

XVIII ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2017

GT-8 – Informação e Tecnologia

PLATAFORMA INFORMACIONAL DO ECOSISTEMA PAULISTA DE INOVAÇÃO: MODELO COMPUTACIONAL E SEMÂNTICO DE APOIO À INOVAÇÃO

Elvis Fusco (Universidade de São Paulo - USP)

Marcos Luis Mucheroni (Universidade de São Paulo- USP)

Caio Saraiva Coneglian (Universidade Estadual Paulista- UNESP)

INFORMATION PLATFORM OF INNOVATION SYSTEMS IN THE STATE OF SÃO PAULO: COMPUTER AND SEMANTICS MODEL OF INNOVATION SUPPORT

Modalidade da Apresentação: Comunicação Oral

Resumo: A inovação como processo é atualmente um dos grandes pilares para que organizações e países possam manter sua sustentabilidade num cenário globalizado de acirrada competição. Uma organização pode desfrutar do resultado de uma inovação por anos, talvez até décadas, considerando que as inovações são capazes de gerar vantagens competitivas, a médio e longo prazo. No Brasil, o poder público ainda é o principal incentivador e financiador de ações de inovação nas organizações, porém, além do governo existe um conjunto de atores que atuam nesse processo, em que se destacam os ambientes de inovação, as universidades, as empresas e as entidades de apoio e fomento. No Estado de São Paulo, o cenário de inovação tem como base o Sistema Paulista de Ambientes de Inovação que institui e regulamenta os Ambientes Formais de Inovação que apoiam as iniciativas de inovação nas organizações. Com o objetivo de ampliar a articulação e apoiar os processos de construção coletiva e colaborativa entre os principais atores de inovação do Estado de São Paulo, esta pesquisa propõe uma Plataforma Informacional do Ecosistema Paulista de Inovação que, a partir de fontes informacionais disponibilizadas pelos atores de inovação e por meio de tecnologias informacionais, semânticas e computacionais, agregue e disponibilize serviços de informação para os atores de inovação. A plataforma informacional utiliza conceitos da Web Semântica e tem como base uma ontologia que representa as informações dos principais atores de inovação do Estado e um conjunto de serviços agregados de informação, formados por agentes computacionais autônomos de extração e serviços de recuperação das informações contidas na plataforma. Por meio dos recursos de extração automática de dados de ambientes informacionais digitais e da disponibilização de serviços informacionais gerados por meio de tecnologias semânticas e computacionais geradas pela plataforma, permite que organizações tenham acesso a insumos informacionais necessários para o fomento a seus projetos de inovação.

Palavras-Chave: Inovação; Web Semântica; Ecosistema de Inovação; Plataforma Informacional.

Abstract: Innovation as a process is currently one of the greatest pillars so that organizations and countries may maintain sustainability in a globalized scenario under fierce competition. An organization can take advantage of the results from an innovation for years, perhaps even for decades, considering that innovations are capable of generating competitive advantages, on medium and long term basis. In Brazil, public authorities are still the main sponsors and funders of innovation actions in organizations, however, in addition to the government, there is a group of performers active in this process where innovation environments are highlighted. These sponsors are universities, companies and entities of development and support. In the state of São Paulo, the innovation scenario is based on the regulations by the *Sistema Paulista de Ambientes de Inovação* (System of Innovation Environments of the State of São Paulo) to all Formal Environments for Innovation that encourage innovation initiatives in organizations. Aiming at expanding the articulation and encouraging collective and collaborative construction processes among the main innovation executors in the state of São Paulo, this research has the purpose of presenting an Information Platform of the Innovation Ecosystem in the State of São Paulo that, from information sources provided by innovation executors and through information technologies, computer and semantics, aggregate and make information services available for innovation performers. The information platform uses Web Semantics concepts and is based on an ontology that represents information about the main innovation executors in the State and a set of aggregated information services, formed by autonomous computer agents of extraction and retrieval services of the data contained in the platform. Through automatic data extraction resources of digital information environments and the provision of information services generated through semantics and data technologies generated by the platform, it allows organizations to access the data inputs needed to foster innovation projects.

Keywords: Innovation; Web Semantics; Innovation Ecosystem; Information Platform.

1 INTRODUÇÃO

O cenário contemporâneo dos ambientes informacionais digitais das organizações caracteriza-se pela dependência do uso da informação nos processos de tomada de decisão. Esse acoplamento entre a tomada de decisão e a informação é uma das consequências da Sociedade da Informação, contexto no qual tecnologia e informação constituem a base organizacional das relações na sociedade.

Dentre os diversos esforços em utilizar as informações para apoiar as ações de apoio à decisão nas organizações está o processo de inovação, que utiliza-se de processos informacionais sistemáticos para sustentar tais ações. Esses processos são representados por fluxos informacionais que caracterizam-se pela busca, acesso, recuperação, transformação, processamento, representação, persistência, transferência, apresentação e uso da informação, tendo como fontes informacionais os ambientes informacionais dos atores da inovação formados principalmente pelo governo, universidades, órgãos de fomento e as empresas.

No contexto dos dados na era do Big Data, observa-se uma disrupção tecnológica pela convergência da colaboração, mobilidade e grande volume de dados. O grande desafio para a pesquisa de ambientes informacionais digitais e para a forma de uso das informações nas organizações é promover a integração destas tecnologias para balancear as necessidades de geração, acesso e controle destas informações, bem como as oportunidades deste comportamento emergente e suas inovações. Ambientes informacionais digitais que utilizam aparatos semânticos e agentes autônomos de extração figuram como possibilidade de se utilizar esse grande volume de informações para sustentar as ações de apoio à decisão em processos de inovação nas organizações, a partir de processos informacionais sistemáticos.

Este cenário em que as organizações estão inseridas exige uma evolução dos processos de gestão e de construção contínua de novos conhecimentos na geração da inovação tecnológica, ou seja, de seus produtos e serviços. Como resultado do uso da informação e da construção do conhecimento na promoção da inovação há a geração de resultados a longo prazo, ou seja, considerando que as inovações são capazes de gerar vantagens competitivas, a médio e longo prazo, inovar torna-se essencial para que as organizações se mantenham estáveis e sustentáveis na sociedade contemporânea.

O sucesso em políticas de inovação é historicamente formado por dois eixos: estímulo e investimento, ambos insuficientes no Brasil. A presença do país na 100ª posição (de 138) no

pilar "Inovação" do Relatório Global de Competitividade¹ do Fórum Econômico Mundial e responder por 0,2% da produção de patentes no mundo, são reflexos de ações insuficientes nesta questão.

No Brasil, o poder público ainda é o principal incentivador e financiador, via agências de fomento. Porém, além do governo, existe um conjunto de atores que atuam nesse processo em que se destacam os ambientes formais de inovação, as universidades, as empresas e os órgãos de fomento. No Estado de São Paulo, o cenário de inovação tem como base o Sistema Paulista de Ambientes de Inovação (SPAI) que institui e regulamenta os ambientes formais de inovação do Estado que apoiam as iniciativas de inovação nas organizações.

Nesse cenário, novas estratégias são necessárias e devem incluir políticas públicas, a articulação entre os ambientes formais de inovação, as universidades e as organizações com uma visão sistêmica para o estímulo e o fomento de criação e o desenvolvimento de ecossistemas de inovação e de mecanismos de geração de empreendimentos e ações de inovação no país. Essas estratégias permitem trazer a economia brasileira para o século XXI, alinhadas com as mais modernas visões de desenvolvimento econômico e social.

Com o objetivo de ampliar a articulação e apoiar os processos de construção coletiva e colaborativa entre os principais atores de inovação do Estado de São Paulo, a partir de fontes de informação disponibilizadas pelos mesmos, esta pesquisa propõe uma Plataforma Informacional do Ecossistema Paulista de Inovação que, por meio de tecnologias informacionais, semânticas e computacionais, agregue e disponibilize serviços de informação para os atores de inovação do Estado de São Paulo.

A plataforma informacional utiliza conceitos da Web Semântica e tem como base uma ontologia que representa as informações dos principais atores de inovação do Estado e um conjunto de serviços agregados de informação, formados por agentes computacionais autônomos de extração e serviços de recuperação das informações contidas na plataforma. Por meio dos recursos de extração automática das informações dos ambientes informacionais dos atores de inovação e da disponibilização de serviços informacionais gerados por meio de

¹ *The Global Competitiveness Report 2016–2017* - <http://reports.weforum.org/global-competitiveness-index>

tecnologias semânticas e APIs² (*Application Programming Interface*) geradas pela plataforma, a plataforma permite que organizações tenham acesso a insumos informacionais necessários para o fomento a seus projetos de inovação.

Na sequência do artigo é apresentada a contextualização da inovação no Brasil com ênfase no ecossistema de inovação do Estado de São Paulo e os conceitos que fundamentam a proposta da plataforma como web semântica, ontologia e agentes computacionais de extração. Ao final, é apresentada a Plataforma Informacional do Ecossistema Paulista de Inovação incluindo sua arquitetura, artefatos, o modelo de domínio, a ontologia e os serviços de informação. Finalmente, são apontados trabalhos correlatos e considerações finais com apontamentos de caminhos para pesquisas futuras.

2 INOVAÇÃO

A inovação é prioridade em países desenvolvidos que têm como estratégia de crescimento e competitividade a diferenciação na pesquisa e desenvolvimento de novas tecnologias. A cultura de inovação de um país ou região depende da articulação dos principais atores da inovação, governo, ambientes formais de inovação, órgãos de fomento, empresas e instituições de ensino e pesquisa, que tem papéis distintos, porém colaborativos.

A inovação é conceituada por muitos autores, porém, por consenso, estabelece-se que esta somente se materializa quando pode ser explorada com sucesso. A inovação em suas diversas facetas e aplicações é fomentada em países desenvolvidos e países emergentes de forma contundente.

Segundo o World Bank (2015), o Brasil investe discretamente em inovação. Em 2011, o país investiu 1,2% do PIB em inovação, enquanto países como EUA e Coréia do Sul investiram 2,8% e 4%, respectivamente.

Em valores nominais essa diferença pode ser notada com maior evidência. Enquanto o Brasil investiu 19,4 bilhões de dólares, os EUA e Coréia do Sul investiram 405,3 e 65,4 por ano, respectivamente.

A conclusão da pesquisa sobre inovação da Confederação Nacional da Indústria (CNI, 2015) brasileira mostra que 62% dos empresários consideram baixo ou muito baixo o grau de investimento em inovação da própria indústria.

² Conjunto de rotinas e padrões de programação para acesso a um aplicativo de software ou plataforma baseado na Web.

O principal motivo para o Brasil estar atrasado em relação a outros países é o reflexo de defasagem tecnológica acumulada ao longo dos anos. A consequência disso é que a indústria, muitas vezes, acaba por importar ou copiar o que é feito em outros países. De acordo com os entrevistados, falta cultura de inovação nas empresas brasileiras em geral. Eles também elencaram como entraves a falta de políticas de incentivo, a dificuldade de interação entre empresas e universidades e o baixo nível de educação dos profissionais (CNI, 2015).

Esta pesquisa da CNI mostra que a indústria brasileira não desenvolveu em seu DNA a cultura da inovação. No entanto, esta deficiência se estende por esferas governamentais e universidades, dificultando a expansão de iniciativas inovadoras que possam obter sucesso.

Um dos pontos negativos também enfatizados pela CNI foi a dificuldade de interação entre as empresas e as universidades, porém deve ser incluído nessa relação os órgãos governamentais, tornando este cenário ainda mais complexo e desafiador.

No entanto é importante destacar que justamente esta relação bem sucedida entre governo, iniciativa privada e universidades é a base da cultura de inovação de países desenvolvidos. Assim, estes são nomeados neste artigo como os atores de inovação.

Desta forma, promover uma cultura inovadora é um legado que deve ser construído com esforço colaborativo entre os atores de inovação.

Destaca-se o cenário atual de inovação no Brasil e como uma metodologia de fluxo informacional apoiada pela Ciência da Computação e a Ciência da Informação podem contribuir para o crescimento do país em relação a sua cultura de inovação.

O fluxo informacional em uma organização ou entre organizações pode ocorrer empiricamente e/ou estabelecido por regras apoiadas em pessoas, organogramas, hierarquias, dependências e relações. Segundo Santo e Valemtim (2014), o contexto dos fluxos de informação possibilitam o estabelecimento das etapas de obtenção, tratamento, armazenamento, distribuição, disseminação e uso da informação no contexto organizacional.

Tratado como elemento estratégico para tomadas de decisões, reorganização e otimização de processos organizacionais, o fluxo informacional pode agregar valor a um conjunto de informações aparentemente correlatas, porém desorganizadas ou não estruturadas.

Os Estados e municípios têm autonomia para criar seus mecanismos de incentivo à inovação. Neste cenário, o Governo do Estado de São Paulo, pelo decreto 60.286, de 25 de

março de 2014, instituiu e regulamentou o Sistema Paulista de Ambientes de Inovação (SPAI). O SPAI (ANPROTEC, 2014) compreende o Sistema Paulista de Parques Tecnológicos (SPTec), a Rede Paulista de Incubadoras de Empresas de Base Tecnológica (RPITec), a Rede Paulista de Centros de Inovação Tecnológica (RPCITec) e a Rede Paulista de Núcleos de Inovação Tecnológica (RPNIT).

A regulamentação do SPAI estabeleceu requisitos para reconhecimento dos Ambientes Formais de Inovação: incubadoras de empresas de base tecnológica, centros de inovação tecnológica, parques tecnológicos e núcleos de inovação tecnológica. Esses ambientes de inovação podem receber apoio do Estado e devem explorar a relação entre os atores de inovação em prol da pesquisa, desenvolvimento e inovação.

O domínio de aplicação da plataforma computacional e informacional desta pesquisa está norteado e apoiado em modelos e estruturas formais, articuladas pelos governos estadual e nacional, como é o caso do SPAI no Estado de São Paulo. Delimitando a abrangência da aplicação da proposta da plataforma aos ambientes formais de inovação, órgãos de fomento e organizações que apoiam processos de inovação nas empresas.

3 WEB SEMÂNTICA

A Web Semântica como uma extensão da atual Web, atribui à informação um significado bem definido, específico, para melhor permitir que os computadores e as pessoas trabalhem em cooperação.

Embora motores de busca ou indexadores atuem de forma a auxiliar as pessoas a encontrarem recursos informacionais, a efetividade dos resultados das buscas geram novas investigações na área da Recuperação da Informação, porque mesmo indexando muito conteúdo da Web, eles ainda não têm total capacidade de selecionar as páginas que o usuário realmente quer ou precisa. E, assim Berners-Lee propôs uma série de mecanismos pelos quais desenvolvedores e autores, individualmente ou em colaborações, pudessem usar auto-descrições e outras técnicas para que os programas de compreensão de contexto, possam encontrar seletivamente o que os usuários realmente desejam.

Um desses mecanismos passa pela introdução de várias normas, padrões ou *standards*, para a padronização de criação e obtenção de dados ou metadados. Essas tecnologias da Web Semântica permitem que as pessoas criem armazenamentos de dados, construam vocabulários e escrevam regras para o tratamento de dados.

No âmbito da Web Semântica, as ontologias são capazes de representar computacionalmente os dados construídos pelas pessoas, possibilitando que agentes computacionais sejam capazes de compreender o contexto e o sentido que os dados são publicados.

3.1. Ontologia

De acordo com Gruber (apud OLIVEIRA et al, 2003) ontologia é uma especificação explícita de uma conceituação do conhecimento em forma de entidades e a especificação é a representação dessa conceituação em uma forma concreta.

Ontologias podem ser classificadas e utilizadas de diversas maneiras, é uma técnica de organização de conceitos e relacionamento que existem dentro desse conceito, um compartilhamento e entendimento comum de algum domínio para facilitar a reutilização de conhecimento. A ontologia, é definida por Silva (2003) como a parte da ciência que estuda o ser e seus relacionamentos. O uso de ontologias é fundamental no processo de desenvolvimento do conhecimento conceitual informacional, dos robôs de busca semântica, sendo aplicada na Ciência da Computação e na Ciência da Informação para condicionar uma busca de maneira mais inteligente e mais próxima do funcionamento do processo cognitivo do usuário de forma que a extração de dados se torne muito mais relevante.

4 AGENTES COMPUTACIONAIS

Os agentes computacionais são elementos que surgiram com maior frequência na literatura da área da Ciência da Computação a partir do final do século XX. Identifica-se tal movimento em bibliografias como Genesereth e Ketchpel (1994), Franklin e Graesser (1996) e Wooldridge (1997).

A princípio verificou-se que o conceito de agentes computacionais estava fortemente vinculado a área de inteligência artificial, em que essa ferramenta se apresentava como um importante requisito na busca de criar aplicações baseadas em inteligência computacional.

Na busca por uma definição de agentes computacionais, encontram-se diversos conceitos utilizados por distintos autores. Uma definição muito citada é dada por Wooldridge (1997), que explicita que agentes computacionais são sistemas capazes de decidir o que deve ser feito em uma determinada situação. Franklin e Graesser (1996) aprofundam na busca pelas definições de agentes computacionais, enumerando diversos conceitos dados pelos principais autores e instituições da época, realizando a partir das mesmas uma definição bastante citada

e utilizada na literatura. Os autores relatam que um agente “é um sistema que se encontra dentro de um ambiente, sentindo e agindo sobre o mesmo ao passar do tempo, na busca de cumprir sua agenda, efetuando o que ele sente no futuro” (FRANKLIN; GRAESSER, 1996, p. 25, tradução nossa).

As definições apresentadas contemplam especialmente características envolvendo a autonomia e a capacidade desses sistemas tomarem decisões, verificando o ambiente em que os mesmos se encontram. Cabe ressaltar que as tomadas de decisão realizadas por agentes podem utilizar-se de diversos tipos de algoritmos computacionais, podendo ser simplesmente sistemas programados para realizarem determinadas tarefas, que se relacione com um ambiente ao se adaptar ao mesmo, ou tomar certa decisão baseada na ocorrência de determinados fatores.

Nessa perspectiva, verifica-se o papel importante dos robôs de extração. Essa tecnologia computacional é definida como uma estrutura que percorre determinadas páginas da Web ou toda a Web recolhendo as informações que apresentam relevância para um determinado sistema. Cada robô apresenta diversos comportamentos, que podem variar conforme cada aplicação, mas em suma, um robô contém uma lista de *websites* que devem ser extraídos, podendo executar também uma extração recursiva, em que o robô acessa primeiramente a página inicial, e navega por todos os links que tal página possui. Destaca-se, que a recuperação das páginas ocorre por meio do protocolo HTTP (DETERS; ADAIME, 2003).

Alvarez (2007) contextualiza o conceito de extração de informação, dizendo que esta é uma área focada em extrair de modo automatizado partes que sejam relevantes, tornando elementos dispostos em linguagem natural, como por exemplo, um texto, em sistemas formais para o entendimento de computadores. O autor ressalva ainda que o processo de extração de informação é determinado por um conjunto de regras e padrões, que pode ser definida por especialistas com graus determinados de automação.

Os dois conceitos apresentados, robô de buscas e extração de informação, estão fortemente interligados. Principalmente em ambientes Web, o processo de extração de informação costuma estar totalmente conectado com robôs de buscas, que realizam a varredura, e extraem, seguindo determinado padrões, as informações que foram previamente estabelecidas como relevantes, em um determinado contexto.

Na perspectiva dessa pesquisa, o termo agente de extração tem como base o conceito de robô de extração, que é capacitado para extrair conteúdos de páginas HTML, com as

propriedades de agentes computacionais, ao expandir as funções realizadas pelos robôs de busca, podendo o mesmo se adaptar e se integrar com diversos módulos de tomadas de decisão. Assim, os agentes computacionais inteligentes na Internet (não necessariamente na Web) ou simplesmente um agente, é um programa que usa uma informação ou realiza um serviço sem uma presença humana direta em um serviço regular, no âmbito da Ciência da Informação pode-se designar a uma busca de informações por meio de algum vocábulo ou informação simples atendendo a algum critério semântico.

Na perspectiva de monitoramento de fontes informacionais, o trabalho de Silva (2003) insere a questão do uso de agentes para realizar esta tarefa, no intuito de automatizar o processo de extração de informação.

Considerando, assim, o conceito de monitoramento de fontes informacionais, cenário no qual as informações estão num ambiente dinâmico e disperso, os agentes computacionais agem como uma ferramenta capaz de se relacionar com o ambiente, para atender a determinadas tarefas. Verifica-se a necessidade de criar agentes computacionais semânticos de extração de informação que sejam capazes de agregar a função de robôs de extração, com características de agentes computacionais e padrões ontológicos para agregar as informações dos ambientes informacionais digitais contidas no Ecosistema de Inovação do Estado de São Paulo.

5 ECOSSISTEMA DE INOVAÇÃO DO ESTADO DE SÃO PAULO

Ao refletir um cenário contemporâneo no qual existem raras representações sistematizadas, como é o proposto em relação ao mapeamento do Ecosistema de Inovação do Estado de São Paulo, foi oportuno valer-se de uma pesquisa exploratória com a utilização do método qualitativo na perspectiva do aprofundamento de um domínio específico de relações do escopo sugerido como parte do método. Esta estratégia de pesquisa é indicada para investigar temas novos e com um aspecto de descrever e analisar resultados e fenômenos.

O conceito de ecossistema de inovação começou a atrair a atenção para órgãos governamentais apenas nos últimos anos.

O termo Ecosistema de inovação é usado para descrever o grande número e de natureza diversa de participantes e recursos que são necessários para a inovação (JACKSON, 2011).

XVIII ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2017
23 a 27 de outubro de 2017 – Marília – SP

Os sistemas de inovação bem-sucedidos são tipicamente caracterizados por uma economia de conhecimento ativa, que compreende atividades de P&D (Pesquisa e Desenvolvimento) e atividades de inovação acadêmica, pública e de negócios com comercialização efetiva e todas apoiadas por mecanismos flexíveis de política pública. Além disso, os ecossistemas de inovação bem sucedidos também precisam de uma cultura de inovação baseada na interação e abertura para oportunidades e mudanças internacionais.

Refletir sobre o ecossistema de inovação desta maneira, sugere algumas características importantes que estão no foco do objetivo desta pesquisa em identificar e representar o Ecossistema Paulista de Inovação.

Um ecossistema é dinâmico e flexível, permitindo que os novos participantes se tornem parte do ecossistema com barreiras de entrada mínimas, permitindo que partes do ecossistema se desvançam e deixem o envolvimento ativo.

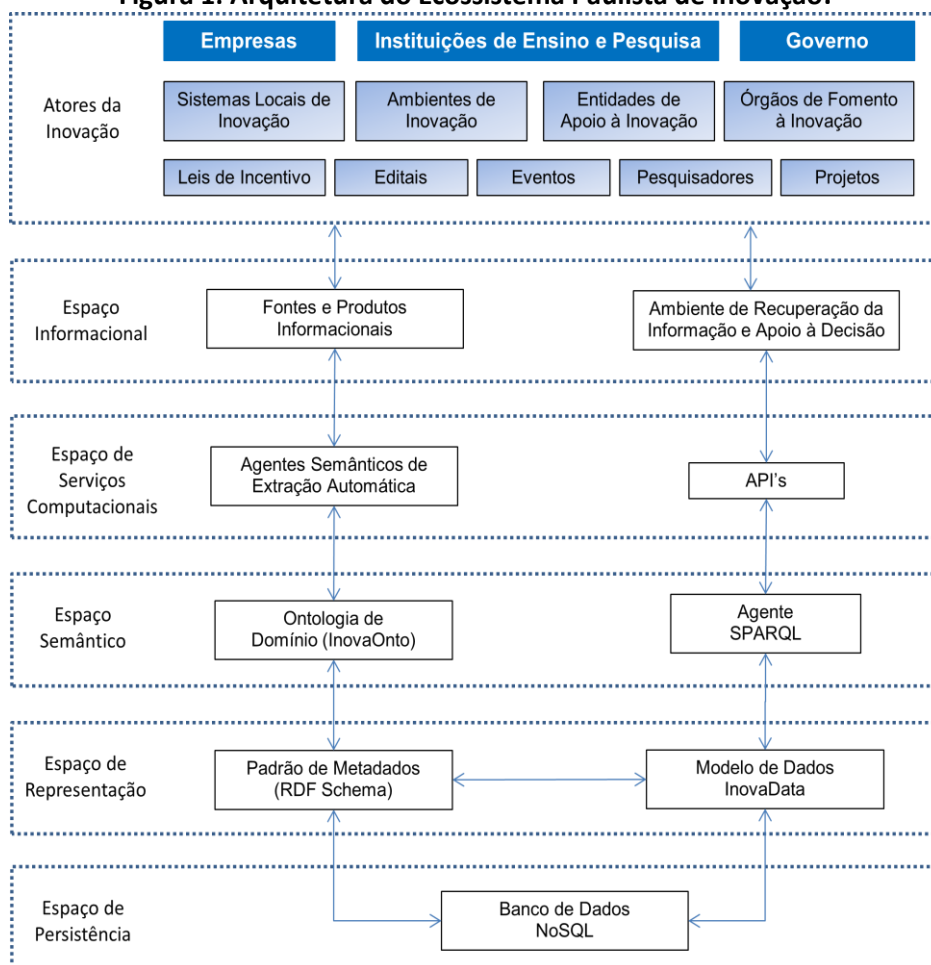
Em segundo lugar, um ecossistema é um sistema aberto sem respeito pelas fronteiras ou geografias jurisdicionais, esse conceito é utilizado nesta pesquisa já que alguns atores de inovação identificados atuam no âmbito nacional.

Em terceiro lugar, um ecossistema não está preocupado com as suas estruturas, mas é mais focado no alcance e na qualidade das interações dentro e entre as estruturas do ecossistema.

O grau de flexibilidade e dinamismo, a abertura do ecossistema e a extensão e qualidade de suas interações e relacionamentos proporcionará evidências importantes da saúde do ecossistema de inovação e da contribuição que o ecossistema de inovação pode, potencialmente, fazer para o Estado de São Paulo e seu desempenho de inovação.

Neste sentido, propõe-se a seguinte arquitetura informacional computacional e semântica que pretende apoiar os principais atores da inovação no Estado de São Paulo por meio da agregação, sistematização e disponibilização das informações para apoiar processos de inovação (Figura 1).

Figura 1: Arquitetura do Ecossistema Paulista de Inovação.



Fonte: Elaborado pelos autores.

O cenário da inovação é composto por diversos atores que disponibilizam e consomem informações dinâmicas e heterogêneas que demandam a utilização de conceitos e tecnologias computacionais, informacionais e semânticas para dar conta das demandas de uso dos usuários da inovação.

A arquitetura proposta considera a premissa da colaboração dos principais atores nos processos da disponibilização da informação para suportar todo o fluxo informacional no qual esta plataforma se baseia.

5.1. Modelo de Domínio do Ecossistema Paulista de Inovação

Modelagem de Domínio é uma forma de estrutura e organização de conhecimento visando possibilitar o entendimento e o mapeamento de quais conceitos são importantes para determinadas tarefas, indicando o que realmente deve ser investigado.

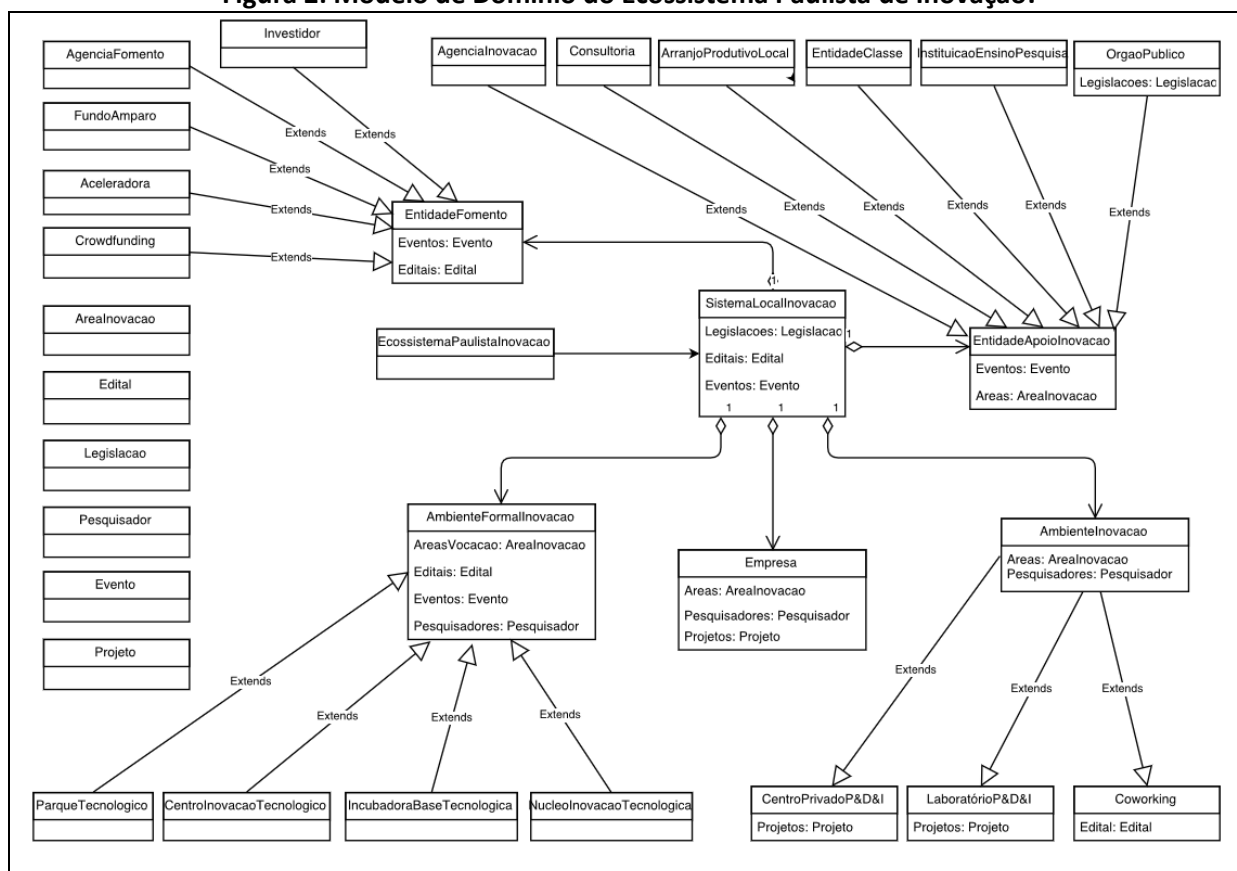
Fouro et. al (2001) consideram Modelagem de Domínio e Ontologias como etapas distintas e complementares que possibilitam a reutilização e que têm como um dos diversos fatores para sua criação, a redução dos custos de desenvolvimento, como consequência do

compartilhamento de custos entre as diversas aplicações de uma mesma área ou domínio. Essas considerações indicam a viabilidade em utilizar a Engenharia de Domínio para a conceituação do domínio de conhecimento do Ecossistema Paulista de Ambientes de Inovação que servirá de base para construção da ontologia que será utilizada para a geração de serviços informacionais na plataforma proposta.

Para ajudar a compreender o domínio de um ecossistema de inovação é útil identificar os principais atores, participantes e contribuintes no ecossistema de inovação. Isso ajudará a focar discussões posteriores em torno das interações e relacionamentos em jogo dentro do ecossistema e, crucialmente, o engajamento e a interação de empresas inovadoras no ecossistema de inovação (THOMAS, 2017).

Na Modelagem de Domínios, pode-se adotar a *Unified Modeling Language* (UML) em diversos campos do conhecimento (HURLBLUT, 1997). Foi utilizado o artefato Diagrama de Classes do UML para representar o domínio das classes do Ecossistema Paulista de Inovação (Figura 2).

Figura 2: Modelo de Domínio do Ecossistema Paulista de Inovação.



Fonte: Elaborado pelos autores.

XVIII ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2017
23 a 27 de outubro de 2017 – Marília – SP

A Figura 2 descreve um Diagrama de Classes do Ecossistema Paulista de Inovação que emergiu da reflexão em identificar os principais atores de inovação do Estado de São Paulo e suas relações. Esse diagrama servirá de base para a construção da ontologia.

O Ecossistema Paulista de Inovação é composto por Sistemas Locais de Inovação que se referenciam à capacidade das localidades de se organizarem de maneira formal com o objetivo de apoiar a inovação nas organizações (GERMANO, 2002).

No contexto dos atores de inovação, os Sistemas Locais de Inovação são articulações entre Entidades de Apoio à Inovação, Entidades de Fomento à Inovação, Ambientes de Inovação e Empresas.

Os ambientes de inovação exercem um papel primordial no apoio aos processos de inovação nas empresas, para Maciel (1997), o conceito de ambiente de inovação em estudos sociais sobre produção tecnológica:

[...] procura dar conta do conjunto de condições – limites, obstáculos, possibilidades, estímulos – da inovação em uma determinada formação social. Ambiente de inovação refere-se, portanto, ao conjunto de fatores políticos, econômicos, sociais e culturais que estimulam ou dificultam a inovação [...]. (MACIEL, 1997, p. 109).

Os Ambientes Formais de Inovação são conceituados desta forma neste artigo baseando-se no Decreto 60.286 de 25 de março de 2014 que instituiu e regulamentou o Sistema Paulista de Ambientes de Inovação que compreende os Parques Tecnológicos, as Incubadoras de Empresas de Base Tecnológica, os Centros de Inovação Tecnológica e os Núcleos de Inovação Tecnológica.

Entre as Entidades de Apoio à Inovação, destacam-se as organizações públicas municipais, estaduais e federais, que têm a responsabilidade de promover políticas públicas de incentivo à inovação.

As Entidades de Fomento têm o objetivo de financiar projetos de inovação nas organizações.

Os Ambientes de Inovação são iniciativas privadas ou públicas de incentivo à geração de projetos de inovação. O foco de um Ecossistema de Inovação é apoiar a cultura da inovação nas empresas de modo que esta possa desenvolver novos produtos, serviços e modelos de negócio com vistas ao desenvolvimento sustentável. Neste contexto, num Ecossistema de Inovação, as empresas são beneficiadas por meio de projetos de inovação ou por Centros

Privados de P&D&I vinculados às empresas ou em Ambientes Formais de Inovação como Parques Tecnológicos ou Centros de Inovação Tecnológica.

6 ONTOLOGIA DO ECOSISTEMA DE INOVAÇÃO DO ESTADO DE SÃO PAULO: INOVAONTO

Nesta pesquisa utilizou-se a ontologia do tipo conceitual, que abrange as classes e as principais relações do modelo de domínio apresentado anteriormente. Utilizou-se a metodologia de Noy e McGuinness (2001) para a construção da ontologia que indica os processos que devem ser seguidos para o desenvolvimento de uma ontologia, desde a sua concepção original, como uma ontologia conceitual, até a implementação desta, com a inserção de instâncias das classes.

Desta forma, esta pesquisa abrange somente a construção da ontologia conceitual, realizando os primeiros quatro passos da metodologia de Noy e McGuinness (2001).

Segundo esta metodologia, o primeiro passo na construção da ontologia é a definição do domínio e escopo. O domínio foi definido pelo Ecosistema de Inovação do Estado de São Paulo, tendo como escopo os sistemas locais de inovação, os ambientes formais de inovação definidos pelo Sistema Paulista de Ambientes de Inovação, ambientes não formais de inovação, as entidades de apoio e fomento e, por fim, as empresas.

O próximo passo trata da utilização ou não de uma ontologia já existente. As autoras entendem que quase sempre vale a pena considerar o que já foi feito e, verificar se é possível refinar e, então, estender as fontes existentes para o domínio particular.

Considerando a reutilização de ontologias existentes, têm-se alguns trabalhos na literatura com proximidade à temática da presente pesquisa. No Quadro 1, são apresentados os trabalhos encontrados.

Quadro 1: Ontologias de representação de ideias de inovação

<i>Iteams Ontology</i> (Ning & al., 2006)	Usado para facilitar a extração, organização e criação de ideias de inovação, concentrando-se em recursos que são centrados nas pessoas.
<i>OntoGate Ontology</i> (Bullinger, 2008)	Apresentada como uma Ontologia de Domínio cujo propósito é descrever a relação das empresas com os processos de inovação. Trata-se de avaliação e seleção de ideias.
<i>Idea Ontology</i> (Riedl & al., 2009)	Ontologia de Aplicação que provê uma linguagem comum para a disseminação de ideias, mas não fornece um modelo para a representação real de uma ideia.
<i>GIZMO Ontology</i> (Westerski, 2012)	Representa uma Ontologia de Domínio que permite o compartilhamento de informações entre sistemas computacionais por meio de tecnologias semânticas.

Fonte: Elaborado pelos autores.

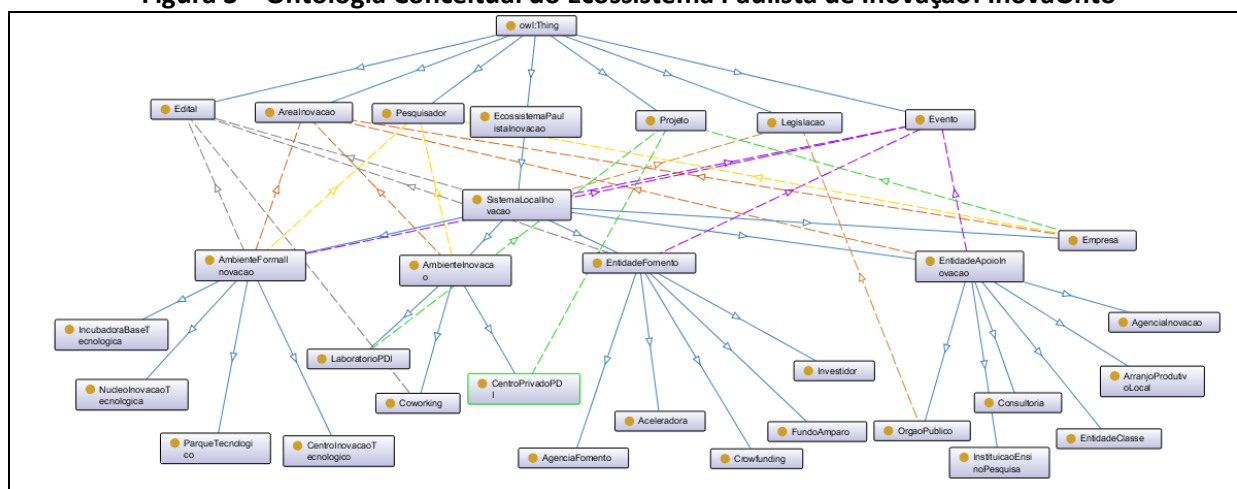
Os trabalhos mencionados tratam do domínio das ideias de inovação, não foram encontradas na literatura ontologias que representassem o domínio de ecossistemas de inovação.

O passo seguinte da metodologia refere-se a identificação de termos importantes no domínio. Baseado nos ambientes informacionais digitais dos principais atores de inovação que abrangem o domínio, identificaram-se os termos correspondentes ao objetivo de apoiar processos de inovação.

E, por fim, o último passo trata da definição de classes e suas relações. Neste caso, foi utilizado o diagrama de classes gerado pelo modelo de domínio apresentado na Figura 2.

A Figura 3 apresenta a ontologia conceitual do Ecosistema Paulista de Inovação que representa as informações dos atores de inovação como: organizações, entidades, pesquisadores, ambientes formais e não formais de inovação que formam os sistemas locais de inovação e geram ações de apoio à inovação nas empresas como eventos, consultorias, editais, leis de incentivo e fomento a projetos de inovação.

Figura 3 – Ontologia Conceitual do Ecosistema Paulista de Inovação: InovaOnto



Fonte: Elaborado pelos autores.

A ontologia foi construída utilizando a linguagem OWL, que é embasada pelos padrões RDF e RDF Schema. Estes padrões fornecem bases para a estruturação dos dados, que permite a disponibilização de informações em formatos de triplas, que em síntese relaciona as informações na forma sujeito-predicado-sujeito, sendo que um sujeito está relacionado a um objeto por meio de um predicado (FERREIRA e SANTOS, 2013).

O RDF e o RDF Schema não fornecem subsídios para descrever um domínio seguindo os princípios de uma ontologia. Desta forma, o OWL é utilizado visando ter a expressividade e

as características que uma ontologia exige, tendo o objetivo de descrever um determinado domínio com propriedades capazes de expressar a semântica inerente a este domínio.

Neste contexto, o OWL apresenta três tipos principais de entidades de dados, as classes, as propriedades e as instâncias. As classes são representações abstratas que descrevem as entidades pertencentes a um domínio. As instâncias são a concretização das classes, sendo os exemplos implantados de uma determinada classe. E as propriedades são as características das classes, possibilitando a relação com outras classes, a inserção de atributos nas próprias classes, além de propriedades especiais que inserem axiomas na descrição de um domínio, permitindo uma posterior inferência.

Um fragmento do OWL gerado está representado na Figura 4.

Figura 4: Fragmento do OWL da Ontologia InovaOnto

```
--<owl:Class rdf:about="http://www.inovamarilia.com.br/ontologia#AmbienteInovacao">
  <rdfs:subClassOf rdf:resource="http://www.inovamarilia.com.br/ontologia#SistemaLocalInovacao"/>
</owl:Class>
-<!--
  http://www.inovamarilia.com.br/ontologia#AreaInovacao
-->
<owl:Class rdf:about="http://www.inovamarilia.com.br/ontologia#AreaInovacao"/>
-<!--
  http://www.inovamarilia.com.br/ontologia#ArranjoProdutivoLocal
-->
<owl:Class rdf:about="http://www.inovamarilia.com.br/ontologia#ArranjoProdutivoLocal">
  <rdfs:subClassOf rdf:resource="http://www.inovamarilia.com.br/ontologia#EntidadeApoioInovacao"/>
</owl:Class>
-<!--
  http://www.inovamarilia.com.br/ontologia#CentroInovacaoTecnologico
-->
<owl:Class rdf:about="http://www.inovamarilia.com.br/ontologia#CentroInovacaoTecnologico">
  <rdfs:subClassOf rdf:resource="http://www.inovamarilia.com.br/ontologia#AmbienteFormalInovacao"/>
</owl:Class>
```

Fonte: Elaborado pelos autores.

A representação das classes é obtida a partir da construção da ontologia na linguagem OWL. O fragmento do OWL apresentando na Figura 4 contempla a definição das classes e subclasses. No entanto, as demais definições são realizadas no restante deste OWL. Como citado anteriormente, as instâncias não foram apresentadas, por se tratar de uma ontologia de domínio. Para as instâncias, será utilizado um modelo de dados em conjunto com um banco de dados NoSQL, que instanciará a ontologia com seus dados.

7 CONSIDERAÇÕES FINAIS

Nos últimos anos inúmeras ações foram tomadas para criar uma cultura de inovação e empreendedorismo no país, chamando a atenção da sociedade para a importância do empreendedorismo e da inovação. Surgiram vários mecanismos de geração de empreendimentos inovadores no Brasil, chamados de Ambientes Formais de Inovação, as

XVIII ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2017
23 a 27 de outubro de 2017 – Marília – SP

incubadoras e posteriormente os centros de inovação, núcleos de inovação e parques tecnológicos. É a partir deste cenário que esta pesquisa se construiu.

Esta pesquisa atingiu os objetivos propostos ao propor uma Plataforma Informacional do Ecosistema Paulista de Inovação que, por meio de tecnologias informacionais, semânticas e computacionais, agregue e disponibilize serviços de informação para os atores de inovação do Estado de São Paulo. O objetivo desta plataforma é possibilitar maior articulação e apoiar os processos de construção coletiva e colaborativa entre os principais atores de inovação do Estado de São Paulo, a partir de fontes de informação disponibilizadas pelos mesmos. Esperamos que a partir desta pesquisa, outros estudos possam ser desenvolvidos para potencializar os ambientes de inovação paulista e do Brasil de modo geral.

Por meio da ontologia proposta, agentes computacionais de extração e serviços podem ser implementados para incorporar o caráter dinâmico que os espaços informacionais exigem. A implementação consiste na integração dos agentes de extração com a ontologia, ou seja, a comunicação das informações que são extraídas, com o intuito de atribuir semântica à busca. Desta maneira, o agente extrai conteúdos das fontes informacionais dos atores inovação e um algoritmo irá avaliar se aquela informação está dentro do contexto da ontologia e se aquela informação de fato será útil aos atores do ecossistema de inovação.

Para validar o modelo, robôs de extração foram desenvolvidos com auxílio da biblioteca Java Jsoup, capazes de realizar a leitura das informações semi-estruturadas encontradas e retiradas dos ambientes informacionais digitais na Web que foram modeladas por meio de triplas RDF, com auxílio do Framework Jena.

A partir das informações extraídas nos ambientes informacionais digitais dos atores de inovação do ecossistema identificado, serviços computacionais informacionais podem ser gerados de acordo com demandas dos processos de inovação nas organizações por meio de tecnologias computacionais como SOA (*Service-Oriented Architecture*) e API (*Application Programming Interface*) que disponibilizam produtos informacionais num ambiente de interoperabilidade com outras plataformas computacionais.

A identificação das principais entidades e organizações e suas relações envolvidas no processo de inovação do estado de São Paulo definidas como o Ecosistema Paulista de Inovação e a utilização da combinação das tecnologias computacionais, informacionais e semânticas contidas na plataforma proposta pretende que os principais atores do domínio da inovação (governo, empresas e universidade) possam ter acesso aos produtos de informação

XVIII ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2017
23 a 27 de outubro de 2017 – Marília – SP

gerados e colaborativamente apoiem-se em processos de construção coletiva e colaborativa de novos conhecimentos e em ações de incentivo à inovação nas organizações.

REFERÊNCIAS

ÁLVAREZ, A. C. **Extração de informação de artigos científicos: uma abordagem baseada em indução de regras de etiquetagem.** 2007. Tese de Doutorado. Universidade de São Paulo.

ANPROTEC. **Criado o Sistema Paulista de Ambientes de Inovação.** 2014. Disponível em: <http://anprotec.org.br/site/2014/04/criado-o-sistema-paulista-de-ambientes-de-inovacao>. Acessado em 20 de julho de 2017.

BULLINGER, A., **Innovation and Ontologies: Structuring the Early Stages of Innovation Management,** Springer, 2008.

CNI. **Pesquisa sobre inovação com 100 líderes empresariais,** 2015. Disponível em: <http://www.portaldaindustria.com.br/publicacoes/2015/05/pesquisa-sobre-inovacao-com-100-lideres-empresariais>. Acessado em 20 de julho de 2017.

DETERS, J. I., E ADAIME, S. F. Um estudo comparativo dos sistemas de busca na web. Anais do V Encontro de Estudantes de Informática do Tocantins. Palmas, TO. 189-200, 2003.

FERREIRA, J. A.; SANTOS, P. L. V. A. C. O modelo de dados resource description framework (RDF) e o seu papel na descrição de recursos. **Informação & Sociedade.** João Pessoa, v. 23, n. 2, p. 13-23, maio/ago. 2013. Disponível em: <<http://www.ies.ufpb.br/ojs/index.php/ies/article/view/15436/9681>> Acesso em: 03 ago. 2017.

FOURO, A. M. M., WERNER, C. M. L. Modelos de domínios ou Ontologias? **RTInfo - Revista Tecnologia da Informação.** 2001.

FRANKLIN, S.; GRAESSER, A. Is it an agent or just a program? a taxonomy from autonomous agents. In: **Third International Workshop On Agent Theories, Architectures And Languages: Intelligent Agents III.** [s.n.], p.21–36, 1996.

GENESERETH, M. R.; KETCHPEL, S. P. Software agents. **Communications of the ACM,** [S.l.], v.38, n.7, p.48–53, 1994.

GERMANO, D. A. **Sistemas locais de inovacao: estudo de estratégias de planejamento regional (parques e pólos tecnológicos intra-urbanos).** Mestrado (Dissertação) Universidade Federal do Rio Grande do Sul. Programa de Pesquisa e Pós-Graduação em Planejamento Urbano e Regional. Porto Alegre, 2002. Disponível em <http://www.lume.ufrgs.br/bitstream/handle/10183/3283/000335288.pdf>.

GRUBER, T. R. A translation approach to portable ontology specifications. **Knowledge acquisition** 5.2. 199-220. 1993.

XVIII ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2017
23 a 27 de outubro de 2017 – Marília – SP

JACKSON, DJ. **What is an Innovation Ecosystem?** National Science Foundation, Arlington, VA, 2011.

HURLBLUT, R. **A survey on approaches for describing and formalizing use-cases.** 1997. Disponível em <http://www.iit.edu/~rhurlbut/xpt-tr-97-03.html>, acessado em 30 de maio de 2001.

MACIEL, M. L. Inovação e conhecimento. In: SOBRAL, F. et al. (Orgs.). **A alavanca de Arquimedes: ciência e tecnologia na virada do século.** Brasília: Paralelo 15, 1997.

NING, K. et al., Semantic innovation management across the extended enterprise. **International Journal of Industrial and Systems Engineering**, 1(1-2), p.109–128, 2006.

NOY, F. N.; McGUINNESS, D. L. **Ontology development 101: a guide to create your first ontology.** 2001. Disponível em: . Acesso em: 03 Outubro 2006.

OLIVEIRA, A. B. F., WERNECK, V.M.B., LANZILOTTI, R.S., SERRÃO, H., SOARES, E.A., PORTELLA, E. S., AVILA, S. S. Ontologia de Domínio da Biodisponibilidade de Ferro: Uma Experiência no Projeto Nutri-Fuzzy-Orixás. In: **Anais do III Workshop de Informática Médica.** Fortaleza. v.1, 2003.

RIEDL, C. and MAY, N., FINZEN, J., STATHEL, S., KAUFMAN, V., KRCCMAR, H., An Idea Ontology for Innovation Management. **International Journal on Semantic Web and Information Systems**, Vol. 5, No. 4, pp. 1-18, 2009.

SANTOS, C.D.; VALENTIM, M. L. P. As interconexões entre a gestão da informação e a gestão do conhecimento para o gerenciamento dos fluxos informacionais. **Perspectivas em Gestão & Conhecimento.** João Pessoa, ISSN: 2236-417X ,v. 4, n. 2, p. 19-33, jul./dez. 2014.

SILVA, T. de M. S. **Extração de informação para busca semântica na web baseada em ontologias.** Mestrado (Dissertação). Universidade Federal de Santa Catarina. Programa de Pós-Graduação em Engenharia Elétrica, 2003.

THOMAS, M. **Innovation ecosystems as drivers of regional innovation - validating the ecosystem.** 2017. Disponível em <http://www.know-hub.eu/knowledge-base/videos/innovation-ecosystems-as-drivers-of-regional-innovation-validating-the-ecosystem.html#footnote1>. Acesso em 29 de julho de 2017.

WESTERSKI,A., **Gi2MO Ontology Specification.** 2012. Disponível em: <http://www.gi2mo.org/ontology>. Acesso em 20 de abril de 2017.

WOOLDRIDGE, M. Agent-based software engineering. **Proceedings...** London, United Kingdom: [s.n.], 1997.

WORLD BANK, **Indicadores, Pesquisa e desenvolvimento (em % do PIB),** 2015. disponível em <http://data.worldbank.org/indicator/GB.XPD.RSDV.GD.ZS>. Acesso em 21 abril de 2017.