito a suas características ou usuários pretendidos. Isto inclui melhorias significativas em especificações técnicas, componentes e materiais, software incorporado, facilidade de uso ou características funcionais. Inovações de produto podem utilizar novos conhecimentos ou tecnologias, ou podem ser baseadas em novos usuários ou combinações de conhecimentos ou tecnologias existentes.
- **Inovação de Processo:** implementação de novos, significativamente melhorados, métodos de produção e de entrega. Isto inclui mudanças significativas em técnicas, equipamentos e/ou software, e pode ter o objetivo de diminuição nos custos de produção ou entrega, de melhoria de qualidade, ou de produção ou entrega de produtos novos ou significativamente melhorados. Exemplos de novos métodos de produção são a implementação de novos equipamentos de automação em uma linha de produção ou a implementação de projeto introdução de um sistema de rastreamento por códigos de barra, ou RFID assistido por computador para o desenvolvimento de produtos.
- **Inovações de marketing:** implementação de um novo método de marketing, tais como mudanças no design e embalagem, em promoções e localização de produtos, e nos métodos de atribuição de preço a produtos e serviços. Este tipo de inovação visa um melhor atendimento das necessidades dos clientes, a abertura de

novos mercados, ou um novo posicionamento dos produtos da empresa no mercado, com o objetivo de aumentar as vendas.

- **Inovação organizacional:** refere-se à implementação de novos métodos nas práticas de negócios, na organização do ambiente de trabalho ou nas relações externas da empresa. Estas inovações podem pretender o aumento do desempenho da empresa pela redução de custos administrativos ou custos de transações, o aumento na satisfação com o ambiente de trabalho (e da produtividade), a melhoria no acesso a conhecimentos externos ou a redução de custos de abastecimento.

Ainda em termos da área de negócio impactada, Tidd; Bessant e Pavit (2008) propõem uma classificação conhecida como 'os 4Ps da inovação', bastante próxima àquela proposta no Manual de Oslo: Inovação de Produto; Inovação de Processo; Inovação de Posição; e, Inovação de Paradigma.

- **Inovação de Produto:** refere-se às mudanças em produtos/serviços que a empresa oferece.
- **Inovação de Processo:** refere-se às mudanças na forma como os produtos/serviços são criados e entregues;
- **Inovação de Posição:** refere-se às mudanças no contexto em que os produtos/serviços são introduzidos no mercado.
- **Inovação de Paradigma:** refere-se às mudanças nos modelos mentais que orientam o que a empresa faz.

Uma segunda dimensão de classificação da mudança leva em conta o grau de novidade envolvido. Há diferentes graus de novidade desde melhorias menores até mudanças realmente radicais que transformam a forma como vemos ou usamos as coisas (BESSANT; TIDD, 2009).

Uma inovação é considerada **incremental** quando consiste de pequenas melhoras em produtos, serviços ou processos já existentes. Inovação **radical**, por sua vez, é aquela em que ocorrem mudanças significativamente diferentes em produtos, serviços ou processos. Para Tidd; Bessant e Pavit (2008), estas mudanças ocorrem em cada um dos "4 Ps", criando um espaço potencial para a inovação.

Finalmente, uma inovação **descontínua** é uma inovação radical em que as "regras do jogo" são modificadas, instaurando-se novas regras, onde novos participantes normalmente levam vantagem. Este é um tema central na teoria original de Schumpeter sobre a inovação, que

ele acreditava envolver um processo de "destruição criativa" (TIDD; BESSANT; PAVIT, 2008).

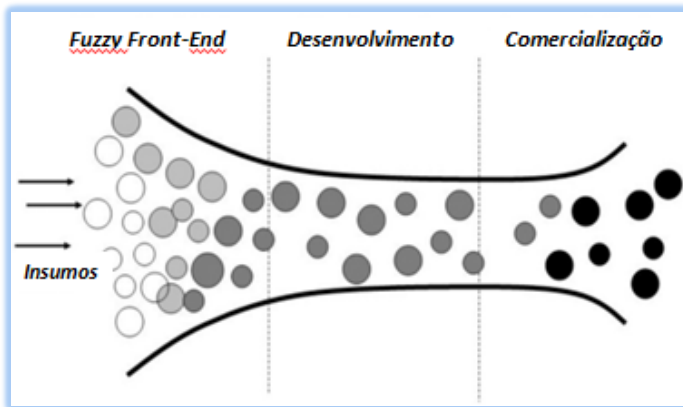
Bessant e Tidd (2009) classificam, ainda, a inovação segundo o grau de controle que a empresa exerce sobre o seu processo. Neste sentido consideram a existência de **inovação fechada**, quando o desenvolvimento é interno, e **inovação aberta** quando se utiliza de caminhos internos e externos para avançar no desenvolvimento de novas tecnologias.

Na chamada inovação fechada, a filosofia é de que sucesso em inovação requer controle. Esta foi a prática comum na indústria durante muitos anos, sob o entendimento de que o investimento interno em P&D era a melhor maneira levar novas ideias ao mercado. As empresas deveriam gerar suas próprias ideias, desenvolvê-las, produzi-las, lançá-las no mercado e distribuí-las (CHESBROUGH, 2003). Os lucros resultantes deveriam, então, ser reinvestidos em P&D, em um ciclo virtuoso de inovação (DOCHERTY, 2006).

No modelo fechado tradicional, as entradas ao modelo vêm tanto de fontes internas quanto externas (informações de clientes, mercado ou de planejamento estratégico) e, com base nelas, as organizações de P&D desempenham suas tarefas de inventar, evoluir ou aperfeiçoar tecnologias para o desenvolvimento imediato ou para serem deixadas "na prateleira" para o desenvolvimento posterior. A representação tradicional deste modelo é feita por meio da analogia do funil de inovação, conforme Figura 3, onde muitos conceitos internos são reduzidos para os que melhor se adequam às necessidades da empresa no momento. O foco está no desenvolvimento interno de tecnologias e produtos para comercialização interna (DOCHERTY, 2006).

Este modelo de inovação, entretanto, passou a sofrer alterações na medida em que o mercado tornou-se mais competitivo e o ciclo de vida dos produtos ficou mais curto. Tais fatores passaram a dificultar os altos investimentos em P&D, dando espaço a um modelo de inovação aberta, onde a fronteira entre a organização e o ambiente à sua volta é porosa, possibilitando uma maior mobilidade das inovações entre os dois (CHESBROUGH, 2003; 2007).

Figura 3 - Modelo de Inovação fechada - Funil da Inovação



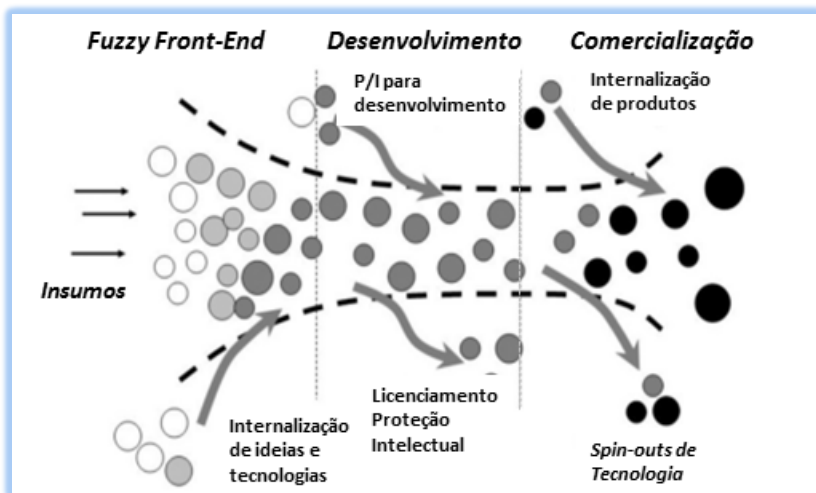
Fonte: adaptado de Docherty (2006)

O processo da inovação tornou-se cada vez mais aberto graças ao crescimento da divisão de trabalho, onde uma companhia desenvolve uma ideia, mas em vez de trazê-la ao mercado, decide associar-se ou vender a ideia para outra companhia que vai comercializá-la, e isso implica em uma mudança nos seus modelos de negócios (Chesbrough, 2007). Para Traitler; Watzke e Saguy (2011), a inovação aberta baseia-se na realidade de que, em um mundo de conhecimento vastamente distribuído e taxa de desenvolvimento acelerado, as companhias não podem mais se dar ao luxo de confiar em sua própria pesquisa e, conseqüentemente, deve utilizar recursos externos e comprar licenças, tecnologias, invenções e soluções.

Em uma definição mais explícita, Chesbrough (2003) afirma que inovação aberta é um paradigma que assume que as empresas podem e devem usar ideias externas assim como internas, e caminhos internos e externos para o mercado, na medida em que buscam avançar suas tecnologias. Em uma redefinição, inovação aberta refere-se ao uso vantajoso de entradas e saídas de conhecimento para acelerar a inovação interna e expandir os mercados para uso externo das inovações, respectivamente (CHESBROUGH; VANHAVERBEKE; WEST, 2006).

O modelo de inovação aberta trata-se, portanto, de uma abordagem mais dinâmica e menos linear da gestão da inovação, no qual as companhias procuram de dentro para fora e de fora para dentro nos três estágios da cadeia de inovação (*Fuzzy Front-End*, Desenvolvimento e Comercialização), criando maior valor ao longo do processo. A Figura 4 representa o funil da inovação no modelo de inovação aberta.

Figura 4 - Modelo de Inovação Aberta - Funil da Inovação



Fonte: Docherty (2006)

No *Fuzzy Front-End*, as empresas não apenas estão buscando externamente por problemas a serem resolvidos, mas agora também inventores, *startups* e outras fontes de tecnologias disponíveis que possam ser usadas como base para um desenvolvimento interno ou conjunto (DOCHERTY, 2006).

Na fase de desenvolvimento, as companhias já estabelecidas podem adquirir inovações externas que já tenham vindo a se tornar produtos ou até comercializadas, mas agora oferecem a oportunidade de desenvolver novas gerações da tecnologia para escala ampliada. Nesta fase as companhias podem também transferir tecnologias e propriedades intelectuais que foram desenvolvidas internamente, mas que são consideradas fora da atividade principal, e melhor desenvolvidos e comercializados por outros (DOCHERTY, 2006).

Na fase de comercialização, as companhias podem transferir tecnologias que já foram comercializadas, mas que podem ter seu valor aumentado em outros lugares, ou então adquirir linhas de produtos já comercializadas ou negócios que possam prover fontes imediatas de novo crescimento para a companhia (DOCHERTY, 2006).

A mudança de paradigma para o da inovação aberta traz como benefício um banco de ideias e tecnologias muito maior para impulsionar o crescimento interno. Além disso, empresas líderes passam a ter uma ferramenta estratégica que permite explorar novas

oportunidades de crescimento com um risco mais baixo. Ainda, por meio da transferência de ideias não utilizadas, além de capturar valor econômico destas, cria-se um importante “senso de urgência” pelos grupos internos para “usá-las ou perde-las”, em relação às tecnologias disponíveis internamente (DOCHERTY, 2006).

Devido às suas características, a inovação aberta tem o potencial de impulsionar o desenvolvimento de tecnologias e parcerias entre academia, setor privado e governo. Indivíduos e regiões inteiras do mundo que, historicamente, nunca iriam interagir, podem agora colaborar em tempo real em um projeto de P&D. Estes sistemas de inovação aberta podem ser utilizados para ampliar as colaborações entre a infraestrutura interna de P&D e a universidade, tendo seus efeitos amplificados pelo aumento do grupo de indivíduos disponíveis para abordar uma dada questão, e pelo decréscimo de tempo para desenvolver soluções, levando a um menor tempo até comercialização (WANG, 2010).

Diante deste cenário, grandes empresas como a P&G, Kraft Foods e Nestlé já implantaram modelos de inovação aberta, obtendo resultados importantes em termos de aquisição de inovações externas, transformando-as em sucesso comercial (KUHN, 2008; TRAITLER; WATZKE; SAGUY, 2011). A adoção de modelos de inovação aberta em pequenas e médias empresas, entretanto, costuma ser mais difícil em função de seus altos custos. Alguns dos custos estão associados ao investimento em recursos humanos necessários para a avaliação, seleção, e negociação com os parceiros externos, além das mudanças organizacionais necessárias para a implantação (TRAITLER; WATZKE; SAGUY, 2011).

Embora ainda seja necessário que as companhias construam sua capacidade interna de P&D, as fontes externas tornam-se cada vez mais importantes, devido à demanda por resposta rápida à mudança e alto nível de especialização. Universidades e instituições de pesquisa revelam-se importantes atores em um sistema dinâmico onde pessoas e instituições devem estar interconectadas para gerar o desenvolvimento tecnológico e econômico.

2.4 SISTEMAS DE INOVAÇÃO

É cada vez mais reconhecido que importantes elementos do processo de inovação tornaram-se regionalizados. Conforme observam Doloreaux e Parto (2004), a concepção da inovação como um fenômeno parcialmente territorial deve-se, em grande parte, aos casos de sucesso

de aglomerações industriais especializadas ou de redes regionalmente concentradas de pequenas e médias empresas e agrupamentos industriais.

A abrangência do que se considera uma região depende do caso em estudo, mas refere-se ao local de interação econômica e inovação. A escala, ou unidade de análise utilizada nos estudos sobre inovação regional varia entre cidades, regiões metropolitanas, unidades territoriais, escalas suprarregionais ou subnacionais (DOLOREAUX; PARTO, 2004).

A necessidade de interação e colaboração entre diferentes tipos de entidades em um processo de inovação fez com que o ambiente em que estas se inserem passasse a fazer parte da análise. A inovação acontece em arranjos de cooperação entre diferentes partes envolvidas, incluindo companhias privadas, organizações públicas, instituições de pesquisa e clientes (JALONEN, 2013).

Quando se passa a considerar a inovação sob o ponto de vista das relações entre os atores nela envolvidos e o ambiente em que se inserem, passa-se a uma visão sistêmica da inovação. O mais prevalente *framework* conceitual para o entendimento dos processos de inovação ao nível da economia é o conceito de Sistemas de Inovação (SI) (MAHROUM; ALSALEH, 2013). O conceito de Sistemas de Inovação foi originado dos trabalhos de Christopher Freeman e Bengt-Ake Lundvall, no final da década de 1980 e início da década de 1990, quando desenvolveram a ideia da inovação como um processo iterativo.

Segundo a definição de (FREEMAN, 1987), um Sistema de Inovação é a rede de instituições do setor público e privado, cujas atividades e interações iniciam, importam, modificam e difundem novas tecnologias. Lundvall (1992) acrescenta que um Sistema de Inovação caracteriza-se pelos elementos e relacionamentos que interagem na produção, difusão e uso de conhecimentos novos e economicamente úteis, estejam ou não localizados dentro dos limites de um estado-nação. Esta definição já traz a referência ao que se chamará de Sistema Nacional de Inovação (SNI).

Posteriormente, o conceito de SNI é complementado por Lundvall:

“O Sistema Nacional de Inovação é um sistema aberto, complexo e em evolução, que abrange relações dentro e entre organizações, instituições e estruturas socioeconômicas que determinam o ritmo e a direção da inovação e do desenvolvimento de competências que emana dos

processos de aprendizado baseados em ciência e em experiência” (LUNDVALL *et al.*, 2009, p.6, tradução nossa).

Os sistemas de inovação costumam ser classificados e explicados na literatura por meio de uma abordagem de componentes, onde busca-se mapear os componentes de um sistema, ou por uma abordagem funcional, segundo a qual são avaliadas as funções e resultados básicos do sistema de inovação (LUNDVALL *et al.*, 2009).

Embora a literatura não apresente uma caracterização rígida em termos dos componentes de um sistema de inovação, algumas obras trazem contribuições no sentido de explicação de sua estrutura fundamental. No trabalho “*Systems of Innovation: Perspectives and Challenges*”, Edquist (2009) faz uma espécie de glossário com os termos principais utilizados na teoria de Sistemas de Inovação, e considera que os constituintes de um SI são seus **componentes** e as **relações** entre eles, sendo que os principais componentes de um SI são as **Organizações** e as **Instituições**.

Organizações são as estruturas formais, que são conscientemente criadas e têm objetivo explícito, sendo chamados de **players** ou **atores**. Algumas das organizações importantes são as empresas, universidades, organizações de *venture capital* e agências públicas responsáveis por políticas de inovação, competição ou regulação de drogas.

Instituições são os conjuntos de hábitos comuns, normas, rotinas, práticas estabelecidas, regras ou leis que regulam as relações e interações entre os indivíduos, grupos e organizações. Tratam-se das regras do jogo. Exemplos de instituições importantes são as leis de patentes, bem como as regras e normas que influenciam as relações entre universidades e empresas (EDQUIST, 2009).

Um Sistema de Inovação tem, ainda, uma **função**, isto é, está desenvolvendo ou buscando atingir algum objetivo. A principal função de um SI é atingir um processo de inovação, ou seja, desenvolver, difundir e usar inovações. As **atividades** em Sistemas de Inovação são os fatores que influenciam no desenvolvimento, difusão e uso das inovações, e são os determinantes da função principal.

Quanto à identificação das organizações (atores) envolvidos em um Sistema de Inovação, há uma referência bastante comum na literatura sobre uma formação tríplice entre universidades, empresas e governo. Etzkowitz e Leydesdorff (2000) lembram que a evolução dos sistemas de inovação e o conflito sobre qual o caminho a ser seguido nas relações universidade-indústria refletem-se nos arranjos institucionais

das relações universidade-indústria-governo. Os autores trazem três configurações que refletem situações históricas do modelo de “Tríplice Hélice” (ETZKOWITZ; LEYDESDORFF, 1995).

Na primeira configuração, o Estado representa o principal ator, englobando a academia e a indústria e direcionando as relações entre eles. Esta configuração era encontrada, ainda que de maneira mais fraca, nas políticas de muitos países latino-americanos. Devido ao reduzido espaço para iniciativas *bottom up* na primeira configuração, a inovação é desencorajada, e o modelo é visto por muitos como um modelo falho (ETZKOWITZ; LEYDESDORFF, 2000).

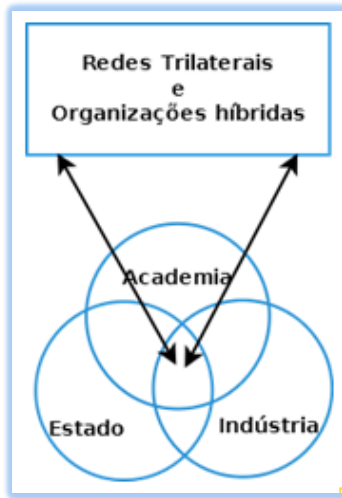
O segundo modelo consiste de esferas institucionais separadas com fortes barreiras separando-as e com relações fortemente circunscritas entre as esferas. Esta configuração implica em uma política de *laissez-faire*, defendida como uma terapia de choque para reduzir o papel do estado na primeira configuração (ETZKOWITZ; LEYDESDORFF, 2000).

Finalmente, a terceira configuração da Hélice Tríplice trata da geração de uma infraestrutura de conhecimento pela sobreposição das esferas institucionais, onde cada um faz o papel do outro, de onde emergem organizações híbridas e interfaces, conforme a Figura 5. Os arranjos são encorajados, mas não controlados, pelo governo que, por meio de "regras do jogo" ou financiamento direto ou indireto promove a inovação (ETZKOWITZ; LEYDESDORFF, 2000).

A base de conhecimento é um resultado dos fluxos de comunicação e informação estabelecidos, restritos e habilitados pelas redes formadas entre os atores (PARK; HONG; LEYDESDORFF, 2005). Muitos países e regiões tentam se adequar a esta realidade por meio da criação de ambientes de inovação que consistem de empresas *spin-off*¹, iniciativas tri-laterais para o desenvolvimento da economia baseada em conhecimento, alianças estratégicas entre empresas, laboratórios públicos e grupos de pesquisa acadêmicos (ETZKOWITZ; LEYDESDORFF, 2000).

¹ O termo *spin-off* caracteriza o processo de geração de novas empresas a partir de organizações, sejam elas empresas ou centros de pesquisa como universidades, laboratórios e institutos (CARAYANNIS *et al.*, 1998).

Figura 5 - O modelo da Trílice Hélice



Fonte: Etzkowitz e Leydesdorff (2000)

Uma consideração importante acerca da Hélice Tríplice é que ela não é estável. Etzkowitz e Leydesdorff (2000) afirmam que a evolução cultural é impulsionada pelos indivíduos e grupos, que tomam decisões conscientes, bem como pelo aparecimento de consequências inesperadas. Assim, as fontes da inovação em uma configuração de Trílice Hélice não são mais sincronizadas a priori, não aparecem em uma ordem previamente definidas, e geram quebra-cabeças para os participantes, analistas e formuladores de políticas resolverem.

O *OECD Science, Technology and Industry Outlook* (OECD, 2014) apresenta uma abordagem para a composição e funcionamento dos sistemas de inovação, que corrobora com a ideia da trílice hélice, por considerar que há uma inter-relação entre Universidades e Instituições Públicas de Pesquisa, Setor de Negócios e Governo, conforme a Figura 6.

Figura 6 - Atores de um Sistema de Inovação



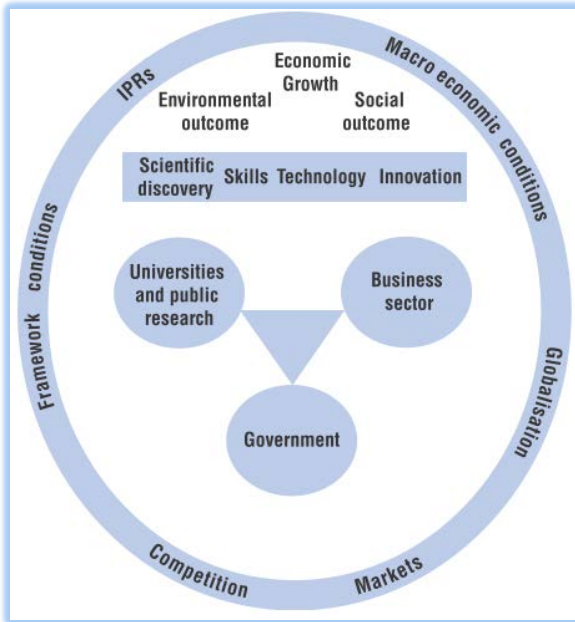
Fonte: Oecd (2014)

O Setor de Negócios responde pela maior parte da pesquisa e desenvolvimento domésticos em muitos países. As empresas são consideradas os principais atores da inovação. Particularmente, as *start-ups* podem explorar conhecimentos não utilizados ou subutilizados e conduzir à emergência de novos mercados.

As Universidades e Instituições Públicas de Pesquisa, por sua vez, desempenham um papel central nos sistemas de inovação, uma vez que provêm novos conhecimentos, especialmente em áreas em que os benefícios econômicos são incertos ou menos imediatos. Além disso, as universidades contribuem para a formação de habilidades e podem inspirar jovens talentosos a adentrarem à carreira de pesquisadores ou de ocupações relacionadas à inovação.

O terceiro principal tipo de ator é o Governo. Primeiro, porque há espaço para a inovação na administração pública, porque é necessário melhorar a entrega dos serviços públicos, seja em termos de seus conteúdos, ou em termos de seus instrumentos, para atender a uma demanda pública crescentemente sofisticada e às crescentes pressões fiscais. Em segundo lugar, porque as decisões de investimento das empresas respondem aos incentivos econômicos e, portanto, às políticas e instituições, e sua propensão à inovação depende do direcionamento eficiente das políticas e iniciativas. Finalmente, porque a agenda pública de pesquisa é projetada em alto nível e o orçamento público de pesquisa formata o cenário de pesquisa nacional (OECD, 2014).

Figura 7 - Componentes de um Sistema de Inovação



Fonte: Oecd (2014)

A Figura 7 apresenta a visão geral dos componentes de um Sistema de Inovação, segundo a abordagem utilizada pela (OECD, 2014). O modelo considera que os atores, que geram e fazem uso do conhecimento, dependem do acesso a: **Desenvolvimento científico; Habilidades; Tecnologia; e Inovações.**

Além disso, são consideradas as condições relativas ao contexto do sistema de Inovação: **Condições do Framework; Direitos de Propriedade Intelectual; Condições Macroeconômicas; Globalização; Mercados; e Competição.**

Como resultados das atividades de inovação, são apontados o **crescimento econômico**, e os **efeitos ambientais e sociais**, uma vez que a inovação deve beneficiar as pessoas e o planeta (OECD, 2014).

2.4.1 Sistemas Regionais de Inovação

O termo Sistemas Regionais de Inovação (SRI) deriva dos SNI, e sua aplicação visa avançar no entendimento do processo de inovação

regional. SRIs são normalmente entendidos como um conjunto de interesses públicos e privados, instituições formais e outras organizações que funcionam de acordo com arranjos institucionais e relações favoráveis para a geração, uso e disseminação do conhecimento (DOLOREUX; PARTO, 2004).

O conceito de SRI foi apresentado por Cooke (1992), que considera que as políticas regionais alavancam a inovação e a competitividade econômica e social.

Doloreaux e Parto (2004) afirmam, ainda, que o sistema regional de inovação é identificado por uma seleção de indicadores-chave em vários aspectos da capacidade organizacional e de infraestrutura, competências e aptidões para a inovação em regiões.

Diferentes nomenclaturas são encontradas na literatura para caracterizar tais arranjos. Sejam chamadas de “redes de inovação”, “clusters de inovação”, “sistemas regionais de inovação” ou “ecossistemas de inovação”, a justificativa para esta cooperação é de que, na medida em que o mundo tornou-se mais complexo, as organizações foram obrigadas a desenvolver novos rumos de ação (JALONEN, 2013).

O sistema dinâmico de instituições e pessoas interconectadas necessárias para impulsionar o desenvolvimento tecnológico e econômico inclui uma gama de atores que compreende academia, indústria, fundações, organizações científicas e econômicas e governos em todos os níveis (WANG, 2010). Wang (2010) afirma, ainda, que a relação universidade-indústria é de grande importância, uma vez que são as universidades que provêm tanto a força de trabalho altamente qualificada, quanto os resultados das pesquisas que a indústria poderá optar por comercializar e levar ao mercado.

Neste mesmo sentido, e complementando a teoria da tríplice hélice, Labiak Junior (2012) observa que as relações regionais de caráter inovador existentes no Brasil são geradas por seis atores principais, formando o que o autor denomina de “hélice sêxtupla”. Esta hélice seria, portanto, composta por atores empresariais, de conhecimento científico e tecnológico, públicos, institucionais, habitats de inovação e de fomento.

As diferentes formas de organização, tipos de atores e suas inter-relações indicam que, embora os papéis desempenhados pelos diferentes atores não sejam rigidamente planejados, a inovação regional não acontece de forma isolada. Ao contrário disso, a existência de um Sistema Regional de Inovação leva em conta um conjunto de fatores, de tal forma que a gestão e avaliação da inovação é uma tarefa complexa.

2.5 INDICADORES DE INOVAÇÃO

Um **indicador** é uma medida, de ordem quantitativa ou qualitativa, derivada de uma série de fatos observados, que possui significado particular e é utilizada com a finalidade de captar informações relevantes dos elementos que compõem o objeto da observação (NARDO *et al.*, 2005; FERREIRA; CASSIOLATO; GONZALEZ, 2009). São, portanto, ferramentas constituídas de variáveis que, associadas em diferentes configurações, expressam significados mais amplos sobre os fenômenos a que se referem (IBGE, 2008).

Saisana e Tarantola (2002) observam, ainda, que indicadores são frequentemente um compromisso entre a precisão científica e as informações disponíveis a um custo razoável. Um **indicador composto** é formado quando indicadores individuais são compilados, combinados ou agregados em um único **índice**, tendo por base um modelo subjacente (SAISANA; TARANTOLA, 2002; NARDO *et al.*, 2005).

Considerando a atenção que vem sendo dispensada à inovação devido ao entendimento de sua importância para o desenvolvimento econômico de uma empresa, região ou país, muitos esforços são feitos no sentido de que se possa medi-la. O levantamento de literatura apresentado por Speroni *et al.* (2015) demonstra a tendência de crescimento no número de artigos científicos publicados sobre o tema nos últimos anos, bem como a distribuição dos pesquisadores ao redor do mundo.

Para Smith (2005), embora exista o argumento de que a inovação seja inerentemente impossível de quantificar e medir, suas características gerais não impossibilitam a mensuração de dimensões-chaves de seus processos e saídas. Entretanto, devido à sua complexidade, um grande desafio para os pesquisadores é compreender os determinantes da inovação (ALCAIDE-MARZAL; TORTAJADA-ESPARZA, 2007; MAKKONEN; HAVE, 2013).

Smith (2005) acredita que uma dificuldade imediata seja o fato de que inovação envolve novidades multidimensionais em aspectos de aprendizado ou organização do conhecimento, que são difíceis de medir ou intrinsecamente não mensuráveis. Assim, o problema-chave em indicadores de inovação diz respeito à conceituação subjacente do objeto a ser medido, o significado do conceito medido, e a viabilidade de diferentes tipos de medidas.

Não há, portanto, um consenso sobre qual indicador deve ser usado para mensurar inovação. O tipo de variável escolhida pode afetar

as conclusões de pesquisa, ou as informações nas quais os decisores políticos baseiam suas decisões (MAHROUM; ALSALEH, 2013; MAKKONEN; HAVE, 2013).

Smith (2005) aponta os principais indicadores utilizados em análises de inovação: em primeiro lugar, os dados de P&D; em segundo os de dados de pedidos, concessões e citações de patentes; em terceiro os dados bibliométricos, ou seja, dados sobre publicações científicas e citações. O autor acrescenta, ainda, outras importantes classes de indicadores:

- Indicadores tecnométricos, que exploram as características técnicas de performance dos produtos.
- Indicadores sintéticos, normalmente desenvolvidos por consultores, para fins de *scoreboards*.
- Bases de dados sobre tópicos específicos, tais como ferramentas de pesquisa para grupos individuais.

Estatísticas de patentes e dispêndios em P&D são informações comumente associadas à mensuração da inovação. O argumento básico para esta associação é que um aumento nos investimento em P&D (fator de entrada), bem como um maior número de patentes (fator de saída) em uma região resultarão em uma maior taxa de invenção. Para Makkonen e Have (2013), contudo, esta relação direta é ilusória, e ambas as medidas, tal como acontece com outras frequentemente adotadas, como licenças, publicações científicas, serviços ou marcas registradas e modelos de utilidades, apresentam suas deficiências.

Makkonen e Have (2013) apresentam uma discussão as forças e fraquezas de indicadores individuais comumente utilizados na mensuração da inovação: Pesquisa e Desenvolvimento (P&D), patentes, despesas totais com inovação, e participação em vendas de produtos inovadores.

As seguintes seções apresentam as características dos principais indicadores utilizados para a mensuração da inovação.

2.5.1 Estatísticas de Pesquisa & Desenvolvimento

Dentre os indicadores de inovação, os esforços em P&D figuram entre os mais populares (SMITH, 2005; MAKKONEN; HAVE, 2013). Normalmente as medidas costumam estar associadas aos investimentos em P&D (em relação ao total de vendas) ou ao número de pessoas envolvidas nas atividades de P&D (em relação ao total de empregados das empresas).

Seguramente, uma referência obrigatória na área de coleta de dados de P&D é o Manual de Frascati (OECD, 2015), um documento redigido por e para especialistas dos países membros que acolheram e difundiram os dados nacionais relativos a P&D e forneceram respostas aos questionários de consulta propostos pela OECD. Na publicação, o P&D compreende tanto a produção de novos conhecimentos como novas aplicações práticas para o conhecimento. O critério básico para a distinção entre P&D e outras atividades é a presença de consideráveis novidades e a resolução de incertezas científicas e/ou tecnológicas.

Um importante ponto positivo em relação à utilização de dados de P&D na mensuração da inovação diz respeito à grande disponibilidade de dados. Desde a década de 1950, a OECD vem coletando e compilando regularmente estes dados, que são utilizados para análises e comparações entre países, indústrias e empresas (KLEINKNECHT; VAN MONTFORT; BROUWER, 2002; SMITH, 2005).

O investimento em P&D é considerado um fator de entrada para a inovação, mas apenas um deles, e sua eficiência depende da sua utilização (KLEINKNECHT; VAN MONTFORT; BROUWER, 2002; MAKKONEN; HAVE, 2013). Balzat (2003) acrescenta, ainda, que a relação entre os gastos em P&D e o sucesso em inovação não é facilmente mensurável, e considera artificial o estabelecimento de um vínculo direto entre este fator de entrada e o desempenho econômico real. Em princípio, P&D não retrata nada sobre os resultados do processo de inovação, tais como a introdução de novos produtos, serviços ou processos ao uso comercial.

Para Nelson (2009), as limitações das estatísticas de P&D estão relacionadas principalmente ao fato de que nem todos os esforços em P&D tornam-se sucesso, ou resultados de inovação comercializados. Kleinknecht; Van Montfort e Brouwer (2002), por sua vez, observam que as medidas de P&D costumam ter um viés para a manufatura, tendendo a subestimar as inovações em serviços, e que os *surveys* tradicionais tendem a subestimar as atividades de P&D em empresas de menor porte. Outro aspecto importante diz respeito às necessidades de sigilo que eventualmente dificultam as pesquisas, resultando em publicações com maior nível de agregação dos dados.

Uma vez que as novas teorias enfatizam sua natureza complexa e sistêmica, as bases teóricas para as políticas de ciência, tecnologia e inovação também mudaram, aumentando a ênfase nos conjuntos de ligações complexas entre as atividades de P&D e o processo de inovação como um todo (MAKKONEN; HAVE, 2013).

2.5.2 Estatísticas de Patentes

Uma patente é um contrato público entre um inventor e um governo, que concede direitos de monopólio por tempo limitado ao solicitante para uso de uma invenção técnica. O titular da patente deve primeiro demonstrar um avanço não óbvio no estado da arte para, então, estabelecer uma relação de vínculo com o Estado. O inventor estabelece que irá revelar informações detalhadas sobre a invenção em troca da proteção do da invenção pelo tempo e área geográfica estabelecidos em contrato (SMITH, 2005).

Os dados relativos a patentes costumam ser utilizados como uma medida de saída (desempenho) para a inovação. Sua utilização é facilitada graças à disponibilidade de séries históricas, cujas bases de dados são publicamente acessíveis e com um considerável grau de estruturação de seus conteúdos. O caráter qualitativo das patentes costuma ser feito por análises de citações, onde é atribuída maior importância àquelas mais citadas (KLEINKNECHT; VAN MONTFORT; BROUWER, 2002; NELSON, 2009).

As deficiências na utilização de estatísticas de patentes assemelham-se às relativas aos dados de P&D. Para Smith (2005), a principal delas diz respeito ao fato de que as patentes são um indicador de invenção, em vez de inovação. Cabe lembrar que nem todas as invenções patenteadas tornam-se inovações (COAD; RAO, 2008). Além disso, alguns setores tecnológicos patenteiam mais do que outros, e a propensão em patentear, está atrelada também ao tamanho das empresas (ARUNDEL; KABLA, 1998; MAKKONEN; HAVE, 2013),

Quando analisado o contexto regional e a localização das patentes, Makkonen e Have (2013) alertam que a distribuição espacial de grandes e pequenas empresas pode distorcer as medidas regionais. A observação de um mero padrão espacial de especialização setorial ou industrial pode desfigurar a inovação regional quando medida através de patentes.

Para Natário *et al.* (2011), Embora os dados de patentes tenham a vantagem de estarem disponíveis por longos períodos de tempo, os pontos fracos apontados na literatura fazem com que os dados de patentes não sejam considerados um indicador preciso sobre o desempenho da inovação, muito menos o único. Exemplificando, a metodologia do *European Innovation Scoreboard*, que anteriormente incluía as patentes como indicadores de saída (output), passou, a partir de 2008, a incluí-las como uma das dimensões utilizadas para capturar o esforço de inovação das empresas (NATÁRIO *et al.*, 2011).

2.5.3 *Surveys* de Inovação

Graças ao entendimento de que a utilização de indicadores individuais mostra-se incompleta quando utilizada para mensurar a inovação, outras abordagens foram desenvolvidas para a criação de indicadores projetados especificamente para este fim. Uma abordagem bastante utilizada é o desenvolvimento e aplicação de *surveys*. Os *surveys* de inovação dividem-se em dois tipos básicos: *subject approach* e *object approach* (OECD, 2005; SMITH, 2005).

Nos *surveys* do tipo *subject approach*, o foco está no agente de inovação. Este tipo de análise inicia pelo comportamento inovador das empresas como um todo, explorando as estratégias, incentivos e barreiras que influenciam este comportamento, o escopo das várias atividades de inovação e examinar as saídas e os efeitos da inovação (OECD, 2005).

Na *subject approach*, a atenção está voltada às atividades de inovação ao nível das empresas, sendo feitas perguntas sobre entradas ou investimentos totais em atividades de inovação (sejam ou não de P&D), e saídas (normalmente inovações de produto) em um determinado período. Estes *surveys* são projetados para serem representativos de todas as indústrias, de maneira que seus resultados possam ser extrapolados para que possam ser feitas comparações entre elas.

Já os *surveys* do tipo *object approach* tem foco nos objetos de saída do processo de inovação. Esta abordagem envolve a coleta de dados sobre inovações específicas, tais como dados descritivos, quantitativos e qualitativos sobre estas (OECD, 2005). A atenção está voltada às inovações tecnológicas significantes, sendo normalmente identificadas pela apreciação de especialistas, ou pelos anúncios de novos produtos.

A título de exemplo, a abordagem do tipo *subject* é a utilizada no Manual de Oslo, que propõe orientações para a coleta e interpretação de dados da inovação.

2.5.4 Indicadores de Inovação Compostos

Como existem muitos indicadores que estão associados à inovação, cada um, individualmente, é apenas um indicativo parcial do esforço total de inovação realizado. Com o objetivo de capturar um cenário mais complexo, os Indicadores de Inovação Compostos (CII)

foram introduzidos, na década de 1990, para a mensuração da inovação nas empresas (MAKKONEN; HAVE, 2013).

Para Grupp e Schubert (2010), desde a definição de inovação de Schumpeter, é muito claro que a inter-relações qualitativas e complexas em um ambiente nacional de inovação não podem ser medidas de uma forma simples. A inovação tem muitas dimensões, e a composição de forças e fraquezas pode resultar em diferentes níveis de inovação.

A complexidade das atividades de inovação faz com que o uso de medidas individuais mostrem-se incompletas. Assim, buscando obter uma compreensão mais aprofundada das atividades de inovação, um tipo de abordagem mais recente considera que o conceito de inovação não é diretamente observável, mas que está latente, e para sua mensuração são utilizados Indicadores de Inovação Compostos (IIC), criados pela agregação de múltiplas medidas ou indicadores (MAKKONEN; HAVE, 2013).

O argumento, portanto, é de que os índices compostos fornecem um melhor resultado do que uma variável individual de inovação. Ao invés da adoção de um indicador de entrada ou um de saída, Booyesen (2002) defende que sejam utilizados fatores de entrada e de saída simultaneamente, sintetizando as informações providas por vários indicadores em um.

De forma prática, um indicador composto é uma métrica de valor real único que é derivada de um conjunto de componentes indicadores por algum método de agregação (GRUPP; SCHUBERT, 2010). Frequentemente, o indicador composto, ou índice, é apresentado como um ranking ordenado ou normalizado em um intervalo fechado.

Para Makkonen e Have (2013), os índices de inovação podem prover uma forma mais compreensiva de referencial do desempenho da inovação nas regiões, para fins de criação de políticas. Dessa forma, tornou-se uma prática comum combinar vários indicadores para ciência, tecnologia e inovação para formar números agregados ou compostos.

Makkonen e Have (2013) observam que os índices são construídos a partir de uma variedade de medidas, comumente divididos entre medidas que capturam as pré-condições ou a capacidade para a atividade de inovação e medidas dos resultados da atividade de inovação ou invenção. Comumente, tal divisão é apresentada na forma de indicadores de entrada (*input*) e de saída (*output*) (ZABALA ITURRIAGAGOITIA; JIMÉNEZ SAEZ; GUTIÉRREZ GRACIA, 2005; CORNELL UNIVERSITY; INSEAD; WIPO, 2013).

Tais indicadores são, portanto, sintéticos, e possuem prós e contras. O Quadro 2 apresenta prós e contras para o uso de índices

compostos, conforme compilados por Saisana e Tarantola (2002), a partir de nota da Comissão Europeia sobre indicadores compostos.

Quadro 2 – Prós e Contras dos indicadores compostos.

Prós

- Indicadores Compostos podem ser usados para **sumarizar questões complexas** ou multidimensionais, a fim de dar suporte a tomadores de decisão.
- Indicadores Compostos fornecem um **panorama geral**, podendo ser mais facilmente interpretado do que quando se tenta encontrar uma tendência em muitos indicadores separados.
- Indicadores Compostos podem ajudar a **atrair interesse público** por fornecerem um sumário segundo o qual é possível comparar o desempenho entre países e seu progresso com o passar do tempo.
- Indicadores Compostos podem ajudar a **reduzir o tamanho** de uma lista de indicadores ou a incluir mais informações.

Contras

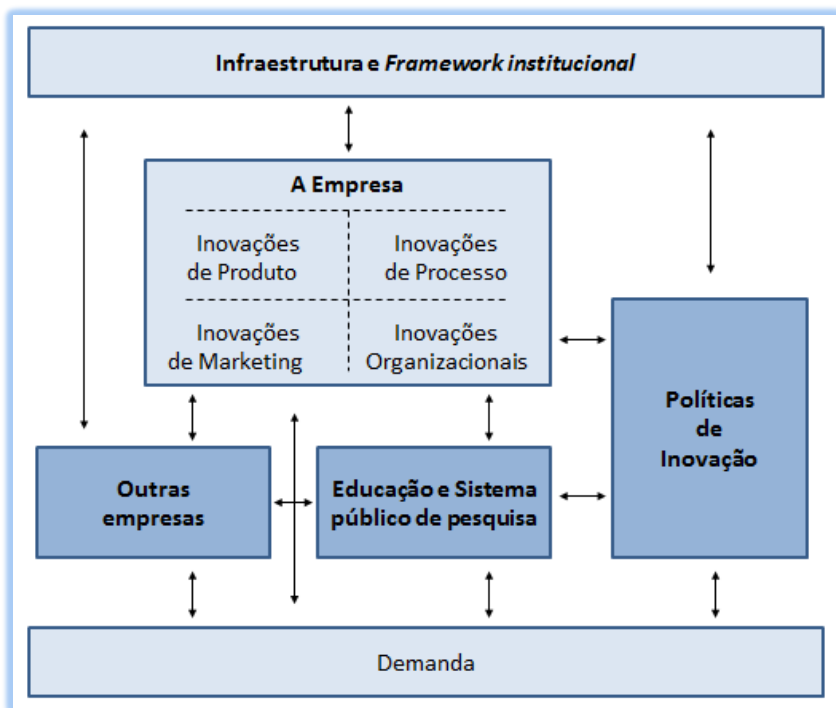
- Indicadores Compostos podem enviar **mensagens políticas ilusórias** ou não-robustas se forem pobremente construídas ou mal interpretadas. Análise de sensibilidade pode ser usada para testar a robustez de indicadores compostos.
- Os resultados de um panorama geral mostrado por indicadores compostos podem levar os decisores a traçar **conclusões políticas simplistas**. Os indicadores compostos devem ser usados em combinação com os sub-indicadores para traçar conclusões políticas sofisticadas.
- A construção de indicadores compostos envolve estágios onde **decisões devem ser tomadas**: a seleção de sub-indicadores, escolha do modelo, definição de peso dos indicadores e tratamento de valores faltantes e etc. Tais decisões devem ser transparentes e baseadas em princípios estatísticos.
- O escopo deve ser mais focado aos Estados Membros do que quando se trata de indicadores individuais. A seleção de sub-indicadores e pesos pode ser o **objetivo de mudança política**.
- Os indicadores compostos aumentam a **quantidade de dados** necessários, uma vez que são necessários dados para todos os sub-indicadores e para uma análise estatisticamente significativa.

Fonte: Saisana e Tarantola (2002)

Uma das mais reconhecidas iniciativas em termos coleta e análise de dados para a inovação é o *Oslo Manual - Proposed guidelines for collecting and interpreting innovation data* (OECD, 1992; 2005), publicado pela *Organisation for Economic Co-operation and Development* (OECD). A título de informação, cabe observar que na revisão de literatura apresentada em (SPERONI *et al.*, 2015), este trabalho foi a referência bibliográfica mais vezes citadas pelos artigos que compuseram o corpo de literatura, demonstrando a importância do documento na área de indicadores de inovação.

Tendo sua primeira edição sido publicada em 1992, e a terceira e mais atual edição no ano de 2005, o Manual de Oslo tem por objetivo orientar e padronizar conceitos, metodologias e construção de estatísticas e indicadores de pesquisa de P&D de países industrializados, e é parte de uma família de manuais voltados à mensuração e interpretação de dados relativos à ciência, tecnologia e inovação. A Figura 8 apresenta uma ilustração do *framework* do Manual de Oslo.

Figura 8 – *Framework* do Manual de Oslo



Fonte: Oecd (2005, p.34)

Enquanto as duas primeiras versões do Manual utilizavam-se da definição de inovação de tecnologia de produto e processo (TPP), refletindo um foco nas empresas, a terceira edição expande o *framework* de mensuração da inovação em três importantes linhas. Primeiro, dá ênfase ao papel das ligações com outras empresas e instituições no processo de inovação. Segundo, reconhece a importância da inovação em indústrias menos intensivas em P&D, tais como serviços e indústrias

de baixa tecnologia. Terceiro, a definição de inovação é expandida para incluir dois tipos adicionais de inovações: inovações organizacionais e inovações de *marketing*.

Uma das características importantes do Manual de Oslo é o fato de considerar que as atividades de inovação compreendem uma diversidade de elementos, não sendo mais vistas como um processo linear e centrado em P&D. Por isso, embora o manual seja voltado especificamente para o fenômeno da inovação nas empresas, traz diretrizes para mensurar a existência e a força das interações tecnológicas sistêmicas entre as empresas e o sistema de inovação onde elas operam (EVANGELISTA *et al.*, 2001).

Outras importantes iniciativas que tratam da mensuração da inovação, desenvolvidas pela União Européia, como o *Community Innovation Survey* (CIS) (EUROSTAT, 2010), o *Innovation Union Scoreboard* (UIS) (EUROPEAN UNION, 2014a) e o *Regional Innovation Scoreboard* (RIS) (EUROPEAN UNION, 2014b) são embasadas no Manual de Oslo.

O CIS é uma ferramenta que tem foco nas inovações de produto e processo e examina principalmente os resultados e as fontes de informação sobre atividades e as despesas com a inovação. Também são abordados os fatores que dificultam a inovação e o uso de direitos de propriedade intelectual. Há, ainda, uma unidade menor que trata das inovações organizacionais e de mercado. Sua primeira versão foi publicada em 1992, e tem periodicidade de quatro anos.

O IUS é uma publicação anual que provê uma avaliação comparativa do desempenho da pesquisa e inovação dos Estados Membros da União Europeia, bem como as forças e fraquezas de seus sistemas de inovação. Sua aplicação ajuda os Estados Membros a avaliar as áreas nas quais precisam concentrar seus esforços para melhorar seu desempenho de inovação (EUROPEAN UNION, 2014a). A primeira edição data do ano de 2001, mas o nome *Innovation Union Scoreboard* somente foi adotado em 2011. Antes disso, era publicada sob o nome *European Innovation Scoreboard*.

No IUS, o desempenho da inovação é medido por meio de um indicador composto – o *Summary Innovation Index* – que sumariza o desempenho de uma série de diferentes indicadores. Os indicadores de inovação são diferenciados em três tipos principais: **Enablers** (Viabilizadores), **Firm activities** (Atividades empresariais) e **Outputs** (Resultados), cada um deles subdividido em “dimensões”.

O RIS é uma publicação que replica a metodologia do IUS, aplicando-a em nível regional. Seu objetivo é proporcionar uma

avaliação comparativa do desempenho da inovação entre 190 regiões da União Europeia, Noruega e Suíça.

Quadro 3 - Comparativo entre indicadores do IUS e RIS

<i>Innovation Union Scoreboard</i>	<i>Regional Innovation Scoreboard</i>
VIABILIZADORES (ENABLERS)	
Recursos Humanos	
<ul style="list-style-type: none"> • Novos doutores (ISCED 6) por 1000 habitantes (25-34 anos) • Porcentagem da população entre 30 e 34 anos que completaram o ensino superior • Porcentagem de jovens entre 20 e 24 anos que concluíram pelo menos o ensino médio 	<ul style="list-style-type: none"> • Dados regionais não disponíveis • Porcentagem da população entre 25 e 64 anos que completaram o ensino superior • Dados regionais não disponíveis
Sistemas de Pesquisa abertos, de excelência e atrativos	
<ul style="list-style-type: none"> • Co-publicações científicas internacionais por milhão de pessoas • Publicações científicas entre as 10% mais citadas mundialmente • Estudantes de doutorado de fora da União Europeia como % de todos os estudantes 	<ul style="list-style-type: none"> • Dados regionais não disponíveis • Dados regionais não disponíveis • Dados regionais não disponíveis
Finanças e apoio	
<ul style="list-style-type: none"> • Despesas com P&D no setor público como % do GDP • Capital de risco (estágio inicial, expansão e reposição) como % do GDP 	<ul style="list-style-type: none"> • Idêntico • Dados regionais não disponíveis
ATIVIDADES EMPRESARIAIS (FIRM ACTIVITIES)	
Investimento das empresas	
<ul style="list-style-type: none"> • Despesas com P&D no setor de negócios como % do GDP • Despesas de inovação que não são de P&D como % do turnover 	<ul style="list-style-type: none"> • Idêntico • Similar (apenas para PMEs)
Links & empreendedorismo	
<ul style="list-style-type: none"> • PMEs inovadoras in-house como % das PMEs • PMEs inovadoras em colaboração com outras (% das PMEs) • Co-publicações público-privada por milhão de pessoas 	<ul style="list-style-type: none"> • Idêntico • Idêntico • Dados regionais não disponíveis
Ativos intelectuais	
<ul style="list-style-type: none"> • Pedidos de patentes PCT por bilhão do GDP (em PPS€) • Pedidos de patentes PCT em desafios 	<ul style="list-style-type: none"> • Pedidos de patentes EPO por bilhão do GDP regional • Dados regionais não disponíveis

societais por bilhão do GDP (em PPS€) • Marcas registradas na Community por bilhão do GDP (em PPS€) • Projetos na Community por bilhão do GDP (em PPS€)	• Dados regionais não disponíveis • Dados regionais não disponíveis
SAÍDAS (OUTPUTS)	
Inovadores	
• PMEs que introduziram inovações de produtos ou processos como % das PMEs • PMEs que introduziram inovações de marketing ou organizacionais como % das PMEs • Emprego em empresas de rápido crescimento do setor inovador	• Idêntico • Idêntico • Dados regionais não disponíveis
Efeitos econômicos	
• Emprego em atividades intensivas em conhecimento (manufatura e serviços) como % do total de empregos • Contribuição da exportação de produtos de média e alta tecnologia para a balança comercial • Exportação de serviços intensivos em conhecimento como total da exportação de serviços • Vendas de inovações novas para o mercado e novas para as empresas como % do turnover • Receitas com licenciamento e patentes como % do GDP	• Emprego em manufaturas de média e alta tecnologia e serviços intensivos em conhecimento como % da força de trabalho total • Dados regionais não disponíveis • Dados regionais não disponíveis • Similar (apenas para PMEs) • Dados regionais não disponíveis

Fonte: European Union (2014b)

As edições anteriores do RIS foram publicadas nos anos de 2002, 2003, 2006, 209 e 2012. As regiões são classificadas em quatro grupos de diferentes desempenhos: Líderes em Inovação, Seguidores em Inovação, Inovadores Moderados e Inovadores Modestos.

Conforme é possível observar no Quadro 3, na edição com dados de 2014 do *Regional Innovation Scoreboard*, são utilizados apenas 11 dos 25 indicadores do IUS. Para os outros 14, a publicação relata não haver dados regionais disponíveis.

Mesmo para os indicadores que têm dados disponíveis em nível regional, esta disponibilidade não é total. Apenas cinco dos indicadores estavam diretamente disponíveis no banco de dados da *Eurostat*. Para os outros seis indicadores foram utilizados os dados do *Community*

Innovation Survey, obtidos por meio de uma requisição especial por dados regionais (EUROPEAN UNION, 2014b).

O Quadro 4 apresenta um comparativo entre as metodologias de indicadores de inovação propostas pela União Europeia.

Quadro 4- Comparação das metodologias da União Europeia

Critério de comparação	<i>European Innovation Survey (E.I.S.)</i>	<i>Regional Innovation Survey (R.I.S.)</i>	<i>Community Innovation Survey (C.I.S.)</i>
Definição	Compara a inovação entre países na UE	Compara a inovação entre regiões na UE	Mensura a inovação em nível doméstico da UE
Data Mining	Empresas na UE que responderam o questionário da CIS	Empresas na EU que responderam o questionário da CIS	Empresas na União Europeia
Escopo (territorial)	União Europeia	Regiões na UE	Estado membro da UE
Periodicidade	Anual	Anual	Quadriannual
Usabilidade	Alta	Média	Muito Alta
Adaptabilidade	Média	Baixa	Alta
Eficiência	Dados do CIS	Dados do CIS	Censo
Efetividade	Config. de políticas para a inovação por centros de tomada de decisão em nível Europeu e doméstico	Config. de políticas para a inovação por centros de tomada de decisão em nível Europeu e doméstico	Config. de políticas para a inovação por centros de tomada de decisão em nível Europeu e doméstico

Fonte: Balatsas; Kyratsis e Mavroeidis (2009)

As publicações EIS, RIS e CIS dependem umas das outras. Os dados são principalmente produzidos pela CIS, de forma que os indicadores do EIS e do RIS não podem levar a resultados sem a ajuda do CIS (BALATSAS; KYRATSIS; MAVROEIDIS, 2009).

Em função da periodicidade das publicações, seus resultados são costumeiramente utilizados como insumos em outros trabalhos. São os casos de Zabala Iturriagoitia; Jiménez Saez e Gutiérrez Gracia (2005) e Calderero-Gutiérrez; Fernández-Macho e Kuittinen (2009), que se baseiam nos dados das publicações EIS, RIS, e CIS, e apresentam análises comparativas de países ou regiões específicas, trazendo classificações ranqueadas destes.

Outras entidades de atuação internacional também contam com publicações que tratam da definição de índices associados à inovação. É o caso do Banco Mundial, com o *World Development Indicators* (WORLD BANK, 2014), e da Universidade de Cornell, INSEAD e Organização Mundial de Proteção Intelectual (WIPO), com o *Global Innovation Index* (CORNELL UNIVERSITY; INSEAD; WIPO, 2013).

O *Global Innovation Index* reconhece o papel chave da inovação como impulsionador do crescimento econômico e prosperidade, e adota uma visão inclusiva e horizontal da inovação, que se aplica tanto a economias desenvolvidas quanto àquelas em desenvolvimento (CORNELL UNIVERSITY; INSEAD; WIPO, 2013). Seu objetivo é capturar os aspectos multidimensionais da inovação e, para isto, baseia-se em um *framework* que considera a visão holística da inovação e que conta com dois sub-índices, um de entrada e outro de saída: *Input Sub-Index* e *Output Sub-Index*.

O *World Development Indicators* (WDI) é uma publicação do Banco Mundial e trata-se da principal compilação de estatísticas internacionalmente comparáveis sobre o desenvolvimento global. Iniciada em 1997, a publicação tem periodicidade anual e conta ferramentas para acesso aos Dados Abertos que publica.

Os indicadores apresentados no WDI são organizados segundo seis temas: Visão mundial; Pessoas; Ambiente; Economia; Estado e mercados; e, Ligações globais (WORLD BANK, 2014).

No Brasil, a Pesquisa de Inovação (PINTEC) é uma publicação do Instituto Brasileiro de Geografia e Estatística, com apoio do Ministério da Ciência, Tecnologia e Inovação e da Financiadora de estudos e Projetos (FINEP). As primeiras edições, publicadas em 2002, 2003, 2005 e 2008 traziam a nomenclatura Pesquisa de Inovação Tecnológica. A partir da edição de 2008, os resultados foram regionalizados, e a partir da edição publicada em 2013 o termo “Tecnológica” foi suprimido do nome da pesquisa para promover um alinhamento aos padrões internacionais (Ibge, 2013).

Quanto ao objetivo da pesquisa, cita-se:

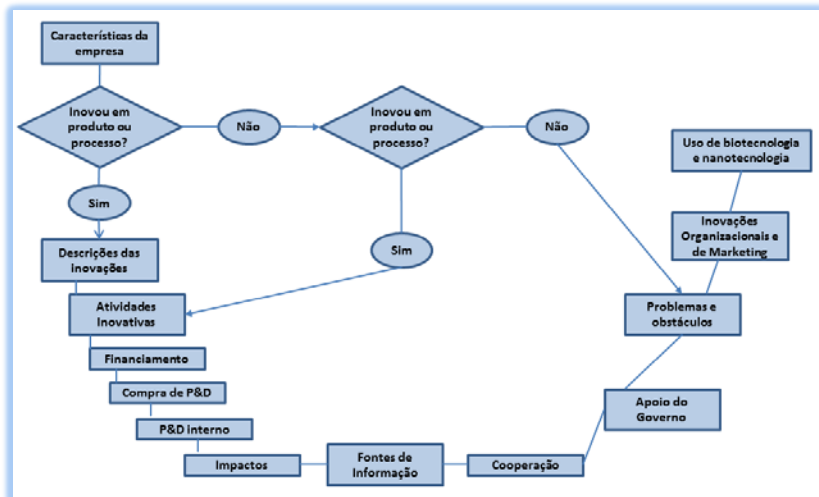
“A Pesquisa de Inovação PINTEC tem por objetivo a construção de indicadores setoriais, nacionais e regionais, das atividades de inovação nas empresas do setor de indústria, e de indicadores nacionais de inovação nas empresas dos setores de Eletricidade e gás e de Serviços selecionados (edição e gravação e edição de música; telecomunicações; atividades dos serviços

de tecnologia e da informação; tratamento de dados; hospedagem na internet e outras atividades relacionadas; serviços de arquitetura e engenharia, testes e análises técnicas; e pesquisa e desenvolvimento, compatíveis com as recomendações internacionais em termos conceituais e metodológicos.” (IBGE, 2013, p.12)

A PINTEC é desenvolvida segundo o referencial conceitual e metodológico do Manual de Oslo (OECD, 2005) e do *Community Innovation Survey* (EUROSTAT, 2010) e, embora incorpore em sua última edição as inovações de natureza organizacional e de *marketing*, tem seu foco nas inovações de produto e processo. Trata-se de um *survey* do tipo *subject-approach*, pois as informações são obtidas são relativas ao “sujeito”, ou seja, comportamento, atividades empreendidas, impactos, incentivos, obstáculos e demais fatores relativos à empresa como um todo.

A captura de dados para a pesquisa é feita por meio de um questionário, que segue uma divisão por blocos, nos quais os temas da pesquisa estão organizados, e cujo fluxo é representado na Figura 9.

Figura 9 - Fluxo de blocos temáticos da PINTEC



Fonte: Ibge (2013)

Os blocos temáticos da PINTEC são:

- **Características das empresas:** incluem variáveis específicas que tem por objetivo identificar: a origem do capital controlador da

empresa e sua localização, no caso de estrangeiro; se a empresa é independente ou parte de um grupo; e, a abrangência geográfica do principal mercado da empresa;

- **Produtos e processos novos ou substancialmente aprimorados:** as empresas que implementaram inovações de produto e processo informam o grau de novidade para o mercado e em termos técnicos, bem como quem desenvolveu a principal inovação;
- **Atividades inovativas:** descrevem os esforços empreendidos para a inovação de produto e processo, compreendendo atividades internas de P&D, aquisição externa de P&D, aquisição de outros conhecimentos externos, aquisição de *software*, aquisição de máquinas e equipamentos, treinamento, introdução de inovações tecnológicas no mercado, projeto industrial e outras perparações técnicas para a produção e distribuição;
- **Financiamento:** busca levantar dados sobre as fontes de financiamento, e compreendem aquelas relativas ao financiamento interno de P&D, desagregadas em financiamento próprio, de terceiros (outras empresas ou público, e do exterior);
- **Compra de serviços de P&D:** coleta informações referentes à distribuição percentual do valor da compra por parte de empresas brasileiras de serviços de P&D, de acordo com o tipo de organização;
- **Atividades internas de P&D:** além dos dispêndios como P&D, solicitam-se, neste bloco, informações sobre a continuidade ou ocasionalidade das atividades, bem como sobre a localização do departamento de P&D da empresa, além do numero de pessoas ocupadas nas atividades de P&D;
- **Impacto das inovações:** visa identificar os impactos associados ao produto, ao mercado, ao processo, aos aspectos relacionados ao meio ambiente, à saúde e segurança e ao enquadramento em regulamentações e normas;
- **Fontes de informação:** compreende informações sobre a origem dos conhecimentos científicos e tecnológicos, se produzidos ou adquiridos por meio de patentes, máquinas, artigos especializados, *softwares*;
- **Relações de cooperação:** identificação da participação ativa da empresa em projetos conjuntos de P&D e outros projetos de inovação com outra organização (empresa ou instituição).
- **Apoio do governo:** compreendem informações acerca de financiamentos, incentivos fiscais, subvenções, participação em

programas públicos voltados para o desenvolvimento tecnológico e científico, etc;

- **Problemas e obstáculos à inovação:** tem por objetivo identificar os motivos pelos quais a empresa não desenvolveu atividades inovativas ou não obteve os resultados esperados;
- **Inovações organizacionais e de *marketing*:** as empresas respondentes devem informar se implementaram inovações organizacionais: novas técnicas de gestão para melhorar rotinas e práticas de trabalho, assim como o uso e troca de informações, de conhecimento e habilidades dentro da empresa; novas técnicas de gestão ambiental; novos métodos de organização do trabalho para melhor distribuir responsabilidades e poder de decisão; e, mudanças significativas nas relações com outras organizações. Inovações de *marketing*: mudanças significativas nos conceitos/estratégias de *marketing*; e, mudanças significativas na estética, desenho ou outras mudanças subjetivas em pelo menos um produto.

2.6 CONSIDERAÇÕES FINAIS

A inovação é reconhecida como um indutor do desenvolvimento econômico e social, e está associada à garantia de competitividade de empresas, regiões ou países. Mais do que invenções, seu conceito está associado à introdução de serviços ou produtos novos, ou significativamente melhorados, no mercado, bem como às alterações nas formas de produção ou rotinas organizacionais, ou mudanças no mercado.

Os chamados Sistemas de Inovação representam o agrupamento de diferentes tipos de organizações que atuam em interação em um mesmo contexto e que dependem das instituições e condições socioambientais. Desta forma, buscam-se desenvolver normas, regulamentações e políticas públicas com o objetivo de potencializar as condições para o desenvolvimento das atividades de inovação. Além disso, há a preocupação em mensurar as condições ou as ações empreendidas, no sentido de avaliá-las.

Diferentes indicadores são utilizados individualmente para medir a inovação, como é o caso das estatísticas de patentes e de P&D. Em outras abordagens, entretanto, tais medidas são consideradas parciais, e são agregadas, juntamente com outros conjuntos são definidos indicadores compostos, criados pela agregação de um conjunto de

medidas, consideradas de *input* e *output*. Muitos destes modelos seguem as orientações metodológicas propostas no Manual de Oslo, um guia de coleta e interpretação de dados de inovação, criado pela OECD, e são criados a partir de *surveys* que têm por objetivo a comparação entre países.

3 ONTOLOGIAS

3.1 CONSIDERAÇÕES INICIAIS

O conhecimento é reconhecido como um recurso vital na economia mundial, formando a base para a inovação e o sucesso econômico (NONAKA; TAKEUCHI, 1997; MEYER; SUGIYAMA, 2007). Em função disso, as organizações buscam novas formas de gestão que atendam suas necessidades, contexto no qual se destaca a Gestão do Conhecimento como uma estratégia para a criação, acesso e suporte a este recurso vital.

Conforme descrição da área de concentração “Engenharia do Conhecimento” do Programa de Pós-Graduação em Engenharia e Gestão do Conhecimento da UFSC, novas disciplinas têm emergido para gerar, codificar e gerir conhecimento organizacional, combinando elementos de mais de uma área do conhecimento, e dentre elas está a Engenharia do Conhecimento (EC) (PPEGC, 2004).

Evoluindo de uma arte, inicialmente dedicada a transferir o conhecimento armazenado na cabeça de especialistas humanos para sistemas capazes de realizar tarefas específicas, a EC atual é uma disciplina que tem o objetivo de fornecer metodologias, técnicas e ferramentas para a construção de Sistemas Baseados em Conhecimento (SBC) de forma sistêmica e controlada (STUDER; BENJAMINS; FENSEL, 1998; NAZÁRIO; DANTAS; TODESCO, 2014).

Para Schreiber *et al.* (1999), a característica que diferencia os SBC dos demais sistemas de *software* é a necessidade de representação explícita do conhecimento do domínio que tratam. Dentre as técnicas para representação de conhecimento, exploram-se aqui as ontologias, cujo uso tem por objetivo a definição de um vocabulário compartilhado que permita a troca e o reuso de informações.

3.2 ENGENHARIA DO CONHECIMENTO

Tradicionalmente, a Engenharia do Conhecimento era regida por um paradigma da Inteligência Artificial para a construção de SBC baseado na transferência do conhecimento. O pressuposto fundamental era o de que o conhecimento necessário para o sistema já existia na cabeça dos especialistas e somente necessitava ser coletado e implementado na forma de Sistemas Especialistas. Os sistemas eram, portanto, desenvolvidos a partir de entrevistas com especialistas, e

implementados com regras de produção, para que estivessem aptos a resolver problemas específicos (STUDER; BENJAMINS; FENSEL, 1998).

Tal entendimento era observável também no conceito de Engenharia do Conhecimento utilizado por Waterman (1986):

“O processo de construção de um sistema especialista é normalmente chamado de ‘Engenharia do Conhecimento’. Tipicamente, ele envolve uma forma de interação especial entre o construtor do sistema especialista, chamado de Engenheiro do Conhecimento, e um ou mais especialistas humanos em alguma área. O Engenheiro do conhecimento ‘extrai’ do especialista humano seus procedimentos, estratégias e regras na resolução de problemas e constrói este conhecimento em um sistema especialista” (WATERMAN, 1986, p.5, tradução nossa).

Posteriormente, como observam Studer et. al. (1998), a tentativa de criação de Sistemas de Resolução de Problemas Gerais demonstrou que a transferência de conhecimento do especialista para os sistemas não era um processo simples, e que muitos conhecimentos eram de difícil explicitação. Além disso, as técnicas utilizadas para a representação do conhecimento nos SBCs misturavam os chamados conhecimentos de domínio com os conhecimentos necessários para a resolução das tarefas, o que tornava a manutenção e reutilização dos SBCs muito difícil.

Um breve histórico da evolução da Engenharia do Conhecimento até a metade da década de 1990 é apresentado por Schreiber *et al.* (1999), que considera que a mudança foi no sentido de uma arte para uma disciplina, conforme representado na Figura 10.

Figura 10 - Evolução da Engenharia do Conhecimento

Motores de busca de uso geral (GPS)	Primeira geração de sist. baseados em regras	Emergência de métodos estruturados (primeiros KADS)	Metodologias maduras (CommonKADS)
1965	1975	1985	1995

Fonte: Schreiber *et al.* (1999, p.14)

Em 1965, as primeiras tentativas tratavam da construção de um Resolvedor Geral de Problemas (GPS), resultando em métodos genéricos para a resolução de problemas, mas perdendo a aplicabilidade. Em 1975, a primeira versão dos SBCs utilizava linguagens de

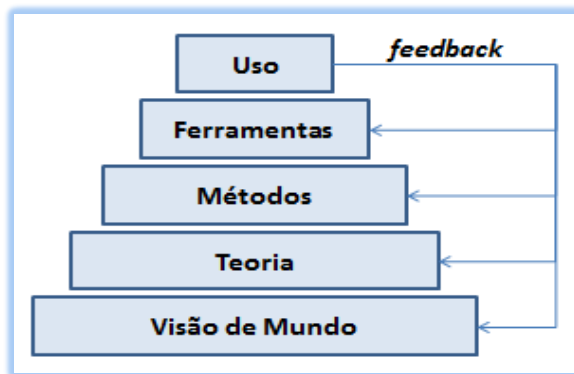
representação de conhecimento simbólico e dedutivo, mas perdeu força devido à escalabilidade e à aplicação a problemas complexos. Em 1985, em uma resposta à necessidade de metodologias e técnicas formais para o desenvolvimento de SBCs surgem os métodos de Aquisição e Documentação Estruturados do Conhecimento (KADS). Em 1995, a metodologia *CommonKADS* representa o amadurecimento destas metodologias estruturadas.

Uma importante mudança de paradigma aconteceu quando o entendimento sobre a forma de conceber Sistemas de Conhecimento mudou. A Engenharia do Conhecimento passou a ser vista não mais como um tipo de “mineração a partir da cabeça dos especialistas”, mas sim como uma atividade de modelagem na qual são construídos diferentes aspectos de modelos do conhecimento humano (STUDER; BENJAMINS; FENSEL, 1998; SCHREIBER *et al.*, 1999).

A nova abordagem permitiu que a Engenharia do Conhecimento, tal qual aconteceu com a Engenharia de *Software*, evoluísse da arte de construção de SBCs até tornar-se uma disciplina de engenharia. Para isso, foram necessárias a análise do processo de construção e manutenção em si, e o desenvolvimento de metodologias apropriadas, linguagens e ferramentas especializadas para este fim (STUDER; BENJAMINS; FENSEL, 1998).

Schreiber *et al.* (1999) propõem uma pirâmide metodológica para representar os elementos de uma metodologia de EC, composta de cinco camadas sobrepostas. As camadas da pirâmide, apresentada na Figura 11, são: Visão de mundo, Teoria, Métodos, Ferramentas e Uso.

Figura 11 - Pirâmide metodológica de Schreiber



Fonte: Schreiber *et al.* (1999)

A visão de mundo, também chamada de *slogans*, deve ser embasada em teorias, métodos, ferramentas e casos práticos de uso, que

correspondem às outras camadas, e pode ser formulada como um conjunto de princípios que formam a base para a abordagem. A teoria envolve os princípios científicos que serão utilizados para embasar os modelos. Os métodos compreendem os procedimentos utilizados para levar às soluções propostas na metodologia. As ferramentas servem para que se apliquem os métodos. O uso compreende as experiências com o uso da metodologia, cujos *feedbacks* são levados até as demais camadas. Schreiber *et al.* (1999) acrescenta que, uma vez que a visão de mundo mude, os fundamentos sob uma abordagem se esvaem, e chega a hora de uma mudança de paradigma.

Dias e Pacheco (2009) apresentam uma análise, descrevendo características gerais, etapas e artefatos gerados de algumas das mais representativas metodologias de Engenharia do Conhecimento: Vital; SPEDE (*Structured Process Elicitation and Demonstration Environment*); MAS-CommonKADS; MIKE (*Model-based and Incremental Knowledge Engineering*); MOKA (*Methodology for Knowledge-Based Engineering Applications*); On-To-Knowledge; XP.K (*eXtreme Programming of Knowledge-based systems*); e CommonKADS, esta última apontada como a mais reconhecida e utilizada.

Conforme observam Rautenberg; Todesco e Steil (2011) a Engenharia do Conhecimento é uma área complementar à Gestão do Conhecimento. Para os autores, a Gestão do Conhecimento se efetiva por meio de seus Instrumentos, e quando um destes instrumentos utiliza tecnologias avançadas, emprega os chamados Agentes Computacionais da Engenharia do Conhecimento, entendidos neste contexto como softwares capazes de executar tarefas intensivas em conhecimento, dado um domínio (SCHREIBER *et al.*, 1999).

Schreiber *et al.* (1999) observam que não há uma linha divisória clara entre os sistemas de conhecimento e os sistemas de *software* “normais”, especialmente porque todos os sistemas modernos contêm algum tipo de conhecimento. A principal distinção, segundo os autores, é que nos sistemas de conhecimento há algum tipo de representação explícita do conhecimento incluído no sistema, e por isso precisam de técnicas de modelagem especiais.

3.3 DEFINIÇÃO DE ONTOLOGIA

O termo ontologia origina-se das palavras gregas *onto* (ser) e *logia* (discurso escrito ou falado). A utilização do termo é originária da filosofia, estando associado à explicação semântica da natureza dos

seres, ou uma explicação sistemática da Existência (GRUBER, 1993; BORST, 1997; SHUE; CHEN; SHIUE, 2009). Posteriormente, nas áreas da computação, inteligência artificial e engenharia do conhecimento, o termo passa a ser utilizado na organização de bases de conhecimento reutilizáveis e compartilháveis (MOREIRA; ALVARENGA; OLIVEIRA, 2004; SHUE; CHEN; SHIUE, 2009).

Em sistemas de Inteligência Artificial, o que "existe" é aquilo que pode ser representado, e quando o conhecimento de um domínio é representado em um formalismo declarativo é chamado de "universo do discurso". Para especificar uma conceitualização é necessário formular axiomas que façam restrições a possíveis interpretações para os termos definidos. Assim, Gruber (1993) define ontologia como sendo "**uma especificação explícita de uma conceitualização**".

De forma semelhante, Uschold e Gruninger (1996) afirmam que ontologia é um termo usado para se referir ao **entendimento compartilhado de algum domínio de interesse**, e deve ser usada como um *framework* de unificação para a resolução de problemas.

As definições enfatizam o fato de que deve haver a concordância na conceitualização que é especificada, pois, caso contrário, se a conceitualização não for comumente aceita, a possibilidade de reuso de uma ontologia é praticamente nula (BORST, 1997). Neste contexto, Borst (1997) estende a definição, afirmando que uma ontologia é "**uma especificação formal e explícita de uma conceitualização compartilhada**".

Guarino e Giaretta (1995) enumeram sete possíveis interpretações de ontologia:

- 1) Ontologia como uma disciplina da filosofia;
- 2) Ontologia como um sistema conceitual informal;
- 3) Ontologia como uma descrição formal;
- 4) Ontologia como uma especificação de uma conceitualização;
- 5) Ontologia como uma representação de um sistema conceitual via teoria lógica;
- 6) Ontologia como o vocabulário usado por uma teoria lógica;
- 7) Ontologia como uma especificação (de meta-nível) de uma teoria lógica.

Shue; Chen e Shiue (2009), por sua vez, explicam, que ontologia é uma palavra que passou a fazer referência à "análise sistemática de conhecimento sobre determinado domínio de interesse, para que possa ser compartilhado por outras pessoas".

A adoção de ontologias permite, portanto, organizar e representar o conhecimento de um determinado domínio, e sua utilização proporciona melhorias na recuperação da informação. Além disso, conforme afirmam Almeida e Bax (2003), as ontologias permitem formas de representação baseadas em lógica, possibilitando criar novo conhecimento pelo uso de mecanismos inferência sobre o conhecimento existente.

O desenvolvimento de ontologias, portanto, leva em conta as características descritas nas definições. O objetivo é explicar um domínio de conhecimento, representando-o formalmente, e levando em consideração a necessidade de que esta representação seja aceita por um conjunto de pessoas. Ainda, a construção de ontologias deve levar em conta o nível de abstração com que o domínio será representado.

Para Gruber (1993), há um conjunto de critérios de projeto para ontologias, que têm por objetivo o compartilhamento de conhecimento e a interoperabilidade entre programas baseados em uma conceitualização compartilhada:

- 1) **Clareza:** Uma ontologia deve comunicar o significado pretendido dos termos definidos. As definições devem ser objetivas. Embora a motivação para a definição de um conceito possa surgir de situações sociais ou de requisitos computacionais, a definição deve ser independente de contextos sociais ou computacionais. O formalismo é uma forma de conseguir isto.
- 2) **Coerência:** Uma ontologia deve ser coerente, ou seja, deve validar inferências que são consistentes com as definições. No mínimo, os axiomas de definição devem ser logicamente consistentes. A coerência deve aplicar-se, também, aos conceitos que são definidos informalmente, tais como aqueles descritos nas documentações em linguagem natural e nos exemplos.
- 3) **Extensibilidade:** Uma ontologia deve ser projetada de forma a antecipar os usos do vocabulário compartilhado. Ela deve oferecer uma fundamentação conceitual para uma gama de tarefas antecipadas, e a representação deve ser construída de maneira que possa ser estendida e especializada. Novos termos podem ser definidos para usos específicos, baseados no vocabulário existente, de forma que não seja necessário revisar as definições existentes.
- 4) **Mínimo compromisso com codificação:** A conceitualização deve ser especificada ao nível do conhecimento, independente de qualquer codificação em particular. A dependência a uma codificação deve ser minimizada porque os agentes de compartilhamento de

conhecimento podem ser implementados em diferentes sistemas e estilos de representação.

- 5) **Mínimo compromisso ontológico**: Uma ontologia deve requerer o mínimo comprometimento ontológico suficiente para dar suporte às atividades de compartilhamento de conhecimento pretendidas. Uma ontologia deve fazer o mínimo de imposições possíveis sobre o mundo que está sendo modelado, permitindo que as partes comprometidas tenham liberdade para especializá-la ou instanciá-la, quando necessário.

As ontologias são compostas por dois princípios essenciais de projeto: os **constructos** e as **regras e axiomas**. Os constructos representam as entidades e sua estrutura (classes, subclasses, atributos e relações), enquanto que as regras e axiomas descrevem os fatos gerais e condições (Reichwald; Bullinger, 2008).

3.4 TIPOS DE ONTOLOGIA

Dependendo do escopo em que o conhecimento é representado, as ontologias recebem classificações quanto ao seu tipo. Esta classificação não é um consenso, e é variam conforme o autor.

Mizoguchi; Vanwelkenhuysen e Mitsuru (1995) propõem uma tipologia de ontologias de conteúdo, segundo a qual, classifica-as em:

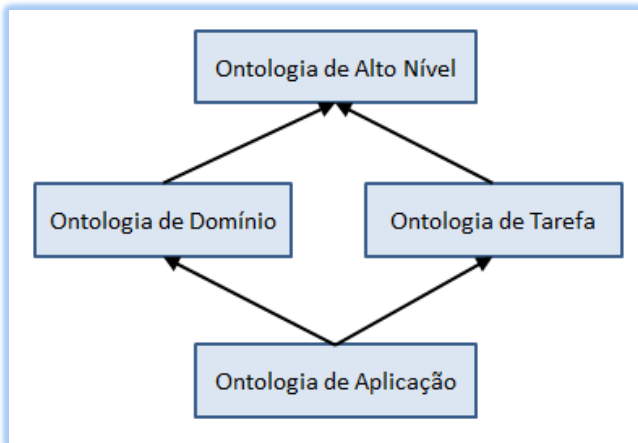
- **Ontologias de domínio** – sistema de vocabulário para descrição do domínio, e este vocabulário divide-se em três categorias: ontologia de objetos; ontologia de atividade; e, ontologia de campo.
- **Ontologias de tarefa** – sistema de vocabulário para descrição da estrutura de solução de problemas de todas as tarefas existentes, independente de domínio. Os autores esclarecem que seu entendimento de tarefa é "uma sequência de passos para a resolução de problemas". Dessa forma, uma ontologia de tarefa prevê primitivas pelas quais o contexto de resolução de problemas pode ser descrito e o conhecimento do domínio pode ser posto no contexto da solução do problema.
- **Ontologias gerais/comuns** – são aquelas usadas para a representação de conhecimentos de senso comum, tais como coisas, eventos, tempo, espaço, causalidade, comportamento e função.

Uschold e Gruninger (1996), por sua vez, apresentam uma classificação de ontologias quanto ao seu uso:

- **Ontologias de comunicação:** por expressar um entendimento compartilhado, possibilitam a comunicação entre pessoas com diferentes necessidades e pontos de vista de diferentes contextos.
- **Ontologias de interoperabilidade:** aplicações em que diferentes usuários precisam trocar dados ou utilizam diferentes ferramentas de software, e ambientes de integração entre diferentes *softwares*.
- **Ontologias de Engenharia de Sistemas:** aplicações de ontologias que dão suporte ao projeto e desenvolvimento de sistemas de *software*, com foco em especificação, confiança e reuso.

Guarino (1998) apresenta uma classificação das ontologias de acordo com seu grau de dependência em uma tarefa em particular ou ponto de vista, organizadas conforme a Figura 12.

Figura 12 - Classificação de ontologias quanto ao grau de dependência em uma tarefa



Fonte: Guarino (1998, p.7)

- **Ontologia de Alto Nível:** este tipo de ontologia descreve conceitos muito gerais, como espaço, tempo, substância, objeto, evento, ação, etc., que são independentes de um problema ou domínio em particular;
- **Ontologia de Domínio e Ontologia de Tarefa:** descrevem, respectivamente, o vocabulário relacionado a um domínio genérico (como medicina ou automóveis) ou uma tarefa ou atividade genérica

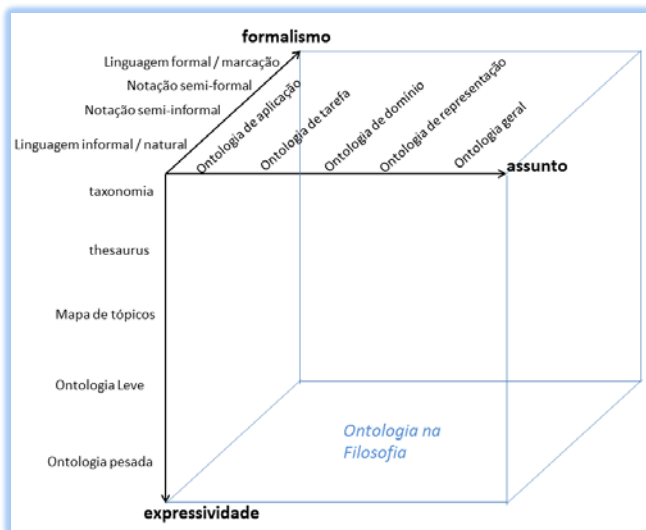
(como diagnóstico ou vendas), pela especialização dos termos introduzidos na ontologia de alto nível;

- **Ontologia de Aplicação:** descreve conceitos que dependem de domínio e tarefa particulares, eu são frequentemente especializações de ambas ontologias. Os conceitos correspondem aos papéis desempenhados pelas entidades do domínio enquanto desempenham uma certa atividade.

Almeida e Bax (2003): classificação quanto à função, ao grau de formalismo, à aplicação, à estrutura, ao conteúdo:

- Quanto à função:** ontologias de domínio; ontologias de tarefa e ontologias gerais
- Quanto ao grau de formalismo:** ontologias altamente informais; ontologias semi-informais; ontologias semi-formais e ontologias rigorosamente formais
- Quanto à aplicação:** ontologia de autoria neutra; ontologia como especificação e ontologia de acesso comum à informação
- Quanto ao conteúdo:** ontologias terminológicas; ontologias de informação; ontologias de modelagem do conhecimento; ontologias de aplicação; ontologias de domínio; ontologias genéricas e ontologias de representação.

Figura 13: OntoCube - um sistema de classificação integrado



Fonte: Reichwald e Bullinger (2008, p.181)

Baseando-se em trabalhos anteriores da comunidade de ontologias, o trabalho de Reichwald e Bullinger (2008) apresenta o framework *OntoCube* para a classificação de ontologias que cobre de forma holística os múltiplos tipos diferentes de ontologias por meio de uma linguagem unificada, e está representado na Figura 13.

Quadro 5 - Descrição dos eixos do *OntoCube*

Eixo	Classificação	Descrição
Assunto	Ontologias de aplicação	Adaptadas a uma aplicação
	Ontologias de tarefa	Ontologias de tarefa: vocabulário e conhecimento sobre uma tarefa
	Ontologias de domínio	Provisão genérica de conhecimento de domínio e entendimento compartilhado
	Ontologias gerais	Noções básicas sobre coisas, eventos, tempo, espaço, causalidade
	Ontologias de representação	Contem primitivas utilizadas para formalizar o conhecimento
Formalismo	Linguagem natural ou informal	Definições livremente expressadas em linguagem natural com objetivo de prover um vocabulário fácil e acessível para facilitar a comunicação
	Notação semi-informal e semi-formal	São expressas em uma forma restrita e estruturada da linguagem natural, aumentando a clareza e reduzindo a ambiguidade
	Notação formal	Definem de forma rígida e sistemática a linguagem para o sistema proposto.
Expressividade	Taxonomia	Descreve um sistema de classificação hierárquica ou um conjunto de relações entre objetos em um certo domínio.
	Thesaurus	Adiciona relações à estrutura hierárquica da taxonomia, tais como similaridade ou relações de homônimo, resultando em uma coleção sistemática de termos de um vocabulário estruturado de um certo domínio.
	Mapa de Tópicos	São criados para um escopo definido em um assunto específico. Consiste de objetos (ou tópicos), associações (relações restritas a tópicos relacionados), e ocorrências para conectar os objetos a recursos de informação externos.
	Ontologias Leves	Ontologias utilizadas para estruturar semanticamente um repositório de informações como base para busca e recuperação. Apontadas como versões prévias das ontologias pesadas
	Ontologias pesadas	Além de classes e relações, traz regras e axiomas e tem maior força de expressão do que as leves.

Fonte: Reichwald e Bullinger (2008)

Reichwald e Bullinger (2008) descrevem as classificações segundo os eixos do *OntoCube*, conforme apresentado no Quadro 5.

Por ser um processo de modelagem, o desenvolvimento de ontologias demanda de formas sistemáticas de produção. A próxima seção apresenta diferentes metodologias de criação de ontologias e suas características.

3.5 METODOLOGIAS DE DESENVOLVIMENTO DE ONTOLOGIAS

Embora haja uma variedade de metodologias, Fernández-López e Gómez-Pérez (2002) observam que muitas delas se preocupam mais com atividades específicas do processo, dando menos atenção a outras. Por isso, algumas atividades são privilegiadas com mais opções de artefatos.

A seguir, serão descritas, de forma sucinta, as metodologias *On-to-Knowledge* (SURE; STUDER, 2003), *METHONTOLOGY* (FERNÁNDEZ; GÓMEZ-PÉREZ; JURISTO, 1997), *101 Ontology Development* (NOY; MCGUINNESS, 2001) e *ontoKEM* (TODESCO *et al.*, 2009).

A *On-to-Knowledge Methodology* (OTKM) é uma metodologia para introdução e manutenção de aplicações de gestão do conhecimento baseadas em ontologias em empresas com foco em Processos e Meta-Processos de Conhecimento (SURE; STUDER, 2003). Para Rautenberg; Todesco e Steil (2011) a metodologia contribui na especificação de requisitos da ontologia, porque emprega questões de competência para confirmar o propósito e o escopo de uma ontologia, o que permite antecipar a identificação de muitos de seus conceitos, propriedades, relações e instâncias.

A *METHONTOLOGY* é uma metodologia bem estruturada utilizada para construir ontologias a partir de rascunhos (FERNÁNDEZ; GÓMEZ-PÉREZ; JURISTO, 1997). A metodologia conta com as fases de Especificação, Aquisição de Conhecimento, Conceitualização, Integração, Implementação, Avaliação e Documentação.

A fase de Especificação tem por objetivo produzir um documento de especificação informal ou formal da ontologia em linguagem natural, que deve incluir, pelo menos, a finalidade, o nível de formalismo e o escopo da ontologia.

A fase de Aquisição do Conhecimento é uma atividade independente, que ocorre de forma simultânea com as outras fases, sendo mais intensa durante a especificação. As fontes de conhecimento

podem ser os especialistas, livros, guias, imagens, tabelas ou mesmo outras ontologias.

A fase de Conceitualização é aquela onde o conhecimento do domínio deve ser estruturado em um modelo conceitual que descreve o problema e sua solução em termos do vocabulário do domínio identificado na fase de especificação.

Na fase de Integração é verificada a possibilidade de reuso de definições já construídas em outras ontologias, em vez de partir dos rascunhos. Isto possibilita uma maior agilidade na construção da ontologia.

A Implementação das ontologies necessita que seja utilizado um ambiente de desenvolvimento que suporte a meta-ontologia e as ontologias selecionadas na fase de integração. Como resultado, tem-se a ontologia codificada em uma linguagem formal.

Avaliação consiste de um julgamento técnico das ontologias, seus ambientes de *software* e documentação com respeito a um quadro de referência, como o documento de especificação de requisitos. Compreende os termos Verificação (processo técnico que garante a correção da ontologia, do ambiente de *software* e da documentação) e Validação (garantia de que a ontologia, o ambiente de *software* e a documentação correspondem ao sistema que representam).

A Documentação deve ser feita durante todo o processo de desenvolvimento da ontologia, e os artefatos resultantes de cada fase, já fazem isso (FERNÁNDEZ; GÓMEZ-PÉREZ; JURISTO, 1997).

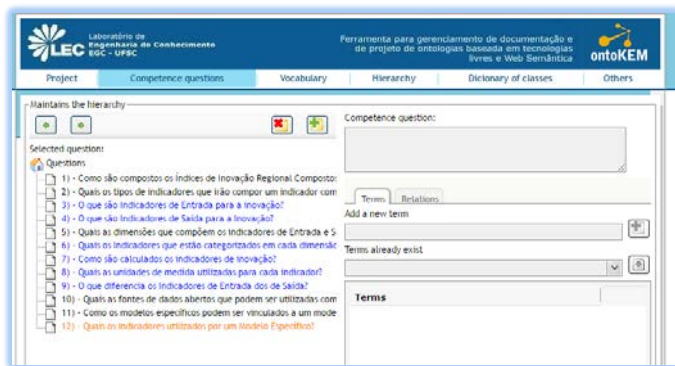
O ***Ontology Development 101*** é um guia desenvolvido na Universidade de Stanford, a partir da experiência com a utilização de editores de ontologia, em especial o *Protége*. O guia descreve uma abordagem iterativa para o desenvolvimento de ontologia, que começa com uma versão grosseira da ontologia, que vai sendo revisada e refinada. Os passos descritos são (NOY; MCGUINNESS, 2001): Determinar o domínio e o escopo da ontologia; Considerar o reuso de ontologias existentes; Enumerar os termos importantes na ontologia; Definir as classes e a hierarquia; Definir as propriedades das classes; Definir as facetas (cardinalidade, tipo, *domain* e *range*) dos *slots*; e Criar as instâncias.

Descrita como uma ferramenta *Web* para a criação e documentação de ontologia, o ***ontoKEM (ontology for Knowledge Engineering and Management)*** (TODESCO *et al.*, 2009) foi desenvolvido para ser utilizado com finalidade de ensino e pesquisa no âmbito do Laboratório de Engenharia do Conhecimento da Universidade

Federal de Santa Catarina, e utiliza conceitos das metodologias OTKM, METHONTOLOGY e 101.

No ontoKEM, as atividades de desenvolvimento de ontologia foram agrupadas em quatro fases: Competência; Vocabulário; Hierarquia; e, Dicionário de Classes(TODESCO *et al.*, 2009).

Figura 14 - Fase de competências no ontoKEM

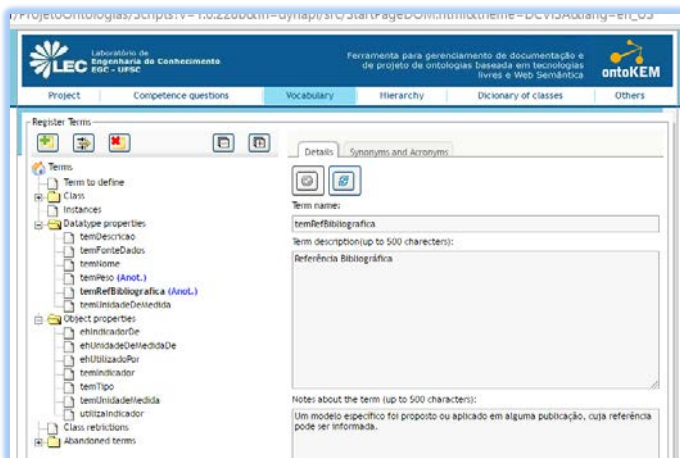


Fonte: o autor.

Na fase de Competência é feita a definição do escopo e do domínio, a partir de entrevistas com especialistas, gerando o documento de especificação de requisitos da ontologia (ORSO). São, então, listadas as questões de competência que a ontologia deve responder. A Figura 14 apresenta a página de cadastro das questões de competência no ontoKEM.

Na segunda fase, de Vocabulário, procede-se a definição precisa dos significados dos termos listados, por meio de uma descrição em linguagem natural. Os termos descritos são, então, classificados segundo o tipo de recurso que irão representar na ontologia. Começa-se, assim, a definição das classes, objetos, propriedades, restrições e sinônimos. Uma representação da página de cadastro de vocabulário no ontoKEM é apresentada na Figura 15.

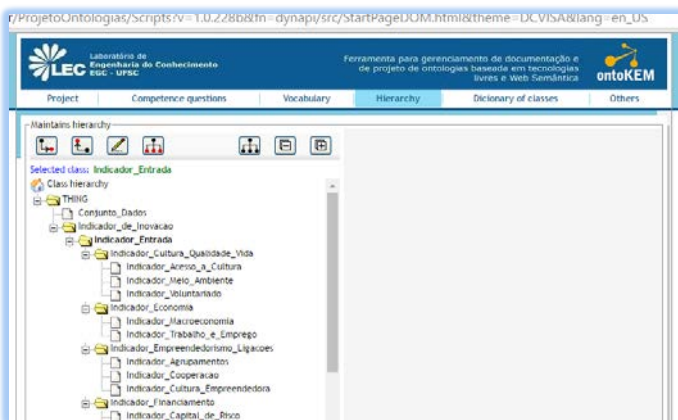
Figura 15 - Fase de Vocabulário no ontoKEM



Fonte: o autor.

Na fase de Hierarquia de Classes, os termos que foram apontados como classes serão organizados. A hierarquia das classes é definida pela informação de superclasses e subclasses de cada uma delas. Há, ainda, uma visualização gráfica desta hierarquia, conforme apresentada na .

Figura 16 - Fase de Hierarquia de classes

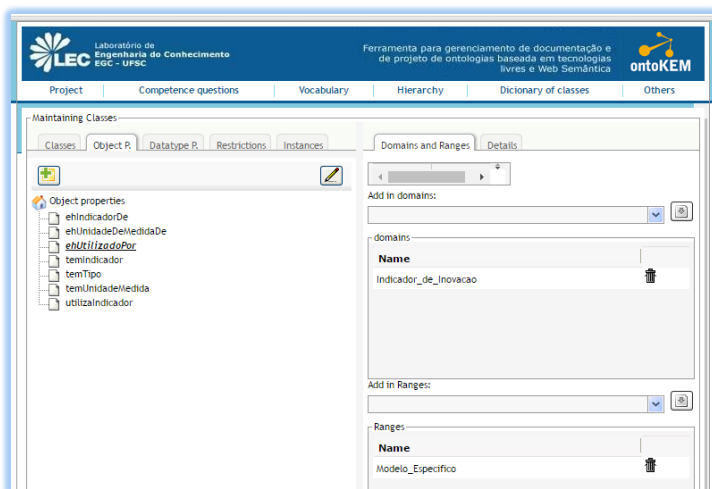


Fonte: o autor.

Na quarta fase, Dicionário de Classes, são definidos *domain* (domínio) e *range* (imagem) para as propriedades de dados e de objetos. Nesta fase são definidas também as restrições para as propriedades, tais

como inversa, funcional, simétrica ou transitiva. A apresenta a página de definição do dicionário de classes no ontokem.

Figura 17 - Fase de Dicionário de classes no ontoKEM



Fonte: o autor.

O ontoKEM foi concebido com o propósito de ser uma ferramenta para utilização em atividades acadêmicas, e um dos requisitos de sua construção foi a facilidade de uso, possibilitando que usuários com pouca ou nenhuma experiência na construção de ontologias pudessem fazer uso. Além disso, o fato de ser uma ferramenta *Web* permite que os usuários trabalhem remotamente, a partir de diferentes locais (TODESCO *et al.*, 2009).

Os artefatos produzidos pelo ontoKEM em cada uma das fases de construção das ontologias documentam o processo de construção. Além disso, a ontologia resultante pode ser exportada em linguagem OWL (*Ontology Web Language*), que é uma das especificações padrão para a *Web Semântica*, podendo ser publicada ou importada em outros editores de ontologia.

3.6 CONSIDERAÇÕES FINAIS

A Engenharia do Conhecimento é uma disciplina que se desenvolveu juntamente com os Sistemas Baseados em Conhecimento. Inicialmente regida por um paradigma de transferência de conhecimento a partir da cabeça dos especialistas, passou a ser vista como uma

disciplina de modelagem de aspectos do conhecimento humano e, por isso, posiciona-se de forma complementar à disciplina da Gestão do Conhecimento, fornecendo métodos e técnicas para a implementação de Sistemas de Conhecimento.

As ontologias fazem parte do ferramental da Engenharia do Conhecimento, especialmente no que concerne a modelagem e representação de conhecimento. Definidas, no âmbito da EC, como “uma especificação formal e explícita de uma conceitualização compartilhada” (BORST, 1997), descrevem domínios de conhecimento sobre os quais se deseja que os sistemas tratem. A formalização do conhecimento proporcionada pelas ontologias possibilita a atribuição de significado bem definido aos conceitos, bem como a definição de seus relacionamentos.

Tais características tornam as ontologias um dos componentes fundamentais da *Web Semântica* (BERNERS-LEE; HENDLER; LASSILA, 2001), pois além de classificar as coisas sobre as quais se está tratando, conta com regras para inferência, o que aumenta as possibilidades de processamento dos dados por máquinas.

O próximo capítulo aborda os Dados Ligados, um padrão de publicação de dados na *Web* que tem por objetivo tornar a *Web Semântica* uma realidade pela criação de um espaço global de dados, também chamado de *Web de Dados*.

4 DADOS LIGADOS

4.1 CONSIDERAÇÕES INICIAIS

Concebida no final da década de 1980, a *World Wide Web*, ou simplesmente *Web*, consiste de uma teia mundial cujos nós são documentos de vários tipos, interligados, e com uma interface de fácil acesso por meio da qual os usuários podem percorrê-los. Esta facilidade de uso fez com que a *Web* tenha vindo a se tornar um grande e importante repositório de informações criadas e consumidas por usuários em todas as partes do mundo.

Graças às suas características, a *Web* tornou-se plataforma para a publicação de conjuntos de dados por parte de organizações públicas e privadas, seguindo preceitos de transparência e participação da sociedade, implementando o conceito de dados abertos. Desta forma, é crescente o volume e a variedade de conjuntos de dados, com disponibilidade de acesso público e que podem ser utilizados conforme as necessidades dos usuários. Isso não garante, entretanto, que os conteúdos sejam amigáveis às máquinas. A falta de padronização e significado explicitamente estruturado faz com que os não estejam apropriados para o consumo por aplicações, dificultando seu melhor aproveitamento.

Considerada uma extensão da *Web* atual, a *Web Semântica* é uma abordagem cujo principal diferencial é que as informações nela publicadas têm seu significado bem definido. Assim, é necessário que os conteúdos atuais da *Web* sejam estruturados, de forma que as máquinas possam processar e compreender os dados que elas atualmente apenas exibem. Além do acesso aos dados, é necessário que estejam disponíveis os relacionamentos entre os dados, de forma a criar uma *Web* de Dados. Esta coleção de conjuntos dados inter-relacionados na *Web* também é chamada *Linked Data* (Dados Ligados).

Este capítulo apresenta, além dos Dados Abertos, os principais conceitos e técnicas relacionados à *Web Semântica* e, conseqüentemente, aos Dados Ligados. São apresentados os padrões basilares, bem como as características desta forma de publicação de dados na *Web*.

4.2 DADOS ABERTOS

Segundo a Open Definition (2009), dados são abertos quando qualquer pessoa pode livremente usá-los, reutilizá-los e redistribuí-los,

estando sujeito, no máximo, à exigência de creditar a sua autoria e compartilhá-los segundo a mesma licença. Devido às suas características, diversas iniciativas incentivam a produção e disponibilização de dados abertos, especialmente no âmbito governamental, como forma de aumentar a transparência e a participação dos cidadãos.

Para que os dados sejam considerados abertos, algumas características devem ser observadas:

- **Disponibilidade e Acesso:** os dados devem estar disponíveis como um todo a um custo não superior ao razoável para uma reprodução, preferencialmente por *download* via *internet*. Os dados devem também estar disponíveis em uma forma conveniente e modificável.
- **Reuso e Redistribuição:** Os dados devem ser providos segundo termos que permitam o reuso e a redistribuição incluindo a mistura com outros conjuntos de dados.
- **Participação Universal:** Todos devem poder usar, reusar e redistribuir os dados, sem que haja discriminação quanto à área de atuação, pessoas ou grupos.

Muitos indivíduos e organizações coletam um grande volume de dados de diferentes tipos para realizarem suas tarefas. No caso dos governos, a quantidade e a centralidade dos dados com que trabalham são ainda mais significativas. Além disso, a maioria dos dados de governo é pública por lei e, portanto, devem ser abertos e disponíveis para o uso de terceiros (Open Knowledge Foundation, 2012).

Em 2009, em seu primeiro dia de mandato, o presidente dos EUA, Barack Obama, emitiu um memorando sobre transparência e governo aberto, vindo a resultar na criação do portal *data.gov*. Muitos outros países seguiram esta tendência, criando seus próprios portais de dados abertos governamentais. A motivação dos governos pela disponibilização de dados abertos é justificada pela esperança do aumento do envolvimento dos cidadãos no governo, do aumento da transparência e da melhoria na tomada de decisão, alinhadas com a ambição de governos mais inteligentes (Bertot; Choi, 2013; Conradi; Choenni, 2014).

São muitos os grupos de pessoas e organizações que podem se beneficiar da disponibilidade de dados abertos governamentais, incluindo os próprios governos. Embora não seja possível prever

exatamente como e onde será criado valor no futuro, algumas áreas são apontadas como beneficiadas (OPEN KNOWLEDGE FOUNDATION, 2012):

Transparência e controle democrático; Participação; Auto-empoderamento; Produtos e serviços novos ou melhorados; Inovação; Melhora na eficiência dos serviços governamentais; Melhora na efetividade dos serviços governamentais; Mensuração de impacto das políticas; Novos conhecimentos a partir de fontes de dados combinadas e padrões em grandes volumes de dados.

Uma importante iniciativa na difusão de Dados Abertos Governamentais é a *Open Government Partnership* (OGP), criada em 2011, que provê uma plataforma internacional para que governo e sociedade civil trabalhem juntos para desenvolver e implementar reformas de governo aberto. Em seu primeiro Plano de Ação para a OGP, o Brasil firmou 32 compromissos que vieram a produzir importantes benefícios como a criação do Portal de Dados Abertos², a organização de uma conferência nacional sobre transparência, e a implementação da Lei de Acesso à Informação. A Lei nº 12.527/2011, conhecida como Lei de Acesso à Informação, que começou a vigorar no Brasil em maio de 2012, estabelece a obrigatoriedade de que os órgãos da administração pública disponibilizem dados abertos na Web, bem como de que respondam a solicitações dos cidadãos por novos conjuntos de dados que ainda não estejam disponíveis. O Artigo 8º da Lei estabelece:

“É dever dos órgãos e entidades públicas, independentemente de requerimentos, a divulgação em local de fácil acesso, no âmbito de suas competências, de informações de interesse coletivo ou geral por eles produzidas ou custodiadas” (BRASIL, 2011).

Mais especificamente em relação à divulgação na *Web*, o parágrafo 2º deste artigo dispõe:

“Para cumprimento do disposto no caput, os órgãos e entidades públicas deverão utilizar de todos os meios e instrumentos legítimos de que dispuserem, sendo obrigatória a divulgação em sítios oficiais da rede mundial de computadores (internet)” (BRASIL, 2011).

A política do governo brasileiro para dados abertos foi instituída pela Instrução Normativa Nº4, de 13 de abril de 2012, e foi nomeada de

² <http://dados.gov.br>

Infraestrutura Nacional de Dados Abertos (INDA). A INDA visa facilitar o acesso pelos cidadãos, pela sociedade e, em especial, pelas diversas instâncias do setor público aos dados e informações produzidas pelo Poder Executivo Federal, e tem por objetivos (MPOG, 2012):

- I. definir, estruturar e coordenar a política de dados abertos, bem como estabelecer o seu modelo de funcionamento;
- II. promover o ordenamento na geração, armazenamento, acesso, e compartilhamento de dados para uso do Poder Executivo federal e da sociedade;
- III. definir e disciplinar os padrões e os aspectos técnicos referentes à disponibilização e disseminação de dados para uso do Poder Executivo federal e da sociedade;
- IV. promover o compartilhamento de recursos de tecnologia da informação e evitar a duplicidade de ações e o desperdício de recursos na disseminação de dados e informações pelos órgãos e entidades do Poder Executivo federal;
- V. apoiar, capacitar e fornecer suporte para a publicação de dados abertos aos órgãos e entidades do Poder Executivo federal ou que aderirem à INDA que não possuem prática, cultura e atribuições finalísticas de disseminação de dados;
- VI. buscar a melhoria contínua da publicação de dados abertos, baseando-se nas melhores práticas concebidas nos cenários nacional e internacional;
- VII. promover a colaboração entre governos dos diferentes níveis da federação e entre o Poder Executivo federal e a sociedade, por meio da publicação e do reuso de dados abertos;
- VIII. promover e apoiar o desenvolvimento da cultura da publicidade de dados e informações na gestão pública;
- IX. disponibilizar tecnologias e apoiar as ações dos órgãos e entidades do Poder Executivo federal ou que aderirem à INDA na implementação da transparência ativa por meios digitais; e
- X. promover a participação social na construção de um ecossistema de reuso e de agregação de valor dos dados públicos.

De acordo com o balanço disponível no portal OGP do Governo Federal³, mais de 90% dos compromissos firmados no primeiro Plano de Ação da OGP foram implementados total ou parcialmente, além de outras 11 iniciativas de governo aberto que não estavam originalmente previstas. Atualmente, o Brasil está executando o segundo Plano de

³ <http://www.governoaberto.cgu.gov.br/no-brasil/planos-de-acao-1>

Ação, que conta com 52 compromissos, tendo 19 deles sido propostos por organizações da sociedade civil, e o envolvimento de 17 órgãos do governo.

4.2.1 Formato e Licenças de Dados Abertos

Para que os dados sejam considerados abertos devem ser estruturados e disponibilizados em um formato reutilizável. Isto significa dizer que os dados devem ser providos em um formato conveniente e modificável de forma tal que não haja obstáculos tecnológicos desnecessários para o bom desempenho dos direitos licenciados.

Especificamente, os dados devem ser legíveis por máquina, disponíveis em massa e providos em formato aberto (isto é, um formato com especificação livremente disponível e que não imponha restrições monetárias ou de qualquer outro tipo quanto ao seu uso) ou que, ao menos possa ser processado com pelo menos uma ferramenta de software livre/open-source (OPEN DEFINITION, 2009).

Exemplos de formatos abertos são *Comma Separated Values* (CSV), *Extensible Markup Language* (XML), *Javascript Object Notation* (JSON) e *Resource Description Framework* (DRF).

A preocupação com a liberdade de uso, cópia e reutilização tem origem nos movimentos de software livre. As licenças que permitem a livre reutilização de objetos dividem-se em licenças de software, conteúdo e dados.

Uma licença é aberta (*Open License*) quando seus termos satisfazem algumas condições, que podem ser obrigatórias ou aceitáveis (OPEN DEFINITION, 2009).

Condições obrigatórias:

- **Uso:** a licença deve permitir o livre uso do objeto licenciado;
- **Redistribuição:** a licença deve permitir a redistribuição do objeto licenciado, seja individualmente ou com parte de uma coleção de objetos de diferentes fontes;
- **Modificação:** a licença deve permitir a criação de derivações do objeto licenciado e permitir a distribuição desses derivados segundo os mesmos termos do objeto original;
- **Separação:** a licença deve permitir que qualquer parte do objeto seja livremente usada, distribuída ou modificada separadamente de qualquer outra parte do objeto ou de qualquer coleção de objetos nos quais tenha sido originalmente distribuída. Todas as partes que

recebam qualquer distribuição de qualquer parte de um objeto com os termos da licença original deve ter os mesmos direitos que são garantidos em conjunção com o objeto original;

- **Compilação:** a licença deve permitir que o objeto licenciado seja distribuído juntamente com outros objetos distintos sem impor restrições a estes objetos;
- **Não-discriminação:** a licença não pode impor discriminação sobre qualquer pessoa ou grupo;
- **Propagação:** os direitos vinculados ao objeto devem aplicar-se a todos a quem sejam redistribuídos sem que seja necessário concordar com qualquer termo legal adicional;
- **Aplicação para qualquer finalidade:** a licença deve permitir o uso, redistribuição, modificação e compilação para qualquer finalidade. A licença não pode restringir qualquer pessoa de fazer uso do objeto em um campo específico;
- **Sem-custo:** a licença não pode impor nenhuma taxa, *royalty*, ou outra compensação ou remuneração monetária como parte de suas condições.

Condições Aceitáveis:

- **Atribuição:** a licença pode requerer que distribuições do objeto incluam atribuição de colaboradores, detentores de direitos, patrocinadores e criadores desde que tais prescrições não sejam onerosas;
- **Integridade:** a licença pode requerer que versões modificadas de um objeto licenciado traga um diferente nome ou número de versão em relação ao objeto original, ou indicar quais modificações foram feitas;
- **Share-alike:** a licença pode requerer que as cópias ou derivados de um objeto licenciado permaneçam sob uma licença igual ou similar à do original;
- **Aviso:** a licença pode requerer a retenção de avisos de direitos autorais e identificação da licença;
- **Fonte:** a licença pode requerer que objetos modificados estejam disponíveis em uma forma preferida para modificações posteriores;
- **Proibição de restrições técnicas:** a licença pode proibir a distribuição do objeto em uma forma onde medidas técnicas imponham restrições para o exercício de outros direitos garantidos;
- **Não-agressão:** a licença pode requerer que os modificadores garantam permissões públicas adicionais (por exemplo, licenças de







patentes) como requisito para o exercício dos direitos garantidos pela licença. A licença pode também condicionar permissões de não agressão de licenciados a respeito do exercício dos direitos garantidos (por exemplo, litígio de patentes).

Embora as licenças de software e conteúdo livre sejam largamente utilizadas, no caso específico dos dados abertos, por serem ativos de naturezas diferentes, foi criado um conjunto de licenças específicas para dados e bancos de dados, conhecido como Open Data Commons, que estabelece três tipos de licença para dados (OPEN DATA COMMONS, 2008):

- ***Open Database License*** (ODbL): aplica-se a bancos de dados, permitindo o compartilhamento, a criação e adaptação. Requer atribuição de autoria e compartilhamento sob a mesma licença; Acrescenta-se, ainda a ***Database Content License*** (DbCL), com as mesmas exigências que a ODbL, mas aplicável ao conteúdo dos bancos de dados;
- ***Open Data Commons Attribution License***: aplicável aos dados e a bancos de dados, mas com exigência apenas de atribuição da autoria;
- ***Public Domain Dedication and License***: a PDDL posiciona o banco de dados e seu conteúdo como sendo de domínio público, renunciando a todos os direitos.

Além das licenças específicas para dados abertos, em muitos casos há a aplicação das licenças *Creative Commons*, quando consideradas adequadas aos dados abertos. *Creative Commons* é uma organização sem fins comerciais que tem por objetivo possibilitar uma forma simples e padronizada de compartilhar e usar criatividade e conhecimento através de ferramentas legais (Creative Commons, 2001). No Quadro 6 são apresentados seus tipos.

Quadro 6 – Tipos de licenças *Creative Commons*

 <p>Attribution</p>	<p>Permite que terceiros distribuam, recontextualizem, adaptem e construam a partir de um objeto, mesmo comercialmente, desde que deem o crédito pela criação original. Esta licença é a mais amigável, sendo recomendada para a máxima disseminação e uso dos materiais licenciados;</p>
 <p>Attribution- NoDerivs</p>	<p>Permite a redistribuição, para uso comercial e não-comercial, desde que seja repassado sem modificações e por completo, dando crédito de autoria;</p>
 <p>Attribution- NonCommercial- ShareAlike</p>	<p>Permite que terceiros recontextualizem, adaptem e construam derivados a partir de um objeto, para uso não comercial, desde que atribuam o crédito pela autoria e licenciem as novas criações sob termos idênticos;</p>
 <p>Attribution- ShareAlike</p>	<p>Permite que terceiros recontextualizem, adaptem e construam derivados a partir de um objeto, mesmo que para fins comerciais, desde que atribuam crédito e licenciem suas criações sob termos idênticos. Esta licença é comparada às licenças de software livre. Todos os novos trabalhos baseados naquele objeto irão carregar a mesma licença, de forma que qualquer derivado irá também permitir uso comercial;</p>
 <p>Attribution- NonCommercial</p>	<p>Permite que terceiros recontextualizem,. Adaptem e construam a partir de um objeto, para fins não-comerciais. Embora os novos trabalhos devam atribuir a autoria do trabalho original e serem não-comerciais, não precisam licenciar os trabalhos derivados sob os mesmos termos do original;</p>
 <p>Attribution- NonCommercial- NoDerivs</p>	<p>É o tipo de licença mais restritiva das seis, permitindo apenas que terceiros façam download de objetos e os compartilhem, desde que atribuam crédito de criação, sem poder fazer alterações de nenhum tipo ou fazer uso comercial.</p>

Fonte: Creative Commons (2001)

Além das características atribuídas aos dados abertos, bem como das licenças existentes, há as três “leis” de dados abertos, propostas por David Eaves na Conferência para Parlamentares – Transparência para a Era Digital (EAVES, 2009):

- Se o dado não puder ser encontrado e indexado na *Web*, ele não existe;

- Se não estiver aberto e disponível em formato compreensível por máquina, ele não pode ser reaproveitado;
- Se algum dispositivo legal não permitir sua replicação, ele não existe.

As leis de Eaves foram propostas para Dados Abertos Governamentais, mas aplicam-se a quaisquer Dados Abertos, e resumem de forma bastante direta o que se espera destes. A preocupação de que os dados sejam compreensíveis por máquinas revela a necessidade cada vez maior na atenção aos padrões e técnicas que visam acrescer significado bem definido aos dados, características da *Web Semântica*.

4.3 WEB SEMÂNTICA

A concepção da *Web* data do final da década de 1980, quando Tim Berners-Lee publicou uma proposta de um sistema para compartilhamento de documentos de pesquisa, chamando-a de *World Wide Web*. Posteriormente, um consórcio foi criado para desenvolver a *Web* ao seu máximo potencial. Este consórcio, chamado *World Wide Web Consortium* (W3C) é responsável pela definição de grupos de trabalho que atuam na criação e/ou reconhecimento de padrões, chamados recomendações, que passam a ser adotados no desenvolvimento de tecnologias para a *Web*.

A *Web* é um espaço de informação na qual os itens de interesse, chamados recursos, são identificados por identificadores globais, os *Uniform Resource Identifiers* (URI). A Identificação, juntamente com a Interação e os Formatos, constituem as três bases arquiteturais da *Web* (JACOBS; WALSH, 2004).

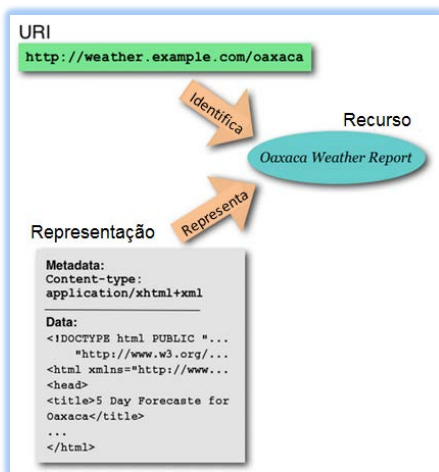
A Identificação é feita por meio de URIs, identificadores únicos para cada recurso. A Interação acontece quando os agentes da *Web* (pessoas ou softwares) se comunicam por meio de protocolos padronizados de troca de mensagens. Quando um recurso é solicitado através de um endereço ou pela seleção de um *hyperlink*, o navegador faz uma requisição pelo URI do recurso via protocolo *HyperText Transfer Protocol* (HTTP) a um servidor, que envia uma mensagem de volta com a representação do recurso. A representação do recurso, disponibilizado pelo servidor, contém dados e metadados codificados segundo um formato específico, comumente o *eXtensible HyperText Markup Language* (XHTML).

De acordo com o RFC 3986 (BERNERS-LEE; FIELDING; MASINTER, 2005) um *Uniform Resource Identifier* (URI) é uma sequência compacta de caracteres que identifica um recurso físico ou abstrato. URI é uma das tecnologias que fundamenta o funcionamento da *Web*, pois fornecem um meio genérico para identificar qualquer entidade que existe, seja um documento ou descrição deste (Bizer, Heath and Berners-Lee, 2009).

O navegador é um aplicativo configurado para interpretar a representação codificada do recurso, recebida como resposta a uma requisição, apresentando seu conteúdo ao usuário. A

Figura 14 ilustra a relação entre um recurso, sua identificação (URI) e sua representação, segundo um formato XHTML.

Figura 18 - Relação entre identificador, recurso e representação na *Web*



Fonte: Jacobs e Walsh (2004)

Embora a arquitetura preveja que os agentes da *Web* podem ser tanto pessoas como programas de computador, a sua evolução mostra que a grande maioria dos conteúdos nela disponibilizados costumam ser codificados de forma que sejam adequados para o entendimento humano. Todavia, embora os computadores possam analisar habilmente páginas *Web* pelo processamento de seu layout, não processam a semântica dos conteúdos.

Proposta por Berners-Lee; Hendler e Lassila (2001), a *Web Semântica* é uma extensão da *Web* tradicional, na qual as informações

recebem significado bem definido, possibilitando que pessoas e computadores trabalhem em cooperação. Para seu funcionamento, os computadores devem ter acesso a coleções de informações e conjuntos de regras de inferência que possam usar para seu raciocínio automatizado. Duas tecnologias importantes para o desenvolvimento da *Web Semântica* são o *Extensible Markup Language* (XML) e o *Resource Description Framework* (RDF).

O XML é uma linguagem de marcação que permite que conteúdos sejam marcados com *tags*, de forma que as páginas *Web* ou parte delas passe a conter anotações. Diferentemente do que acontece no HTML, as marcações no XML não se destinam à apresentação dos conteúdos, mas sim à sua estrutura. Embora os conteúdos possam ser estruturados de acordo com a necessidade do desenvolvedor, as marcações não trazem nenhuma indicação sobre o seu significado (BOSAK; BRAY, 1999).

Figura 19 - Trecho de documento XML contendo informações sobre artigo científico

```
<ref-type name="Journal Article">17</ref-type>
▼ <contributors>
  ▼ <authors>
    ▼ <author>
      <style face="normal" font="default" size="100%">Adams, Richard</style>
    </author>
    ▼ <author>
      <style face="normal" font="default" size="100%">Bessant, John</style>
    </author>
    ▼ <author>
      <style face="normal" font="default" size="100%">Phelps, Robert</style>
    </author>
  </authors>
</contributors>
▼ <titles>
  ▼ <title>
    <style face="normal" font="default" size="100%">Innovation management measurement: A review</style>
  </title>
  ▼ <secondary-title>
    <style face="normal" font="default" size="100%">International Journal of Management Reviews</style>
  </secondary-title>
</titles>
▼ <periodical>
  ▼ <full-title>
    <style face="normal" font="default" size="100%">International Journal of Management Reviews</style>
  </full-title>
</periodical>
▼ <pages>
  <style face="normal" font="default" size="100%">27</style>
```

Fonte: O Autor

A Figura 15 apresenta o exemplo de um trecho de documento contendo dados sobre artigos científicos, representado em sintaxe XML, onde as *tags* são representadas pelos elementos identificados pelos

caracteres < e >. No exemplo, as *tags* <author> e </author> delimitam um trecho de conteúdo que representa o nome do autor. Já as *tags* <title> e </title> delimitam um trecho de conteúdo que representa o título do artigo.

Desde 1996, o XML é desenvolvido por um grupo de trabalho que é parte do *World Wide Web Consortium*, passando a ser uma recomendação W3C no ano de 1998. Dez objetivos de projeto para o XML são descritos no documento de recomendação (BRAY *et al.*, 2008):

1. XML deve ser diretamente utilizável na internet;
2. XML deve suportar uma ampla variedade de aplicações;
3. XML deve ser compatível com *Standard Generalized Markup Language* (SGML);
4. Deve ser fácil escrever programas que processam XML;
5. O número de recursos opcionais em XML deve ser mantido ao mínimo absoluto, idealmente zero;
6. Os documentos XML devem ser legíveis por humanos e razoavelmente claros;
7. O projeto XML deve ser preparado rapidamente;
8. O projeto do XML deve ser formal e conciso;
9. Os documentos XML devem ser fáceis de criar;
10. A concisão é de mínima importância nas marcações XML.

Os objetivos de projeto demonstram, portanto a preocupação no desenvolvimento de uma linguagem que se mantivesse simples e de fácil utilização, o que a tornou um importante padrão para o desenvolvimento de aplicações *Web*.

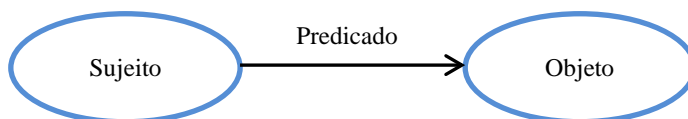
Considerada outra tecnologia fundamental para a *Web* semântica, o RDF é um *framework* para representação de informações na *Web*. O RDF permite representar grafos genéricos baseados no modelo de dados com os quais é possível descrever a estrutura de dados e links usada para representar as coisas no mundo (MANOLA; MILLER, 2004; MANOLA; MILLER; MCBRIDE, 2014).

A estrutura central da sintaxe do RDF é um conjunto de triplas, cada uma consistindo de um sujeito, um predicado e um objeto. Um conjunto destas triplas é chamado de grafo. Um grafo RDF pode ser visualizado como um diagrama de nós e arcos direcionais, onde cada tripla é representada como uma ligação nó-arco-nó, conforme representação da Figura 16. Os nós em um grafo RDF podem ser de três tipos: IRI, literal ou nó em branco (W3C, 2014).

Os *Internationalized Resource Identifiers* (IRIs) são uma generalização das URIs que permitem uma gama maior de caracteres

Unicode. No RDF, qualquer IRI ou literal denota um recurso ou entidade, que pode ser qualquer coisa, incluindo coisas físicas, documentos, conceitos abstratos, números e textos. O recurso denotado por uma IRI é o seu “referente”, enquanto que o recurso denotado por um literal é o seu “valor literal”. A asserção de uma tripla RDF representa que existe alguma relação, indicada pelo **predicado**, entre os recursos denotados pelo **sujeito** e **objeto** (W3C, 2014).

Figura 20 - Grafo RDF com dois nós (Sujeito e Objeto) e conexão entre eles (Predicado)



Fonte: Adaptado de W3c (2014)

Uma tripla RDF consiste, portanto, de três componentes, que são convencionalmente escritos na seguinte ordem:

- O **sujeito**, que é uma IRI ou um nó em branco
- O **predicado**, que é uma IRI
- O **objeto**, que é uma IRI, um literal ou um nó em branco

O **sujeito** e o **objeto** representam os dois recursos que se relacionam, e o **predicado** representa a natureza do relacionamento. O relacionamento pode ser descrito por uma frase de forma direcional, do sujeito para o objeto, e é chamada em RDF de **propriedade** (Manola; Miller; McBride, 2014). A Figura 17 apresenta um exemplo de triplas, informalmente escritas em pseudocódigo.

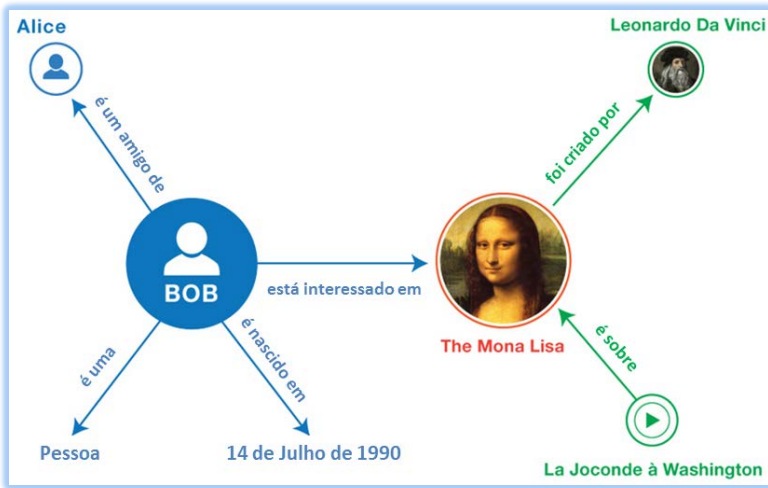
Figura 21 - Exemplo de triplas RDF escritas em pseudocódigo

```
<Bob> <é uma> <Pessoa> .
<Bob> <é um amigo de> <Alice> .
<Bob> <é nascido em> <14 de Julho de 1990> .
<Bob> <está interessado em> <the Mona Lisa> .
<the Mona Lisa> <foi criado por> <Leonardo da Vinci> .
<the vídeo 'La Jaconde à Washington' <é sobre> <the Mona Lisa>
```

Fonte: Manola; Miller e McBride (2014)

Um mesmo recurso pode ser referenciado múltiplas vezes nas diferentes triplas. Esta característica permite que um recurso conste como sujeito em uma tripla e como objeto em outra, como é o caso do recurso representado por “<the Mona Lisa>” no exemplo, o que torna possível encontrar as conexões entre as triplas. O mesmo exemplo pode ser representado por um grafo, conforme representado na Figura 18.

Figura 22 - Exemplo de grafo RDF



Fonte: Manola; Miller e McBride (2014)

O conjunto de nós de um grafo RDF é o conjunto de sujeitos e objetos das triplas no grafo. Além disso, é possível que uma IRI de um predicado também seja um nó no mesmo grafo. Um *dataset* RDF, por sua vez, é uma coleção de grafos RDF.

Além da representação gráfica, o RDF pode ser serializado em diferentes formatos. A serialização consiste no armazenamento, seja em arquivo ou em memória, ou transmissão por meio de conexão de rede, em forma binária ou de arquivo de texto. Os formatos RDF/XML e RDFa são considerados padrões para serialização de RDF pela W3C (Heath; Bizer, 2011). Além destes, citam-se os formatos Turtle, N-Triple e RDF/JSON como alternativas para serialização.

Além do XML e do RDF, citam-se, ainda, como tecnologias semânticas: as linguagens de representação de conhecimento, especificamente RDFS e OWL; os mecanismos de regras e intercâmbio; os *web services* semânticos; e as consultas semânticas (KONSTANTINOU *et al.*, 2010).

O *Resource Description Framework Schema* (RDFS ou RDF *Schema*) provê um vocabulário para modelagem de dados, sendo uma extensão semântica para o RDF básico. O vocabulário conta com mecanismos para descrever grupos de recursos relacionadas e as relações entre eles. Estes recursos são usados para determinar características de outros recursos, tais como *domain* e *range* das propriedades de classes e objetos, também definidos no RDF *Schema* (BRICKLEY; GUHA, 2014).

A *Ontology Web Language* (OWL) (OWL WORKING GROUP, 2012), por sua vez, é uma linguagem de *Web Semântica* projetada para representar conhecimentos ricos e complexos sobre coisas, grupos de coisas e relações entre elas. Trata-se de uma linguagem computacional baseada em lógica que permite que o conhecimento nela expresso possa ser explorado por programas computacionais, de forma a verificar sua consistência ou realizar inferências.

A linguagem OWL é o padrão recomendado para a serialização de documentos de ontologias, de forma que podem ser publicados na *Web*, interligado com outras ontologias, bem como serem ligados a conjuntos de dados, de forma a descrevê-los.

4.4 DADOS LIGADOS (*Linked Data*)

O objetivo primordial da *Web Semântica* é a construção de uma *Web* de Dados. Para isso, é necessário que existam dados em grande quantidade disponíveis na *Web* em um formato padrão, acessíveis, interligados e gerenciáveis por ferramentas de *Web Semântica*. Esta coleção de conjuntos de dados inter-relacionados na *Web* pode também ser chamada de Dados Ligados⁴ (*Linked Data*).

Uma revisão sistemática de literatura sobre Dados Ligados foi realizada no âmbito desta tese, com o objetivo de identificar as mais relevantes publicações científicas, bem como seus autores, grupos de pesquisa e localização, além das principais referências bibliográficas adotadas. Os resultados desta revisão são apresentados em Dal Pizzol *et al.* (2015), onde são listadas, ainda, as principais temáticas abordada nas publicações.

O termo Dados Ligados, proposto por Berners-Lee (2006), refere-se a um estilo de publicar e interligar dados estruturados de diferentes

⁴ Outras formas de tradução para *Linked Data* encontradas na literatura são: “Dados *Linkados*” e “Dados *Vinculados*”. Neste trabalho, optou-se pela utilização de “Dados *Ligados*”, por ser a tradução mais difundida até este momento.

fontes na *Web*. O assunto vem despertando considerável interesse acadêmico (BIZER; HEATH; BERNERS-LEE, 2009; BRADLEY, 2009; HEATH, 2011) e diversas ações estão sendo desenvolvidas no sentido da criação de repositórios de dados ligados. O objetivo é a construção de um espaço global de dados baseado em padrões abertos, a *Web* de Dados, cuja proposta é ser uma extensão da *Web* atual, através de um paradigma de publicação, não apenas documentos, mas também de dados (Heath; Bizer, 2011)

Para que seja possível atingir o objetivo da *Linked Data*, de forma que os dados passem a fazer parte de um repositório global, Berners-Lee (2006) apresenta quatro regras, conhecidas como os princípios da *Linked Data*:

- 1) Deve-se usar URIs (*Uniform Resource Identifiers*) como nomes para as coisas;
- 2) Utilizar HTTP URIs (*Hyper Text Transfer Protocol*) de modo que as pessoas possam procurar por esses nomes;
- 3) Quando alguém procurar um URI, fornecer informações úteis usando os padrões RDF (*Resource Description Framework*) e SPARQL (linguagem de consulta); e
- 4) Incluir links para outros URIs, a fim de que se possa descobrir mais coisas.

As quatro regras fornecem, portanto, uma receita básica para a publicação e conexão de dados usando a infraestrutura da web em aderência à sua arquitetura e padrões. Diferentemente da web de hipertextos, entretanto, onde as relações são ligações em documentos escritos em HTML, na web de dados as ligações são estabelecidas entre "coisas" arbitrárias descritas por RDF e as URIs identificam qualquer tipo de objeto ou conceito (Berners-Lee, 2006).

As práticas de *Linked Data* oferecem um mecanismo simples para combinação de múltiplas fontes na *Web* (HYLAND e WOOD, 2011), fornecendo uma gama de padrões internacionais e melhores práticas para a publicação, divulgação e reutilização de dados estruturados. A *Web* de dados ligados pode ser vista como uma camada adicional que é estreitamente entrelaçada com a *Web* clássica de documentos e tem muitas propriedades em comum (BIZER, 2009):

- Qualquer um pode publicar dados na *Web* de dados ligados;
- *Links* conectam entidades, criando um grafo de dados global que se estende e possibilita a descoberta de novas fontes de dados;

- Os dados são auto-descritos. Se uma aplicação encontra dados representados usando um vocabulário que não lhe é familiar, pode resolver as URIs que identificam os termos do vocabulário para encontrar suas definições;
- A *Web* de dados ligados é aberta, significando que aplicações podem descobrir novas fontes de dados seguindo os *links* durante sua execução.

Há muitos usos internos importantes para dados ligados, e para uso em dados pessoais e de grupos (BERNERS-LEE, 2006). Assim como muitos dados abertos não são dados ligados, os dados publicados na forma de *Linked Data* não precisam, necessariamente, ser abertos.

4.4.1 Dados Abertos Ligados (*Linked Open Data*)

O termo *Linked Open Data* (LOD), segundo Berners-Lee (2006), refere-se a dados ligados (*Linked Data*) que são disponibilizados sob uma licença aberta, que não impede o seu livre reuso.

No ano de 2010, Tim Berners-Lee acrescentou ao documento de *Design Issues* (BERNERS-LEE, 2006), um sistema de avaliação dos dados com estrelas. O objetivo, segundo o autor, era encorajar as pessoas, especialmente os responsáveis por dados governamentais, ao caminho do bom *Linked Data*. Pelo sistema de estrela, detalhado no Quadro 7, os dados são classificados segundo seu poder e facilidade de uso para as pessoas.

Dentre as iniciativas relacionadas à publicação de dados ligados abertos, uma das mais importantes é o projeto *Linking Open Data*, uma comunidade fundada em 2007 com o objetivo contínuo de dar início à *Web* de Dados por meio da identificação dos conjuntos de dados existentes disponíveis sob licenças abertas, convertendo-os em RDF de acordo com os princípios de *Linked Data*, e publicando-os na *Web*. Desde então, um número crescente de provedores de dados passaram a adotar estes princípios, conduzindo à criação de um espaço global de dados contendo bilhões de asserções sobre localizações geográficas, pessoas, companhias, livros, publicações científicas, filmes, música, programas de rádio e televisão, genes, proteínas, medicamentos e experimentos médicos, comunidades online, dados estatísticos, resultados de censos e publicações (BIZER, 2009).

Quadro 7 - As cinco estrelas dos Dados Abertos

★	Disponível na <i>web</i> (independente do formato) mas com uma licença aberta, para ser <i>Open Data</i> .
★★	Disponível como dado estruturado legível por máquinas (ex. excel ao invés de uma imagem digitalizada de uma tabela).
★★★	Mesmo que duas estrelas, mas com formato não proprietário (ex. CSV ao invés de excel).
★★★★	Todas as anteriores, mais o uso de padrões abertos da W3C (RDF e SPARQL) para identificar as coisas, de maneira que as pessoas possam apontar para elas.
★★★★★	Todas as anteriores, mais: Ligue seus dados aos dados de outras pessoas para prover contexto.

Fonte: Berners-Lee (2006)

No âmbito do projeto *Linking Open Data* o grafo de dados interligados é representado pela nuvem de dados ligados *LOD Cloud*, representada na Figura 19. Segundo as estatísticas do projeto, em Setembro de 2011 esta nuvem contava com 295 *datasets*, de sete diferentes domínios, contendo mais de 31 bilhões de triplas (JENTZSCH; CYGANIAK; BIZER, 2011). Em Agosto de 2014, já eram 570 *datasets*. Destes 374 foram descritos pelos próprios provedores dos dados no catálogo *datahub*⁵. Os outros 196 foram descobertos por meio de uma iniciativa de busca por ferramentas de *crawling*, gerando novas estatísticas sobre os dados (SCHMACHTENBERG; BIZER; PAULHEIM, 2014).

Para que alguém venha a se tornar um participante do projeto, basta que publique um conjunto de dados seguindo os princípios de *Linked Data* e que seus dados tenham interligação com os conjuntos de dados existentes na nuvem.

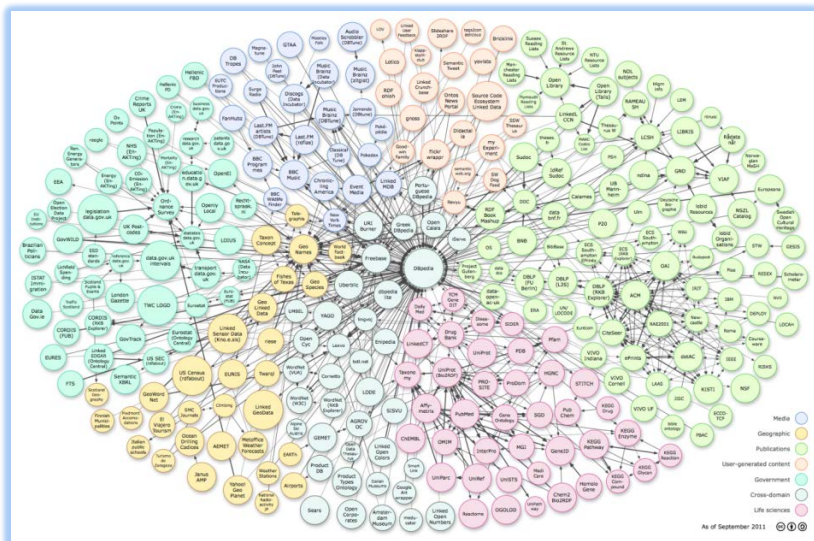
Dependendo das características dos dados ou de sua aplicação, os dados ligados recebem diferentes denominações na literatura. É o caso de termos como *Linked Government Data* (HYLAND; WOOD, 2011; VILLAZÓN-TERRAZAS *et al.*, 2011), *Enterprise Linked Data* (SERVANT, 2008; HU; SVENSSON, 2010; ALLEMANG, 2011), *Statistical Linked Data* (KÄMPGEN; O'RAIN; HARTH, 2012; ZANCANARO *et al.*, 2013; CYGANIAK; REYNOLDS; TENNISON, 2014), *Sensor Linked Data* (PAGE *et al.*, 2009; PATNI; HENSON; SHETH, 2010).

Referida como *Linked Government Data*, a aplicação dos princípios de *Linked Data* a bases de dados do governo traz um enorme

⁵ <http://datahub.io>

potencial (HEATH; BIZER, 2011), Para Cyganiak; Maali e Peristeras (2010), no entanto, esse potencial é pouco explorado principalmente por causa da falta de recursos necessários para transformar em grande escala os dados brutos de alta qualidade. É neste contexto que *linked data* são fundamentais, permitindo um esquema aberto e extensível com informações explicativas sobre o domínio, interligado com dados relacionados de outras fontes (HYLAND; WOOD, 2011).

Figura 23 – Diagrama da nuvem *Linking Open Data*



Fonte: Jentzsch; Cyganiak e Bizer (2011)

Villazón-Terrazas *et al.* (2011) apresentam um processo, na forma de um modelo cíclico incremental, com diretrizes, para o processo de publicação de *Linked Open Government Data*, que conta com as seguintes atividades principais: 1. especificação, 2. modelagem, 3. geração, 4. a publicação, e 5. exploração. Cada atividade é decomposta em uma ou mais tarefas, e algumas técnicas e ferramentas são necessárias para sua realização.

Governos, como os do Canadá, Estados Unidos, Reino Unido, França, Espanha, Suécia, Itália, Hong Kong, Austrália, Nova Zelândia e Brasil já possuem sites dedicados a tornar acessíveis dados públicos, e alguns desses já estão disponíveis como *Linked Data*.

Na medida em que a iniciativa *Linked Data* passou a demonstrar seu valor em projetos acadêmicos e com dados públicos, as histórias de

sucesso passaram a chamar a atenção de grandes companhias (HU; SVENSSON, 2010). Segundo Allemang (2011), as experiências da *Web* sugerem que há um caminho para uma empresa construir uma arquitetura sustentável de informações empresariais, transformando-se em uma “*Linked Enterprise Data*”, na qual o ato de criação da informação está intimamente ligado com o ato de compartilhamento da informação, e o compartilhamento de dados é tão importante como sua produção.

Uma revisão sistemática de literatura, mais específica, sobre o tema “*Linked Enterprise Data*” é descrita por Speroni; Trindade; *et al.* (2014), onde é apresentado um levantamento sobre as publicações científicas sobre o assunto, principais autores e referências utilizadas.

Abordagens como as apresentadas por Servant (2008) e Hu e Svensson (2010) descrevem casos de aplicação de *Linked Data* em grandes companhias, através da utilização de tecnologias de informação e comunicação já utilizadas pelas empresas, da publicação de seus dados internos na forma de *Linked Data*, promovendo a ligação dos mesmos com fontes externas, além da implementação de ferramentas que fazem acesso aos dados, apresentando novas possibilidades de visualização dos mesmos.

4.4.2 Publicação de Dados Ligados

O principal diferencial dos Dados Ligados é justamente o conjunto de padrões e práticas adotadas para a sua publicação, e sua simplicidade é abordada por Tim Berners-Lee:

"Linked data é essencial para conectar efetivamente a Web semântica. É bastante fácil de fazer sem pensar muito, e torna-se natural. Várias considerações de senso comum determinam quando fazer um link e quando não fazê-lo" (Berners-Lee, 2006).

Publicar um conjunto de dados ligados na *Web* envolve as seguintes etapas básicas (BIZER; HEATH; BERNERS-LEE, 2009):

1. **Atribuir URIs** às entidades descritas pelo conjunto de dados;
2. Definir **ligações** (*links*) RDF para outras fontes de dados na *Web*, possibilitando navegar pela rede de dados, seguindo os *links* RDF;

3. Fornecer **metadados** sobre os dados publicados, de modo que se possa avaliar a qualidade dos dados publicados e escolher entre diferentes meios de acesso.

A atribuição de URIs às entidades consiste na definição de um identificador para cada uma delas, que podem ser expressas na forma de URIs 303 ou URIs *hash* (SAUERMAN; CYGANIAK, 2008). Em um ambiente aberto como a *Web*, entretanto é comum que diferentes provedores de informações publiquem dados sobre uma mesma entidade ou recurso, como uma localização geográfica ou uma celebridade, definindo diferentes URIs para identificar a mesma entidade. Como ambas as URIs referem-se à mesma entidade, são chamadas de *aliases* URIs (BIZER; HEATH; BERNERS-LEE, 2009).

Para Bizer; Heath e Berners-Lee (2009), as *aliases* URIs provêm uma importante função social na *Web* de Dados, uma vez que levam a diferentes descrições da entidade, possibilitando diferentes visões e opiniões expressas na *Web*. Para garantir que seja possível rastrear diferentes provedores de informação falando de uma mesma entidade, é prática comum o uso de ligações do tipo *owl:sameAs* para *aliases* conhecidos.

Além da interligação de dados, é uma boa prática a reutilização de **vocabulários** sempre que possível. Conforme lembra (ALLEMANG, 2011), as empresas geralmente têm vocabulários refinados para seu uso interno, mas algumas têm **vocabulários, ontologias ou taxonomias** que podem ser úteis a outras, e que estes devem ser convertidos para formatos que sejam padrões internacionais. Para Bizer; Heath e Berners-Lee (2009) citam os vocabulários FOAF, SIOC, SKOS, DOAP, vCard, *Dublin Core*, OAI-ORE e *GoodRelations* como exemplos de vocabulários conhecidos.

A recomendação é de que apenas nos casos em que estes vocabulários já conhecidos não provejam os termos requeridos, é que os publicadores de dados devem definir novas terminologias específicas (BIZER; CYGANIAK; HEATH, 2008). Se uma nova terminologia for definida, deve ser auto descritiva, com URIs “derefenciáveis”⁶ na

⁶ O termo *dereference* é utilizado na computação para obter o endereço de um item de dado a partir de um ponteiro. Em *linked data*, uma URI *dereferenceable* é um mecanismo de recuperação de recurso que utiliza um dos protocolos da internet para obter uma cópia ou representação do recurso que ela identifica.

Web, de forma que os clientes possam recuperar o RDF *Schema* ou as definições dos termos, além de mapeá-los para outros vocabulários.

Para ser parte de *Web* de Dados, portanto, as fontes de dados devem manter *links* RDF para entidades relacionadas em outras fontes de dados. Os *links* RDF possibilitam que aplicações cliente naveguem entre fontes de dados e descubram dados adicionais. Normalmente, a criação dos *links* RDF é feita de forma automatizada ou semi-automatizada (BIZER; HEATH; BERNERS-LEE, 2009).

A disponibilização de metadados, por sua vez, tem o objetivo de possibilitar aos clientes a verificação da qualidade dos dados publicados e de determinar se os dados são confiáveis. São comumente utilizados metadados relativos à autoria, método, data de criação e proveniência dos dados, além de metadados técnicos sobre as formas de acesso (BIZER; HEATH; BERNERS-LEE, 2009).

Além das etapas recomendadas, envolvendo a identificação dos dados com URIs, a interligação com outras fontes e a criação de metadados, a publicação de *Linked Data* requer que estes dados estejam disponíveis aos clientes. Para isto, ferramentas específicas foram desenvolvidas com o objetivo de disponibilizar na *Web* conteúdos na forma de *Linked Data*.

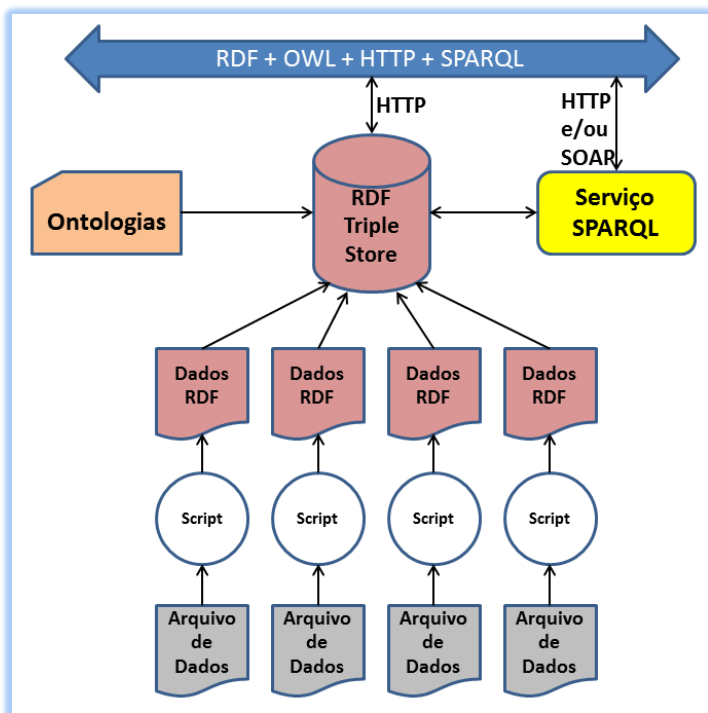
As ferramentas podem tanto servir conteúdos RDF armazenados na forma de *Linked Data* como prover visões de *Linked Data* a partir de bases de dados legadas não-RDF, facilitando o trabalho dos publicadores de conteúdo por não precisarem lidar com detalhes técnicos tais como negociação de conteúdo, provendo descrições RDF para as URIs (*dereference*) e consultas SPARQL (BIZER; CYGANIAK; HEATH, 2008; SAUERMANN; CYGANIAK, 2008; BIZER; HEATH; BERNERS-LEE, 2009).

Em *Linked Data*, a representação dos dados em RDF corresponde a um grafo, onde cada sequência nó-arco-nó é chamada de tripla, que representa uma relação sujeito-predicado-objeto. Um ***Triple Store*** é banco de dados projetado para armazenar e recuperar entidades que são construídas por coleções de triplas (RUSHER, 2003).

Os dados transformados em RDF, juntamente com as ontologias que representam as classes e propriedades do domínio de conhecimento, são armazenados em *triple stores*, ficando disponíveis para acesso via navegador *Web*, ou via aplicações, diretamente ou por meio de consultas a um *endpoint* SPARQL. A representação de um exemplo de arquitetura

de publicação dos dados RDF em um *triple store* é apresentada na Figura 20.

Figura 24 - Arquitetura de publicação de dados RDF em um Triple Store



Fonte: Berners-Lee (2005)

Há diversas ferramentas para armazenamento de dados RDF, como o *Virtuoso Universal Server*⁷, Jena⁸, Sesame⁹, 4Store¹⁰, YARS¹¹ e OWLIM¹² (VILLAZÓN-TERRAZAS *et al.*, 2011). Algumas delas já contam com recursos como *endpoint* SPARQL, que permite realizar consultas, e *Linked Data frontend*, que possibilita a navegação nos dados.

⁷ <http://virtuoso.openlinksw.com/>

⁸ <http://jena.sourceforge.net/>

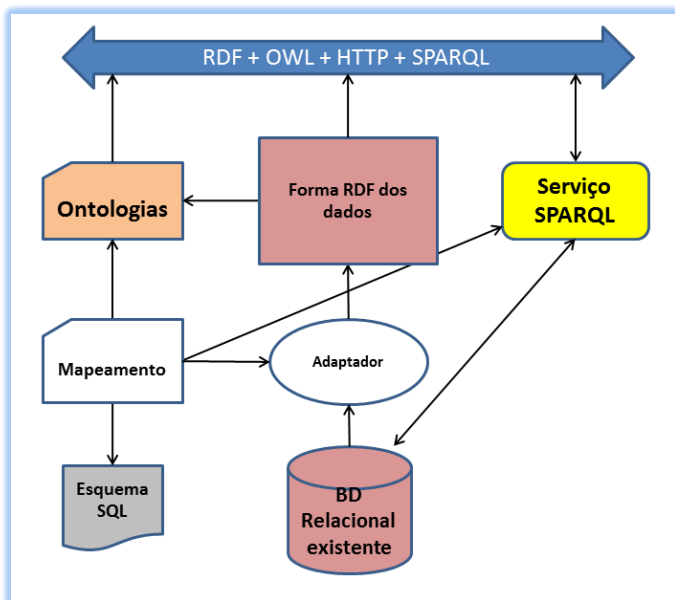
⁹ <http://www.openrdf.org/>

¹⁰ <http://4store.org/>

¹¹ <http://sw.deri.org/2004/06/yars/>

¹² <http://www.ontotext.com/owlim>

Figura 25 - Exemplo de arquitetura de publicação com mapeamento entre relacional e RDF



Fonte: Berners-Lee (2005)

Além da publicação de dados já codificados em RDF nos *triple stores*, há a opção de ferramentas que fazem o mapeamento de bancos de dados relacionais para RDF, tais como o *Triplify*, o *D2R* e o módulo *RDF Views* do *Virtuoso*. O W3C conta com uma recomendação para o mapeamento direto de dados relacionais para RDF (ARENAS *et al.*, 2012).

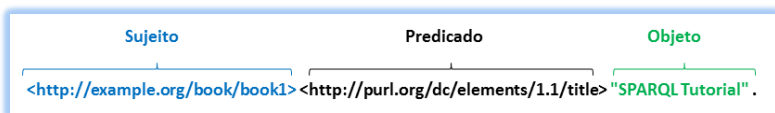
Neste tipo de abordagem, representada na Figura 21, o banco de dados relacional existente é mantido, sendo criado um mapeamento entre o esquema do banco de dados e as ontologias, de forma que, quando os dados sejam solicitados no formato RDF, ou por consultas SPARQL, o mapeamento permite que os dados sejam transformados para a entrega.

Da mesma forma como acontece com bancos de dados relacionais, a *Web* de Dados necessita de uma linguagem de consulta específica para RDF. A linguagem de consulta SPARQL possibilita que os consumidores de dados possam extrair informações possivelmente complexas, tais como referências a recursos e seus relacionamentos, que

são retornadas, por exemplo, em um formato tabular, que pode incorporada em outras páginas *Web*.

A maior parte das consultas SPARQL contém um conjunto de padrões de triplas onde sujeito, predicado e objeto podem ser variáveis. O padrão é comparado com o grafo RDF, e o sub-grafo correspondente aos critérios estabelecidos no padrão de triplas é o resultado da busca (PRUD'HOMMEAUX; SEABORNE, 2008). A Figura 22 apresenta o exemplo de uma tripla RDF contendo dados de um livro, especificamente a indicação de seu título.

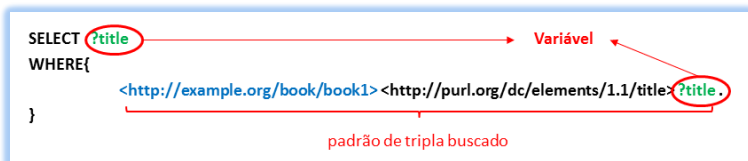
Figura 26 - Exemplo de tripla RDF



Fonte: O autor

A sintaxe de uma consulta que utiliza um padrão de tripla, onde o sujeito e o predicado são definidos, e o objeto é uma variável, indicada pelo caractere "?". A execução da consulta é feita pela comparação do padrão com as triplas RDF.

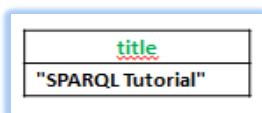
Figura 27 - Exemplo de consulta SPARQL



Fonte: O autor

A consulta representada na Figura 23 solicita que o resultado seja composto por apenas uma coluna, representada no exemplo pela variável `?title`. O resultado da execução da consulta corresponde, portanto, ao(s) objeto(s) da(s) tripla(s) cujo sujeito e predicado correspondam ao padrão de tripla, conforme a Figura 24.

Figura 28 - Resultado da consulta SPARQL



Fonte: O autor

Os padrões de busca podem ser mais complexos, refletindo a necessidade da explicitação de relações entre os recursos das triplas, com o acréscimo de mais variáveis de busca, de tal forma que o resultado venha a ser apresentado em um formato tabular semelhante ao que acontece com consultas SQL.

A linguagem SPARQL conta, ainda, com operações de ordenação, agregações e, a partir da versão 1.1, com *Update*, projetada para realizar modificações, tais como inserir triplas em grafos, excluir triplas de grafos, atualizar dados em triplas, criar ou excluir nos grafos RDF (GROUP, 2013). Estas características a tornam uma linguagem de consulta e manipulação de dados simples para a utilização em aplicações dos mais variados tipos.

4.5 DADOS ESTATÍSTICOS LIGADOS

Considerando a crescente disponibilização de dados, especialmente por parte de órgãos da administração pública, observa-se que, em sua grande maioria, tais dados são heterogêneos e não passam por um tratamento estatístico integrado (ZAPILKO; MATHIAK, 2011). Além disso, da forma como costumeiramente são publicados, não são fornecidos meios para expor, compartilhar ou interliga-los, de tal forma que possam ser acrescentadas mais informações (MILOŠEVIĆ *et al.*, 2012).

Como característica, estes dados costumam ser resultantes de processos de coleta e publicação que seguem métodos estatísticos, e com periodicidades específicas de atualização. Além disso, os dados estatísticos são importantes fontes de informação para agentes das áreas de governo, ciência e negócios, pois caracterizam-se como ferramentas para a verificação hipóteses adotadas na tomada de decisões (ZANCANARO *et al.*, 2013).

Salas *et al.* (2012) consideram de fundamental importância que os dados estatísticos sejam semanticamente ligados a ontologias ou bases de conhecimento. A publicação de Dados Estatísticos Ligados possibilita a melhor aplicação de técnicas de análise de dados com o objetivo de auxiliar o processo de tomada de decisão.

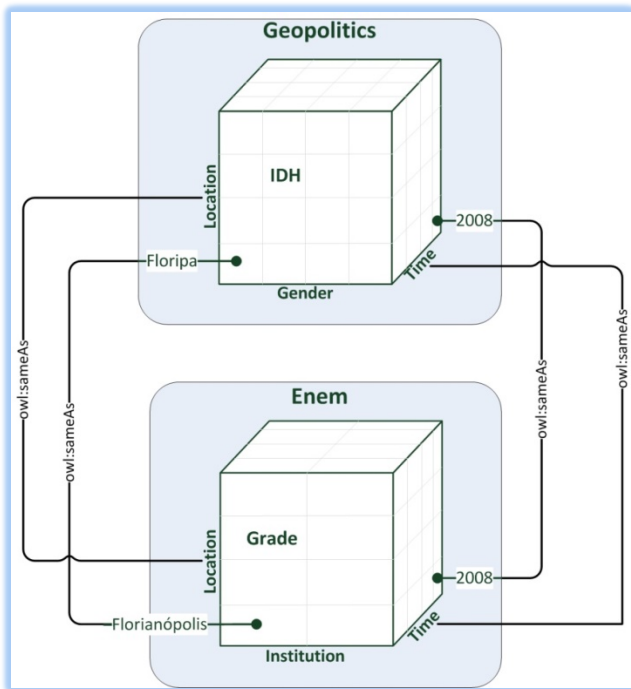
De forma semelhante ao que acontece em sistemas de bancos de dados, a publicação de dados estatísticos em RDF será representada segundo um modelo lógico. No caso de dados estatísticos, há vasta literatura recomendando o uso dos cubos lógicos OLAP (KÄMPGEN;

O’RAIN; HARTH, 2012; ZANCANARO *et al.*, 2013; CYGANIAK; REYNOLDS; TENNISON, 2014). Sistemas OLAP (*Online Analytical Processing*) são uma categoria de tecnologias de software que permite que analistas, gestores e executivos tenham uma visão sobre os dados por meio de acesso rápido, consistente e interativo a uma vasta variedade de visões de informações geradas a partir de dados brutos (CODD; CODD; SALLEY, 1993).

Dois dos principais desafios no uso da OLAP em Dados Ligados são (KÄMPGEN; O’RAIN; HARTH, 2012): a) OLAP requer um modelo de cubo de dados, com dimensões e medidas; b) As consultas OLAP são complexas e requerem modelos dedados especializados, tais como o modelo estrela em bancos de dados relacionais, para que funcionem eficientemente.

A Figura 25 ilustra um exemplo contendo dois cubos lógicos multidimensionais, onde cada um deles representa um conjunto de dados estatísticos ligados. Um dos cubos representa dados geopolíticos brasileiros e o outro representa dados do ENEM. Ambos utilizam a mesma dimensão para tempo, representada pelo literal “2008”. Além disso ambos têm uma dimensão que denota uma entidade geográfica e um membro que denota a cidade de Florianópolis. Se há a indicação de que estes membros são um só (representado na ilustração pela declaração *owl:sameAs*), os dois cubos podem ser representados como um multicubo.

Figura 29 - Exemplo de multicubo



Fonte: Adaptado de Zancanaro *et al.* (2013)

A publicação vários conjuntos de dados estatísticos na forma de cubos lógicos multidimensionais justifica-se pelo fato de que, em OLAP, as operações são padronizadas. Na Figura 26 estão apresentadas as principais operações OLAP.

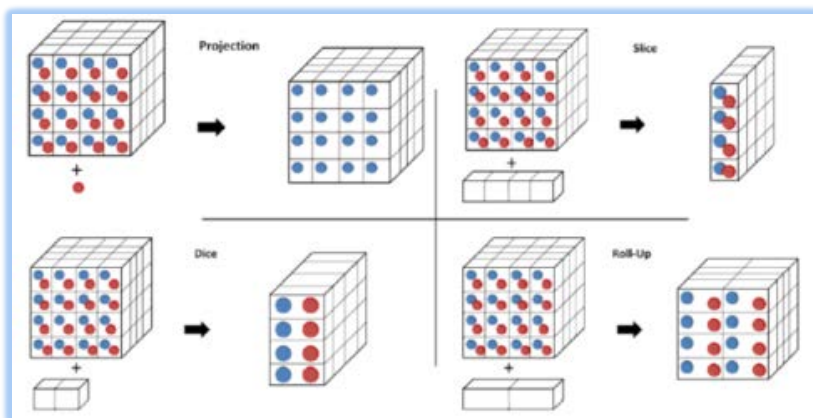
As operações são descritas por Kämpgen; O'rain e Harth (2012):

- Projection: DataCube x Measure* → *DataCube* – remove uma medida de um dado cubo, permitindo ver apenas uma medida em específico. No exemplo da figura, todas as triplas referentes a uma medida foram removidas, resultando em um subcubo.
- Slice: DataCube x Dimension* → *DataCube* – remove uma dimensão de um dado cubo e de todos seus conteúdos.
- Dice: DataCube x Dimension x Value* → *DataCube* – permite filtrar e agregar por certos membros de uma dimensão. Note-se

que *Dice* não é uma operação de seleção, mas um filtro combinado com uma operação de *Slice*.

- d) *Roll-up: DataCube x Dimension x Level* → *DataCube* – permite que se crie um cubo que contenha dados de um maior nível de agregação. A operação de *Drill-Down*, por sua vez, cria um cubo com um nível menor de agregação, sendo, portanto, uma operação inversa ao *Roll-up*.

Figura 30 - Representação das operações OLAP



Fonte: Kämpgen; O’rain e Harth (2012)

As operações *Roll-up* e *Drill-down* permitem alterar a visão dos dados para um diferente nível de agregação. Como exemplo, pode-se imaginar uma consulta em que os dados de municípios possam ser analisados nos níveis de agregação de microrregiões, macrorregiões, ou estados.

A utilização destes cubos lógicos permite representar de forma simples as entidades do mundo real. Além disso, tende a facilitar a análise das medidas, a definição de quais dimensões e atributos podem representar dados significativos, bem como a organização das dimensões em um escopo de níveis e hierarquias (ZANCANARO *et al.*, 2013).

4.6 CONSIDERAÇÕES FINAIS

As tecnologias da Web Semântica têm por objetivo transformar a Web tradicional, de documentos interligados, em uma Web de Dados,

onde os dados têm significado e contexto bem definidos, de forma que os conteúdos estejam preparados não apenas para a compreensão de humanos, mas também para o processamento por máquinas. Este é um grande diferencial para aquela que se pode referenciar como a *Web* tradicional, cuja disponibilidade de conteúdos é muito grande, mas sem que haja, necessariamente, uma preocupação em termos de padronização e estruturação de conteúdos.

Quando os dados publicados são de livre uso, reuso e redistribuição para qualquer pessoa, estando sujeitos, no máximo, à exigência crédito de sua autoria e compartilhamento segundo a mesma licença, são chamados de dados abertos. Diversas iniciativas pelo mundo todo, muitas delas governamentais, tratam da criação e disponibilização de conjuntos de dados abertos, promovendo o livre acesso e participação da sociedade.

As características dos dados abertos não garantem, entretanto, que estes dados sejam amigáveis às máquinas, ou seja, não são apropriados para o processamento automatizado, em função da falta de padronização e estrutura. O termo Dados Ligados (*Linked Data*) é utilizado para descrever um conjunto de técnicas para publicação de dados, utilizando padrões e em um formato tal que estejam interligados com outros conjuntos de dados de outras fontes, materializando a *Web* de Dados.

Neste contexto, as ontologias desempenham um importante papel, pois fornecem vocabulários, com semântica formalmente definida. Assim, os dados são ligados aos conceitos nela representados, passando a fazer parte de um grafo único de dados, cuja exploração possibilita que sejam descobertos novos dados.

5 MODELO DE REFERÊNCIA PARA INDICADORES DE INOVAÇÃO REGIONAL SUPOTADOS POR DADOS LIGADOS

5.1 CONSIDERAÇÕES INICIAIS

Esta seção apresenta o modelo desenvolvido neste trabalho, que tem por objetivo caracterizar, segundo a literatura, os constructos utilizados nos modelos de indicadores de inovação, seus componentes e formas de organização.

5.2 ABORDAGEM DO MODELO PROPOSTO

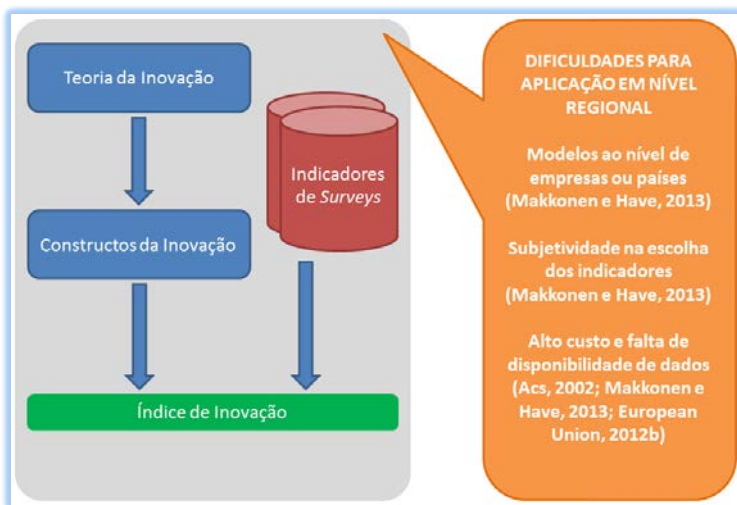
A definição de indicadores de inovação é uma atividade desenvolvida por diversas entidades, que coletam e interpretam dados sobre a inovação e divulgam seus resultados na forma de publicações com periodicidades diferentes periodicidades, com o objetivo de possibilitar a avaliação das atividades de inovação desenvolvidas em determinados contextos. Comumente, deseja-se estabelecer comparações entre países, sendo definidos índices de inovação que agregam conjuntos de indicadores (CORNELL UNIVERSITY; INSEAD; WIPO, 2013; EUROPEAN UNION, 2014a).

Em sua concepção, tais publicações baseiam-se nas teorias de inovação e desenvolvem modelos de indicadores compostos, cujos dados são coletados por meio da realização de *surveys* específicos que seguem orientações de guias de coleta e interpretação de dados da inovação, como os manuais de Oslo e Frascati (OECD, 2005; 2015). Em outros casos, como no *Global Innovation Index*, no *European Innovation Scoreboard* e *Regional Innovation Scoreboard*, são desenvolvidos modelos que utilizam como insumos os indicadores produzidos por publicações de terceiros, como as do Banco Mundial, OECD, FMI (CORNELL UNIVERSITY; INSEAD; WIPO, 2013; EUROPEAN UNION, 2014a).

A Figura 31 apresenta uma representação da abordagem normalmente utilizada para a concepção de indicadores de inovação. A partir de indicadores oriundos de *surveys* próprios ou de terceiros, índices específicos são formulados. Conforme observam (MAKKONEN; HAVE, 2013), tais modelos são bastante subjetivos na escolha de métodos e variáveis. Além disso, a realização de *surveys* cujo foco está principalmente no nível de empresas ou de países, faz com que sua

aplicação em nível regional seja dificultada pela falta de disponibilidade dos dados (EUROPEAN UNION, 2012; MAKKONEN; HAVE, 2013).

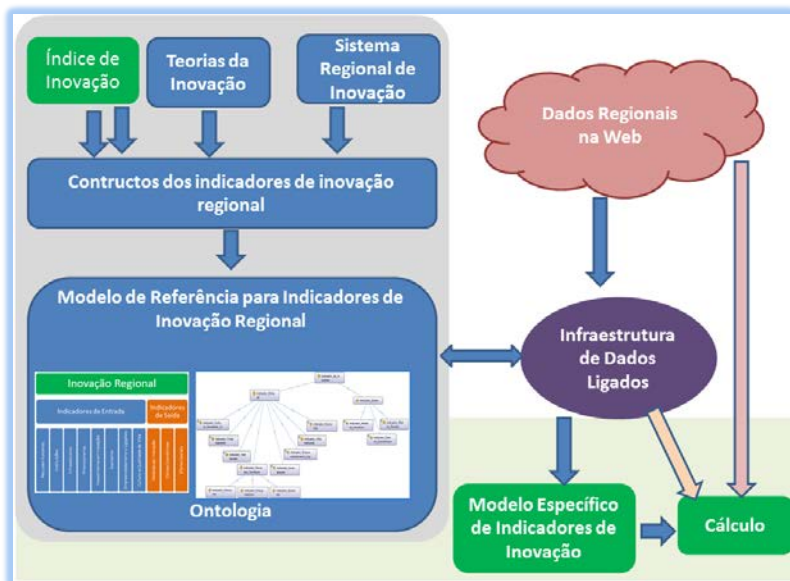
Figura 31 - Abordagem usual para definição de índices de inovação



Fonte: o autor

A abordagem proposta neste trabalho tem por objetivo criar um modelo de referência para indicadores de inovação regional. Embasado na literatura sobre inovação, no conceito de sistemas regionais de inovação e em modelos de índices de inovação regional coletados na literatura, o modelo de referência deverá servir de base para a definição de modelos específicos de índices de inovação regional suportados por Dados Ligados.

Figura 32 - Modelo de Referência para Indicadores de Inovação Regional



Fonte: O autor

A Figura 32 apresenta uma representação conceitual da abordagem na qual se insere o modelo de referência proposto, considerando seu uso como base a definição de índices específicos usando da tecnologia de Dados Ligados.

Além da representação conceitual do modelo, o mesmo é representado na forma de uma ontologia, de tal forma que todos os constructos componentes do modelo de referência, bem como os demais conceitos associados e suas interrelações são explicitados em uma linguagem de representação de ontologias e publicados em uma infraestrutura específica, de forma que estejam acessíveis na *Web* e aptos para o processamento automatizado.

O cenário de utilização previsto para o modelo de referência proposto prevê que diferentes modelos indicadores compostos para a inovação regional, a exemplo dos encontrados na literatura, possam ser nele representados. Ao adotar-se um modelo específico para a mensuração da inovação regional, é possível, portanto, instanciá-lo no modelo de referência.

A instanciação de diferentes modelos específicos, juntamente com seus dados, na forma de Dados Ligados permite que um mesmo conjunto de dado possa vir a ser reutilizado por mais de um modelo.

Além disso, no caso da opção pela definição de um novo modelo específico de indicadores para a inovação regional, o modelo ora proposto visa servir de ponto de partida, como uma referência dos constructos a serem abordados.

Uma vez que os conjuntos de dados estejam associados a um determinado modelo específico, e publicados na forma de dados ligados, os processamentos necessários para os cálculos de índices de inovação definidos no modelo podem ser efetuados.

5.3 MODELO CONCEITUAL DE INDICADORES DE INOVAÇÃO REGIONAL

Para a concepção dos constructos que embasam o modelo de referência, foram utilizadas referências coletadas durante a revisão de literatura sobre inovação e, mais especificamente sobre os indicadores de inovação compostos. Além disso, uma busca sistemática de literatura foi conduzida com o objetivo de identificar os modelos de indicadores utilizados para a avaliação da inovação regional.

Quanto à forma de concepção dos indicadores de inovação, sua estrutura e componentes, a literatura apresenta diferentes abordagens. Conforme observam Zabala Iturriagoitia; Jiménez Saez e Gutiérrez Gracia (2005), entretanto, a maioria dos indicadores disponíveis e estatísticas de inovação consideram os sistemas de inovação como sistemas de **entrada-saída** (*input-output*), com ênfase particular nos recursos empregados. A título de exemplo, esta é a classificação utilizada no *global Innovation Index* (CORNELL UNIVERSITY; INSEAD; WIPO, 2013).

A classificação entre entrada e saída é alternativamente caracterizada pela distinção entre **insumos e produtos** (FAPESP, 2011), ou **insumos e resultados** (GUSMÃO; RAMOS, 2006). Já os modelos da União Europeia (EUROPEAN UNION, 2014a; 2014b), por sua vez, subdividem os indicadores em: **Viabilizadores** (*Enablers*), **Atividades das Empresas** (*Firm Activities*) e Saída (*Output*).

De forma geral, pode-se dizer que os modelos de indicadores utilizados para avaliar a inovação costumam subdividir os indicadores entre aqueles de **entrada**, que definem as condições, ou as capacidades existentes em um sistema de inovação, e os de **saída**, que representam os resultados obtidos pelo sistema de inovação. A Figura 33 representa a primeira subdivisão proposta para a classificação dos indicadores para Inovação Regional, conforme apresentado por Speroni; Macedo e Gauthier (2016a).

Figura 33 – Modelo de Índice de Inovação Regional: Níveis zero e um



Fonte: (SPERONI; MACEDO; GAUTHIER, 2016a)

Os Indicadores de Entrada são aqueles que representam as condições que levam um sistema de inovação a desempenhar seu papel. Por isto, são localizados neste agrupamento os dados referentes ao ambiente de inovação, aquelas condições que dizem respeito às características da região quanto aos recursos humanos, economia, instituições; e os dados referentes aos esforços diretamente empreendidos para a realização da inovação, tais como os dispêndios realizados em Pesquisa e Desenvolvimento, e as parcerias estabelecidas.

Com o objetivo de identificar os constructos utilizados nos modelos existentes de índices para a inovação regional, uma busca sistemática de literatura foi realizada no mês de Dezembro de 2014 nas bases científicas *Scopus*¹³ e *Web Of Science*¹⁴, utilizando como termos de busca **“regional innovation” AND (indicator OR index)**. Dos 199 artigos resultantes da busca, 96 foram selecionados após a aplicação dos critérios de exclusão: a) exclusão de duplicidades; b) exclusão de artigos sem texto completo disponível para *download* gratuito; c) exclusão de artigos que não se encaixavam no contexto. A busca e seus resultados são relatados em maiores detalhes em Speroni; Macedo e Gauthier (2016b).

Os 96 artigos selecionados foram analisados, buscando identificar aqueles que utilizavam, ou que propunham algum modelo de índice de inovação, pelo agrupamento de múltiplos indicadores. Destes, 52 artigos foram selecionados, e os tipos de indicadores por eles utilizados foram catalogados, e suas características estão sendo levadas em consideração para a construção do modelo de referência aqui proposto.

A partir da observação das características dos indicadores catalogados nas obras selecionadas, foram elencadas as classes que representam os principais constructos abordados nos modelos. Tal classificação levou em consideração os diferentes níveis e agrupamentos

¹³ <http://www.scopus.com>

¹⁴ <http://webofscience.com>

de indicadores observados. O Quadro 8 apresenta as classes propostas para os indicadores de entrada, e indica referências bibliográficas que apresentam indicadores destas classificações. Foram consideradas, para este quadro, as obras dos últimos cinco anos.

A seguir, as classificações apresentadas no Quadro 8 são descritas:

Recursos Humanos: compreende indicadores relativos à educação e a disponibilidade de força de trabalho altamente qualificada e educada (EUROPEAN UNION, 2014a).

Instituições: agrupa indicadores associados às condições de um sistema de inovação de prover uma boa governança e níveis proteção e incentivos essenciais à inovação (CORNELL UNIVERSITY; INSEAD; WIPO, 2013).

Ambiente de Inovação / Infraestrutura: agrupa indicadores relativos à infraestrutura que dá suporte à inovação, tais como a disponibilidade tecnologias de informação e comunicação (TICs), a cobertura de estradas pavimentadas, instituições de ensino e de institutos de pesquisa (CHEN; WANG, 2012; CORNELL UNIVERSITY; INSEAD; WIPO, 2013).

Financiamento / Apoio à Inovação: agrupa indicadores que dizem respeito à disponibilidade de crédito, fundos de investimentos e fontes de financiamento para as atividades de inovação, bem como do apoio governamental (CORNELL UNIVERSITY; INSEAD; WIPO, 2013; EUROPEAN UNION, 2014a).

Investimento em Inovação: compreende os indicadores que tratam da quantidade de recursos investidos nas atividades de inovação. São exemplos o dispêndio público ou privado em Pesquisa e Desenvolvimento, o pessoal alocado em atividades de P&D, a compra de tecnologia.

Economia e Mercado de Trabalho: as variáveis deste conjunto estão relacionadas às condições econômicas da região. Dentre elas, citam-se indicadores como o GDP, ou o PIB, ou o nível de emprego.

Empreendedorismo e Ligações: esta categoria engloba indicadores que representam as características da região em relação ao desenvolvimento do empreendedorismo e de arranjos e da colaboração entre os diferentes atores do sistema, tais como treinamentos para o empreendedorismo e as relações universidade-indústria (CHANG *et al.*, 2012; CORNELL UNIVERSITY; INSEAD; WIPO, 2013).

Cultura e Qualidade de vida: compreende indicadores que descrevem as características de uma região em relação ao saneamento ambiental, qualidade de vida dos residentes, cultura e participação social. Exemplos deste tipo são o consumo industrial de água, consumo de energia,

reservas de áreas, investimento em cultura, existência de teatros, ou participação em atividades voluntárias (AFFORTUNATO *et al.*, 2010; CHEN; JU, 2014).

Quadro 8- Classificação dos Indicadores de Entrada utilizados na literatura

Autores	Recursos Humanos	Instituições	Infraestrutura	Financiamento	Investimento em Inovação	Economia e Mercado de Trabalho	Ligações e Empreendedorismo	Cultura / Qualidade de Vida
(EUROPEAN UNION, 2014a)	X			X	X		X	
(CHEN; JU, 2014)	X		X		X	X		X
(GLEBOVA; KOTENKOVA, 2014)	X		X	X	X			
(PONS; MARTÍNS; PARRILLI, 2014)	X			X	X	X	X	
(QI; LIU, 2014)					X	X		
(SÁNCHEZ TOVAR; GARCÍA FERNÁNDEZ; MENDOZA FLORES, 2014)	X	X				X	X	X
(SUN, 2014)	X		X		X		X	
(CORNELL UNIVERSITY; INSEAD; WIPO, 2013)	X	X	X	X	X	X	X	
(GOGODZE, 2013)	X		X		X			X
(NIKULINA <i>et al.</i> , 2013)				X		X		X
(MAKKONEN; HAVE, 2013)	X	X	X			X		
(TU; SONG, 2013)	X	X			X	X		
(ZHENG; ZHANG, 2013)	X		X	X	X	X		X
(MARKOWSKA; STRAHL, 2012)	X		X	X	X		X	
(WANG <i>et al.</i> , 2012)	X				X	X		
(CHEN; GUAN, 2011)	X	X		X	X	X	X	
(GUO; REN, 2011)	X			X	X			
(HÁJKOVÁ; HÁJEK, 2011)	X				X	X		
(LI, 2011)	X	X	X	X				
(LIU; ZHANG; ZHANG,	X		X	X	X			

2011)								
(NING; CHUN, 2011)	X		X		X	X		
(TIAN; TONG, 2011)	X	X	X	X	X	X		X
(XIAO; YANG, 2011)	X			X	X			
(YAM <i>et al.</i> , 2011)	X	X			X	X		
(YIN; DIAO; LI, 2011)	X		X		X	X		
(ZHANG; LIU; ZHANG, 2011)			X	X	X			
(AFFORTUNATO <i>et al.</i> , 2010)	X	X	X	X	X	X		X
(BUESA; HEIJS; BAUMERT, 2010)	X	X		X	X	X	X	X
(CHANG <i>et al.</i> , 2012)	X		X	X		X	X	
(LI, 2010)	X		X	X		X	X	
(PINTO; GUERREIRO, 2010)	X			X		X		X

Fonte: (SPERONI; MACEDO; GAUTHIER, 2016a)

De forma semelhante, Quadro 9 apresenta a classificação dos Indicadores de Saída e os trabalhos nos quais estão presentes, considerando os últimos cinco anos.

Medidas de Inovação: estão incluídos neste grupo os indicadores que dizem respeito aos resultados de conhecimentos gerados pelas atividades de inovação, tais como a quantidade de empresas que inovaram em produtos ou serviços, a quantidade de empresas que inovaram *marketing* ou em processos organizacionais, o número de patentes solicitadas e concedidas, e o número de publicações científicas (CORNELL UNIVERSITY; INSEAD; WIPO, 2013; EUROPEAN UNION, 2014a).

Efeitos Econômicos: compreende os indicadores que capturam o sucesso econômico da inovação, tais como o a porcentagem de vendas relativas a produtos inovadores ou a exportação de serviços inovadores (EUROPEAN UNION, 2014a).

Efeitos Sociais: Uma vez que os resultados da inovação trazem benefícios às pessoas e ao planeta, há um reconhecimento cada vez maior sobre os efeitos sociais da inovação (OECD, 2014).

Quadro 9- Classificação dos Indicadores de Saída utilizados na literatura

Autores	Medidas de Inovação	Efeitos Econômicos	Efeitos Sociais
(EUROPEAN UNION, 2014a)	X	X	
(CHEN; JU, 2014)		X	X
(GLEBOVA; KOTENKOVA, 2014)		X	X
(PONS; MARTÍNS; PARRILLI, 2014)		X	
(QI; LIU, 2014)	X		
(SÁNCHEZ TOVAR; GARCÍA FERNÁNDEZ; MENDOZA FLORES, 2014)	X	X	X
(SUN, 2014)	X		X
(CORNELL UNIVERSITY; INSEAD; WIPO, 2013)	X	X	
(GOGODZE, 2013)	X	X	
(NIKULINA <i>et al.</i> , 2013)		X	
(MAKKONEN; HAVE, 2013)	X	X	X
(TU; SONG, 2013)	X	X	
(ZHENG; ZHANG, 2013)	X	X	
(WANG <i>et al.</i> , 2012)	X	X	
(MARKOWSKA; STRAHL, 2012)	X	X	
(GUO; REN, 2011)	X	X	
(HÁJKOVÁ; HÁJEK, 2011)	X		
(LI, 2011)	X	X	
(LIU; ZHANG; ZHANG, 2011)	X		
(NING; CHUN, 2011)	X	X	X
(TIAN; TONG, 2011)		X	X
(XIAO; YANG, 2011)	X	X	
(YAM <i>et al.</i> , 2011)	X	X	
(YIN; DIAO; LI, 2011)	X	X	
(ZHANG; LIU; ZHANG, 2011)	X		
(AFFORTUNATO <i>et al.</i> , 2010)	X		
(BUESA; HEJIS; BAUMERT, 2010)		X	X
(LI, 2010)	X	X	
(PINTO; GUERREIRO, 2010)	X	X	

Fonte: (SPERONI; MACEDO; GAUTHIER, 2016a)

A Figura 34 apresenta a classificação dos Indicadores, nos níveis zero, um e dois, segundo o levantamento feito com base na literatura. No nível dois estão representadas as classificações que irão compor os Indicadores de Entrada e os de Saída.

Figura 34 – Classificação dos Indicadores de Inovação Regional



Fonte: (SPERONI; MACEDO; GAUTHIER, 2016a)

Os constructos do modelo conceitual representado na Figura 30 são, ainda, subdivididos e servem como base para a construção de uma ontologia que tem por objetivo a representação do conhecimento relativo à forma como se organizam os indicadores, de forma a possibilitar a criação de modelos de indicadores de inovação regional suportados por Dados Ligados.

O Quadro 10 apresenta uma lista, feita a partir do levantamento bibliográfico, dos assuntos abordados pelos indicadores classificados no tipo “Indicadores de Entrada”. Para cada um deles, há diferentes medidas utilizadas, tais como valores percentuais, ou valores sobre 10.000 pessoas, ou diferentes moedas.

Quadro 10 – Assuntos abordados pelos indicadores das dimensões de entrada

Nível 2	Nível 3	Assuntos abordados pelos indicadores
Recursos Humanos	Educação	<ul style="list-style-type: none"> • Estudantes de graduação (TIAN; TONG, 2011) • População com curso superior (ZABALA-ITURRIAGAGOITIA <i>et al.</i>, 2007) • Estudantes de Pós-Graduação (ZABALA-ITURRIAGAGOITIA <i>et al.</i>, 2007) • Mestres e Doutores (GUSMÃO; RAMOS, 2006; WANG; LI, 2009) • Taxa de Escolaridade (GUSMÃO; RAMOS, 2006) • Educação continuada (HÁJKOVÁ; HÁJEK, 2011)
	Pesquisadores	<ul style="list-style-type: none"> • Pesquisadores em P&D (WANG <i>et al.</i>, 2012) • Grupos de Pesquisa (GUSMÃO; RAMOS, 2006)
	Docentes	<ul style="list-style-type: none"> • Docentes em tempo integral (LIU; ZHANG; ZHANG, 2011) • Docentes doutores (GUSMÃO; RAMOS, 2006) • Relação aluno/docente (GUSMÃO; RAMOS, 2006)
Instituições	Impostos	<ul style="list-style-type: none"> • Taxas de impostos (AFFORTUNATO <i>et al.</i>, 2010) • Impostos sobre as atividades de inovação (YAN; YAN; MA, 2009)
	Ambiente Político	<ul style="list-style-type: none"> • Ambiente político (CORNELL UNIVERSITY; INSEAD; WIPO, 2013) • Disponibilidade de serviços governamentais (WANG; LIU; TIAN, 2006) • Crimes contra administração pública (AFFORTUNATO <i>et al.</i>, 2010)
	Legislação	<ul style="list-style-type: none"> • Estado de Direito (CORNELL UNIVERSITY; INSEAD; WIPO, 2013) • Leis para a inovação (TIAN; TONG, 2011)

	Burocracia	<ul style="list-style-type: none"> • Custo de abertura de negócios (AFFORTUNATO <i>et al.</i>, 2010) • Dias para a abertura de um negócio (CORNELL UNIVERSITY; INSEAD; WIPO, 2013)
Infraestrutura	Tecnologias de Informação e Comunicação (TICs)	<ul style="list-style-type: none"> • Acesso à Internet banda-larga (MARKOWSKA; STRAHL, 2012; ZHENG; ZHANG, 2013) • Internet nas escolas (AFFORTUNATO <i>et al.</i>, 2010) • Internet nas residências (ZHENG; ZHANG, 2013) • Uso de TICs (e-gov; telecomunicações) (GLEBOVA; KOTENKOVA, 2014)
	Logística	<ul style="list-style-type: none"> • Cobertura de rodovias asfaltadas (TIAN; TONG, 2011) • Cobertura de linhas férreas (WANG <i>et al.</i>, 2012) • Transporte de cargas (TIAN; TONG, 2011) • Transporte de passageiros (TIAN; TONG, 2011)
	Ensino e Pesquisa	<ul style="list-style-type: none"> • Disponibilidade de Universidades (LIU; ZHANG; ZHANG, 2011) • Disponibilidade de Institutos de Pesquisa (CHANG <i>et al.</i>, 2012) • Revistas Científicas (AFFORTUNATO <i>et al.</i>, 2010)
	Empresas	<ul style="list-style-type: none"> • Disponibilidade de Empresas de alta-tecnologia (LIU; ZHANG; ZHANG, 2011) • Empresas com departamento de P&D (WANG; LI, 2009)
Financiamento	Orçamento Público	<ul style="list-style-type: none"> • Orçamento público para atividades de inovação (HÁJKOVÁ; HÁJEK, 2011) • Proporção do Financiamento de P&D na economia local (MARKOWSKA; STRAHL, 2012)
	Capital de Risco	<ul style="list-style-type: none"> • Disponibilidade de Capital de Risco (CHANG <i>et al.</i>, 2012) • Investimento de capital de risco (ZABALA-ITURRIAGAGOITIA <i>et al.</i>, 2007)
	Fontes de Financiamento	<ul style="list-style-type: none"> • Fontes de financiamento (CORNELL UNIVERSITY; INSEAD; WIPO, 2013)
Investimen-	Dispêndios em P&D	<ul style="list-style-type: none"> • Dispêndios privados em P&D (HÁJKOVÁ; HÁJEK, 2011)

to em Inovação		<ul style="list-style-type: none"> • Dispendícios públicos em P&D (HÁJKOVÁ; HÁJEK, 2011)
	Renovação Técnica	<ul style="list-style-type: none"> • Dispendícios na renovação técnica em indústrias de alta tecnologia (GLEBOVA; KOTENKOVA, 2014)
	Aquisição de Inovação	<ul style="list-style-type: none"> • Aquisição direta de inovação (CHEN; GUAN, 2011) • Aquisição indireta de inovação (CHEN; GUAN, 2011)
	Importação de Tecnologia	<ul style="list-style-type: none"> • Importação de produtos de alta tecnologia (AFFORTUNATO <i>et al.</i>, 2010) • Importação da indústria intensiva em P&D (WANG <i>et al.</i>, 2012)
Economia	Macroeconomia	<ul style="list-style-type: none"> • Produto Interno Bruto (ZHENG; ZHANG, 2013) • Taxa de crescimento econômico (BARKLEY; HENRY; NAIR, 2006) • Rendimento anual per capita (AFFORTUNATO <i>et al.</i>, 2010) • Participação das indústrias do 1º, 2º e 3º setor (TIAN; TONG, 2011)
	Trabalho e Emprego	<ul style="list-style-type: none"> • Taxa de desemprego (HÁJKOVÁ; HÁJEK, 2011) • Empregos em empresas de alta-tecnologia (BARKLEY; HENRY; NAIR, 2006) • Taxa de emprego de técnicos altamente especializados (GOGODZE, 2013)
Empreendedorismo e Ligações	Cultura empreendedora	<ul style="list-style-type: none"> • Treinamentos e educação para o empreendedorismo (CHANG <i>et al.</i>, 2012) • <i>Habitats</i> de inovação (CHANG <i>et al.</i>, 2012) • Disponibilidade de <i>Startups</i> (CHANG <i>et al.</i>, 2012)
	Cooperação	<ul style="list-style-type: none"> • Disponibilidade de agências intermediárias (CHANG <i>et al.</i>, 2012) • Cooperações técnicas entre universidades, institutos de pesquisa e empresas (YAN; YAN; MA, 2009)
	Agrupamentos	<ul style="list-style-type: none"> • Existência de <i>clusters</i> e agrupamentos (CHANG <i>et al.</i>, 2012)
Cultura e Qualidade de	Acesso à cultura	<ul style="list-style-type: none"> • Disponibilidade de bibliotecas (públicas; reservas) (NING; CHUN, 2011) • Disponibilidade de teatros e cinemas

Vida		(AFFORTUNATO <i>et al.</i> , 2010) <ul style="list-style-type: none"> • Inclusão cultural (CHANG <i>et al.</i>, 2012) • Investimento em cultura (AFFORTUNATO <i>et al.</i>, 2010)
	Voluntariado	<ul style="list-style-type: none"> • Envolvimento da população em atividades voluntárias (AFFORTUNATO <i>et al.</i>, 2010)
	Meio-Ambiente	<ul style="list-style-type: none"> • Sanidade ambiental (CHEN; JU, 2014) • Consumo de água (CHEN; JU, 2014) • Consumo de energia (XIAO; YANG, 2011) • Emissão de poluentes (CHEN; JU, 2014) • Expectativa de vida da população (NIKULINA <i>et al.</i>, 2013)

Fonte: (SPERONI; MACEDO; GAUTHIER, 2016a)

O Quadro 11 apresenta os assuntos abordados pelos indicadores classificados no tipo “Indicadores de Saída”.

Quadro 11 – Assuntos abordados pelos indicadores das dimensões de saída

Nível 2	Nível 3	Assuntos abordados pelos indicadores
Medidas de Inovação	Inovações	<ul style="list-style-type: none"> • Inovações de Produto e Serviço (MARKOWSKA; STRAHL, 2012) • Inovações de Processo e Organizacionais (MARKOWSKA; STRAHL, 2012) • Empresas que inovaram (MARKOWSKA; STRAHL, 2012) • Marcas Registradas e padrões (CHEN; GUAN, 2011)
	Patentes	<ul style="list-style-type: none"> • Pedidos de Patentes (GUO; REN, 2011) • Patentes concedidas (ZHANG; LIU; ZHANG, 2011)
	Publicações Científicas	<ul style="list-style-type: none"> • Publicações científicas (ZHENG; ZHANG, 2013) • Prêmios científicos (LIU; ZHANG; ZHANG, 2011)
Efeitos Econômicos	Receitas	<ul style="list-style-type: none"> • Vendas de novos produtos e serviços (LU; ZHANG; PENG, 2009) • Proporção das empresas de alta-tecnologia para o PIB (WANG <i>et al.</i>, 2012) • Crescimento do valor de produção (LI; LEI, 2009)
	Exportações	<ul style="list-style-type: none"> • Exportações de <i>commodities</i> (LU; ZHANG; PENG, 2009) • Exportação de produtos de alta tecnologia

		(NING; CHUN, 2011)
Efeitos Sociais	Padrão de vida	<ul style="list-style-type: none"> • Bem-estar (CHEN; JU, 2014) • Padrões de consumo (TIAN; TONG, 2011)
	Ambiental	<ul style="list-style-type: none"> • Benefícios ambientais (CHEN; JU, 2014) • Sustentabilidade (TIAN; TONG, 2011)

Fonte: (SPERONI; MACEDO; GAUTHIER, 2016a)

Cabe salientar que as listas de indicadores apresentados no Quadro 10 e no Quadro 11 não representam à exaustão todos os indicadores utilizados nos modelos observados.

A próxima seção descreve os passos da construção de uma ontologia para representação do modelo de referência proposto, de forma que o mesmo possa ser mais facilmente reutilizado para a construção de modelos específicos suportados por Dados Ligados.

5.4 MODELO DE REFERÊNCIA SUPORTADO POR DADOS LIGADOS

5.4.1 Construção da Ontologia

Além de uma organização conceitual para a classificação dos indicadores de inovação regional, pretende-se que o modelo aqui proposto possibilite a publicação destes indicadores segundo os preceitos dos dados ligados. Para isto, é necessário que os conceitos explicitados sejam codificados em linguagem de representação semântica, garantindo significado e estrutura a estes dados.

Uma das características dos dados ligados é que os dados são auto descritivos. Conforme observam Bizer; Heath e Berners-Lee (2009), se uma aplicação encontra dados que estão descritos em um vocabulário não familiar, pode consultar pelas URIs que identificam os termos do vocabulário para encontrar suas definições. Assim, o modelo conceitual aqui proposto deve fornecer um vocabulário, representado na forma de uma ontologia, para descrição e dos dados.

A construção de uma ontologia é um processo de modelagem, e diferentes metodologias podem ser aplicadas para este fim. Neste trabalho, está sendo utilizada como apoio para o desenvolvimento a ferramenta ontoKEM¹⁵ (TODESCO *et al.*, 2009). O ontoKEM utiliza conceitos oriundos de diferentes metodologias, proporcionando um

¹⁵ <http://ontokem.egc.ufsc.br/>

ambiente *Web*, onde as atividades são agrupadas em quatro fases: competência, vocabulário, hierarquia de classes e dicionário de classes.

A Figura 35 apresenta as questões de competência registradas no ontoKEM. A partir do registro das questões de competência, são informados os termos e relações que estão associadas a estas questões. Foram registradas dez questões de competência. O artefato gerado pelo ontoKEM é apresentado no Apêndice A.

Figura 35 - Questões de competência registradas no ontoKEM

Open the report in PDF format

Report of Competence questions

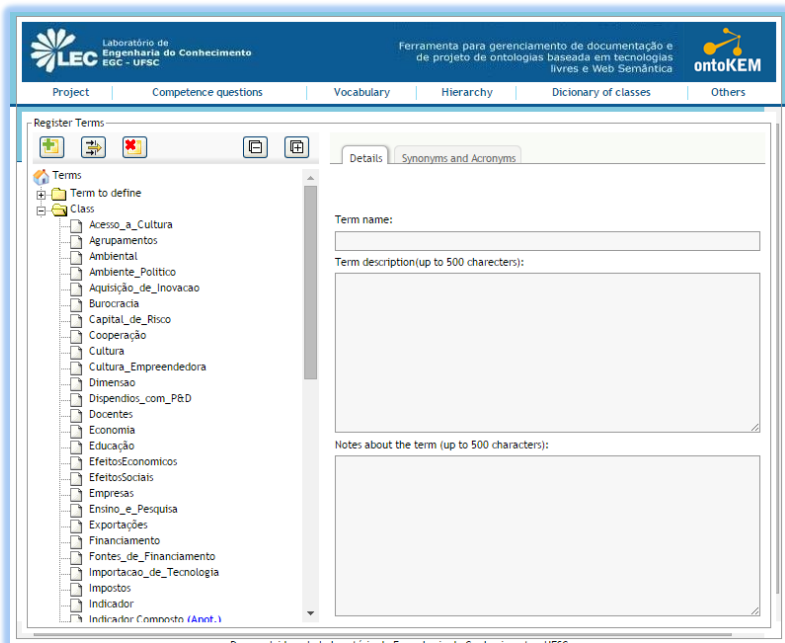
DOCUMENT
Competence questions
(version 1.0)
25-05-2015

- 1) Como são calculados os indicadores de inovação?
Suggested terms: Cálculo
Suggested relations:
- 2) Como são compostos os Indicadores de Inovação Regional?
Suggested terms: Indicador Composto; Indicador Simples; Indicador; Regional Innovation
Suggested relations: ehSubdivisaoDe; temSubdivisao
- 3) O que diferencia os indicadores simples dos compostos?
Suggested terms: Indicadores Compostos; Indicadores Simples
Suggested relations: compoe; ehCompostoPor
- 4) O que são Indicadores de Entrada para a inovação?
Suggested terms: Entrada
Suggested relations:
- 5) O que são Indicadores de Saída para a inovação?
Suggested terms: Saída
Suggested relations:
- 6) Quais as dimensões que compõem os indicadores de Entrada e Saída?
Suggested terms: compõem; Dimensões; Entrada; Saída
Suggested relations: ehDimensaoDe; temDimensao
- 7) Quais as unidades de medida utilizadas para cada indicador?
Suggested terms: ehUnidadeDeMedidaDe; temUnidadeDeMedida; Unidade de Medida
Suggested relations:
- 8) Quais os Indicadores que estão categorizados em cada dimensão?
Suggested terms: Categorizado
Suggested relations:

Fonte: o autor

Na etapa seguinte, os termos e relações sugeridos durante o registro das questões de competência são exportados para utilização na etapa de vocabulário. Nesta segunda etapa, os termos são categorizados segundo seu tipo na ontologia: Classes, Relações entre classes (Object Properties), Atributos (Data Properties). A Figura 36 apresenta a interface do ontoKEM onde são categorizados os termos. A lista das classes definidas na primeira versão da ontologia é apresentada no Apêndice B.

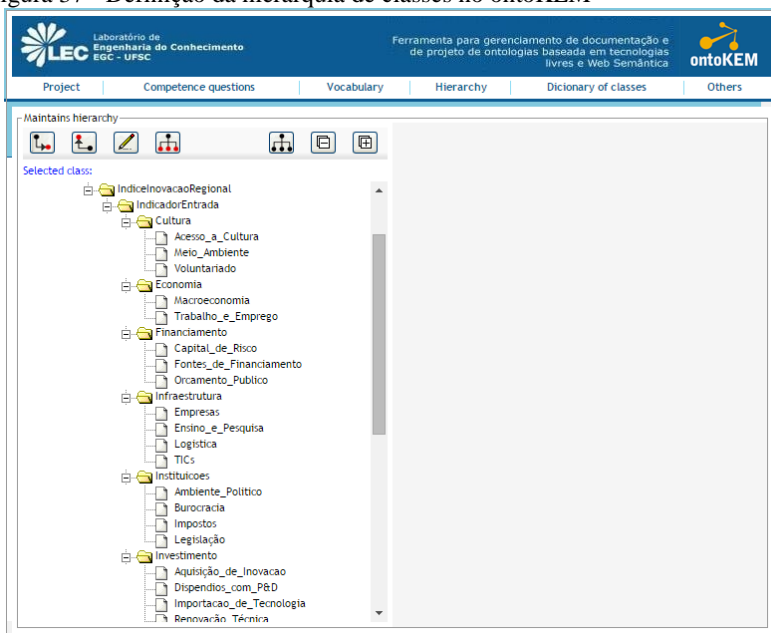
Figura 36 – Registro e categorização de termos no ontoKEM



Fonte: o autor

A terceira etapa da construção da ontologia consiste da definição da hierarquia das classes. Neste passo, os termos, já classificados como classes, são categorizados segundo uma hierarquia onde são definidas as superclasses e subclasses. A Figura 33 apresenta a interface do ontoKEM onde é feita a definição e manutenção da hierarquia de classes.

Figura 37 - Definição da hierarquia de classes no ontoKEM

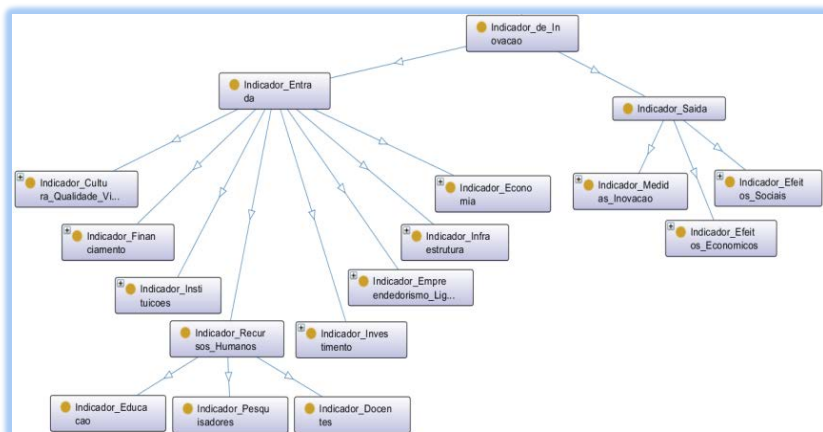


Fonte: O autor

A quarta etapa prevista no ontoKEM consiste da definição do dicionário de classes, onde são definidas as características das classes e das propriedades (relações e atributos).

Na ontologia, as classificações propostas no modelo conceitual são representadas em uma hierarquia de classes. A Figura 34 exibe uma representação parcial da estrutura de classes que compõem a ontologia do modelo de referência, onde apenas uma das classes de so segundo nível está sendo representada com suas subclasses. As demais estão fechadas.

Figura 38 - Representação parcial das classes da ontologia



Fonte: O autor

Com a conclusão de sua primeira versão no ontoKEM, a ontologia é exportada, sendo gerado um documento codificado em padrão RDF, que pode ser publicado em um servidor de Dados Ligados ou passar por mais ajustes em ferramentas de edição de ontologia. A próxima seção aborda a publicação desta ontologia em uma infraestrutura de dados ligados, bem como exemplifica sua instanciação na representação de modelos específicos de indicadores.

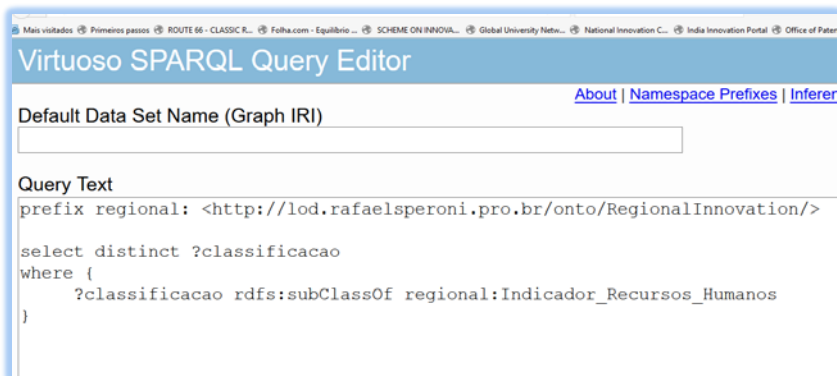
5.4.2 Publicação e uso da Ontologia

A ontologia criada tem por finalidade explicitar a representação do modelo, tornando-o um vocabulário aplicável para o mapeamento de indicadores de diferentes modelos específicos para a inovação regional. Sua codificação na linguagem RDF a torna apta a ser publicada em uma infraestrutura de dados ligados.

De forma semelhante ao que acontece com qualquer conjunto de dados ligados, quando carregada em um servidor de dados ligados, a ontologia passa a estar disponível para consultas SPARQL. Dessa forma, a representação dos recursos nela descritos está acessível também via navegador *Web*.

A título de exemplo, uma consulta SPARQL pode solicitar ao servidor a lista das classificações de indicadores que são subclasses de “Indicadores de Recursos Humanos”.

Figura 39 - Consulta SPARQL: lista de subclasses



Fonte: O autor

Como resultado da execução da consulta, uma lista de três subclasses é apresentada, conforme a Figura 36.

Figura 40 - Resultado da consulta por subclasses

classificacao
http://lod.rafaelsperoni.pro.br/onto/RegionalInnovation/Indicador_Docentes
http://lod.rafaelsperoni.pro.br/onto/RegionalInnovation/Indicador_Educao
http://lod.rafaelsperoni.pro.br/onto/RegionalInnovation/Indicador_Pesquisadores

Fonte: O autor

A efetiva utilização da ontologia enquanto modelo de referência acontece quando os recursos nela descritos passam a ser mencionados na representação de modelos específicos de indicadores. As classes da ontologia são instanciadas à medida que os indicadores individuais são catalogados. Dessa forma, o modelo específico é catalogado levando em conta as anotações semânticas que classificam seus indicadores de acordo com as classes do modelo de referência.

Para exemplificar a criação de instâncias de indicadores, será utilizado o modelo específico do *Regional Innovation Scoreboard* (EUROPEAN UNION, 2014b). O RIS conta com onze indicadores classificados segundo um modelo de dois níveis de classificação. A lista dos indicadores é apresentada no Quadro 12.

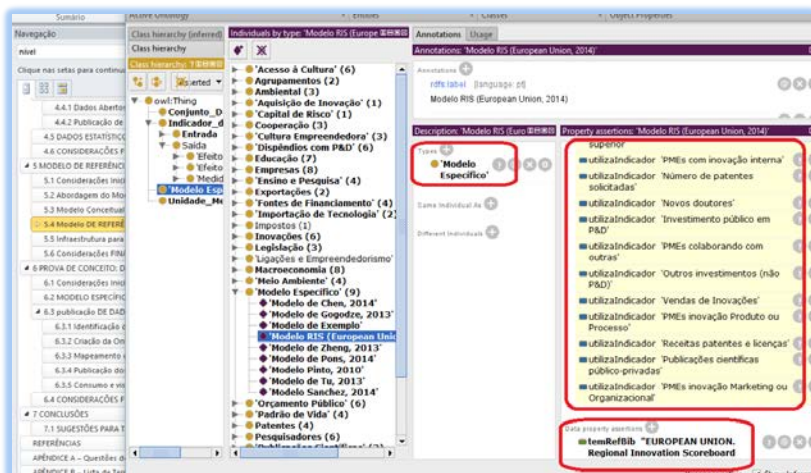
Quadro 12 - Lista de indicadores utilizados no *Regional Innovation Scoreboard*

Ord	Indicador
1	Porcentagem da população entre 25 e 64 anos com ensino superior completo
2	Despesas em P&D no setor público (% do PIB)
3	Despesas em P&D no setor privado (% do PIB)
4	Despesas com inovação que não sejam P&D (% do valor dos negócios)
5	PMEs com inovação interna (% das PMEs)
6	PMEs inovadoras que colaboraram com outras (% das PMEs)
7	Pedidos de patentes EPO por bilhão do PIB regional
8	PMEs que introduziram inovação de produto ou processo (% das PMEs)
9	PMEs que introduziram inovações de <i>marketing</i> ou organizacionais (% das PMEs)
10	Ocupação em manufaturas de média/alta tecnologia e em serviços intensivos em conhecimento (% da força de trabalho)
11	Vendas de inovações novas para o mercado e novas para a empresa (% do valor dos negócios)

Fonte: Adaptado de European Union (2014b)

A instanciação do modelo específico do exemplo é feita no editor *Protégé* e de seus indicadores. A Figura 37 exibe a interface do editor com o modelo específico selecionado, a indicação de seu tipo e de sua referência bibliográfica. Além disso são listados os indicadores que estão vinculados ao modelo pela propriedade “*utilizaIndicador*”.

Figura 41 - Instância de modelo específico

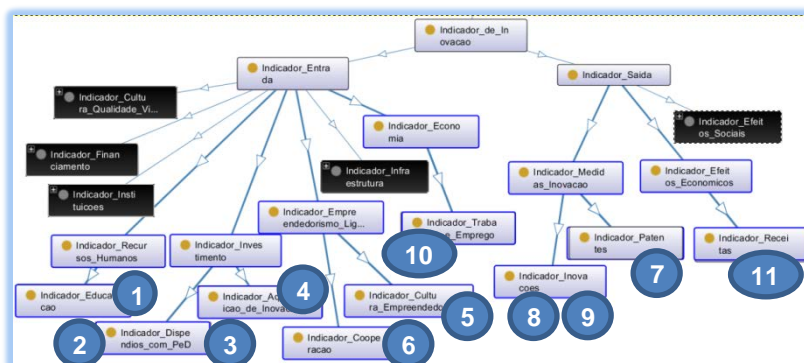


Fonte: O autor

Com o objetivo de ilustrar a vinculação do modelo específico ao modelo de referência, a Figura 38 exibe parte da hierarquia das classes da ontologia, com as indicações dos indicadores correspondentes instanciados de acordo com a numeração em destaque. Conforme é possível observar, as classes representadas em cor mais escura não tiveram suas subclasses instanciadas.

Os números exibidos na Figura 38 representam os indicadores listados no Quadro 12 e estão posicionados sobre as classes às quais as instâncias estão ligadas. A representação indica, ainda, que nem todas as classes do segundo e terceiro níveis do modelo de referência precisam, obrigatoriamente, ser instanciadas para um modelo específico. Isso não impede, entretanto, que outros modelos específicos que venham a ser catalogados, tenham indicadores como instâncias em qualquer uma das classes.

Figura 42 - Classes instanciadas com os indicadores do RIS



Fonte: O autor

Embora os modelos específicos apresentem diferentes estruturas, em termos de quantidade de classes, bem como de níveis de classificação, a catalogação segundo este modelo de referência possibilita que se crie um rol de indicadores vinculados a cada classe, que crescerá à medida que o número de modelos catalogados cresça.

5.5 INFRAESTRUTURA PARA A PUBLICAÇÃO DE DADOS ESTATÍSTICOS LIGADOS SOBRE A INOVAÇÃO REGIONAL

O Modelo de Referência proposto tem por característica a exploração do potencial de utilização de dados disponíveis na Web. Pela aplicação das técnicas de dados ligados, os dados oriundos de diferentes

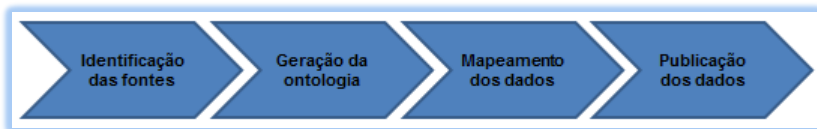
fontes e publicadores podem ganhar nova finalidade ao serem vinculados ao modelo ora proposto.

Grande parte dos dados com potencial utilização para os indicadores de inovação são disponibilizados por órgãos públicos em função de iniciativas de transparência e legislações específicas, tais como a Lei de Acesso à Informação (BRASIL, 2011). Assim, embora haja uma suposta disponibilidade de dados, eles são, tipicamente, heterogêneos e não têm um tratamento estatístico integrado, além de não haver uma forma de expor, compartilhar ou interligar os dados para que possam ser acrescentadas mais informações (ZAPILKO; MATHIAK, 2011; MILOŠEVIĆ *et al.*, 2012; ZANCANARO *et al.*, 2013).

Na revisão de literatura sobre Dados Ligados realizada no âmbito desta tese, e apresentada em Dal Pizzol *et al.* (2015), uma série de microtemas são apresentados para a classificação dos trabalhos publicados na área. Dentre estes, está o microtema referente à publicação de dados estatísticos. Esta temática é explorada por Zancanaro *et al.* (2013), onde o modelo de cubos de dados multidimensionais, característicos dos sistemas OLAP (*Online Analytical Processing*), é apresentado como uma abordagem adequada para a publicação de Dados Ligados estatísticos.

Zancanaro *et al.* (2013) apresentam, ainda, um processo automatizado para construção destes cubos de dados, composto por quatro passos: Identificação das fontes; Geração da ontologia; Mapeamento dos dados; e, Publicação dos dados. A Figura 39 representa a sequência dos passos.

Figura 43 - Passos do processo de publicação de dados ligados



Fonte: Zancanaro *et al.* (2013)

O primeiro passo consiste da identificação de fontes de dados, onde devem ser observados critérios como a precisão, temporalidade, completude e consistência dos dados. As principais fontes de dados são as páginas de órgãos governamentais, que, por força de lei, devem publicar seus conjuntos de dados para acesso público.

No segundo passo, a criação ou reutilização da(s) ontologia(s) tem por objetivo prover uma especificação explícita da conceitualização, por meio de uma representação semântica das dimensões e observações

que descrevem os dados. Neste passo, em específico, a ontologia do Modelo de Referência já está disponível e será utilizada no mapeamento para publicação dos dados.

A etapa de mapeamento dos dados consiste na análise dos dados oriundos das fontes selecionadas e sua interligação com os conceitos representados na(s) ontologia(s). Esta interligação é feita pelo mapeamento dos dados com as URIs, da(s) ontologia(s) ou de outros Dados Ligados.

Finalmente, a etapa de publicação consiste na transformação dos conjuntos de dados, previamente mapeados, em um modelo de cubo multidimensional OLAP. O vocabulário adotado para a representação dos cubos é o *RDF Data Cube* (CYGANIAK; REYNOLDS; TENNISON, 2014). Esta transformação é feita de forma automatizada por meio da ferramenta apresentada em Zancanaro *et al.* (2013), e o conjunto de dados resultante em formato RDF pode, então, ser carregado em um servidor de Dados Ligados do tipo *Triple Store*.

A publicação dos dados em servidores do tipo *Triple Store* possibilita que estejam acessíveis por meio de consultas SPARQL. Além disso, pelo fato de estarem disponibilizados segundo uma estrutura padronizada de cubos de Dados Estatísticos Ligados, usando o vocabulário *RDF Cube*, as operações de consulta também são padronizadas.

Figura 44 - Visualização da URI de uma classe do modelo no navegador

The screenshot shows a web browser window displaying the URI of a class in a model. The page is titled "About: Recursos Humanos" and is part of the "OPEN LINK SOFTWARE" project, powered by "VIRTUOSO". The page has three tabs: "Description", "Usage Statistics", and "Settings". The main content area shows the following information:

About: Recursos Humanos
 An Entity of Type : [Class](#), within Data Space : [lod.rafaelsperoni.pro.br](#)
[Start faceted browsing from this Type](#)

DEFINIÇÃO: Compreende indicadores relativos à educação e a disponibilidade de força de trabalho altamente qualificada e educada (EUROPEAN UNION, 2014).

Attributes	Values
type	owl:Class
label	Recursos Humanos
comment	DEFINIÇÃO: Compreende indicadores relativos à educação e a disponibilidade de força de trabalho altamente qualificada e educada (EUROPEAN UNION, 2014).
subClassOf	http://lod.rafaelsperoni.pro.br/onto/RegionalInnovation/Indicador_Entrada
is type of	http://lod.rafaelsperoni.pro.br/resource/newDoctors http://lod.rafaelsperoni.pro.br/resource/personnelReD http://lod.rafaelsperoni.pro.br/resource/lifelongLearning http://lod.rafaelsperoni.pro.br/resource/populationTertiaryEducation http://lod.rafaelsperoni.pro.br/resource/doctorResearchFirm

Fonte: O autor

A Figura 44 apresenta a representação de um recurso correspondente a uma classe do modelo, já publicada na forma de dados ligados. Desta forma, quando sua URI (no exemplo, http://lod.rafaelsperoni.pro.br/onto/RegionalInnovation/Indicador_Recur_sos_Humanos) é solicitada por um navegador, o servidor executa uma consulta SPARQL e o resultado é exibido em uma página *Web*.

O mesmo conteúdo exibido em formato legível para os usuários pode ser acessado programaticamente. A Figura 45 exibe um exemplo de consulta SPARQL solicitando a descrição de um recurso identificado pela URI especificada.

Figura 45 - Exemplo de consulta no *endpoint* SPARQL

Virtuoso SPARQL Query Editor

[About](#) | [Namespace Prefixes](#) | [Inference rules](#)

Default Data Set Name (Graph IRI)

Query Text

```
describe <http://lod.rafaelsperoni.pro.br/onto/RegionalInnovation/Indicador_Recur_sos_Humanos>
```

(Security restrictions of this server do not allow you to retrieve remote RDF data, see [details](#).)

Results Format: N3/Turtle

Execution timeout: 0 milliseconds (values less than 1000 are ignored)

Options: Strict checking of void variables

(The result can only be sent back to browser, not saved on the server, see [details](#))

Run Query Reset

Fonte: O autor

Esta mesma consulta pode ter seu resultado gerado em diferentes formatos além do HTML exibido na página da Figura 44. A Figura 46 apresenta exemplos de resultados em formatos adequados ao processamento, como JSON e RDF/XML.

Figura 46 - Resultados de consulta em diferentes formatos

```

{ "head": { "link": {}, "vars": [ "s", "p", "o" ] },
  "results": {
    { "s": {
      "value": "http://lod.rafael.speroni.pro.br/ontologia/indicadores/indicadores_relacionados_religiao",
      "uri": "http://lod.rafael.speroni.pro.br/ontologia/indicadores/indicadores_relacionados_religiao",
      "lang": "pt-BR"
    },
    { "p": {
      "value": "rdf:type",
      "uri": "http://www.w3.org/2002/07/owl#Class"
    },
    { "o": {
      "value": "http://lod.rafael.speroni.pro.br/ontologia/indicadores/indicadores_relacionados_religiao",
      "uri": "http://lod.rafael.speroni.pro.br/ontologia/indicadores/indicadores_relacionados_religiao",
      "lang": "pt-BR"
    }
  }
}

```

Fonte: O autor

Com o uso de bibliotecas específicas, aplicações computacionais podem realizar consultas SPARQL que serão executadas em servidores de dados ligados, recebendo os dados estruturados em resposta e fazendo sua manipulação conforme necessário. Dessa forma, tanto os dados referentes aos indicadores catalogados quanto a estrutura do modelo estão acessíveis por meio de consultas.

5.6 CONSIDERAÇÕES FINAIS

A inovação regional costuma ser analisada na literatura como um processo sistêmico, que acontece pela interação entre diferentes tipos de atores e é diretamente influenciada pelas pré-condições do ambiente que os cerca. Dessa forma, a mensuração da inovação regional deve levar em conta os múltiplos aspectos desse Sistema Regional de Inovação, e necessita de indicadores que os descrevam.

De forma semelhante ao que acontece com os Sistemas Nacionais de Inovação, a abordagem comumente adotada para a mensuração da inovação regional envolve a adoção de modelos de indicadores compostos, cujas escolhas de variáveis e formas de organização variam entre os autores. Este capítulo apresentou a concepção de um modelo de referência baseado em uma busca de literatura, onde foram identificadas 52 publicações, que foram catalogadas.

A partir dos dados levantados, foram identificados os indicadores utilizados pelos autores e, de acordo com suas características, foram agrupados em classes que representam os constructos que estão relacionados com a inovação regional. As classes foram organizadas em

um modelo conceitual com estrutura hierárquica composta por três níveis.

O modelo conceitual foi representado na forma de uma ontologia, onde os conceitos representados foram descritos e interligados para que, posteriormente, fosse publicada na forma em uma infraestrutura específica para dados ligados. A ontologia publicada neste formato torna-se acessível tanto via navegador *Web*, quanto via interface de consulta SPARQL, possibilitando seu consumo por aplicações computacionais.

Por suas características, os dados ligados da ontologia podem ser facilmente consultados como fonte de referência para a definição de modelos específicos de indicadores da inovação regional. Os indicadores catalogados e ligados ao modelo de referência podem servir como um ponto de partida para a escolha de indicadores a serem utilizados.

O próximo capítulo apresenta uma prova de conceito desenvolvida para demonstrar a viabilidade do modelo proposto como referência para a representação de indicadores da inovação regional. Um protótipo foi desenvolvido, demonstrando o processo de coleta e publicação dos dados ligados, além de seu consumo por uma aplicação *Web* com uma interface que propicia a análise comparativa entre diferentes regiões.

6 PROVA DE CONCEITO: DADOS LIGADOS SOBRE A INOVAÇÃO NOS MUNICÍPIOS DE SANTA CATARINA

6.1 CONSIDERAÇÕES INICIAIS

Este capítulo apresenta um protótipo desenvolvido como prova de conceito, termo aplicável à realização de uma ideia que tem por objetivo demonstrar sua viabilidade (PAPADOPOULOS *et al.*, 2016). São descritos o processo publicação de dados ligados, bem como sua utilização em uma aplicação *Web* embasada no modelo de referência proposto.

O Instituto Brasileiro de Geografia e Estatística (IBGE) subdivide os estados brasileiros em mesorregiões e microrregiões, nas quais os municípios são agrupados segundo o processo social, o quadro natural, a rede de comunicação, a estrutura de produção e a interação pessoal (IBGE, 1990). O estado de Santa Catarina é subdividido em seis mesorregiões e 20 microrregiões.

A prova de conceito desenvolvida neste trabalho descreve a coleta de dados referentes a duas mesorregiões do estado de Santa Catarina, buscando-se os dados disponíveis em nível de municípios. Foram escolhidas as mesorregiões Norte e Grande Florianópolis. Um dos pressupostos para a obtenção dos dados foi a identificação de conjuntos de dados que estivessem disponíveis como dados abertos, em páginas *Web* cuja origem pudesse ser verificada, mas sem que houvesse a necessidade de solicitação explícita ou formal por qualquer conjunto de dados.

6.2 MODELO ESPECÍFICO DE INDICADORES DE INOVAÇÃO

Tendo como ponto de partida o Modelo de Referência para Indicadores da Inovação Regional suportado por Dados Ligados, a prova de conceito aqui apresentada pretende demonstrar a utilização de um modelo específico, que estará vinculado ao modelo de referência. No caso em análise, foram catalogados os indicadores utilizados por oito modelos específicos (PINTO; GUERREIRO, 2010; EUROPEAN UNION, 2012; GOGODZE, 2013; TU; SONG, 2013; ZHENG; ZHANG, 2013; CHEN; JU, 2014; PONS; MARTÍNS; PARRILLI, 2014; SÁNCHEZ TOVAR; GARCÍA FERNÁNDEZ; MENDOZA FLORES, 2014).

Pela catalogação dos oito modelos, foram identificados 97 indicadores diferentes, que foram classificados segundo a ontologia do Modelo de Referência.

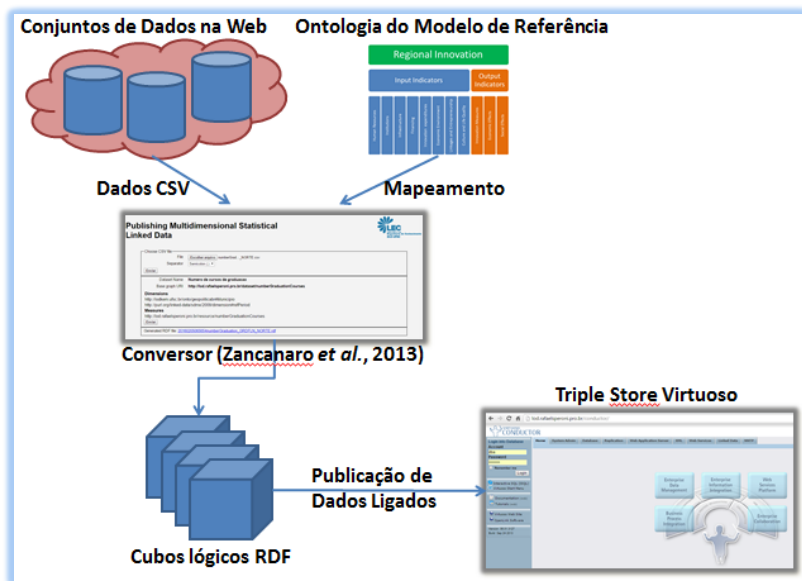
6.3 DADOS LIGADOS SOBRE A INOVAÇÃO REGIONAL

A publicação dos Dados Ligados consiste da transformação de dados já existentes, disponíveis em outros formatos, e sua disponibilização segundo formatos e técnicas específicas que estabelecem vínculos entre diferentes conjuntos de dados, bem como com suas descrições.

Com base no processo descrito por Zancanaro *et al.* (2013), a publicação dos dados estatísticos ligados segue uma sequência de quatro passos: Identificação das fontes; Geração da ontologia; Mapeamento dos dados; e, Publicação dos dados.

A Figura 47 apresenta a abordagem utilizada na obtenção, mapeamento e publicação dos dados para o caso em estudo.

Figura 47 – Publicação de Dados Ligados sobre a Inovação Regional



Fonte: O autor

A identificação das fontes de dados é feita pela busca em páginas *Web* por conjuntos de dados disponíveis para acesso público. A ontologia criada para a representação do modelo de referência é utilizada na fase de mapeamento, onde os dados são ligados aos conceitos nela descritos.

Uma vez mapeados em relação à ontologia, os dados são transformados para uma representação de cubo lógico dimensional, no vocabulário *RDF Data Cube*, que são, posteriormente carregados no servidor do tipo *Triple Store*.

As etapas da publicação dos dados são descritas nas próximas seções.

6.3.1 Identificação das Fontes de Dados

O primeiro passo, identificação das fontes de dados, consiste da procura pelos conjuntos de dados que correspondem ao modelo. Assim, foram buscados conjuntos de dados disponíveis em páginas *Web*, com acesso público.

A busca consistiu no acesso às páginas *Web* de órgãos, programas e iniciativas, sendo elas governamentais ou não, visando identificar conjuntos de dados cujos conteúdos fossem os mais próximos às necessidades das variáveis contidas nos modelos. O Quadro 13 apresenta a lista dos conjuntos de dados identificados, bem como suas fontes e os formatos em que os dados estavam disponíveis.

Quadro 13 - Lista de fontes e conjuntos de dados utilizados no protótipo

Conjunto de dados	Nível de agregação	Fonte	Formato disponível
Informações geopolíticas sobre os Municípios do Estado de SC	Município	Endpoint SPARQL do LODKEM http://lodkem.led.ufsc.br:8890/sparql	RDF (consultas via SPARQL)
Número de Instituições de Ensino Superior	Município	Censo da Educação Superior 2013 http://portal.inep.gov.br/web/censo-da-educacao-superior	Página <i>Web</i> com exportação de planilha
Número de estudantes matriculados no Ensino Superior	Município	Censo da Educação Superior 2013 http://portal.inep.gov.br/web/censo-da-educacao-superior	Página <i>Web</i> com exportação de planilha
Total de Habitantes e Densidade populacional	Município	Cidades IBGE (Censo 2013) http://cidades.ibge.gov.br/	Página <i>Web</i> com exportação para .CSV

Renda per Capita	Município	Cidades IBGE (Censo 2013) http://cidades.ibge.gov.br/	Página Web com exportação para .CSV
Número de doutores	Município	Observatório de Egressos da Pós-Graduação (2014) http://ovegressos.com.br/	Página Web
Número de empregos formais	Município	Portal do Trabalho e Emprego (TEM) (2015) http://acesso.mte.gov.br/porta-pdet/	Página Web com exportação para .CSV
Número de Bibliotecas públicas	Município	Sistema Nacional de Bibliotecas Públicas (Ministério da Cultura) http://bibliotecas.cultura.gov.br/	Página Web (Mapa)
Número de ONGs	Município	ONGs Brasil (2016) http://ongsbrasil.com.br/	Página Web
Número de empregos no Setor de Serviços	Município	Portal do Trabalho e Emprego (TEM) (2015) http://acesso.mte.gov.br/porta-pdet/	Página Web com exportação para .CSV
PIB per Capita	Município	Cidades IBGE (Censo 2013) http://cidades.ibge.gov.br/	Página Web com exportação para .CSV
População com abastecimento de água	Município	Sistema Nacional de Informações sobre Saneamento (2015) http://www.snis.gov.br/	Página Web com exportação para .CSV
População atendida com esgotamento sanitário	Município	Sistema Nacional de Informações sobre Saneamento (2015) http://www.snis.gov.br/	Página Web com exportação para .CSV

Fonte: o autor

Da descrição dos conjuntos de dados obtidos, é possível observar que a maioria estava disponível para *download* em formato CSV, um

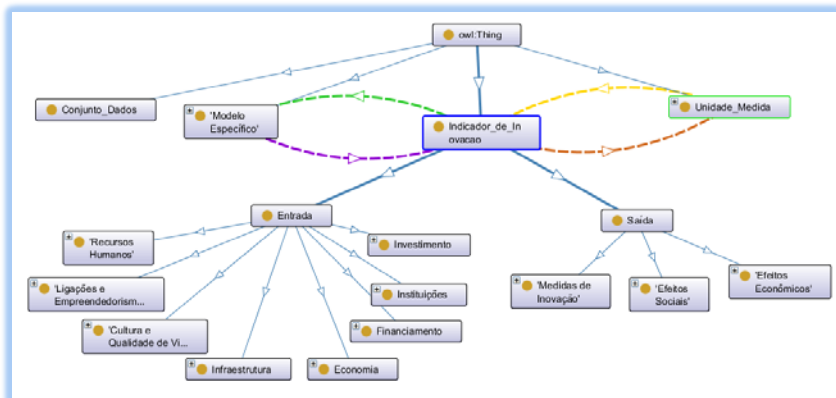
padrão aberto. Embora esta característica torne a manipulação dos dados mais acessível, o fato de serem disponibilizados por fontes diferentes, e com diferentes finalidades, gera a necessidade de um tratamento específico, no sentido de promover sua integração, corroborando o entendimento de Zapilko e Mathiak (2011) e Milošević *et al.* (2012) sobre a necessidade de um tratamento estatístico integrado e de exposição, compartilhamento ou interligação dos dados.

A integração dos diferentes conjuntos de dados é feita pela contextualização dos dados neles contidos enquanto variáveis que representam características específicas do processo de inovação. Assim, a segunda etapa, de criação da(s) ontologia(s), tem como finalidade a vinculação dos dados aos conceitos descritos semanticamente na forma de ontologias.

6.3.2 Criação da Ontologia

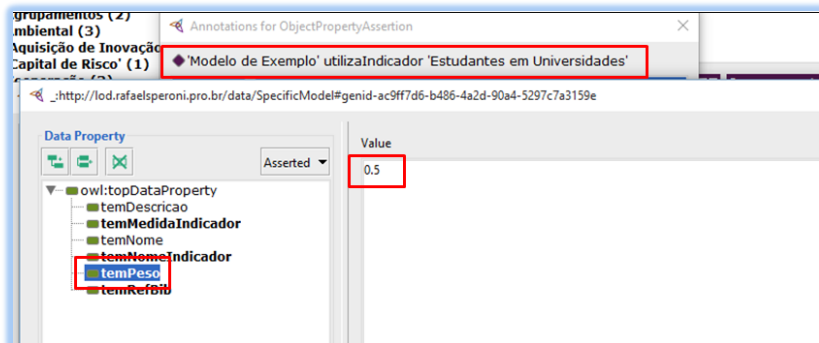
Além da ontologia descrita na seção 5.4.1, foi criada uma ontologia para a representação de modelos específicos de indicadores de inovação regional. Nesta ontologia, os modelos específicos são descritos e instanciados, fazendo a ligação de cada indicador com o modelo de referência, classificando-os segundo a hierarquia de classes proposta. A Figura 44 apresenta uma visualização de parte da hierarquia das classes da ontologia.

Figura 48 - Hierarquia de classes da ontologia



Fonte: O autor

A hierarquia das classes traz a representação da organização dos indicadores do modelo conceitual. Além disso, uma classe representa os modelos específicos que virão a ser representados pela sua instanciação. Figura 49 - Atribuição de peso a um indicador em um modelo



Fonte: O autor

Ao instanciar um modelo específico, os indicadores específicos são instanciados de acordo com sua classificação, e podem ter um peso a eles atribuídos, de forma que a posterior agregação dos dados seja feita de forma ponderada. Na Figura 45, um peso de 0.5 é atribuído ao indicador ‘Estudantes em Universidades’, quando associado ao ‘Modelo de Exemplo’.

6.3.3 Mapeamento dos Dados

A etapa de mapeamento dos dados tem por finalidade explicitar a semântica dos dados, por meio da ligação de cada dado a um conceito ou indivíduo representado na ontologia. A Figura 46 apresenta um trecho de um arquivo contendo dados sobre a Renda per Capita dos municípios de Santa Catarina, no formato CSV, onde o separador de colunas é representado por um ponto e vírgula (;). Desta forma, a primeira coluna indica um nome de município, a segunda indica o ano de referência, e a terceira contém o valor da medida.

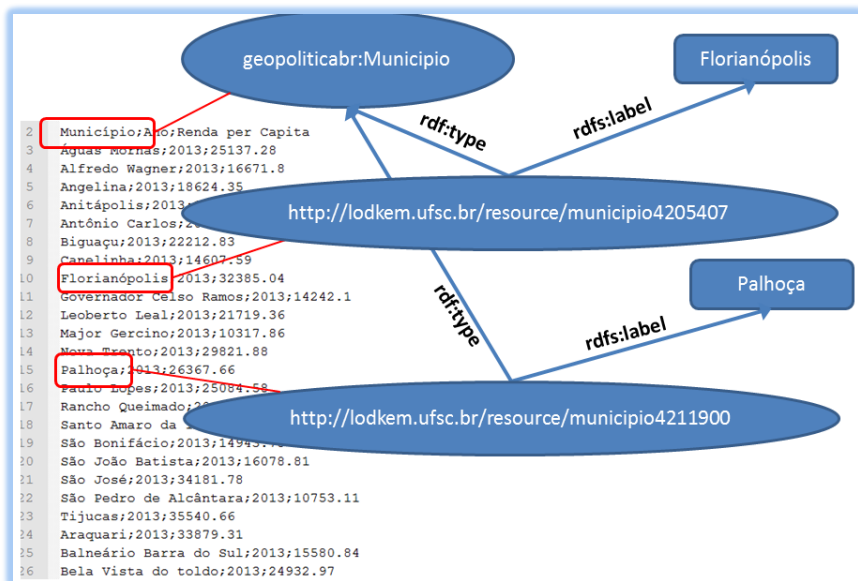
Figura 50 - Trecho de arquivo em formato CSV

```
1 Município;Ano;Renda per Capita
2 Águas Mornas;2010;25137.28
3 Alfredo Wagner;2010;16671.8
4 Angelina;2010;18624.35
5 Anitápolis;2010;16869.1
6 Antônio Carlos;2010;60936.38
7 Biguaçu;2010;22212.83
8 Canelinha;2010;14607.59
9 Florianópolis;2010;32385.04
10 Governador Celso Ramos;2010;14242.1
11 Leoberto Leal;2010;21719.36
12 Major Gercino;2010;10317.86
13 Nova Trento;2010;29821.88
14 Palhoça;2010;26367.66
15 Paulo Lopes;2010;25084.58
16 Rancho Queimado;2010;32484.85
17 Santo Amaro da Imperatriz;2010;29967.34
18 São Bonifácio;2010;14943.78
19 São João Batista;2010;16078.81
```

Fonte: O autor

A etapa recebe o nome de mapeamento, pois trata do estabelecimento das ligações existentes entre cada dado contido no conjunto de dados e os recursos descritos na ontologia, e nos demais dados ligados já existentes. Assim, trata-se da identificação de URIs que representem cada dado. A Figura 47 ilustra uma parte do mapeamento que deve ser feito para que os dados contidos no arquivo CSV do exemplo sejam ligados à classe que representa o conceito “Município” e às instâncias que representam cada um dos municípios.

Figura 51 - Ilustração do mapeamento de municípios



Fonte: O autor

No mapeamento apresentado no exemplo, o título de coluna “Município” é mapeado para a URI da classe “[geopoliticabr:Municipio](#)”. De forma semelhante, o dado “Florianópolis” é mapeado para a URI “[http://lodkem.ufsc.br/resource/municipio4205407](#)”, assim como o dado “Palhoça” é mapeado para a URI “[http://lodkem.ufsc.br/resource/municipio4211900](#)”.

O mesmo princípio valerá para os demais títulos de coluna e aqueles dados que estejam representados na ontologia ou em outros conjuntos de dados ligados. A Figura 52 apresenta o arquivo CSV, desta vez visualizado em um editor de planilha eletrônica, contendo os dados após o mapeamento.

Figura 52 - Arquivo CSV após o mapeamento

	A	B	C	D	E	F
1	Município	Ano	Renda Per Capta			
2	http://lodkem.ufsc.br/onto/geopoliticabr#Município	http://purl.org/linked-data/sdmx/2009/dimension#refPeriod	http://lod.rafaelsperoni.pro.br/resource/revenuePerCapita			
3	http://lodkem.ufsc.br/resource/municipio4200606	http://reference.data.gov.uk/id/year/2010	668.5			
4	http://lodkem.ufsc.br/resource/municipio4200706	http://reference.data.gov.uk/id/year/2010	610			
5	http://lodkem.ufsc.br/resource/municipio4200803	http://reference.data.gov.uk/id/year/2010	740			
6	http://lodkem.ufsc.br/resource/municipio4201109	http://reference.data.gov.uk/id/year/2010	616.67			
7	http://lodkem.ufsc.br/resource/municipio4201208	http://reference.data.gov.uk/id/year/2010	784			
8	http://lodkem.ufsc.br/resource/municipio4202305	http://reference.data.gov.uk/id/year/2010	650			
9	http://lodkem.ufsc.br/resource/municipio4203709	http://reference.data.gov.uk/id/year/2010	600			
10	http://lodkem.ufsc.br/resource/municipio4205407	http://reference.data.gov.uk/id/year/2010	1166.67			
11	http://lodkem.ufsc.br/resource/municipio4206009	http://reference.data.gov.uk/id/year/2010	605			
12	http://lodkem.ufsc.br/resource/municipio4209606	http://reference.data.gov.uk/id/year/2010	617.5			
13	http://lodkem.ufsc.br/resource/municipio4210209	http://reference.data.gov.uk/id/year/2010	700			
14	http://lodkem.ufsc.br/resource/municipio4211504	http://reference.data.gov.uk/id/year/2010	765			
15	http://lodkem.ufsc.br/resource/municipio4211900	http://reference.data.gov.uk/id/year/2010	740			
16	http://lodkem.ufsc.br/resource/municipio4212304	http://reference.data.gov.uk/id/year/2010	537.5			
17	http://lodkem.ufsc.br/resource/municipio4214300	http://reference.data.gov.uk/id/year/2010	765			
18	http://lodkem.ufsc.br/resource/municipio4215703	http://reference.data.gov.uk/id/year/2010	757.5			
19	http://lodkem.ufsc.br/resource/municipio4215901	http://reference.data.gov.uk/id/year/2010	750			
20	http://lodkem.ufsc.br/resource/municipio4216306	http://reference.data.gov.uk/id/year/2010	725			
21	http://lodkem.ufsc.br/resource/municipio4216602	http://reference.data.gov.uk/id/year/2010	900			
22	http://lodkem.ufsc.br/resource/municipio4217253	http://reference.data.gov.uk/id/year/2010	745			

Fonte: O autor

Os arquivos CSV gerados com os dados mapeados são, posteriormente, submetidos a um processo de transformação, do qual resultarão cubos lógicos dimensionais codificados no vocabulário *RDF Data Cube*. Para execução de tal transformação, utilizou-se a ferramenta desenvolvida e apresentada no artigo “*Publishing Multidimensional Statistical Linked Data*” (ZANCANARO *et al.*, 2013), de coautoria do autor.

Figura 53 - Carga dos dados e parametrização do Cubo RDF

Publishing Multidimensional Statistical Linked Data

Choose CSV file

File: revenuePerC..._NORTE.csv

Separator: ▼

Dataset Name:

Graph URI:

Dimensions Measures

<http://lodkem.ufsc.br/onto/geopoliticabr#Município>

<http://purl.org/linked-data/sdmx/2009/dimension#refPeriod>

<http://lod.rafaelsperoni.pro.br/resource/revenuePerCapita>

Fonte: O autor

Na Figura 53 é apresentada a interface da aplicação para transformação dos dados. São informados o arquivo CSV contendo os dados mapeados, o tipo de separador de colunas nele utilizado, um nome e uma URI base para identificação do conjunto de dados resultante, além de serem listadas colunas do arquivo, para que sejam informados quais deles serão dimensões e quais serão medidas no cubo gerado.

Figura 54 - Trecho de dados transformados em *RDF Data Cube*

```

</rdf:Description>
- <qh:Observation rdf:about="http://lod.rafaelsperoni.pro.br/dataset/revenuePerCapita#observation9">
  <qb:dataSet rdf:resource="http://lod.rafaelsperoni.pro.br/dataset/revenuePerCapita#datasetRendapercapita"/>
  <geopoliticabr:Municipio rdf:resource="http://lodkem.ufsc.br/resource/municipio4205407"/>
</qb:Observation>
<geopoliticabr:Municipio rdf:about="http://lodkem.ufsc.br/resource/municipio4205407">
- <rdf:Description rdf:about="http://lod.rafaelsperoni.pro.br/dataset/revenuePerCapita#observations9">
  <ns1:refPeriod rdf:resource="http://reference.data.gov.uk/id/year/2010"/>
</rdf:Description>
<ns1:refPeriod rdf:about="http://reference.data.gov.uk/id/year/2010">
- <rdf:Description rdf:about="http://lod.rafaelsperoni.pro.br/dataset/revenuePerCapita#observation9">
  <ns2:revenuePerCapita>1166.67</ns2:revenuePerCapita>
</rdf:Description>
- <qh:Observation rdf:about="http://lod.rafaelsperoni.pro.br/dataset/revenuePerCapita#observation10">
  <qb:dataSet rdf:resource="http://lod.rafaelsperoni.pro.br/dataset/revenuePerCapita#datasetRendapercapita"/>
  <geopoliticabr:Municipio rdf:resource="http://lodkem.ufsc.br/resource/municipio4206009"/>
</qb:Observation>
<geopoliticabr:Municipio rdf:about="http://lodkem.ufsc.br/resource/municipio4206009">
- <rdf:Description rdf:about="http://lod.rafaelsperoni.pro.br/dataset/revenuePerCapita#observation10">
  <ns1:refPeriod rdf:resource="http://reference.data.gov.uk/id/year/2010"/>
</rdf:Description>
<ns1:refPeriod rdf:about="http://reference.data.gov.uk/id/year/2010">
- <rdf:Description rdf:about="http://lod.rafaelsperoni.pro.br/dataset/revenuePerCapita#observation10">

```

Fonte: O autor

O trecho de apresentado na Figura 54 demonstra uma porção de dados já transformados para RDF, descritos segundo o vocabulário *RDF Data Cube*. No trecho, a URI do município Florianópolis, em destaque, está vinculada a uma observação, que conta, ainda, com a dimensão tempo definida para o ano de 2010, e tem um valor de 1166.67.

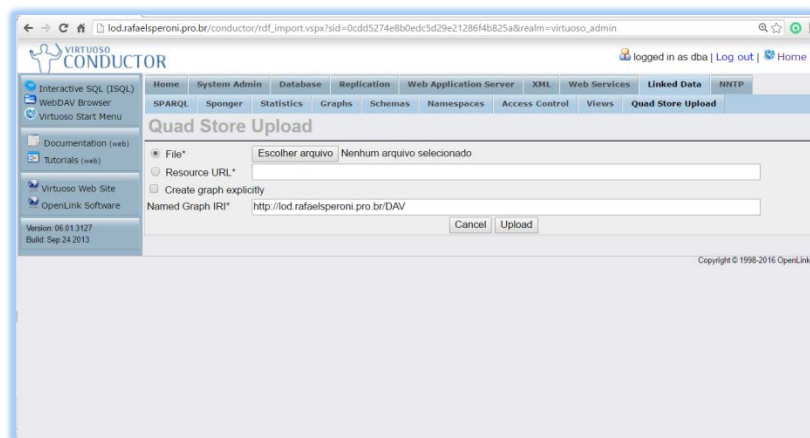
6.3.4 Publicação dos Dados

A etapa de publicação dos dados consiste em fazer a carga dos dados RDF em um servidor com a capacidade de torna-los acessíveis, seja via navegador *Web*, seja via consultas SPARQL. Na infraestrutura preparada para o desenvolvimento da prova de conceito, utilizou-se como servidor de dados ligados o *Virtuoso Universal Server* em sua distribuição de código aberto.

Uma vez configurado o servidor, a publicação dos dados consiste no envio dos arquivos, devidamente codificados em RDF, para o servidor. Por meio do gerenciador via *Web*, denominado *Conductor*, o

envio dos dados é feito na aba *Quad Store Upload*, apresentada na Figura 55, onde são indicados o arquivo, bem como uma URI pela qual o grafo correspondente ao conjunto de dados será identificado.

Figura 55 - Interface para upload dos dados ligados



Fonte: O autor

À medida que os conjuntos de dados são carregados no servidor, tornam-se disponíveis para que sejam consumidos por usuários de forma direta, via navegador, ou de forma indireta, por meio de aplicações.

6.3.5 Consumo e visualização dos Dados

Esta seção descreve um protótipo desenvolvido para acesso aos dados ligados publicados, tendo por objetivo prover uma interface de visualização tabular e gráfica dos dados, agregados em diferentes níveis de divisões regionais. Todas as visualizações são construídas a partir de dados obtidos por meio de consultas ao *endpoint* SPARQL¹⁶.

Uma página *Web*¹⁷ foi construída para disponibilizar informações sobre o modelo, apresentando uma representação gráfica das principais classes, e *hyperlinks* que apontam para as URIs que as descrevem, já publicadas em Dados Ligados. A página de descrição do modelo é apresentada na Figura 56.

¹⁶ <http://lod.rafaelsperoni.pro.br/sparql>

¹⁷ <http://lod.rafaelsperoni.pro.br/inovacao/>

Figura 56 - Página *Web* de apresentação do modelo

Fonte: O autor

Além da representação gráfica do modelo, os *hyperlinks* levam à descrição dos recursos armazenados no *Triple Store*, conforme apresentado na Figura 57, onde, além da descrição da URI referente à classe “Educação”, são apresentados os indicadores já catalogados e que foram classificados segundo este tipo. Este tipo de apresentação de conteúdo é construída pelo *Triple Store*, para que, além de serem acessíveis por máquinas, os dados sejam legíveis por humanos.

Figura 57 - Apresentação de dados ligados no navegador

Fonte: O autor

Da mesma forma que os dados ligados são visíveis via navegador *Web*, são acessíveis programaticamente por meio de consultas ao *endpoint* SPARQL. Assim, de forma semelhante ao que acontece com sistemas de bancos de dados relacionais, os dados armazenados em um servidor de dados ligados podem ser consumidos por aplicações desenvolvidas para fins específicos.

A Figura 58 apresenta um exemplo de consulta SPARQL que busca um conjunto de dados, codificado no vocabulário *RDF Data Cube*, com os valores agrupados por microrregião.

Figura 58 - Consulta SPARQL por um conjunto de dados, agregado por microrregiões

```

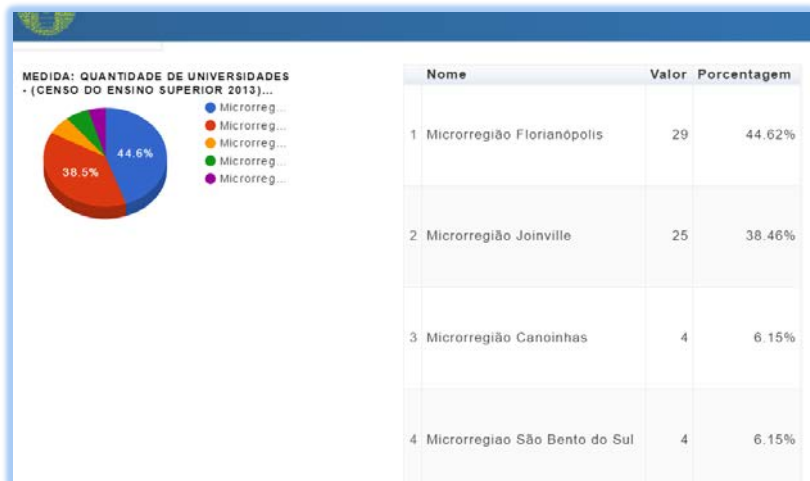
1 select
2   distinct ?uri_local ?nome_local sum(xsd:double(?valor)) as ?valor
3 where {
4   ?uri_dataset a qb:DataSet .
5   ?uri_dataset rdfs:label ?nm_dataset .
6   ?uri_dataset qb:structure ?dsd .
7   ?dsd qb:component ?uriMeasure .
8   ?uriMeasure qb:measure ?conceptMeasure .
9   ?obs ?conceptMeasure ?valor .
10  ?obs qb:dataSet ?uri_dataset .
11  ?obs geopolitical:Municipio ?uri_municipio .
12  ?uri_municipio <http://lod.rafaelsperoni.pro.br/resource/bertenceMicro> ?uri_local .
13  ?uri_local a <http://lod.rafaelsperoni.pro.br/resource/Microrregiao> .
14  ?uri_local rdfs:label ?nome_local
15  filter (?uri_dataset = <http://lod.rafaelsperoni.pro.br/dataset/numberIes>)
16 }
17 group by ?nome_local ?uri_local
18 order by desc(?valor)

```

Fonte: O autor

A consulta do exemplo é utilizada na aplicação desenvolvida para a representação tabular e gráfica de um conjunto de dados específico, conforme apresentado na Figura 59. A página apresenta uma tabela contendo o nome da microrregião, a quantidade de Instituições de Ensino Superior, e a porcentagem que esta quantidade representa do total. É exibido, ainda, um gráfico da proporção representada por cada microrregião.

Figura 59 - Visualização de tabular e gráfica de conjunto de dados, agregado por microrregiões



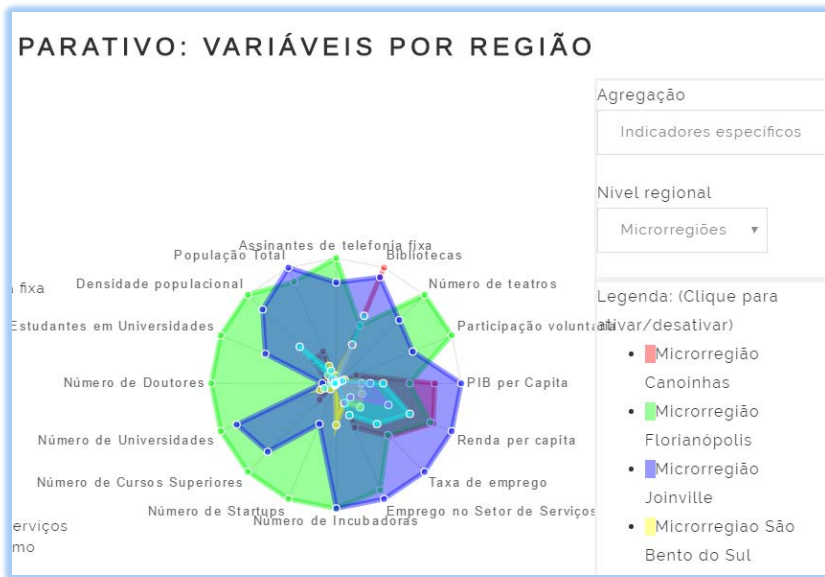
Fonte: O autor

Conforme é possível observar dos dados, extraídos do Censo da Educação Superior 2013: A microrregião de Florianópolis conta com 29 IES, representando 44,6% do total; a microrregião Joinville, conta com 25 IES, correspondendo a 38,5% do total; as microrregiões Canoinhas e São Bento do Sul contam com quatro IES, representando 6,15% do total, cada; a Microrregião Tijucas tem três IES, representando 4,62% do total; e, a microrregião Tabuleiro não conta com nenhuma IES. Diante deste resultado, é razoável supor que o componente “Infraestrutura”, especificamente no quesito “Educação”, no Modelo de Referência, tenha um desempenho mais expressivo nas microrregiões Florianópolis e Joinville, do que nas microrregiões Tijucas e Tabuleiro.

Outro tipo de visualização é apresentado na Figura 60. Neste exemplo, onze indicadores específicos são apresentados em um único gráfico do tipo radar, onde cada microrregião é representada por um conjunto de dados de cor diferente.

A observação dos dados apresentados no gráfico de radar permite uma análise comparativa entre as diferentes regiões. Os valores para cada indicador específico são normalizados segundo a abordagem da distância em relação ao melhor desempenho (*distance to the best performer*) (NARDO *et al.*, 2005), para que diferentes grandezas possam visualizadas em relação à região com melhor desempenho naquela medida.

Figura 60 - Visualização comparativa entre diferentes regiões



Fonte: O autor

As microrregiões com maior destaque no gráfico são Florianópolis, representada na cor verde, e Joinville, representada na cor roxa. Das dez medidas apresentadas, a microrregião Joinville, embora conte com maior número total de habitantes (842.709), apresenta maior deficiência no número de Doutores (414), enquanto que a microrregião Florianópolis, com um total de 740.926 habitantes, conta com 3.732 doutores.

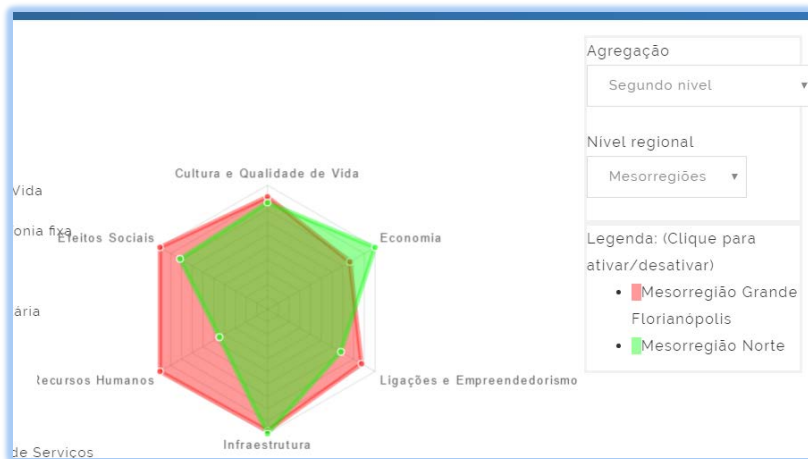
A análise do gráfico apresentado no exemplo indica que a microrregião Florianópolis destaca-se, em relação à microrregião Joinville, em quesitos relacionados à Educação, enquanto que tem desempenho inferior em quesitos relacionados com o Emprego e Renda.

A visualização dos dados pode ser modificada em termos do nível de agregação em que se deseja realizar a análise. Assim, a dimensão que descreve as regiões pode ter o seu nível de agregação aumentado (*drill up*) ou diminuído (*drill down*). De forma semelhante, a estrutura de classificação dos indicadores pode ser agregada segundo os diferentes níveis apresentados no modelo.

A Figura 61 apresenta um gráfico de radar contendo dois conjuntos de dados, sendo estes relativos às mesorregiões Grande

Florianópolis e Norte. Quanto aos eixos apresentados, a agregação se dá, agora, ao terceiro nível de classificação dos indicadores utilizados no modelo.

Figura 61 – Visualização comparativa entre regiões utilizando maiores níveis de agregação



Fonte: O autor

Conforme é possível observar, a Mesorregião Grande Florianópolis apresenta um desempenho superior nos eixos Educação, Voluntariado, e Cultura Empreendedora, enquanto que a Mesorregião Norte.

6.4 CONSIDERAÇÕES FINAIS

Considerando a proposta de que o Modelo de Referência seja suportado por dados ligados, este capítulo apresentou, como prova de conceito, a aplicação de um processo de coleta e publicação de dados ligados sobre a inovação regional, que contempla, ainda, o desenvolvimento de um protótipo para visualização e análise dos dados.

O pressuposto para esta prova de conceito foi a busca e reutilização de dados abertos sobre os municípios do Estado de Santa Catarina, que estivessem disponíveis na web para acesso público. Os dados passaram por um processo de mapeamento e transformação, até que fossem publicados na forma de dados ligados, usando o vocabulário *RDF Data Cube*, um formato adequado para a publicação e análise de dados estatísticos.

A utilização do *RDF Data Cube* garante escalabilidade na publicação de novos dados, que poderão ser mapeados para ontologia construída. Além disso, como se trata de uma estrutura padrão para dados estatísticos, a capacidade de análise sobre os novos dados é mantida. Ainda, outras formas de visualização e consultas SPARQL podem ser criadas para potencializar a capacidade de análise dos dados.

7 CONCLUSÕES

As características da sociedade do conhecimento, com uma economia dinâmica e fortemente ligada às condições científicas e tecnológicas, fazem com que a necessidade de adequação em resposta às mudanças seja uma preocupação constante dos gestores públicos ou privados. Para que uma empresa, município, região ou país mantenha-se em condição de competitividade e desenvolvimento econômico, são necessárias constantes intervenções voltadas à criação de soluções que os coloquem em posição de vantagem competitiva.

É neste cenário que a inovação tem ganhado destaque. Mais do que uma palavra largamente utilizada de forma promocional, a inovação, conforme a visão de Schumpeter, está associada às mudanças nas formas de se fazer determinadas atividades, que revolucionam constantemente a estrutura econômica pelo surgimento de novas tecnologias, produtos e empresas que substituem as antigas.

Quando considerada sobre o ponto de vista regional, a literatura aponta a necessidade de que a inovação seja analisada segundo um enfoque sistêmico. A partir do conceito dos Sistemas Nacionais de Inovação, o processo de inovação regional passa a ser analisado sob a perspectiva dos Sistemas Regionais de Inovação (SRI), onde se considera a importância dos diferentes atores envolvidos no processo, bem como seus interesses e interações na geração, uso e disseminação do conhecimento.

Para que a inovação aconteça em uma região, é necessário que sejam criadas condições favoráveis a estas interações. Os diferentes tipos de atores devem estar presentes e, mais do que isso, devem ser encorajados à interação. A título de exemplo, as interações universidade-empresa representam papel fundamental neste processo, mas, para tal, devem ser incentivadas, por meio da diminuição da burocracia, da disponibilidade de financiamento, e da formação de recursos humanos que atendam às necessidades regionais. Conhecer o cenário como um todo permite identificar os pontos positivos e as fragilidades, a fim de planejar ações de gestão da inovação.

Assim, a complexidade do processo da inovação faz com que a escolha por indicadores para sua mensuração não seja uma tarefa trivial. Costumeiramente, estatísticas de patentes e P&D são utilizadas como indicadores de inovação, mas sua aplicação de forma individual é apontada como insuficiente, podendo levar a análises equivocadas. Em outro tipo de abordagem, entretanto, os autores optam pela utilização de indicadores de inovação compostos, cuja concepção leva em conta uma

série de variáveis que retratam os diferentes aspectos do processo de inovação.

Indicadores compostos, portanto, são aqueles formados pela compilação, combinação ou agregação de indicadores individuais, estes últimos caracterizados como medidas de ordem quantitativa ou qualitativa, derivadas de uma série de fatos observados. A partir de uma busca sistemática de literatura, foram catalogadas e analisadas 52 publicações científicas que utilizam modelos de indicadores compostos para a inovação regional. De maneira geral, os autores utilizam classificações hierárquicas, nas quais os indicadores são agrupados segundo os aspectos da inovação a que seus objetos de observação estão relacionados.

A análise dos indicadores presentes na literatura, constatou-se que estão inseridos como objetos de observação uma série de componentes que dizem respeito a múltiplas facetas do ambiente regional, tais como características socioeconômicas, científicas e tecnológicas, educacionais, legais e ambientais. Tal constatação vai ao encontro do entendimento encontrado na literatura, que considera que os Sistemas Regionais de Inovação são identificados por uma seleção de indicadores-chave em vários aspectos da capacidade organizacional e de infraestrutura, competências e aptidões para a inovação em regiões.

Pela catalogação dos indicadores encontrados na literatura, buscou-se identificar os constructos da inovação a que estão associados e, a partir destes, desenvolveu-se um modelo conceitual que classifica os indicadores de inovação regional em três níveis hierarquicamente organizados. O primeiro nível é composto por duas, o segundo por onze, e o terceiro por 33 classes, que foram definidas em função das semelhanças nas características dos indicadores catalogados, visando fornecer um referencial para a escolha de indicadores em trabalhos futuros.

Os argumentos relatados como dificultadores na aplicação dos modelos ao nível regional foram tomados como uma lacuna de pesquisa a ser explorada. Para tal, desejou-se investigar a possibilidade de exploração dos dados abertos regionais disponibilizados na Web como alternativa de abordagem para a falta de disponibilidade de dados. Buscou-se verificar a hipótese de que tais dados, mesmo não tendo sido gerados para esta finalidade, possam ser utilizados na composição de indicadores de inovação regional.

Ao considerar o caso dos SRI e a disponibilidade de dados abertos que retratem a inovação, é possível supor que sejam necessárias múltiplas fontes de dados, de diferentes órgãos produtores e que,

consequentemente, estes dados não tragam nenhum tratamento estatístico integrado. Assim, a pluralidade de aspectos a avaliar necessita de uma estratégia de integração, onde os dados disponíveis sejam adequados ao contexto de sua participação enquanto indicadores de inovação regional.

Adotando-se o uso de tecnologias semânticas, uma técnica de publicação de dados na Web, chamada Dados Ligados, utiliza-se de um padrão onde os dados são contextualizados segundo uma representação semântica do domínio de aplicação em que se inserem, tornando-se, portanto, auto-descritivos e sendo armazenados segundo uma infraestrutura de publicação que possibilita seu consumo por meio de linguagem de consultas.

Uma vez definido o modelo de referência, partiu-se para sua representação na forma de uma ontologia, visando à representação do modelo em uma linguagem de tecnologia semântica, processável por máquina. Foi utilizada a metodologia do ontoKEM para a construção da primeira versão e, posteriormente, o editor *Protégé* para a finalização da ontologia e instanciação do modelo.

Uma importante contribuição deste trabalho diz respeito justamente à explicitação do modelo de referência, publicado em uma infraestrutura de dados ligados, uma vez que este formato de publicação possibilita a visibilidade do modelo não apenas em uma linguagem legível por humanos, mas também legível por máquinas. Assim, a estrutura do modelo e as inter-relações existentes entre os conceitos nele representados são processáveis por máquinas.

Foi apresentada, ainda uma de conceito. No protótipo de aplicação Web desenvolvido, as informações apresentadas são todas resultantes de consultas a uma infraestrutura de dados ligados. O protótipo teve por objetivo demonstrar ser viável a construção de uma ferramenta analítica, com visualização gráfica, onde os modelos e indicadores escolhidos são apresentados de forma a estabelecer um comparativo entre diferentes regiões.

Foi estabelecido como requisito para a obtenção de dados para o protótipo a utilização de conjuntos de dados abertos que estivessem disponíveis publicamente em páginas Web, sem que houvesse a necessidade de solicitação de qualquer tipo. Tal estratégia, embora não tenha sido executada à exaustão do modelo, demonstrou ser possível a obtenção, transformação e utilização de dados gerados por diferentes órgãos, com distintas finalidades e que, pelo mapeamento realizado para com o modelo de referência, passaram a integrar um mesmo contexto relacionado à inovação regional.

A busca pelos dados para o protótipo demonstrou que, embora os órgãos estejam publicando seus dados abertos, há pouca preocupação em relação à facilidade de processamento destes dados, uma vez que os formatos utilizados dificultam a identificação automatizada do significado dos dados.

Quanto aos dados, verificou-se que a opção pela utilização de uma estrutura de cubos multidimensionais, além de propiciar escalabilidade na publicação, em função do uso de ferramentas automatizadas para a transformação, garante um maior poder de análise, em função das operações de análise pré-definidas que permitem diferentes formas de “fatiamento” dos cubos, bem como a visualização dos dados sob diferentes níveis de agregação.

Embora a busca por dados com granularidade de municípios do estado de Santa Catarina tenha se mostrado uma dificuldade, ratificando os relatos de falta de disponibilidade de dados aplicáveis em nível regional, cabe salientar que a própria legislação pode ser uma aliada neste sentido. A Lei brasileira de Acesso à Informação, de 2011, regula e estabelece critérios para que o cidadão solicite dados e informações detidos pelo governo. Esta é, certamente, uma característica que torna os dados abertos tão interessantes, pois, além de estarem cientes da obrigação de disponibilizar dados em formato aberto, tais órgãos têm a oportunidade de saber quais os dados são demandados pela sociedade, podendo orientar suas ações de transparência, e tornando-se fontes para a obtenção de dados para diversas finalidades, inclusive a gestão da inovação regional.

A proposta de um modelo de referência fornece um ponto de partida para aqueles pesquisadores que desejem estabelecer modelos específicos de indicadores para a mensuração da inovação regional. Sua representação na forma de dados ligados traz um diferencial no que diz respeito à sua publicação na Web, de tal forma que sua estrutura e descrição estão disponíveis juntamente com os dados.

A instanciação de modelos específicos encontrados na literatura possibilita, ainda, que estes dados sirvam como fonte de referência para a definição de novos modelos específicos, de maneira que o conjunto de indicadores já instanciados serve de sugestão para reutilização ou aproximação para composição de novos modelos específicos. Como diferencial, tais referenciais são acessíveis também programaticamente.

A conclusão desta tese, obviamente, não esgota o assunto. O trabalho, na verdade, abre possibilidades de pesquisa que venham a explorar as características dos dados ligados e sua aplicação no campo

da inovação regional. A seguir, são apresentadas algumas sugestões para estudos futuros.

7.1 SUGESTÕES PARA TRABALHOS FUTUROS

Esta seção apresenta algumas das oportunidades que se vislumbram a partir da conclusão desta tese. Além da publicação de artigos científicos descrevendo a concepção do modelo de referência, bem como a prova de conceito desenvolvida.

Sugere-se, inicialmente, a complementação do protótipo, pela aquisição de dados de todos os municípios de Santa Catarina, a fim de englobar todas as suas seis mesorregiões. Além disso, pode-se explorar um maior número de tipos de visualização dos dados apresentados, a fim de aumentar seu valor como ferramenta de análise para o auxílio à decisão. Diferentes representações gráficas podem ser acrescentadas, bem como pode ser explorado o uso de produção textual automatizada para confecção de relatórios a partir dos dados exibidos. O uso da estrutura de cubos multidimensionais permite, ainda, a exploração da dimensão tempo, já presente nos dados, para a análise da evolução temporal dos indicadores.

Outra oportunidade de pesquisa a ser explorada diz respeito à definição de modelos de referência suportados por dados ligados aplicados a outros domínios de aplicação, além da inovação. Sugere-se, a título de exemplo, que a uma abordagem semelhante àquela aqui apresentada seja aplicada ao domínio da avaliação do ensino superior.

De forma semelhante ao que acontece com a inovação regional, a avaliação do ensino superior segue modelos compostos, cujos dados são oriundos das informações prestadas pelas Instituições de Ensino Superior. Uma vez que as instituições devem fornecer tais dados a cada processo de avaliação, é possível que se imagine uma interface para publicação de dados ligados em cada instituição, e uma ferramenta que se utilize desta tecnologia de interligação para realizar as consultas necessárias, de forma centralizada.

Outra oportunidade de pesquisa diz respeito ao fluxo de trabalho de publicação. O processo de publicação dos dados ligados é uma atividade complicada, demorada e que demanda muitos recursos. Trata-se, da obtenção dos dados, anotações e transformações necessárias para que um conjunto de dados passe de seu formato inicial, a exemplo de uma planilha em formato CSV, até tornar-se um conjunto de dados ligados. Sugere-se, portanto, a investigação de um sistema de Gestão de Fluxo de Trabalho de Dados Ligados, que propicia que os mapeamentos

e transformações dos dados sejam feitos de forma planejada, sistemática e seguindo um conjunto de boas práticas.

Além da necessária transformação de dados estáticos, há que se explorar as alternativas existentes para que os órgãos produtores de dados abertos passem a disponibilizar, também, dados ligados abertos a partir dos dados manipulados em seus sistemas de informação. Uma vez que se tenha um modelo de referência suportado por dados ligados, é possível adotá-lo como um vocabulário padrão, de tal forma que os dados já sejam publicados com anotações semânticas. Para isto, sugere-se a abordagem de “triplificação”, pela qual os dados armazenados em sistemas de bancos de dados passam por um mapeamento pré-definido e são disponibilizados na forma de dados ligados, podendo este processo de transformação ocorrer *online*.

A possibilidade da adoção da “triplificação” para que os órgãos que publicam dados abertos passem também a publicar dados ligados requer, entretanto, a utilização de tecnologias semânticas, bem como da infraestrutura de suporte aos dados ligados. Assim, uma última sugestão de trabalho futuro diz respeito ao desenvolvimento de ferramentas que automatizem as configurações e mapeamento semântico dos dados, de modo a facilitar a estes órgãos sua experiência com tais tecnologias.

As sugestões de pesquisa apresentadas possibilitam que, além do desenvolvimento de produtos para casos específicos, sejam criadas as condições para o aprendizado e difusão das tecnologias semânticas. Assim, o envolvimento de acadêmicos e de profissionais de órgãos públicos ou de empresas privadas, que tenham algum tipo de participação na produção, publicação ou consumo de dados, representará um importante passo na direção de difundir a publicação dos dados ligados.

Para que se viabilize tal difusão, é importante que sejam adotadas iniciativas de capacitação, no sentido de estimular o funcionamento do Sistema Regional de Inovação no qual estes indivíduos se inserem. Esta é uma oportunidade para que as Universidades tomem frente neste processo de formação, promovendo a interação com os demais atores e potencializando suas próprias atividades de pesquisa e desenvolvimento.

REFERÊNCIAS

- ACS, Zoltan. **Innovation and the Growth of Cities**. Cheltenham, UK: Edwar Elgar, 2002.
- ADAMS, Richard; BESSANT, John; PHELPS, Robert. Innovation management measurement: A review. **International Journal of Management Reviews**, v. 8, n. 1, p. 27, 2006.
- AFFORTUNATO, F. *et al.* Assessing local knowledge dynamics: Regional knowledge economy indicators. **World Academy of Science, Engineering and Technology**, v. 67, p. 176-181, 2010.
- ALCAIDE-MARZAL, Jorge; TORTAJADA-ESPARZA, Enrique. Innovation assessment in traditional industries. A proposal of aesthetic innovation indicators. **Scientometrics**, v. 72, n. 1, p. 33-57, // 2007.
- ALLEMANG, Dean. Semantic Web and The Linked Data Enterprise. In: WOOD, David (Org.). **Linking Enterprise Data**. Fredericksburg, Virginia, USA: Springer, 2011. cap. Part I Why Link Enterprise Data? , p.
- ALMEIDA, Mauricio B.; BAX, Marcello P. Uma visão geral sobre ontologias: pesquisa sobre definições, tipo e aplicações, métodos de avaliação e construção. **Ciência da Informação**, v. 32, n. 3, p. 14, set/dez 2003 2003.
- ARENAS, Marcelo *et al.* **A Direct Mapping of Relational Data to RDF. W3C Recommendation** 2012.
- ARUNDEL, A.; KABLA, I. What percentage of innovations are patented? Empirical estimates for European firms. **Research Policy**, v. 27, n. 2, p. 15, 1998.
- BALATSAS, Aggelos; KYRATISIS, Panagiotis; MAVROEIDIS, Vasileios. Review of modern regional innovation measurement tools and proposal of a business excellence barometer-implementation and conclusions for the Region of Western Greece. In: ISADS 2009: 2009 INTERNATIONAL SYMPOSIUM ON AUTONOMOUS DECENTRALIZED SYSTEMS, PROCEEDINGS, 2009. 2009. p. 9-15.
- BAREGHEH, A.; ROWLEY, J.; SAMBROOK, S. Towards a multidisciplinary definition of innovation. **Management Decision**, v. 47, n. 8, p. 17, 2009.
- BARKLEY, D. L.; HENRY, M. S.; NAIR, S. Regional innovation systems: Implications for nonmetropolitan areas and workers in the south. **Growth and Change**, v. 37, n. 2, p. 278-306, Jun 2006.
- BEI, Wu; ZENGYUAN, Wu. The theory and empirical study of regional technological innovation capability. In: 2006 IEEE INTERNATIONAL ENGINEERING MANAGEMENT CONFERENCE, 2006. 2006. p. 453-457.

BERNERS-LEE, Tim. **Putting the Web back in Semantic Web.** International Semantic Web Conference 2005 2005.

_____. **Linked Data: Design Issues** 2006.

_____. **Q&A with Tim Berners-Lee.** Bloomberg Business 2007.

BERNERS-LEE, Tim; FIELDING, Roy T.; MASINTER, Larry. **Uniform Resource Identifier (URI): Generic Syntax.** 2005. Disponível em: <http://tools.ietf.org/html/rfc3986>. Acesso em: 06 Dez 2013.

BERNERS-LEE, Tim; HENDLER, L.; LASSILA, O. **The Semantic WEB: A new form of WEB content that is meaningful to computers will unleash a revolution of new possibilities.** Scientific American 2001.

BERTOT, John Carlo; CHOI, Heeyoon. **Big Data and e-Government: Issues, Policies, and Recommendations.** 14th Annual International Conference on Digital Government Research. Quebec, Canada: ACM 2013.

BESSANT, John; TIDD, Joe. **Inovação e empreendedorismo.** Porto Alegre: Bookman, 2009.

BIZER, Chris. The emerging web of linked data. **IEEE Intelligent Systems**, v. 24, n. 5, p. 87-92, 2009.

BIZER, Chris; CYGANIAK, Richard; HEATH, Tom. **How to publish Linked Data on the Web.** 7th International Semantic Web Conference (ISWC2008) 2008.

BIZER, Chris; HEATH, Tom; BERNERS-LEE, Tim. Linked data - The story so far. **International Journal on Semantic Web and Information Systems**, v. 5, n. 3, p. 22, 2009.

BORST, Willem Nico. **CONSTRUCTION OF ENGINEERING ONTOLOGIES FOR KNOWLEDGE SHARING AND REUSE.** 1997. 243 f. (Doctorate) - CIP-DATA KONINKLIJKE BIBLIOTHEEK, Enschede, The Netherlands, 1997.

BOSAK, Jon; BRAY, Tim. XML and the Second-Generation Web. **Scientific American**, 1999.

BRADLEY, Fiona. Discovering Linked Data. **Library Journal**, v. 134, n. 7, p. 3, 2009.

BRASIL. **Lei Nº 12.527, de 18 de Novembro de 2011. Regula o acesso a informações previsto no inciso XXXIII do art. 5º, no inciso II do § 3º do art. 37 e no § 2º do art. 216 da Constituição Federal; altera a Lei no 8.112, de 11 de**

dezembro de 1990; revoga a Lei no 11.111, de 5 de maio de 2005, e dispositivos da Lei no 8.159, de 8 de janeiro de 1991; e dá outras providências. 2011.

BRAY, Tim *et al.* **Extensible Marku Language (XML) 1.0 (Fifth Edition)** 2008.

BRICKLEY, Dan; GUHA, R.V. **RDF Schema 1.1. W3C Recommendation** 2014.

BRITO, Ronnie Fagundes de. **Modelo de Referência para Desenvolvimento de Artefatos de Apoio ao Acesso dos Surdos ao Audiovisual.** 2012. f. Tese de Doutorado - Departamento de Engenharia do Conhecimento, Universidade Federal de Santa Catarina, Florianópolis, 2012.

BUESA, Mikel; HEIJS, Joost; BAUMERT, Thomas. The determinants of regional innovation in Europe: A combined factorial and regression knowledge production function approach. **Research Policy**, v. 39, n. 6, p. 722-735, Jul 2010.

CADORI, AluÍzia Aparecida. **A gestão do conhecimento aplicada ao processo de transferência de resultados de pesquisa de instituições federais de ciência e tecnologia para o setor produtivo: processo mediado pelo núcleo de inovação tecnológica.** 2013. f. Tese de Doutorado (Doutorado) - Departamento de Engenharia do Conhecimento, Universidade Federal de Santa Catarina, Florianópolis, 2013.

CALDERERO-GUTIÉRREZ, A.; FERNÁNDEZ-MACHO, J.; KUITTINEN, H. European regional innovation. A methodological and updated alternative for RIS. **Dyna (Spain)**, v. 84, n. 6, p. 501-516, 2009.

CARAYANNIS, E. G.; CAMPBELL, D. F. J. Open Innovation Diplomacy and a 21st Century Fractal Research, Education and Innovation (FREIE) Ecosystem: Building on the Quadruple and Quintuple Helix Innovation Concepts and the "Mode 3" Knowledge Production System. **Journal of the Knowledge Economy**, v. 2, n. 3, p. 327-372, 2011.

CARAYANNIS, Elias G. *et al.* High technology spin-offs from government R&D laboratories and research institutes. **Technovation**, v. 18, n. 1, p. 10, 1998.

CASTRO, Luciano Patrício Souza de. **Modelo de referência para comunicação da marca em interfaces gráfico-digitais.** 2010. f. Tese de Doutorado - Departamento de Engenharia do Conhecimento, Universidade Federal de Santa Catarina, Florianópolis, 2010.

CGEE. **Doutores 2010: Estudos da demografi a da base técnico-científí ca brasileira.** Brasília, DF: Centro de Gestão e Estudos Estratégicos, 2010.

CHANG, Yuan-Chieh *et al.* Measuring Regional Innovation and Entrepreneurship Capabilities: The Case of Taiwan Science Parks. **Journal of the Knowledge Economy**, v. 3, n. 2, p. 90-108, 2012.

CHEN, J.; JU, X. F. Establishing a measurement framework indicating regional technology innovation ability. In: 21TH ANNUAL INTERNATIONAL CONFERENCE ON MANAGEMENT SCIENCE AND ENGINEERING, ICMSE 2014, 2014. **International Conference on Management Science and Engineering - Annual Conference Proceedings**. IEEE Computer Society, 2014. p. 1570-1575.

CHEN, Kaihua; GUAN, Jiancheng. Mapping the functionality of China's regional innovation systems: A structural approach. **China Economic Review**, v. 22, n. 1, p. 11-27, Mar 2011.

CHEN, Wei-li; WANG, Cai-jie. Study of the construction of assessment system on regional innovation capacity in knowledge management. In: 2012 INTERNATIONAL WORKSHOP ON INFORMATION AND ELECTRONICS ENGINEERING, IWIEE 2012, 2012. Harbin. **Procedia Engineering**. Harbin, 2012. p. 1830-1834.

CHESBROUGH, H. W. **Open Innovation: The New Imperative for Creating and Profiting from Technology** 2003.

_____. Why Companies Should Have Open Business Models. **Mit Sloan Management Review**, v. 48, n. 2, Winter, 2007. 2007.

CHESBROUGH, H. W. ; VANHAVERBEKE, W; WEST, J. **2006. Open innovation: researching a new paradigm**. . Oxford, U.K.: Oxford Univ. Press., 2006.

COAD, Alex; RAO, Rekha. Innovation and firm growth in high-tech sectors: A quantile regression approach. **Research Policy**, v. 37, n. 4, p. 633-648, 2008.

CODD, E.F.; CODD, S.B.; SALLEY, C.T. **Providing OLAP to User-Analysts: An IT Mandate**. 1993.

CONRADIE, Peter; CHOENNI, Sunil. On the barriers for local government releasing open data. **Government Information Quarterly**, 2014.

COOKE, PHILIP. Regional Innovation Systems: Competitive Regulation in the New Europe. **Geoforum**, v. 23, n. 3, p. 18, 1992.

CORNELL UNIVERSITY; INSEAD; WIPO. **The Global Innovation Index 2013: The Local Dynamics of Innovation**. Geneva, Switzerland: 2013. ISBN 978-2-9522210-3-0.

CREATIVE COMMONS. **Creative Commons Website**. 2001. Disponível em: <<http://creativecommons.org/>>. Acesso em: 03 de Março de 2015.

CROSSAN, Mary M.; APAYDIN, Marina. A Multi-Dimensional Framework of Organizational Innovation: A Systematic Review of the Literature. **Journal of Management Studies**, v. 47, n. 6, p. 1154-1191, 2010.

CUPANI, Alberto. La peculiaridad del conocimiento tecnológico. **ScientiaeStudia, São Paulo**, v. v. 4, n. n. 3, p. p. 353-71, 2006.

CYGANIAK, R.; MAALI, F.; PERISTERAS, V. **Self-service linked government data with dcat and gridworks**. 6th International Conference on Semantic Systems, I-SEMANTICS '10; ACM: 3 p. 2010.

CYGANIAK, Richard; REYNOLDS, D.; TENNISON, J. **The rdf data cube vocabulary**. W3C Recommendation 2014.

DAL PIZZOL, Leandro *et al.* ANÁLISE BIBLIOMÉTRICA DA PRODUÇÃO CIENTÍFICA SOBRE LINKED DATA. **Informação e Informação**, v. 20, n. 3, p. 36, Set/Dez 2015 2015.

DALL'AGNOL, Roberto Mauro. **A Gestão Da Inovação Nas Universidades: O Capital Social E A Institucionalização De Unidades De Inovação No Ambiente Acadêmico**. 2010. f. Tese (Doutorado) - Departamento de Engenharia do Conhecimento, Universidade Federal de Santa Catarina, Florianópolis, 2010.

DIAS, Maria Madalena; PACHECO, Roberto Carlos dos Santos. Uma visão geral de metodologias para desenvolvimento de sistemas baseados em conhecimento. **Datagramazero - Revista de Ciências da Informação**, v. 10, n. 5, 2009.

DOCHERTY, Michael. Primer of Open Innovation: Principles & Practice. **PDMA Visions**, v. XXX, n. 2, p. 5, 2006.

DOLOREAU, David; PARTO, Saeed. **Regional Innovation Systems: A Critical Review**. XLème Colloque de L'ASRDLE. Bruxelles 2004.

DOROW, Patrícia. **O Processo de Geração de Ideias para Inovação: Estudo de Caso em uma Empresa Náutica**. 2013. f. (Mestrado) - Departamento de Engenharia do Conhecimento, UFSC, Florianópolis, 2013.

DOU, W. Z. *et al.* Urban competitive pattern and its changes in China. **Chinese Geographical Science**, v. 10, n. 2, p. 105-112, 2000.

DU PREEZ, N.; LOUW, L.; ESSMANN, H. An Innovation Process Model for Improving Innovation Capability. **Journal of High Technology Management Research**, p. 24, 2006.

EAVES, David. **The Three Laws of Open Data**. 2009. Disponível em: <<http://eaves.ca/2009/09/30/three-law-of-open-government-data/>>. Acesso em: 10 de Março de 2015.

EDQUIST, Charles. **The Systems of Innovation Approach and Innovation Policy: An account of the state of the art.** DRUID Conference. Aalborg 2001.

_____. **Systems of Innovation: Perspectives and Challenges.** 2009. 24 p.

ETZKOWITZ, Henry; LEYDESDORFF, Loet. The Triple Helix---University-Industry-Government Relations: A Laboratory for Knowledge-Based Economic Development. **EASST Review**, v. 14, p. 6, 1995.

_____. The dynamics of innovation: from National Systems and ‘‘Mode2’’ to a Triple Helix of university–industry–government relations. **Research Policy**, v. 29, p. 15, 2000.

EUROPEAN UNION. **Innovation Union Scoreboard 2014** 2014a.

_____. Regional Innovation Scoreboard 2014. 2014b.

EUROPEAN UNION, European Union. **European Innovation Scoreboard (EIS) 2009.** Belgium: 2010. ISBN 978-92-79-14222-2.

_____. **Regional Innovation Scoreboard 2012.** Belgium: European Union, 2012.

EUROSTAT. **Community Innovation Survey.** 2010. Disponível em: <<http://ec.europa.eu/eurostat/web/microdata/community-innovation-survey>>. Acesso em: 15 de Dezembro de 2015.

EVANGELISTA, R. *et al.* Measuring the regional dimension of innovation. Lessons from the Italian Innovation Survey. **Technovation**, v. 21, n. 11, p. 733-745, 2001.

FACHIN, Gleisy Regina Bories. **Ontologia de referência para periódico científico digital.** 2011. f. Tese de Doutorado - Departamento de Engenharia do Conhecimento, Universidade Federal de Santa Catarina, Florianópolis, 2011.

FAN, Xia; ZHU, Guilong. Regional Innovation Network Construction and Development On the Basis of Date Envelopment Analysis: Evidence from Guangdong Province, PRC. In: PROCEEDINGS OF THE INTERNATIONAL CONFERENCE ON INFORMATION MANAGEMENT, INNOVATION MANAGEMENT AND INDUSTRIAL ENGINEERING, VOL 1, 2008. 2008. p. 321-324.

FAPESC. Política Catarinense de Ciência, Tecnologia e Inovação. 2010.

FAPESP. **Indicadores de Ciência, Tecnologia e Inovação em São Paulo.** 2011.

FERNÁNDEZ-LÓPEZ, M.; GÓMEZ-PÉREZ, A. Overview and analysis of methodologies for building ontologies. **The Knowledge Engineering Review**, v. 17, n. n. 2, p. 28, 2002.

FERNÁNDEZ, Mariano; GÓMEZ-PÉREZ, A.; JURISTO, Natalia. **METHONTOLOGY: From Ontological Art Towards Ontological Engineering. AAAI Technical Report**, 1997.

FERREIRA, H.; CASSIOLATO, M.; GONZALEZ, R. **Uma experiência de desenvolvimento metodológico para avaliação de programas: o modelo lógico do programa segundo tempo**. 1369, Texto para discussão. Brasília: IPEA 2009.

FIATES, José Eduardo Azevedo. **Influência dos Ecossistemas de Empreendedorismo Inovador na Indústria de Venture Capital: Estratégias de Apoio às Empresas Inovadoras**. 2014. f. Tese de Doutorado (Doutorado) - Departamento de Engenharia do Conhecimento, Universidade Federal de Santa Catarina, Florianópolis, 2014.

FREEMAN, C. . **Technology and Economic Performance: Lessons from Japan**. London: Pinter, 1987.

GIL, Antônio Carlos. **Métodos e Técnicas de Pesquisa Social**. 6. Sao Paulo: Editora Atlas, 2008. ISBN 978-85-224-5142-5

GIUGLIANI, Eduardo. **Modelo de Governança para Parques Científicos e Tecnológicos no Brasil**. 2011. f. (Doutorado) - Departamento de Engenharia do Conhecimento, UFSC, Florianópolis, 2011.

GLEBOVA, Irina; KOTENKOVA, Svetlana. Evaluation of Regional Innovation Potential in Russia. **International Conference on Applied Economics (Icoae 2014)**, v. 14, p. 230-235, 2014 2014.

GOGODZE, Joseph. Measuring innovative capacities of the Georgia regions. **Journal of Technology Management and Innovation**, v. 8, n. 3, p. 116-126, 2013.

GROUP, W3C SPARQL Working. **SPARQL 1.1 Overview. W3C Recommendation** 2013.

GRUBER, Thomas R. Toward Principles for the Design of Ontologies Used for Knowledge Sharing. **International Journal Human-Computer Studies**, v. 43, p. 23, 1993.

GRUPP, Hariolf; SCHUBERT, Torben. Review and new evidence on composite innovation indicators for evaluating national performance. **Research Policy**, v. 39, n. 1, p. 67-78, // 2010.

GUARINO, Nicola. **Formal Ontology and Information Systems. Proceedings of FOIS'98**. Trento, Italy 1998.

GUBIANI, Juçara Salete. **Modelo para Diagnosticar a Influência do capital Intelectual no Potencial de Inovação nas Universidades**. 2011. f. (Doutorado) - Departamento de Engenharia do Conhecimento, UFSC, Florianópolis, 2011.

GUEMBAROVSKI, Ricardo Hauss;. **Um Modelo de Referência Orientado ao Conhecimento para o Processo de Planejamento de Sistemas de Distribuição de Média Tensão**. 2014. f. Tese de Doutorado - Departamento de Engenharia do Conhecimento, Universidade Federal de Santa Catarina, Florianópolis, 2014.

GUO, W.; REN, L. Regional innovation evaluation of northwest based on Grey relational analysis. In: 2011 6TH IEEE JOINT INTERNATIONAL INFORMATION TECHNOLOGY AND ARTIFICIAL INTELLIGENCE CONFERENCE, ITAIC 2011, 2011. Chongqing. **Proceedings - 2011 6th IEEE Joint International Information Technology and Artificial Intelligence Conference, ITAIC 2011**. Chongqing, 2011. p. 292-294.

GUO, Weidong. Research on innovation ecosystem in IT industry. 2009. Guilin. Guilin, 2009. p. 6004-6007.

GUSMÃO, R.; RAMOS, M. Y. Concentração regional da C&T no Brasil: Perfil da liderança paulista no cenário nacional. **Sao Paulo em Perspectiva**, v. 20, n. 3, p. 120-141, 2006.

HÁJKOVÁ, V.; HÁJEK, P. Typology of regional innovation systems in Europe - A neural network approach. **International Journal of Mathematical Models and Methods in Applied Sciences**, v. 5, n. 3, p. 463-471, 2011.

HANSEN, S. O.; WAKONEN, J. . Innovation, a winning solution? **International Journal of Technology Management**, v. 13, p. 14, 1997.

HEATH, Tom. Linked Data - Welcome to the Data Network. **IEEE Internet Computing**, v. 15, n. 6, p. 4, 2011.

HEATH, Tom; BIZER, Christian. **Linked Data: Evolving the Web into a Global Data Space** Morgan & Claypool: 136 p. 2011.

HU, B.; SVENSSON, G. **A Case Study of Linked Enterprise Data**. ISCW2010. Shanghai, China. 2010.

HYLAND, Bernadette; WOOD, David. The Joy of Data - A Cookbook for Publishing Linked Government Data on the Web. In: WOOD, David (Org.). **Linking Government Data**. Fredericksburg, VA, USA: Springer, 2011. p.

IBGE. **Divisão do Brasil em mesorregiões e microrregiões geográficas**. Rio de Janeiro: Departamento de Geografia IBGE, 1990.

_____. **Síntese dos indicadores sociais: uma análise das condições de vida da população brasileira.** Rio de Janeiro: IBGE, 2008.

IBGE, Coordenação de Indústria. **Pesquisa de inovação tecnológica: 2008.** Rio de Janeiro: IBGE, 2010. ISBN 978-85-240-4149-5.

_____. **Pesquisa de Inovação 2011.** Rio de Janeiro: IBGE, 2013.

JACOBS, Ian; WALSH, Norman. **Architecture of the World Wide Web, Volume One. W3C Recommendation** 2004.

JALONEN, H. The knowledge-innovation nexus in the welfare service ecosystem. **International Journal of Knowledge-Based Development**, v. 4, n. 1, p. 34-49, 2013.

JENTZSCH, Anja; CYGANIAK, Richard; BIZER, Christian. **State of the LOD Cloud.** 2011.

KÄMPGEN, B.; O'RAIN, S.; HARTH, A. **Interacting with Statistical Linked Data via OLAP Operations.** Proceedings of the International Workshop on Interacting with Linked Data: 36-49. p. 2012.

KINCELER, Lucia Morais. **UM FRAMEWORK BASEADO EM ONTOLOGIA DE APOIO À GESTÃO ESTRATÉGICA DA INOVAÇÃO EM ORGANIZAÇÕES DE P&D+i.** 2013. f. (Doutorado) - Departamento de Engenharia do Conhecimento, UFSC, Florianópolis, 2013.

KLEINKNECHT, ALFRED; VAN MONTFORT, KEES ; BROUWER, ERIK. THE NON-TRIVIAL CHOICE BETWEEN INNOVATION INDICATORS. **Econ. Innov. New Techn.**, v. 11, n. 2, p. 13, 2002.

KONSTANTINOU, N. *et al.* Technically approaching the semantic web bottleneck. **International Journal of Web Engineering and Technology**, v. 6, n. 1, p. 83-111, 2010.

KOSLOSKY, Marco Neiva; SPERONI, Rafael de Moura; GAUTHIER, Fernando A. Ostuni. Ecosistemas de inovação – Uma revisão sistemática da literatura. **Espacios (Caracas)**, v. 36, n. 3, 2015.

KUHN, M. E. Driving growth through open innovation. **Food Technol**, v. 62, p. 7, 2008.

LABIAK JUNIOR, Silvestre. **MÉTODO DE ANÁLISE DOS FLUXOS DE CONHECIMENTO EM SISTEMAS REGIONAIS DE INOVAÇÃO.** 2012. 264 f. (Doutorado) - Departamento de Engenharia do Conhecimento, UFSC, Florianópolis, 2012.

LEI, T.; CHEN, X. D. Stage differences of continuous innovation abilities in China's biotechnology industry. In: INTERNATIONAL CONFERENCE ON MANAGEMENT AND SERVICE SCIENCE, MASS 2009, 2009. Wuhan. **Proceedings - International Conference on Management and Service Science, MASS 2009**. Wuhan, 2009. p.

LI, Bai-zhou; ZHU, Xiao-xia. The regional research on the innovative driving impetus & innovative performance of RIS. In: PROCEEDINGS OF 2007 INTERNATIONAL CONFERENCE ON MANAGEMENT SCIENCE & ENGINEERING, 2007. 2007. p. 2158-2163.

LI, C. B.; LEI, M. Evaluation and empirical study on the capacity of regional technical innovation. In: 1ST INTERNATIONAL CONFERENCE ON INFORMATION SCIENCE AND ENGINEERING, ICISE2009, 2009. Nanjing. **2009 1st International Conference on Information Science and Engineering, ICISE 2009**. Nanjing, 2009. p. 3777-3780.

LI, Xibao. China's regional innovation capacity in transition: An empirical approach. **Research Policy**, v. 38, n. 2, p. 338-357, Mar 2009.

LI, Z. Research on evaluation index system of regional innovation. In: 2011 INTERNATIONAL CONFERENCE ON MANAGEMENT SCIENCE AND INDUSTRIAL ENGINEERING, MSIE 2011, 2011. Harbin. **2011 International Conference on Management Science and Industrial Engineering, MSIE 2011**. Harbin, 2011. p. 1269-1273.

LI, Zhao. Research on the construction of innovative province, cities and regional innovations system - A case of Heilongjiang province. In: 2010 INTERNATIONAL CONFERENCE ON LOGISTICS SYSTEMS AND INTELLIGENT MANAGEMENT, ICLSIM 2010, 2010. Harbin. **2010 International Conference on Logistics Systems and Intelligent Management, ICLSIM 2010**. Harbin, 2010. p. 1980-1985.

LIU, H.; ZHANG, X.; ZHANG, F. Regional innovation system efficiency evaluation based on the triple helix model. In: 2011 INTERNATIONAL CONFERENCE ON BUSINESS COMPUTING AND GLOBAL INFORMATIZATION, BCGIN 2011, 2011. Shanghai. **Proceedings of the 2011 International Conference on Business Computing and Global Informatization, BCGIn 2011**. Shanghai, 2011. p. 154-157.

LIU, Shu-guang; CHEN, Cai. REGIONAL INNOVATION SYSTEM: THEORETICAL APPROACH AND EMPIRICAL STUDY OF CHINA. **Chinese Geographical Science**, v. 13, n. 3, p. 193-198, Sep 2003.

LORENZO, M. Applying open innovation principles for triggering and accelerating innovations: The experience of Ericsson Spain, 2004 through 2007. In: LEON, G., *et al* (Org.), v.287, 2008. p. 553-560. ISBN 15715736 (ISSN); 9780387875026 (ISBN).

LU, Yanqiu; ZHANG, Gongyi; PENG, Peng. The Research on Efficiency Mode of Scientific and Technological Innovation System of Three Provinces in Northeast China. In: 2009 INTERNATIONAL CONFERENCE ON INFORMATION MANAGEMENT, INNOVATION MANAGEMENT AND INDUSTRIAL ENGINEERING, VOL 2, PROCEEDINGS, 2009. 2009. p. 215-219.

LUNDEVALL, B.-Å. **National Innovation Systems: Towards a Theory of Innovation and Interactive Learning**. London: Pinter, 1992.

LUNDEVALL, B.-Å. *et al.* **Handbook of Innovation Systems and Developing Countries: Building Domestic Capabilities in a Global Setting**. Reino Unido: Edward Elgar Publishing Limited, 2009.

MAHROUM, Sami; ALSALEH, Yasser. Towards a functional framework for measuring national innovation efficacy. **Technovation**, // 2013.

MAKKONEN, Teemu; HAVE, Robert P. Benchmarking regional innovative performance: composite measures and direct innovation counts. **Scientometrics**, v. 94, n. 1, p. 247-262, 2013.

MALDONADO, Mauricio Uriona. **Dinâmica de Sistemas Setoriais de Inovação: Um Modelo de Simulação Aplicado no Setor Brasileiro de Software**. 2012. f. (Doutorado) - Departamento de Engenharia do Conhecimento, UFSC, Florianópolis, 2012.

MANHÃES, Maurício C. **A INOVAÇÃO EM SERVIÇOS E O PROCESSO DE CRIAÇÃO DO CONHECIMENTO: uma proposta de método para o design de serviço**. 2010. f. Dissertação (Mestrado) - Departamento de Engenharia do Conhecimento, Universidade Federal de Santa Catarina, Florianópolis, 2010.

MANOLA, Frank; MILLER, Eric. **RDF Primer**. W3C Recommendation 2004.

MANOLA, Frank; MILLER, Eric; MCBRIDE, Brian. **RDF 1.1 Primer**. W3C Recommendation 2014.

MARKOWSKA, Malgorzata; STRAHL, Danuta. EVALUATION OF THE EUROPEAN UNION REGIONS CONVERGENCE REGARDING INNOVATION. **Argumenta Oeconomica**, v. 28, n. 1, p. 41-67, 2012 2012.

MCTI. **Estratégia Nacional de Ciência, Tecnologia e Inovação 2012 - 2015**. Brasília, DF: 2012.

MEYER, Bertolt; SUGIYAMA, Kozo. The concept of knowledge in KM: a dimensional model. **Journal of Knowledge Management**, v. 11, n. 1, p. 17-35, 2007.

MIGUEZ, Viviane Brandão. **Uma Abordagem de Geração de Ideias para o Processo de Inovação**. 2012. f. (Mestrado) - Departamento de Engenharia do Conhecimento, UFSC, Florianópolis, 2012.

MILOŠEVIĆ, U. *et al.* Publishing Statistical Data As Linked Open Data. In: 2ND INTERNATIONAL CONFERENCE ON INFORMATION SOCIETY 2012. Serbia. Serbia: Technology Information Society of the Republic of Serbia, 2012. p.

MINAYO, Maria Cecília de Souza. **O Desafio do Conhecimento - Pesquisa Qualitativa em Saúde**. Rio de Janeiro: HUCITEC, 2007.

MIZOGUCHI, Riichio; VANWELKENHUYSEN, Johan; MITSURU, Ikeda. **Task Ontology for Reuse of Problem Solving Knowledge**. Knowledge Building & Knowledge Sharing 1995 (KB&KS'95) (2nd International Conference on Very Large-Scale Knowledge Bases). Enschede, The Netherlands: 14 p. 1995.

MOREIRA, Alexandra; ALVARENGA, Lúdia; OLIVEIRA, Alcione P. O nível do conhecimento e os instrumentos de representação: tesouros e ontologias. **Datagramazero - Revista de Ciências da Informação**, v. 5, n. 6, 2004.

MOTA, Jéssica Romeiro. **A Proteção do Conhecimento Resultante da Parceria de Pesquisa, Desenvolvimento & Inovação Originado da Relação Universidade e Empresa**. 2011. f. Dissertação de Mestrado (Mestrado) - Departamento de Engenharia do Conhecimento, Universidade Federal de Santa Catarina, Florianópolis, 2011.

MPOG. **Instrução Normativa N°4**. MINISTÉRIO DO PLANEJAMENTO, Orçamento e Gestão, Brasil. Brasília, DF 2012.

NAMBISAN, S.; BARON, R. A. Entrepreneurship in innovation ecosystems: Entrepreneurs' self-regulatory processes and their implications for new venture success. **Entrepreneurship: Theory and Practice**, v. 37, n. 5, p. 1071-1097, 2013.

NARDO, Michela *et al.* **Handbook on Constructing Composite Indicators**. 2005.

NATÁRIO, Maria Manuela Santos *et al.* Evaluating The Determinants Of National Innovative Capacity Among European Countries. **Global Journal of Management and Business Research**, v. 11, n. 1, p. 13, 2011.

NATIONAL INNOVATION COUNCIL. **Introduction**. 2010. Disponível em: <http://www.innovationcouncil.gov.in/index.php?option=com_content&view=article&id=26&Itemid=5>. Acesso em: 20 Dec 2013.

NAZÁRIO, D.C.; DANTAS, M.A.R.; TODESCO, José Leomar. Knowledge Engineering: Survey of Methodologies, Techniques and Tools. **IEEE Latin America Transactions**, v. 12, n. 8, p. 7, 2014.

NELSON, A. Measuring knowledge spillovers: What patents, licenses and publications reveal about innovation diffusion. **Research Policy**, v. 38, n. 6, p. 12, 2009.

NIKULINA, Y. *et al.* The formation of priority directions of social and economic development of the region. **World Applied Sciences Journal**, v. 22, n. 5, p. 608-615, 2013.

NING, Cai; CHUN, Huang. The Construction and Evaluation of the Regional Innovation System of Zhejiang Province. In: INNOVATIVE COMPUTING AND INFORMATION, ICCIC 2011, PT I, 2011. 2011. p. 36-45.

NONAKA, Ikujiro; TAKEUCHI, Hirotaka. **Criação de Conhecimento na Empresa - Como as Empresas Japonesas Geram a Dinâmica da Inovação**. 20ª edição. Rio de Janeiro, Brasil: 1997.

NOY, Natalia F.; MCGUINNESS, Deborah L. **Ontology Development 101: A Guide to Creating Your First Ontology**. 2001.

OECD. **Oslo Manual - Guidelines for Collecting and Interpreting Innovation Data**. Paris, France: OECD, 1992.

_____. **Oslo Manual - Guidelines for Collecting and Interpreting Innovation Data**. Paris, France: OECD, 2005.

_____. **OECD Science, Technology and Industry Outlook 2014**: OECD Publishing 2014.

_____. **Frascati Manual: Guidelines for Collecting and Reporting Data on Research and Experimental Development**. Paris: 2015.

OPEN DATA COMMONS. **Open Data Commons - Legal Tools for Open Data**. 2008. Disponível em: <<http://opendatacommons.org/>>.

OPEN DEFINITION. **Open Definition - version 1.1**. 2009. Disponível em: <<http://opendefinition.org/od/>>. Acesso em: 10 de Dezembro de 2013.

OPEN KNOWLEDGE FOUNDATION. **Open Data Handbook Documentation** 2012.

OROFINO, Maria Augusta. **Técnicas de criação do conhecimento no desenvolvimento de modelos de negócio**. 2011. f. - Departamento de Engenharia do Conhecimento, UFSC, Florianópolis, 2011.

OWL WORKING GROUP. **Web Ontology Language (OWL)** 2012.

PAGE, K.R. *et al.* **Linked Sensor Data: RESTfully serving RDF and GML.** International Workshop on Semantic Sensor Networks 2009. 522: 49–63 p. 2009.

PAPADOPOULOS, Georgios Z. *et al.* Performance evaluation methods in ad hoc and wireless sensor networks: a literature study. **IEEE Communications Magazine**, v. 54, n. 1, p. 122-128, 2016.

PATNI, Harshal; HENSON, Cory; SHETH, Amit. **Linked sensor data.** International Symposium on Collaborative Technologies and Systems (CTS): IEEE: 362-370 p. 2010.

PINTO, Hugo; GUERREIRO, João. Innovation regional planning and latent dimensions: the case of the Algarve region. **The Annals of Regional Science**, v. 44, n. 2, p. 315-329, 2010.

PONS, X. A.; MARTÍNS, J. J. G.; PARRILLI, M. D. Apprising fragmentation in regional innovation systems: A typology for Spain. **Investigaciones Regionales**, n. 28, p. 7-35, 2014.

PPEGC. **Engenharia do Conhecimento - Descrição das áreas de Concentração.** 2004. Disponível em: <<http://www.egc.ufsc.br/index.php/egc/pos-graduacao/areas-de-concentracao#ec>>. Acesso em: 23 de Outubro de 2014.

PRADA, Charles A. **Proposta de modelo para o gerenciamento de portfólio de inovação: modelagem do conhecimento na geração de ideias.** 2009. f. (Mestrado) - Departamento de Engenharia do Conhecimento, UFSC, Florianópolis, 2009.

PRUD'HOMMEAUX, Eric; SEABORNE, Andy. **SPARQL Query Language for RDF.** W3C, 2008.

QI, Yanru; LIU, Yun. Empirical Research about the Regional Innovation Capability Based on China's Patent Application Activities. In: 2014 PORTLAND INTERNATIONAL CONFERENCE ON MANAGEMENT OF ENGINEERING & TECHNOLOGY (PICMET), 2014. 2014. p. 365-369.

QUINTANE, Eric *et al.* Innovation as a knowledge-based outcome. **Journal of Knowledge Management**, v. 15, n. 6, 2011.

RATANAWARAHANA, A.; POLENSKE, K. R. Measuring the geography of innovation: A literature review. In: POLENSKE, K. (Org.). **The economic geography of innovation**: Cambridge Cambridge University Press, 2007. p.

RAUTENBERG, Sandro *et al.* **LODFlow - a Workflow Management System for Linked Data Processing.** 12th International Conference on Semantics Systems. Leipzig, Germany 2015.

RAUTENBERG, Sandro;; TODESCO, José Leomar; STEIL, Andrea Valéria Steil. UMA ONTOLOGIA PARA INSTRUMENTOS DA GESTÃO DO CONHECIMENTO E AGENTES DA ENGENHARIA DO CONHECIMENTO. **Informação e Sociedade: Estudos**, v. 21, n. 1, 2011.

REICHWALD, Ralf; BULLINGER, Angelika. Innovation and Ontologies: Structuring the Early Stages of Innovation Management. 2008.

RUSHER, Jack. **Triple Store**. Rhetorical Device 2003.

SAISANA, M.; TARANTOLA, S. **State-of-the-Art Report on Current Methodologies and Practices for Composite Indicator Development**. Ispra: European Commission–Joint Research Centre, 2002.

SALAS, P. E. R. *et al.* Publishing Statistical Data on the Web. In: 6TH INTERNATIONAL CONFERENCE ON SEMANTIC COMPUTING, 2012. 2012. p. 285–292.

SÁNCHEZ TOVAR, Yesenia; GARCÍA FERNÁNDEZ, Francisco; MENDOZA FLORES, Esteban. Determinantes de la capacidad de innovación regional en México: Una tipología de las regiones. **Región y sociedad**, v. 26, n. 61, p. 118-158, 2014-12 2014.

SANTA CATARINA. **LEI Nº 14.328 (Lei Catarinense de Inovação)** 2008.

SARTORI, Rejane. **Governança em Agentes de Fomento dos Sistemas Regionais de CT&I**. 2011. f. Tese de Doutorado (Doutorado) - Departamento de Engenharia do Conhecimento, Universidade Federal de Santa Catarina, Florianópolis, 2011.

SAUERMANN, L.; CYGANIAK, R. **Cool URIs for the Semantic Web**. W3C Interest Group Note 2008.

SCHMACHTENBERG, Max; BIZER, Christian; PAULHEIM, Heiko. **Adoption of the Linked Data Best Practices in Different Topical Domains**. International Semantic Web Conference 2014. Trentino, Italy 2014.

SCHMITT, Maurilio Tiago Brüning. **Inteligência Competitiva na Web: Um Framework Conceitual para Aquisição de Ativos de Conhecimento no Contexto do Front-End da Inovação**. 2013. f. Dissertação de Mestrado (Mestrado) - Departamento de Engenharia do Conhecimento, Universidade Federal de Santa Catarina, Florianópolis, 2013.

SCHMOCH, Ulrich; GAUCH, Stephan. Service marks as indicators for innovation in knowledge-based services. **Research Evaluation**, v. 18, n. 4, p. 323-335, // 2009.

SCHREIBER, Guus *et al.* **Knowledge Engineering and Management - The CommonKADS Methodology**. MIT Press, 1999. 471 p.

SCHUMPETER, Joseph Alouis. **Capitalismo, Socialismo e Democracia**. Rio de Janeiro: Editora Fundo de Cultura, 1961.

_____. **The Theory of Economic Development: An Inquiry into Profits, Capital, Credit, Interest, and the Business Cycle**. 16. New Jersey, USA: Transaction Publishers, 2012.

SERVANT, François-Paul. **Linking enterprise data. "Linked Data on Web" in Workshop at the 17th World Wide Web Conference**. Beijing, China 2008.

SHUE, Li-Yen; CHEN, Ching-Wen; SHIUE, Weissor. The development of an ontology-based expert system for corporate financial rating. **Expert Systems with Applications**, v. 36, n. 2, p. 2130-2142, 2009.

SILVA, Arleide Rosa da. **Análise da relação entre a gestão do conhecimento e o ambiente de inovação em uma instituição de ensino profissionalizante**. 2011. f. Tese (Doutorado) - Departamento de Engenharia do Conhecimento, Universidade Federal de Santa Catarina, Florianópolis, 2011.

SMITH, Keith. Measuring Innovation. In: JAN FAGERBERG, David C. Mowery and Richard R. Nelson (Org.). **The Oxford Handbook of Innovation**. New York, USA: Oxford University Press, 2005. cap. 6, p.

SPERONI, Rafael de Moura *et al.* Estado da Arte da Produção Científica sobre Indicadores e Índices de Inovação. **RAI: Revista de Administração e Inovação**, v. 12, n. 4, p. 27, 2015.

SPERONI, Rafael de Moura; MACEDO, Marcelo; GAUTHIER, Fernando A. Ostuni. Modelo de Referência para Indicadores de Inovação Regional. **Espacios (Caracas)**, v. 37, n. 10, 2016a.

_____. Revisão Sistemática de Literatura sobre Indicadores da Inovação Regional. **Espacios (Caracas)**, v. 37, n. 05, 2016b.

SPERONI, Rafael de Moura *et al.* **Avaliação do Ensino Superior Público no Brasil: protótipo de aplicação de Linked Data**. [Linked Open Data Brasil](#). Florianópolis, Brasil: Universidade Federal de Santa Catarina 2014.

SPERONI, Rafael de Moura *et al.* Avaliação da Produção Científica sobre Enterprise Linked Data. In: LINKED OPEN DATA BRASIL 2014, 2014. Florianópolis, SC, Brasil. Florianópolis, SC, Brasil: UFSC, 2014. p.

STUDER, Rudi; BENJAMINS, V. Richard; FENSEL, Dieter. Knowledge Engineering: Principles and Methods. **Data & Knowledge Engineering**, v. 25, n. 1, p. 38, 1998.

SUN, Dezhong. Regional Intellectual Capital Integration Performance Evaluation Based on Two-phase Model. In: 2014 SIXTH INTERNATIONAL CONFERENCE ON MEASURING TECHNOLOGY AND MECHATRONICS AUTOMATION (ICMTMA), 2014. 2014. p. 467-471.

SURE, Y.; STUDER, R. A Methodology for Ontology-based Knowledge Management. In: DAVIES, John; et al (eds). (Org.). **Towards The Semantic Web: Ontology-Driven Knowledge Management**. Chichester: John Wiley & Sons,, 2003. p.

TIAN, H.; TONG, G. The model of efficacy measurement for the innovation system in Chinese resource urban. In: 2011 ASIA-PACIFIC POWER AND ENERGY ENGINEERING CONFERENCE, APPEEC 2011, 2011. Wuhan. **Asia-Pacific Power and Energy Engineering Conference, APPEEC**. Wuhan, 2011. p.

TIDD, Joe; BESSANT, John; PAVIT, K. **Gestão da Inovação**. 3. Porto Alegre: Bookman, 2008.

TODESCO, José Leomar; *et al.* **ontoKEM: A web tool for ontologies' construction and documentation**. International Conference on Information & Knowledge Engineering, 2009; IKE 09. Las Vegas: 1-19 p. 2009.

TOSTA, Kelly Cristina Benetti Tonani. **A Universidade como Catalisadora da Inovação Tecnológica Baseada em Conhecimento**. 2012. f. Tese (Doutorado) - Departamento de Engenharia do Conhecimento, Universidade Federal de Santa Catarina, Florianópolis, 2012.

TRAITLER, Helmut; WATZKE, Heribert J.; SAGUY, I. Sam. Reinventing R&D in an Open Innovation Ecosystem. **Journal of Food Science**, v. 76, n. 2, p. R62-R68, 2011.

TROTT, Paul J. **Gestão da inovação e desenvolvimento de novos produtos**. 4. Porto Alegre: Bookman, 2012.

TU, W.; SONG, X. Research on evaluation of the competitiveness of regional Intellectual Property Rights and enhancing strategies. In: 2013 6TH INTERNATIONAL CONFERENCE ON INFORMATION MANAGEMENT, INNOVATION MANAGEMENT AND INDUSTRIAL ENGINEERING, ICIII 2013, 2013. Xi'an. **Proceedings of 2013 6th International Conference on Information Management, Innovation Management and Industrial Engineering, ICIII 2013**. Xi'an, 2013. p. 299-304.

USCHOLD, Mike; GRUNINGER, Michael. Ontologies Principles Methods and Applications. **Knowledge Engineering Review**, v. 11, n. 2, 1996.

VAN KLEEF, J. A. G.; ROOME, N. J. Developing capabilities and competence for sustainable business management as innovation: a research agenda. **Journal of Cleaner Production**, v. 15, n. 1, p. 38-51, 2007.

VARRICCHIO, Pollyana *et al.* Collaborative Networks and sustainable business: a case study in the Brazilian System of Innovation. In: DURRANI, T.; TJAKRAATMADJA, J. H. ; DHEWANTO, W. (Org.). **10th Triple Helix Conference 2012**, v.52, 2012. p. 90-99. (Procedia Social and Behavioral Sciences). ISBN 1877-0428.

VILLAZÓN-TERRAZAS, Boris *et al.* Methodological Guidelines for Publishing Government Linked Data. In: WOOD, David (Org.). **Linking Government Data**. Fredericksburg, VA, USA: Springer, 2011. p.

W3C. **RDF 1.1 Concepts and Abstract Syntax**. W3C Recommendation: W3C 2014.

WANG, Ai-feng; LI, Shao-bo. Analysis of Technology Innovation Diffusion Effect in Regional Innovation System by Unascertained Measurement. In: 2009 INTERNATIONAL CONFERENCE ON MANAGEMENT SCIENCE & ENGINEERING, 2009. 2009. p. 1873-1878.

WANG, Bei *et al.* Spatial disparity and efficiency of science and technology resources in China. **Chinese Geographical Science**, v. 22, n. 6, p. 730-741, Dec 2012.

WANG, Guo-Zhen; LIU, Xiu-Jun; TIAN, Yong-Ying. Research on evaluation of regional innovation ability based on gray related degree model. In: PROCEEDINGS OF 2006 INTERNATIONAL CONFERENCE ON MACHINE LEARNING AND CYBERNETICS, VOLS 1-7, 2006. 2006. p. 1696-1701.

WANG, J. F. Framework for university-industry cooperation innovation ecosystem: Factors and countermeasure. 2010. Wuhan. Wuhan, 2010. p. 303-306.

WANG, Wenjuan *et al.* The Evaluation and Application Research about Regional Innovation Capability Based on Rough Set and BP Neural Network. In: ICIC 2009: SECOND INTERNATIONAL CONFERENCE ON INFORMATION AND COMPUTING SCIENCE, VOL 3, PROCEEDINGS, 2009. 2009. p. 308-311.

WATERMAN, Donald A. **A Guide to Expert Systems**. Pearson Education, 1986. ISBN 9788131713310. Disponível em: <
<http://books.google.com.br/books?id=PpvxFGwIzAQC>>.

WORLD BANK. **World Development Indicators**. Washington, DC: 2014.

XIAO, X. Y.; YANG, L. W. A comparative study of regional innovation capability of China based on TOPSIS method. In: 2011 INTERNATIONAL CONFERENCE ON BUSINESS MANAGEMENT AND ELECTRONIC INFORMATION, BMEI 2011, 2011. Guangzhou. **BMEI 2011 - Proceedings 2011 International Conference on**

Business Management and Electronic Information. Guangzhou, 2011. p. 579-582.

YAM, Richard C. M. *et al.* Analysis of sources of innovation, technological innovation capabilities, and performance: An empirical study of Hong Kong manufacturing industries. **Research Policy**, v. 40, n. 3, p. 391-402, Apr 2011.

YAN, Huizhe; YAN, Pingjian; MA, Lihua. The Regional Innovation Capability Based on Comprehensive Evaluation of AHM and Unascertained Method. In: 2009 INTERNATIONAL CONFERENCE ON INFORMATION MANAGEMENT, INNOVATION MANAGEMENT AND INDUSTRIAL ENGINEERING, VOL 3, PROCEEDINGS, 2009. 2009. p. 251-254.

YIN, J.; DIAO, Z.; LI, L. Classification and evaluation for the midwest regional innovation capability based on principal component analysis and self-organizing neural network. In: 2011 4TH INTERNATIONAL SYMPOSIUM ON COMPUTATIONAL INTELLIGENCE AND DESIGN, ISCID 2011, 2011. Hangzhou. **Proceedings - 2011 4th International Symposium on Computational Intelligence and Design, ISCID 2011.** Hangzhou, 2011. p. 22-25.

ZABALA-ITURRIAGAGOITIA, JON MIKEL *et al.* What indicators do (or do not) tell us about Regional Innovation Systems. **Scientometrics**, v. 70, n. 1, p. 22, 2007.

ZABALA ITURRIAGAGOITIA, J. M.; JIMÉNEZ SAEZ, F.; GUTIÉRREZ GRACIA, A. Analysis and measurement of interactions in regional innovation systems: Need to define new network indicators. In: 15TH INTERNATIONAL CONFERENCE ON ENGINEERING DESIGN, ICED 05, 2005. Melbourne, VIC. **Proceedings ICED 05, the 15th International Conference on Engineering Design.** Melbourne, VIC, 2005. p.

ZANCANARO, Airton *et al.* **Publishing Multidimensional Statistical Linked Data.** eKNOW 2013, The Fifth International Conference on Information, Process, and Knowledge Management. Nice, France: 7 p. 2013.

ZAPILKO, B.; MATHIAK, B. Performing Statistical Methods on Linked Data. In: INTERNATIONAL CONFERENCE ON DUBLIN CORE AND METADATA APPLICATIONS, 2011. The Hague, Netherlands. The Hague, Netherlands, 2011. p. 10.

ZHANG, X.; LIU, Z.; ZHANG, F. Study on regional innovation system efficiency evaluation of Shandong province based on entropy and DEA. In: 2011 2ND INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE, MANAGEMENT SCIENCE AND ELECTRONIC COMMERCE, AIMSEC 2011, 2011. Zhengzhou. **2011 2nd International Conference on Artificial Intelligence, Management Science and Electronic Commerce, AIMSEC 2011 - Proceedings.** Zhengzhou, 2011. p. 3387-3390.

ZHENG, Y.; ZHANG, Y. Measurements and evaluation of regional innovation capacity and spatial difference. In: 2013 10TH INTERNATIONAL CONFERENCE ON SERVICE SYSTEMS AND SERVICE MANAGEMENT, ICSSSM 2013, 2013. Hong Kong. **2013 10th International Conference on Service Systems and Service Management - Proceedings of ICSSSM 2013**. Hong Kong, 2013. p. 583-588.

APÊNDICE A – Questões de Competência da Ontologia (ontoKEM)

ONTOLOGY

RegionallInnovation

DOCUMENT

Competence questions

(version 1.1)

25-05-2015

1) Como os modelos específicos podem ser vinculados a um modelo de referência?

Suggested terms: Modelo Especifico

Suggested relations: temRefBibliografica

2) Como são calculados os indicadores de inovação?

Suggested terms: Calculo

Suggested relations:

3) Como são compostos os Índices de Inovação Regional Compostos?

Suggested terms: Indicador; Indicador_Composto; Indicador_Individual; IndiceInovacaoRegional

Suggested relations: ehSubdivisaoDe; temSubdivisao

4) O que diferencia os Indicadores de Entrada dos de Saída?

Suggested terms: Indicador_Entrada; Indicador_Saida

Suggested relations:

5) O que são Indicadores de Entrada para a inovação?

Suggested terms: Indicador_Entrada; temFuncao

Suggested relations:

6) O que são Indicadores de Saída para a Inovação?

Suggested terms: Indicador_Saida

Suggested relations:

7) Quais as dimensões que compõem os indicadores de Entrada e Saída?

Suggested terms: compoeFuncao; Dimensao; Indicador_Cultura_Qualidade_Vida;

Indicador_Economia; Indicador_Efeitos_Economicos;

Indicador_Efeitos_Sociais;
 Indicador_Empreendedorismo_Ligacoes;
 Indicador_Entrada; Indicador_Financiamento;
 Indicador_Infraestrutura; Indicador_Instituicoes;
 Indicador_Investimento; Indicador_Medidas_Inovacao;
 Indicador_Recursos_Humanos;
Indicador_Saida; QualidadeVida; temNome

Suggested relations: ehDimensaoDe; temDimensao

8) Quais as fontes de dados abertos que podem ser utilizadas como subsídio para indicadores da inovação regional?

Suggested terms: Conjunto_Dados; Fonte_de_Dados;
 utiliza

Suggested relations: temFonteDados

9) Quais as unidades de medida utilizadas para cada indicador?

Suggested terms: ehUnidadeDeMedidaDe;
 temUnidadeDeMedida; Unidade_de_Medida

Suggested relations:

10) Quais os Indicadores que estão categorizados em cada dimensão?

Suggested terms: temIndicador

Suggested relations:

11) Quais os indicadores utilizados por um Modelo Especifico?

Suggested terms:

Suggested relations: ehUtilizadoPor; utilizaIndicador

12) Quais os tipos de indicadores que irão compor um indicador composto ou índice de inovação regional?

Suggested terms: Indicadores Compostos;
 Indicadores Simples; Indicador_Entrada;
 Indicador_Saida; temTipo

Suggested relations: compoe; ehCompostoPor;
 temMedida

APÊNDICE B – Termos categorizados como Classes na Ontologia (ontoKEM)

ONTOLOGY

RegionallInnovation

DOCUMENT

Definition of classes

(version 1.0)

25-05-2015

1) Conjunto_Dados

Type: Class

Description: Descreve um conjunto de dados que diz respeito a algum dos indicadores. Por exemplo, no caso de séries temporais, pode representar os dados obtidos para determinado indicador em um ano específico.

2) Indicador_Acesso_a_Cultura

Type: Class

Description: Classe que descreve os indicadores de entrada, classificados como Cultura e Qualidade de Vida, referentes ao acesso da população à cultura. Compreende, por exemplo, a quantidade de dispositivos destinados ao acontecimento de eventos culturais.

3) Indicador_Agrupamentos

Type: Class

Description: Classe que descreve os indicadores de entrada, classificados como Empreendedorismo e Ligações, referentes à existência de clusters e agrupamentos de empresas.

4) Indicador_Ambiental

Type: Class

Description: Classe que descreve os indicadores de saída, classificados como Efeitos Sociais, referentes aos resultados nas condições ambientais. Compreende indicadores que apontem os benefícios ambientais decorrentes de ações empreendidas.

5) Indicador_Ambiente_Politico

Type: Class

Description: Classe que descreve os indicadores de entrada, classificados como Instituições, referentes ao Ambiente Político.

6) Indicador_Aquisicao_de_Inovacao

Type: Class

Description: Classe que descreve os indicadores de entrada, classificados como Investimentos, referentes à aquisição direta ou indireta de inovação.

7) Indicador_Burocracia

Type: Class

Description: Classe que descreve os indicadores de entrada, classificados como Instituições, referentes à burocracia para a abertura de negócios.

8) Indicador_Capital_de_Risco

Type: Class

Description: Classe que descreve os indicadores de entrada, classificados como Financiamento, referentes à disponibilidade de capital de risco.

9) Indicador_Cooperacao

Type: Class

Description: Classe que descreve os indicadores de entrada, classificados como Empreendedorismo e Ligações, referentes às condições existentes para a cooperação entre os diferentes atores do sistema de inovação.

10) Indicador_Cultura_Empreendedora

Type: Class

Description: compreende indicadores que descrevem as características de uma região em relação ao saneamento ambiental, qualidade de vida dos residentes, cultura e participação social. Exemplos deste tipo de indicadores dizem respeito ao consumo industrial de água, consumo de energia, reservas de áreas, investimento em cultura, a existência de teatros, ou a participação em atividades voluntárias (AFFORTUNATO et al., 2010; CHEN; JU, 2014).

11) Indicador_Cultura_Qualidade_Vida

Type: Class

Description: Conjunto de indicadores de entrada que descrevem as características de uma região em relação ao saneamento ambiental, qualidade de vida dos residentes, cultura e participação social. (AFFORTUNATO et al., 2010; CHEN e JU, 2014).

12) Indicador_de_Inovacao

Type: Class

Description:

13) Indicador_Dispêndios_com_PeD

Type: Class

Description: Classe que descreve os indicadores de entrada, classificados como Investimentos, referentes aos dispêndios em P&D realizados para a inovação.

14) Indicador_Docentes

Type: Class

Description: Classe que descreve os indicadores de entrada, classificados como Recursos Humanos, referentes aos Docentes.

15) Indicador_Economia

Type: Class

Description: Conjunto de indicadores de entrada que estão associados às condições econômicas regionais. Compreendem, por exemplo, indicadores sobre o mercado de trabalho, bem como indicadores econômicos, como o Produto Interno Bruto.

16) Indicador_Educacao

Type: Class

Description: Classe que descreve o conjunto de indicadores de entrada, associados a recursos humanos, especificamente ligados à educação.

17) Indicador_Efeitos_Economicos

Type: Class

Description: Conjunto de indicadores de saída que estão associados aos resultados econômicos das atividades de inovação. Visam capturar o sucesso econômico da inovação, tais como a porcentagem de vendas relativas a produtos inovadores ou a exportação de serviços inovadores (EUROPEAN UNION, 2014).

18) Indicador_Efeitos_Sociais

Type: Class

Description: Conjunto de indicadores de saída que estão associados aos resultados sociais das atividades de inovação. Indicadores que visam capturar os benefícios às pessoas e ao planeta (OECD, 2014).

19) Indicador_Empreendedorismo_Ligacoes

Type: Class

Description: Conjunto de indicadores de entrada que representam as características da região em relação ao desenvolvimento do empreendedorismo e de arranjos e da colaboração entre os diferentes atores do sistema (CHANG et al., 2010; CORNELL UNIVERSITY et al., 2013).

20) Indicador_Empresas

Type: Class

Description: Classe que descreve os indicadores de entrada, classificados como infraestrutura, referentes à disponibilidade de empresas de alta tecnologia.

21) Indicador_Ensino_e_Pesquisa

Type: Class

Description: Classe que descreve os indicadores de entrada, classificados como Infraestrutura, referentes à disponibilidade de instituições de ensino e pesquisa.

22) Indicador_Entrada

Type: Class

Description: Aqueles indicadores que representam as pré-condições atribuídas à capacidade de inovação são caracterizados como Indicadores de Entrada.

Anotation/Observation: A maioria dos indicadores e estatísticas de inovação consideram os sistemas de inovação como sistemas de entrada-saída (input-output) (ZABALA ITURRIAGAGOITIA et al., 2005).

23) Indicador_Exportacoes

Type: Class

Description: Classe que descreve os indicadores de saída, classificados como Efeitos Econômicos, referentes às exportações.

24) Indicador_Financiamento

Type: Class

Description: Conjunto de indicadores de entrada que estão associados à disponibilidade de crédito e financiamento para as atividades de inovação, bom como do apoio governamental (CORNELL UNIVERSITY et al., 2013; EUROPEAN UNION, 2014).

25) Indicador_Fontes_de_Financiamento

Type: Class

Description: Classe que descreve os indicadores de entrada, classificados como

Financiamento, referentes a outras fontes de financiamento.

26) Indicador_Importacao_de_Tecnologia

Type: Class

Description: Classe que descreve os indicadores de entrada, classificados como investimentos para a Inovação, referentes à importação de tecnologia.

27) Indicador_Impostos

Type: Class

Description: Classe que descreve os indicadores de entrada, classificados em Instituições, referentes aos impostos.

28) Indicador_Infraestrutura

Type: Class

Description: Conjunto de indicadores de entrada que estão associados à infraestrutura existente na região. Agrupa indicadores relativos à infraestrutura que dá suporte à inovação, tais como a disponibilidade de tecnologias de informação e comunicação (TICs), a cobertura de estradas pavimentadas, instituições de ensino e de institutos de pesquisa (CHEN e GUAN, 2011; CORNELL UNIVERSITY et al., 2013).

29) Indicador_Inovacoes

Type: Class

Description: Classe que descreve os indicadores de saída, classificados como medidas de inovação, referentes aos quantitativos de inovação.

30) Indicador_Instituicoes

Type: Class

Description: Classe que descreve o conjunto de indicadores de entrada que agrupa indicadores associados às condições de um sistema de inovação de prover uma boa governança e níveis de proteção e incentivos essenciais à inovação (CORNELL UNIVERSITY et al., 2013).

31) Indicador_Investimento

Type: Class

Description: Conjunto de indicadores de entrada que estão associados aos investimentos para a inovação. Compreende os indicadores que tratam da quantidade de recursos investidos nas atividades de inovação.

32) Indicador_Legislacao

Type: Class

Description: Classe que descreve os indicadores de entrada, classificados como Instituições, referentes à legislação.

33) Indicador_Logistica

Type: Class

Description: Classe que descreve os indicadores de entrada, classificados como Infraestrutura, referentes às condições de logística.

34) Indicador_Macroeconomia

Type: Class

Description: Classe que descreve os indicadores de entrada, classificados como Economia, referentes às condições macroeconômicas.

35) Indicador_Medidas_Inovacao

Type: Class

Description: Conjunto de indicadores de saída associados aos resultados de conhecimentos gerados pelas atividades de inovação, tais como a quantidade de empresas que inovaram em produtos ou serviços, marketing ou processos, o número de patentes e de publicações científicas (CORNELL UNIVERSITY et al., 2013; EUROPEAN UNION, 2014).

36) Indicador_Meio_Ambiente

Type: Class

Description: Classe que descreve os indicadores de entrada, classificados como Cultura e Qualidade de Vida, referentes às condições do meio-ambiente, e sua exploração.

37) Indicador_Orçamento_Publico

Type: Class

Description: Classe que descreve os indicadores de entrada, classificados como Financiamento, referentes ao Orçamento Público reservado para a inovação.

38) Indicador_Padrao_de_Vida

Type: Class

Description: Classe que descreve os indicadores de saída, classificados como Efeitos Sociais, referentes às melhorias no padrão de vida da população.

39) Indicador_Patentes

Type: Class

Description: Classe que descreve os indicadores de saída, classificados como Medidas de Inovação, referentes às patentes.

40) Indicador_Pesquisadores

Type: Class

Description: Classe que descreve os indicadores de entrada, classificados como de Recursos Humanos, referentes aos Pesquisadores.

41) Indicador_Publicacoes_Cientificas

Type: Class

Description: Classe que descreve os indicadores de saída, classificados como Medidas de Inovação, referentes às publicações científicas.

42) Indicador_Receitas

Type: Class

Description: Classe que descreve os indicadores de saída, classificados como Efeitos Econômicos, referentes às receitas obtidas com as inovações.

43) Indicador_Recursos_Humanos

Type: Class

Description: Compreende indicadores relativos à educação e a disponibilidade de força de trabalho altamente qualificada e educada (EUROPEAN UNION, 2014).

44) Indicador_Renovacao_Tecnica

Type: Class

Description: Classe que descreve os indicadores de entrada, classificados como investimentos, referentes às despesas com a renovação técnica.

45) Indicador_Saida

Type: Class

Description: Indicadores de Saída são aqueles que estão associados aos resultados das atividades de inovação.

Anotation/Observation: A maioria dos indicadores e estatísticas de inovação consideram os sistemas de inovação como sistemas de entrada-saída (input-output) (ZABALA ITURRIAGAGOITIA et al., 2005).

46) Indicador_TICs

Type: Class

Description: Classe que descreve os indicadores de entrada, classificados como Infraestrutura, referentes às tecnologias de informação e comunicação.

47) Indicador_Trabalho_e_Emprego

Type: Class

Description: Classe que descreve os indicadores de entrada, classificados como Economia, referentes às condições de trabalho e emprego.

48) Indicador_Voluntariado

Type: Class

Description: Classe que descreve os indicadores de entrada, classificados como Cultura e Qualidade de Vida, referentes à participação da população em atividades de voluntariado.

49) Modelo_Especifico

Type: Class

Description: Modelo Específico de Indicadores de Inovação

Anotation/Observation: Modelos específicos apresentados na literatura para a representação dos indicadores de inovação.

50) Thing

Type: Class

Description: Classe raiz de todas.

51) Unidade_de_Medida

Type: Class

Description: Classe que representa as unidades de medida utilizadas para os indicadores.

APÊNDICE C – Levantamento dos modelos e indicadores utilizados na literatura de referência

Referência	Classificação / Indicador	
(AFFORTUNATO et al., 2010)	Overall Performance of the Economy	GDP per capita
		HDI
		Index of inequality
		Competitiveness
		Available income per capita
		Suffering banking customers use than ordinary
		Exports in goods and services
		Imports in goods and services
	The Economic and institutional regime	Bankruptcy of enterprises
		Companies registered per inhabitant
		Enterprise discontinued per capita
		Relationship between the number of employees and the total of firms
		Number of firms in a given field k than the national average for the respective field
		Cost to register a business
		Number of days required to start a business
		Gross credit to the private sector
		Exchange rate between banks
		Role of Law
		Offenses against public administration
		Taxation
	Innovation System	foreign direct investment
		Domestic direct investment
		Number of subscribers to the science on the total number of subscribers
		Employees in R&D
		Spending on research and development
		Number of employees in basic research
		Collaborations between research institutions and business world
		Number of innovative enterprises
		Internet and business

		Broadband
	Education	Schooling
		Internet access in schools
		Regional spending on education
		Training
		Spread of foreign language
		Number of businesses in education
		Educational qualifications for employment
		Investment in training per capita
		Contents of envelope structures for education
	Information and communication technology	Degree of spread of the Internet in households
		Allocation of personal computers
		Dissemination Service of e-government
		ICTs
	culture and capital social	Number of people involved in volunteering
		Contents of envelope structures for culture
		Number of non-school books products
		Investing in culture
		Number of companies operating in the field of culture
		Participation in community activities
		Working days lost to strike
(BARKLEY; HENRY; NAIR, 2006)	Innovative activity	Number of patents issued per 1,000 population (USPTO 1990–1999)
		Academic R&D expenditures per capita (NSF 1998–2000)
		Doctorates awarded in science and engineering per 1,000 population (NSF 1998–2000)
		Graduate science and engineering students per 1,000 population (NSF 1998–2000)
		Percentage of employment in technical professions—computer science engineering except civil natural, physical, and social science (BLS 2000)
	Labor force quality	Percentage of adult population (25+) that were high school graduates (Census 2000)
		Percentage of adult population (25+) that were college graduates (Census 2000)
		Percentage of population (age 16–64) that were

		employed (Census 2000)	
	Entrepreneurial environment	Percentage change in number of establishments (CBP 1990–2000)	
		Percentage of establishments with fewer than 20 employees (BLS 2000)	
		Number of INC 500 companies per 100,000 population (www.inc500.com, 1990–2000)	
		Venture capital investments (\$) per capita (Price Waterhouse Coopers 2000)	
		Percentage of employment in managerial and business professions (BLS 2000)	
		Percentage of employment in SIC 73, business services (CBP 2002)	
		Percentage of establishments in SIC 73, business services (CBP 2002)	
	Localization economics	Percentage of employment in high-technology industries (CBP, 2002)	
		Percentage of establishments in high-technology industries (CBP 2002)	
		Percentage of employment in information technology industries (CBP 2002)	
		Percentage of establishments in information technology industries (CBP 2002)	
	Competitiveness in global economy	Exports as a percent of gross metropolitan product, metro areas ranked in quartiles (DOC 1999)	
(BEI; ZENGYUAN, 2006)	Knowledge innovation capacity	Knowledge produced capability	R&D input
			R&D institution
			number of R&D people
			scientific research papers
			patent
		Knowledge scattered capability	FDI
			technology diffusion
			technology cooperation
			technology absorption and communication
			technology results application
	Enterprise	Technology	number of all kinds of professional

	Technology innovation capability	development capability	technological people	
			Investment of technological development	
			number of renewing and reconstruction projects	
			number of key science&technology harvest	
			number of national scientific advance awards	
		Technology application capability	Rate of technology commercial-	
			Rate of scientific re sults transforming to market	
			Rate of science&technology advanced contribution	
			Number of patent application	
			Situation of Technology cooperation	
	Network innovation capability	Network cluster degree	Number of enterprise per space	
			Proportion between dominant industrial production and industrial production	
			Widely usage of telephone	
			Number of people using internet	
			Total telecommunication services per person	
Network performance		Proportion between product selling income and development fee		
		Proportion between product selling income and technology introducing fee		
		Proportion between FDI ad GDP		
		proportion between GDP and amount of technology trade		
		Proportion between GDP and production output		
		Proportion between production selling income and enterprise's technoogy purchasing fee		
		Innovation Environment	Humanity environment	Synthetic index of innovation culture
				Synthetic index of innovation

			Service environment	talented person	
				Synthetic index of government service ability	
				Synthetic index of financial service ability	
			Infrastructure	Traffic environment	
				Information environment	
			Innovation Performance	Economic foundation	GDP per person
					Labor production rate
				Industry structure superiority	rate of the third industry
					rate of the industrial high-tech products
					Rate of information industry products
				Civil living level	Civil revenue per person
					Civil income per person
			Employment	Employment rate of city and countryside-	
			Opening Degree	Actual using foreign currency value per person	
(BUESA; HEIJS; BAUMERT, 2010)			Regional Innovation Environment	Size of the region	Gross fixed capital formation
					Number of people employed
					Gross domestic production
					Gross value added
					Wages
			Human Resources	HR in C&T High-Technology (total)	
				HR in C&T Services sector	
				HR in C&T (tech intensive services)	
			National Environment	Investment capital (seed and startup) with regard to Gross Domestic Product (%)	
				Investment capital (development) with regard do GDP (%)	
				Penetration of TICs	
			Innovatory Firms	Firms' staff in R&D (HC) with regard to total number employed (%)	
				Firms' staff in R&D (FTE) with	

		University	regard to total number employed (%)	
			Firms' R&D expenditure with regard to GDP (%)	
			University staff in R&D (HC) with regard to total number employed (%)	
			University staff in R&D expenditure with regard to GDP (%)	
			3rd cycle students with regard to the population	
			PA staff in R&D (HC) with regard to total number employed (%)	
		Public administrations	PA staff in R&D (FTE) with regard to total number employed (%)	
			PA R&D expenditure with regard to GDP (%)	
(CHANG et al., 2012)	Regional financing	Abundant Regional venture capital		
		Government R&D innovative grants incentives		
		Regional information and communication infrastructure expenditure		
	Industrial cluster	Regional universities and research institutes innovative support		
		Abundant regional talents pool		
		Regional technology transfer pipeline diversity		
		High-quality life function		
	Regional entrepreneurship	Regional entrepreneurship education courses and training opportunities		
		Assistance of Regional Innovation Incubator		
		Start-up number		
	Regional culture	Encourage self-entrepreneurship		
		Encouraged to take risks		
		Encourage the spirit of lifelong learning		
		Cultural inclusiveness		
	(CHEN; JU, 2014)	Regional technology innovation potential	Human resource potential	Student number in university(million)
				Education expenditure(100 million yuan)
Economic		Gross domestic production(100		

		condition	million yuan)
			The total numbers of attracting foreign investment (million) X4
		Knowledge innovation potential	Scientific research and comprehensive technical services industry newly increased fixed assets accounted for the total newly increased fixed assets proportion(%)
			Employment patent applications (project/million)
	The numbers of Internet users (users/ million)		
	Regional innovation input indicators	Human capital investment	Engaged in research and development technical personnel (person /million)
			R&D institutions R&D scientists and engineers full-time equivalent (million years)
			The employment of high tech enterprises (person)
		Innovation expenditure	R&D expenditure (100 million yuan)
			R&D funds and GDP ratio(%)
			The average funding of scientific research personnel (yuan/person)
			Local science and technology funding accounted for the local fiscal expenditure proportion (%)
			Large and medium-sized enterprises R&D internal expenditure accounted for the main business income ratio (%)
		Technology innovation direct	Scientific Paper published by Science and technology personnel (paper/million)
			Number of winning the national service and technology achievement award
Numbers of three kinds of patents granted (project/million)			

	Regional innovation output indicators	Technology indirect achievement	Technology market employment turnover (million yuan/million)
			The value added of the hi-tech industry accounted for the proportion of industrial added value (%)
			Exports of high-tech products accounted for the proportion of industrial manufactured goods exports (%)
			Sales revenue of new products accounted for the proportion of all products sales income (%)
		Economy development	Per capita GDP (yuan/person)
			Labour productivity of employees (yuan/person)
			Investment in new GDP (billion yuan/billion)
			Output rate of comprehensive energy consumption (yuan/kg)
	Regional innovation environment indicators	Environmental sanitation	Target rate of industrial wastewater discharge
			Numbers of health officer (person/million)
			Comprehensive utilization rate of industrial solid waste (%)
		Resident living quality	Fixed telephone subscribers (person/million)
			Post and Telecommunications (yuan/person)
			Mobile phone users (users/million)
			Area savings (thousand/person)
(CHEN; GUAN, 2011)	Innovative output	Incremental innovative output	International papers
			Domestic papers
			Domestic patents
			Value added of high-technology industries

	Innovative resources	Cumulative knowledge stock	GDP per capita
			Cumulative international papers
			Cumulative domestic papers
			Cumulative domestic patents
		Incremental innovative input	Basic R&D
			Applied and experiment R&D
			FTE S&E
			Government funds for great S&T
		External knowledge acquisition	S&T collaboration
			Technology transfer
			Foreign direct investment
		Innovative environment	Innovation sophistication
	Manufacturing and production capacity		
	Quality of employees		
	Common innovation environment		Communication environment
			Marketing environment
			Financial environment
			Public research institutions performance
	Cluster-specific environment		Specification for high-technology industry R&D performance
			Specification for traditional manufacturing industry R&D performance
	Innovation linkages		University R&D performance
Venture level			
(Dou et al., 2000)	Urban development level		Economic base
		Growth rate of GDP	
		Labor production efficiency	
		Urbanization level	Urbanization level
			Population density
		Inter-city linkages	Passenger traffic
	Freight traffic		
	Production factors	National resources	Urban built-up areas per capita
Urban location indicator			

			Water consuming per unit production		
		Human and knowledge resources	Number of professionals per 10000 employees		
			Number of students in high schools and universities per 10000 person		
		Capital resource	Local government revenue		
			The ratio of residents savings to GDP		
			Utilization of foreign capitals per capita		
		Infrastructure	Electricity consuming amount per 100 yuan production value		
			Number of public facilities		
			Density of roads		
			Number of telephones per capita		
			Beds in hospitals per capita		
		Demands	Consuming demands	Retail sales of consumer goods	
				Housing areas per capita	
			Investment demands	Investment in fixed assets per capita	
			Government purchasing demands	Government expenditures	
		Export demands	Rate of exports to GDP		
		Industrial and Business performance	Industrial structure	The percentage of secondary sector in GDP	
				The percentage of the third sector in GDP	
			Industrial development	Industrial output values per capita	
				Revenues and profits per 100 yuan output value	
			Real estate development	Investments on real estate per capita	
		OPEN AN UNIO N, 201	Enablers	Human resources	Percentage population aged 25-64 having completed tertiary education

		Finance and support	R&D expenditure in the public sector as % of GDP
	Firm Activities	Firm investments	R&D expenditure in the business sector as % of GDP
			Non-R&D innovation expenditures as % of turnover
		Linkages & entrepreneurship	SMEs innovating in-house as % of SMEs
			Innovative SMEs collaborating with others as % of SMEs
	Intellectual assets	EPO patent applications per billion regional GDP (PPSE)	
	Outputs	Innovators	SMEs introducing product or process innovations as % of SMEs
			SMEs introducing marketing or organisational innovations as % of SMEs
		Economic effects	Employment in medium-high/high-tech manufacturing and knowledge-intensive services as % of total workforce
			Sales of new to market and new to firm innovations as % of turnover
(EVANGELISTA et al., 2001)	Firm's strategy	Innovation activity	% of innovation costs devoted to R&D
			% of innovation costs devoted to design
			% of innovation costs devoted to marketing
			% of innovation costs devoted to patents
			% of innovation costs devoted to trial production
			% of innovation costs devoted to investment
		Process/product orientation	process innovations
			product innovations
			$\frac{((Proc+Proc&d)}{(Proc+Proc&d)+(Prod+Proc&d))}$ \square
		Innovation objectives	creating new markets
			extending the product range

			improving production flexibility	
			lowering production costs	
			replacing products	
	Firm's performance	Input		Innovation expenditure per employee
				Innovative investments per employee
		Output		new for the firm
				new for the sector
				new in absolute terms
		System's performance	Diffusion of innovation	
	Interregional flows			% of innovation expenditure sustained by resident firms in the region of residence on the total regional expenditure
				% of innovation expenditure sustained by firms resident in the region on the total expenditure sustained by the same firm
	Information sources			competitors
				conferences and exhibitions
				consultant firm
				customers
				internal information sources
				patents
				suppliers
	Obstacles to innovation			universities and public research institutes
			economic constraints	
		risk to be imitated		
		lack of information on technologies and markets		
		legislative constraints		
	scarce responsiveness of customers			

			lack of technological opportunities
			lack of external technological services
		Innovation policy	financial incentives from EC
			financial incentives from the public sector
			other financial incentives
			public research orders
			suppliers orders
			reaserch services
	technological services		
	Industrial Structure	Firm Size	Total no. of employees out of total no.of firm
		Productivity	Total sales per employee
		Relevance of science based industries	% of sales in science based industries
		Relevance of specialized suppliers industries	% of sales in specialised suppliers industries
		Relevance of Scale intensive industries	% of sales in scale intensive industries
Relevance of suppliers dominated industries		% of sales in suppliers dominated industries	
(FAN; ZHU, 2008)	The input of information and technology links	Technology trade income in this year	
		total number of enterprises and technological exchanges with foreign enterprises	
	The input of project cooperation links	The number of cooperation projects	
		the number of joint scientific research institutions	
		The proportion of production and research funding in all project funds	

	The input of personnel motivation links	The number of technology professional's introduction	
		The number of technology professionals outflow in this year	
		The number of experts at home and abroad engaged in technological development	
(GLEBOVA; KOTENKOVA, 2014)	Innovative Potential	Social Potential	Staffing
		Economic Potential	level of economic development of the region, industry
		Financial Potential	private and budget allocation of financial resources for the creation of innovation
		Science and Technology Potential	client level and progressive engineering and technology
		Legal Potential	well conceived regulatory framework and the development of special institutions support innovation
(GOGODZE, 2013)	Educational Level		
	Infrastructure		
	Governmental Support		
	Social Network		
	Knowledge Generation		
	Knowledge Intensive Production		
	Competitive Capacity		
(GUO; REN, 2011)	Innovation Input	R & D personnel inputs (million per years)	
		R & D expenditure-GDP ratio(%)	
		Investment in science and technology personnel per million population (person)	
		Expenditure-GDP ratio of technology(%)	
		The proportion of total local fiscal expenditure (%)	
	Innovation Output	The number of domestic Chinese journal articles (articles)	
		Patent applications (items)	
		Per capita GDP, (yuan)	
		High-tech product exports (million U.S. dollars)	
	High-tech	The proportion of high-tech industry output value of	

	industry	above-scale enterprises (%)	
		The proportion of high-tech industry added value of large-scale enterprises (%)	
	Innovation environment	the proportion of college academics above 6 years old (%)	
		Investment in expenditure on education accounted for the proportion of GDP (%)	
		Average years of schooling (years)	
		the number of Internet users of Per million people (households. per million people)	
(HAJKOVÁ; HAJEK, 2011)	Economy	Regional GDP per capita	
		Real growth rate of regional GDP	
		Employment rate	
		Long-term unemployment share	
	R&D	Patent applications per capita	
		Public R&D expenditure	
		Private R&D expenditure	
			R&D employment
	Education	Population with secondary education	
		Population with tertiary education	
		Participation in life-long learning	
(LEI; CHEN, 2009)	Innovation activities	funs for S&T activities	
		Internal expenditure for S&T	
		Internal expenditure for R&D	
		New product development fees	
		R&D fees / the main income	
		R&D fees / added value	
		Costs of digestion and absorption	
		Applying patent number	
		Innovation patents number	
		Output value of new products	
		Nes product revenue / main business income	
		Industrial added value ratio	
		New product development of new productys / sales revenues	

	Innovation Resources	Invention patent number / R&D expenses
		Number of scientific institutions
		S&T activities staff
		Scientists and engineers accounted for S&T staff
		Full-time personnel in R&D
		Industrial added value
		Government funds in S&T
		Government funds / S&T fees
		Product market occupied rate
		Total output / the whole country
		New product sales revenue / the whole country
(Li; ZHU, 2007)	Government's innovation support	
	The role of SMEs	
	SMEs' R&D	
	Cooperation among industries and universities as well as research	
(Li; LEI, 2009)	Innovation Base	The percentage of graduate on employment %
		The percentage of senior professional on employment %
		The percentage of high-grade technician on employment %
		The number of students of higher education per 10000 inhabitants
		The number of employment on R&D
	Innovation Environment	Gross Regional Product
		Per Capita GRP, Yuan
		The number of using internet computer
		Education Funds
		Foreign Direct Investment
	Innovation Input	Expenditure on R&D
		Growth rate of expenditure of R&D %
		Expenditure on new product development
		The number of employment on science and technology
		The number of scientist and engineer
		Local government appropriates funds on science and technology

		Expenditure for Science and Technology Promotion	
	Innovation Output	Growth rate of total production value %	
		New Products Value	
		Rate of New Products %	
		Three Kinds of Patents Granted	
		Transaction Value in Technical Market	
(Li, 2009)	Innovation effort		
	Interactions		
	Support from government		
	Interaction between knowledge users		
	Innovation environments		
(Li, 2011)	Innovation Environment	innovation infrastructure	
		market environment	
		qualities of workers	
		innovation funds	
		levels of entrepreneurship	
	Innovation Subject	comprehensive competitiveness	
		research and development capability	
		independent innovation capacity	
		the profitability of corporate brands	
	Knowledge Acquisition Capability	The level of cooperation	
		the level of technology transfer	
		foreign direct investment	
	Knowledge Creation Capability	input of research and development	
		the number of patents	
		the number of scientific papers	
		the ratio of input-output in S&T	
	Innovation Performance	macroeconomic	
		industrial structure	
		international competitiveness of products	
		resident income level	
employment level			
U; CH EN; 20 03	RIS milieu index	Location	Location attractions
			Communication facilities

			Location potentiality
			Location quotient
		Environment quality	Pollutant density
			Carrying capacity per capita
			Air pollution
		Economic level	GDPper capita
		Population quality	Education level
			Science and technology capability
		Social progress	Social civilization
			Social security
	Social development motivation		
	RIS element index	Talent input	R&D staff quantity
			R&D staff/total pop ulation
		Capital input	R&D capital expenditure
R&D capital expenditure total expenditure			
RIS un# index	Enterprises	Innovative production level	
		Innovation input-output ratio	
	Universities	Innovative talent training level	
		Innovation talent input-output ratio	
	Independent R&D institutes	Innovative knowledge & technology production level	
		Innovation knowledge & technology input-output ratio	
	Local government	Quality & efficiency of regional innovation coordination	
	Services	Scale of service industries	
Services contribution to regional innovation			
RIS structure index	Organization structure	Organizational centralization	
		Organizational openness	
		Organizational connection	
		Organizational hierarchy	
	Spatial structure	Spatial centralization	
		Spatial openness	
Spatial connection			

(LU; ZHANG; PENG, 2009)	RISfunction index	Innovation output	Spatial hierarchy	
			High-tech products export	
			Patents application	
			Scale of technology market	
		Innovation impact	Publication of research papers	
			Degree of science and economy integration	
			Economic growth	
			Environment protection	
	Input	Total quantity of R&D personnel		
		Total quantity of R&D personnel in each million		
		Interior expenditures of the R&D funds accounts for GDP (%)		
		The government science and technology collection accounts for GDP(%)		
		The enterprise R&D funds accounts for total sales of industrial enterprises (%)		
		R&D institutions accounts for expenditures of science and technology activities (%)		
		Output	Output of scientific and technological achievements	The number of three domestic and foreign patents admissibility (pieces)
				The new product sales revenue accounts for the product sales revenue(%)
				Per capita turnover of technology market (yuan)
			Output of economic efficiency	Per capita GDP (yuan)
High-tech industry output value accounts for the total output value (%)				
The commodity export accounts for GDP in the region.(%)				

		Output of social efficiency	Social labor productivity (million/people)
			The proportion of employed population in existing cities.
			Per capita disposable income of Urban households(yuan)
(MARKOWSKA; STRAHL, 2012)	Innovation stimulators	Human resources	percentage of university graduates aged 25-64,
			share in life-long learning per 100 people aged 25-64,
		Financing and support for innovation	share of public expenditure on R&D in GDP (in %),
			broadband access to the Internet (as % of total households)
	Enterprises activities	Investments	share of enterprises expenditure on R&D in GDP (in %),
			share of companies expenditure on innovation other than R&D in total expenditure
		External relations and entrepreneurship	share of SME implementing their own innovations in the total number of SME
			share of SME cooperating in innovation activities in the total number of SME
		Intermediary effects	number of patents granted by EPO per 1 million of population
	Enterprise performance results	Innovators	technological innovators (innovations related to products, services, processes) as % of total SME number
			non-technical innovators (marketing, organizational innovation) as % of total SME number,
			innovators in company resources efficiency
	Economic effects of innovation	share of workforce in mid-high and high-tech sectors in the total number employed in industry and service sectors	

		share of workforce in services based on expert knowledge as % of total workforce,	
		share of new or modernized (marketwise) products' sales in total sales	
		share of new or modernized (for enterprises) products' sales in total sales	
(Qi; Liu, 2014)	Innovation Environment	Per capita GRP	
		Proportion of Expenditure on Education to Government expenditure	
		Proportion of Expenditure on Science to Government expenditure	
		Scientific Research and development Institution	
		Internet Penetration Rate	
	Knowledge Innovation	Full-time teachers in regular Schools of Higher Education	
		Number of Enrolled Postgraduated	
		Application for Patents Accepted	
		Papers Taken by Major Foreign Referencing Systems	
	Technology Innovation	Proportion of Expenditure on R&D to GDP	
		Personnel Engaged in S&T activities	
		Funding for S&T activities	
		Proportion of new product value to Gross Output Value	
		Transaction Value in Technical Market	
	System Management Innovation	State industrialization Projects	
		Intellectual relative Treatment	
		Management of Knowledge Park	
	(SÁNCHEZ TOVAR; GARCÍA FERNÁNDEZ; MENDOZA FLORES, 2014)	Resources related to innovation	University Centers
			Firms enrolled in R&D
			Technological research centers
Technological parks			
Patents per 1000 inhabitants			
Employment in manufacturing sectors of medium-high technology			
Employment in knowledge intensive sectors			

	Socioeconomic and productive structure	Population density	
		GDP rate generated by the state	
		Employment in manufacturing sector	
		Employment in service sector	
		Population with Higher education	
		Employment rate in big firms	
		Specialization index	
(SUN, 2014)	Inputs	Resource input capability	
		Environment capability	
		Integration implementation capacity	
	Outputs	Human Capital	
		Relation Capital	
		Structure Capital	
		Innovation Capital	
(TIAN; TONG, 2011)	Flow of innovation resources	knowledge media	Number of public libraries
			Total reserves of public libraries (Per 10,000)
			Circulation of newspapers and magazines (Per 10,000)
			Broadcast coverage rate of population (%)
			TV coverage rate of population (%)
		Foreign design investment	Projects of foreign investment agreements signed (-)
			Amount of foreign investment agreements (Per 10,000 dollars)
	Foreign capital utilized in practice (Per 10,000 dollars)		
	Knowledge Creation	Innovation main body	Number of scientific research institutions
			Number of institutions of higher learning (-)
			Number of scientific and technological institutions owned by industrial enterprises of large and medium scale (-)
			Number of staff involved in scientific and technological activities in industrial enterprises

		of large and medium scale (Per capita)
	Innovation resources	Total expenses for scientific and technological activities (Per 10,000 yuan)
		Ratio of R&D expenditures in GDP (%)
		Total income of scientific and technological activities (Per 10,000 yuan)
		Comprehensive technical personnel for scientific research (Per capita)
Innovation environment	Infrastructure	Subscribers for the fixed telephones at the end of year (Per 10,000 households)
		Coverage rate of telephones in urban (Per 100 people)
		Length of the road paved at the end of the year (Per kilometer)
		Size of road paved at the end of year (Per million square meters)
		Number of public buses taken by per 10,000 people on average (--)
		Total annual quantity of freight transported (both by rail and by road included) (Per 10,000 tons)
		Total annual quantity of passengers transported (both by rail and by road included) (Per million population)
	Quality of workers	Local fiscal and education expenditure (Per 0.1 billion yuan)
		Number of undergraduates per 10,000 people
		Ratio of undergraduates in the total number of students per 10,000 people(%)
		Number of post-graduates
	Market environment	Total retail sales of consumer goods(Per 0.1 billion yuan)

		Creation of innovative environment by government	Preferential policies for high-tech industries, education and technological development(Extent)
			Support of laws and local regulations for innovation(Extent)
Pertinent treatment of intellectuals(Extent)			
Managerial level of knowledge zone or new and high technology area(Extent)			
Economic and social performance of innovation system	industrial structure		Ratio of primary industry in GDP (%)
			Ratio of secondary industry in GDP (%)
			Ratio of tertiary industry in GDP (%)
	Leading industries		Energy production (Per 10,000 tons of standard coal)
			Energy consumption (Per 10,000 tons of standard coal)
	Macroeconomics		GRP per capita (per yuan)
	Living standards		Disposable income of urban residents per capita(Per yuan)
			Consumption expenditure of urban residents per capita(Per yuan)
	Comprehensive employment situation		Unemployment rate registered in towns and rural areas(%)
			Number of people employed in extractive industries at the end of year(--)
	Regional sustainable development		Coverage rate of water in urban (%)
			Size of public green areas owned by per 10,000 population in urban (Per hectare)
			Energy consumed by per 10,000 yuan gross regional production (standard coal equivalent)(Per ton)
			Discharge quantity of industrial wastewater(Per 10,000 tons)

		Discharge quantity of industrial waste gas(Per 0.1 billion standard cubic meters)
(Tu; Song, 2013)	Intellectual Property Creation Ability	the R&D funds proportion in GDP
		the number of R&D personnel in ten thousands people
		the number of professional technical personnel in ten thousands people
		the annual patents application quantity in ten thousands people
		the proportion of authorized patent
		the quantity of invention patent in ten thousands people
		the quantity of effective registrations trademark in ten thousands people
		the annual application quantity of PCT patent in per million people
		the quantity of geographical indications in per million people
	Intellectual Property Application	contract turnover in technological market
		new product sales revenue
		the profit rate of output value of Enterprises above designated size with patent application
	Intellectual Property Service Ability	the number of legal services unit concerning intellectual property
		the main business income of legal service units
		the institution of intellectual property rights protection
(WANG; LI, 2009)	Innovation Structure System	Proportion of employees with Master's degree
		Enterprise innovation strategy
		Regional industrial structure
		Quantity of High-tech enterprises in region
		Region enterprise specialty correlation
	Innovation Supply and Demand system	Region Infrastructure
		Quantity of region innovation sources
		Region technology trade turnover
		Enterprise technology level gap
	Innovation	Technology trade turnover outside region
Innovation	Region university quantity	

	Support System	Region key laboratory quantity		
		Enterprises with R&D department quantity		
		R&D fund amount		
	Innovation assistant System	Policy environment		
		Government innovation prompt measure		
		Technology tradeplatform		
		Innovation encourage fund		
		Dissemination organization		
		Region human resource flowing rate		
	(WANG et al., 2012)	Input of S&T resources	Total R&D expenditure	
Full-time equivalent person-year of R&D personnel				
Number of academic papers				
Output of S&T resources		Number of invention patent granted		
		Output of high-tech products		
		Number of technological transaction contracts		
(WANG; LIU; TIAN, 2006)	Knowledge produced ability	Investment of scientific activities	Number of research and development organization above county belonged per 10,000 people	
			Investment strengthening of scientific activities people	
			Investment strengthening of scientific activities funds	
		Scientific and technological result	Number of patent per one million people	
			Number of issuing articles per 10,000 scientific and technological people	
	Knowledge scattered ability	Technology cooperation	Degree of cooperation of producing studying and researching	
		Introducing	Proportion of absorbed introduced technology	
		Scattering technology	Rate of applying and exchanging of industrial technological result	
	Enterprise Technology innovation capability	Developing ability	Proportion of number of scientific activities organization covering number of enterprises	
			Proportion of R&D people covering employment	

			Proportion of R&D funds covering sales revenue
			Proportion of new product developments covering scientific activities funds
		Designing ability	Number of applying for patent about practical new type and outward design per 10,000 people
		Manufacturing ability	Proportion of microelectronics controlling equipment equipment
			Proportion of funds of transforming technology
	Selling ability	Proportion of new product sales revenue covering sales revenue	
	Innovation Environment	Humanity environment	Synthetic index of innovation culture
			Synthetic index of innovation talented people
		Service environment	Synthetic index of governmental service ability
			Synthetic index of financial service ability
Infrastructure		Synthetic index of traffic environment	
		Synthetic index of information environment	
Innovation benefit	Economic strength	GDP per person	
		growth rate of social productive rate	
	Industry structure superiority	Contribution rate of the third industry	
		Rate of industrial high-tech products	
		Increasing rate of total income of trade industry and technology of private scientific enterprises	
	International competition of industry	Proportion of import and export value covering import and export total value of the whole province	
Degree of open up	Actual using foreign currency value per capita		

		Living standard	Controllable revenue per resident	
			Net income per farmer	
(WANG et al., 2009)	Innovation environment	Innovation Infrastructure		
		Market Demand		
		Labor Quality		
		Financial Environment		
		Levels of Entrepreneurship		
	Knowledge creation	Research and development Input		
		Patent		
		Research Papers		
		The Efficiency of Input-Output		
	Knowledge Acquisition	Technical Cooperation		
		Technology Transfer		
		Foreign Direct Investment		
	Enterprise Innovation Capability	Research Investment		
		Design Capability		
		Manufacturing Capability		
		Output Value of New Products		
	The Economic benefits of innovation	Macroeconomic		
		Industrial Structure		
		The International Competitiveness		
		People's Income Level		
Employment				
(XIAO; YANG, 2011)	Knowledge creativity	Investment in research & exploitation	Number of the researchers	
			Sci-tech investment of the government	
			The proportion of sci-tech investment in GDP	
		Output of science and technology	Number of patented invention	
			number of sci-tech thesis in foreign periodicals	
		Technology transformation	money involved in concluded contract in technological market	
			money involved in concluding the import of foreign technology	
		Foreign investment	actual utilization of foreign investment	
		Technological	The scale of	number of researchers for each

	innovation capacity of enterprises	research investment for enterprise	enterprise
			the invested funds for research
		The output of innovation for enterprise	number of nationally qualified technical centers
			value of new output
	Environment for technological innovation	Infrastructure	per capita household electricity consumption
			per capita highways ownership
		Market demand	fiscal expenditure of local government
			Per capita nonproductive expenditure of urban citizen
			The proportion of educational expenditure in GDP
			the number of undergraduates of higher learning in each 10 thousand
	Innovated economic performance	Macro-economy	per capita book consumption
			per capita GDP
International competitiveness of industry		The proportion of tertiary industry in primary industry	
		export sales of commodities	
(YAM <i>et al.</i> , 2011)	Technological innovation capabilities	Learning capability	
		R&D capability	
		Resource Allocation capability	
		Manufacturing capability	
		Marketing capability	
		Organization capability	
	Strategic planning capability		
	Technological Innovations	Sales performance	
(YAN; YAN; MA, 2009)	The capability in Innovation of investment	The R&D funds in GDP proportion	
		The scientific and technical activities funds in GDP proportion	

(YIN; DIAO; LI, 2011)		The number of research and development insitute
		The number of Colleges and Universities
	The allocation capability in innovation	The University-Industry Cooperation of standard
		The cooperation of the enterprises
		The turnover of technical market in GDP proportion
		The the popular rate of Home PC
		The Service Level of sci-tech intermediate organizations
		The Capita Added Value of Tertiary Industry in GDP proportion
		The household savings in GDP proportion
	The innovation management	The tax policy for innovation
		The financial policy for innovation
		The fundamental knowledge and industrial general technology
		The marketization level
	The production ability in Innovation	The economic growth rate
		The the scientific and technological progress of contribution rate to GDP growth
		The per capita disposable income
		The per capita GDP
	Knowledge Innovation capability	Full-time equivalent of R&D personnel
		Scientists and engineers in S&T activities
		Ratio of Scientists and engineers in S&T activities
Ratio of government grant in total funds obtained for S&T activities		
The number of patent application		
Owning inventive patent		
Enterprise Technology creation capability		Original value of productive equipment
		Equipment controlled by micro-electronics
		Ratio of enterprise self-financed in total funds obtained for S&T activities
		Ratio of expenditure on the purchase of instruments in instrumental expenditure on S&T activities
	Expenditure for new products	
	Intramural expenditure of R&D	
	Expenditure on technology absorption	

	Innovation environment	Number of S&T institutes in enterprises	
		Personnel in S&T institutes	
		Spending for S&T institutes	
		Total foreign investment	
	Innovation economy performance	Output value for new products	
		Revenue from new products	
		GDP in region	
	(ZABALA-ITURRIAGAGOITIA <i>et al.</i>, 2007)	Regional and Productive Environment for Innovation	Gross Value Added high and medium technology firms
			Gross Value Added low technology firms
			Employment high and medium technology firms
Employment low technology firms			
Exports High and medium-high technology firms			
Exports Medium-low technology firms			
Exports Low technology firms			
Spanish patents			
National projects funded by the CDTI			
European Patents			
GDP			
Role of Universities		Internal University R&D expenditure (% of GDP)	
		Internal University personnel (FTE) in R&D	
		University researchers (FTE) in R&D	
		Students enrolled in tertiary education	
		Graduated Students	
		Students enrolled in postgraduate courses	
		Defended PhD thesis	
		Research quality indicator of university	
Role of Civil Service		Government Expenditure on R&D (% of GDP)	
		Government personnel (FTE) in R&D	
		Government researchers (FTE) in R&D	
		Scientific capital stock in R&D	
		Venture capital investment	
Role of Innovating Firms		Firms internal R&D expenditure (% of GDP)	
		Internal personnel of firms(FTE) in R&D	
		R&D researchers (FTE) of firms	
		Firm's technological capital stock in R&D	

		Regional distribution of technology centres	
		Annual income of technology centres	
		Innovation Expenditures	
(ZHANG; LIU; ZHANG, 2011)	Input	University	Number of colleges and universities
			ratio of cost of scientific research that comes from industries
			ratio of cost of scientific research that comes from governments
		Industry	proportion of personnel engaged in R & D activities accounted for the overall scientific and technical personnel
			the number of high-tech enterprises
			the proportion of R & D expenditure of the main business income
		Government	proportion of science and technology spending in local finance
			the number of research and development institutions
			governmental funding and expenditures for scientific and technological activities to industrial enterprises above designated size
	Output	Granted invention patent	
		Per capita GDP	
		the amount of technology market clinches a deal the contracts	
(ZHENG; ZHANG, 2013)	Innovation Input	Internal expenditure of R&D to GDP	
		Coefficient of fixed assets investment	
		Percentage of government expenditure to internal expenditures	
		R&D funding Per capita	
		Intensity of R&D	
		Number of R&D personnel	
		R&D personnel Per million	
		Full-time equivalent of R&D personnel	

	Innovation Output	Number of high-tech enterprises
		Proportion of the high-tech enterprises output value to GDP
		Internal expenditures of R&D of high-tech enterprises
		Transactions of technical market
		Amount of the introduction of foreign technology
		Amount of contract by the introduction of foreign technology
		Number of authorized patents
		Total amount of patent applications
		Total quantity of scientific papers retrieved
		New product sales of large and medium-sized enterprises
		Innovation environment
	Annual disposable income per capita	
	Revenue Per capita	
	Ratio of Local fiscal revenue to GDP	
	Number of telephone subscribers	
	Number of Internet users	
	Number of mobile users	
	Business volume of Telecom	
	Ratio of Non-state-owned economy fixed assets investment to total	
Percentage of government expenditure to Internal Expenditures		
Ratio of fund of universities and research institutions of science and technology activities to the enterprise funds		
Percentage of Education expenditure to the budget expenditure		
Proportion of population of college education aged 15 and over		
University		
Library		

Fonte: o autor