

ObserveUnB: Observatório da Rede de Colaboração Científica da Universidade de Brasília

ObserveUnB: Observatory of the Scientific Collaboration Network of the University of Brasilia

ObserveUnB: Observatorio de la Red de Colaboración Científica de la Universidad de Brasilia

Célia Ghedini Ralha, PhD em Ciência da Computação pela *University of Leeds*, Inglaterra. Endereço: Campus Darcy Ribeiro, s/n. Caixa Postal 4466. CEP: 70904-970 – Brasília, DF. Telefone: (61) 3107-3675. E-mail: ghedini@cic.unb.br.

Denise Bomtempo Birche de Carvalho, doutora em Ciências Sociais e Sociologia pela *Université Paris 1, Panthéon-Sorbonne*, França. Endereço: Campus Universitário Darcy Ribeiro, gleba A, prédio da Reitoria – Asa Norte. CEP: 70910-970 – Brasília, DF. Telefone: (61) 3107-0264/0263. E-mail: denisebomtempo@terra.com.br.

Gabriela Netto Guerra, mestre em Ciência da Computação pela Universidade de Brasília. Endereço: SQN 214, bl. J, apto. 109 – Asa Norte. CEP: 70873-100 – Brasília, DF. Telefone: (61) 8493-9947. E-mail: gabrielag.guerra@gmail.com.

Resumo

Este artigo apresenta a definição e o desenvolvimento do ObserveUnB, um observatório da rede de colaboração científica e acadêmica na Universidade de Brasília (UnB). O foco do ObserveUnB é representar a rede de colaboração científica e potencializar sua ampliação por meio da descoberta de potenciais parcerias. O ObserveUnB utiliza informações existentes nos Currículos Lattes dos pesquisadores da UnB, nos grupos de pesquisa certificados pela

instituição no Diretório de Grupos de Pesquisa do CNPq e nas bases de dados dos sistemas acadêmicos e de pessoal. O projeto do ObserveUnB contempla ainda a extensão da ontologia *Conscientias* (base do Lattes) e a definição de um modelo de reputação dos participantes da rede.

Palavras-chave: Observatório. Colaboração Científica. Programa de Pós-Graduação. Ontologia. Modelo de Reputação. ObserveUnB.

Abstract

This article presents the definition and implementation of ObserveUnB, an observatory of the scientific and academic collaboration network in the University of Brasilia (UnB). ObserveUnB focuses on representing the network and aims to assist the expansion of scientific collaboration through partnership discovery. The ObserveUnB uses information from the Lattes curriculum of UnB researchers, from the research groups certified by UnB listed in the Directory of Research Groups of the CNPq, and from institutional databases. The ObserveUnB project also includes an extension of the Conscientias ontology (a Lattes base) and the definition of a reputational model for network participants.

Keywords: Observatory. Scientific Collaboration. Graduate Program. Ontology. Reputational Model. ObserveUnB.

Resumen

En este artículo se presentan la definición y el desarrollo del ObserveUnB, un observatorio de la red de colaboración científica y académica en la Universidad de Brasilia (UnB). El foco del ObserveUnB es representar y ayudar en la ampliación de la red de colaboración científica a través del descubrimiento de posibles asociaciones. El ObserveUnB utiliza información de los Currículos Lattes de los investigadores de la UnB, de los grupos de investigación certificados por la institución en el Directorio de Grupos de Investigación del CNPq y de las bases de datos

de los sistemas académicos y de personal. El proyecto del ObserveUnB también contempla una extensión de la ontología Conscientias (base del Lattes) y la definición de un modelo de reputación de los miembros de la red.

Palabras clave: Observatorio. Colaboración Científica. Programa de Posgrado. Ontología. Modelo de Reputación. ObserveUnB.

Introdução

Novos desafios estão postos à educação e à pós-graduação, bem como às atividades de pesquisa, desenvolvimento e prestação de serviços científicos e tecnológicos no mundo globalizado. Nesse cenário competitivo, os atores percebem a necessidade da formação de redes de colaboração de pesquisa e pós-graduação para fazer frente aos desafios globais de inovação e demais formas de apropriação do conhecimento científico e tecnológico.

Os desafios de formação de redes de colaboração são agravados pela crescente complexidade do desenvolvimento científico e tecnológico, demandando uma organização em rede com equipes maiores, mais completas, com competências mais diversificadas para o desenvolvimento de pesquisas mais complexas e em evolução. Aliada a essa complexidade vem a crescente necessidade de desenvolvimento de projetos realizados por equipes geograficamente dispersas, vinculados a diferentes grupos de pesquisa e programas de pós-graduação das diversas instituições de ensino e pesquisa do País e do mundo.

No caso brasileiro, o cenário de formação de redes de colaboração científica e tecnológica é composto por uma extensa comunidade, principalmente considerando os membros da comunidade acadêmica vinculados às universidades públicas e federais. Associado ao grande volume de informação dos membros da comunidade, existe o fato de os dados encontrarem-se distribuídos de forma heterogênea em diferentes bases institucionais, dificultando a identificação dos pares no importante processo de formação de redes.

Como um exemplo inserido no cenário nacional, podemos citar a comunidade acadêmica da Universidade de Brasília (UnB), composta por mais de 30 mil integrantes registrados nas bases de dados dos sistemas acadêmicos institucionais. Além dos registros internos, existem os externos relativos às bases de dados nacionais, por exemplo, a base de Currículo Lattes e de Diretório de Grupos de Pesquisa do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), bem como os dados dos programas de pós-graduação da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes). Portanto, diante do grande volume de informações distribuídas em diversos repositórios, surge a necessidade de definir instrumentos que auxiliem a comunidade acadêmica científica no processo de identificação de parcerias para a formação de redes de colaboração, bem como facilitem o processo de gestão de informações acadêmicas e científicas.

Conforme cenário descrito, neste trabalho apresentamos um instrumento útil para a identificação de parcerias em redes de colaboração da comunidade, que pretende auxiliar no processo de gestão institucional da informação científica, tecnológica e artística. O ObserveUnB¹, apresentado neste artigo, deve ser entendido como um observatório da rede de colaboração científica da UnB, pois disponibiliza informações dos atores da comunidade em arranjos de pessoas e conexões entre os integrantes da rede acadêmica. Acreditamos que a disponibilização da informação por meio do conceito de rede, em que as pessoas ou os grupos são os atores (nós da rede) e as vinculações acadêmicas, bem como os projetos de pesquisa e programas de pós-graduação, são as conexões, seja intuitiva e enseje a formação de equipes maiores, aptas a enfrentar os desafios de complexidade crescente do desenvolvimento científico e tecnológico globalizado.

Para viabilizar a identificação dos atores na rede, o ObserveUnB integra dados dos diversos repositórios internos e externos à instituição em uma base única. No momento, estão sendo utilizadas as informações dos Currículos Lattes dos docentes disponíveis na base do CNPq, os dados dos programas de pós-graduação da UnB mantidos pela Capes e os dados obtidos de bases institucionais.

¹ <http://observeunb.unb.br/>.

O ObserveUnB utiliza uma ontologia de aplicação, que foi estendida da ontologia de Currículo Lattes do Conscientias (CONSCIENTIAS, 2011). O uso da ontologia, além de permitir a organização do domínio de conhecimento de forma adequada, auxilia na recuperação da informação semântica, uma vez que representa formalmente o conhecimento em lógica descritiva, um fragmento decidível de lógica de primeira ordem (BAADER; NUTT, 2010). O uso de ontologias se mostra eficiente na definição de um vocabulário comum em que o conhecimento a ser compartilhado pode ser representado (JARDIM; PALAZZO, 2009). Além da ontologia, a proposta apresenta um modelo de reputação aplicado aos membros da rede, que é baseado na produção científica e acadêmica descrita no Currículo Lattes dos membros (CERVIL; GALANTE; OLIVEIRA, 2011). Espera-se que a reputação possa auxiliar na busca por potenciais parcerias na ampliação das redes de colaboração de pesquisa e pós-graduação.

Neste trabalho, buscaremos descrever, na seção 2, os conceitos relacionados à pesquisa, enquanto na seção 3 serão apresentadas algumas iniciativas de redes de colaboração científica. Para uma melhor compreensão do ObserveUnB, na seção 4 serão apresentados o projeto, a arquitetura e os principais módulos que compõem a ferramenta. Os resultados obtidos serão descritos na seção 5; e, finalmente, na seção 6, serão apresentadas as conclusões e os direcionamentos futuros.

Conceitos relacionados

Este projeto envolve diferentes conceitos que se interligam para formar um instrumento útil de identificação de redes de colaboração científica, os quais serão apresentados nesta seção. Vale ressaltar que foi adotada, neste trabalho, a premissa de que os relacionamentos existentes entre os conceitos da ontologia constituem a base para a formação da própria rede de colaboração, uma vez que representam formas de produção e interação entre os atores.

a. Rede de colaboração científica

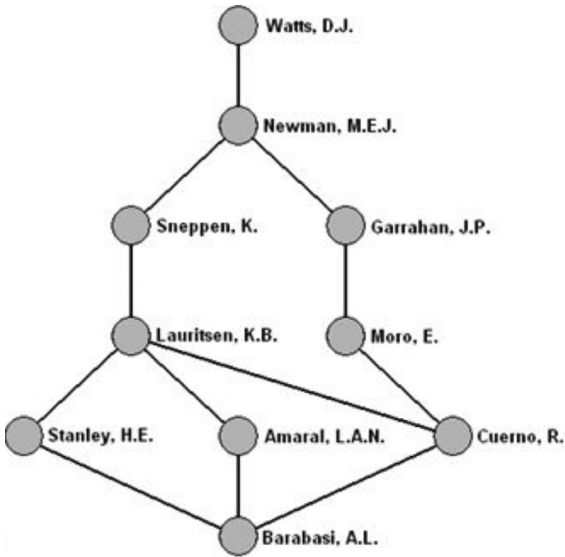
A hipótese que sustenta o ObserveUnB é o fato de a colaboração científica ser um empreendimento cooperativo que envolve metas

comuns, esforço coordenado e resultados ou produtos científicos com responsabilidade e mérito compartilhados. Assim, a colaboração científica oferece uma fonte de apoio para melhorar o resultado e maximizar o potencial da produção científica (BALANCIERI et al., 2005).

Segundo Katz e Martin (1997), a colaboração científica pode ser definida como o trabalho conjunto de pesquisadores para alcançar um objetivo comum de produzir novos conhecimentos científicos. Para Castells (1999), rede pode ser definida como uma estrutura aberta capaz de se expandir de forma ilimitada, integrando novos nós, desde que consigam comunicar-se dentro da rede, ou seja, desde que compartilhem os mesmos códigos de comunicação. Dessa forma, o conceito de rede de colaboração científica adotado neste artigo compreende uma comunidade de atores que se relacionam de forma a permitir o compartilhamento de conhecimento e aumentar a produção científica e tecnológica resultante.

Uma rede científica pode ser representada por meio de um grafo. Segundo Cormen et al. (2002), um grafo é um conjunto de pontos, chamados vértices, conectados por linhas, chamadas de arestas. No caso de uma rede de colaboração científica, o vértice pode ser chamado de ator ou entidade social. São exemplos de vértices: pessoa, organização, grupo, conceito. Já a aresta, no contexto social, pode ser chamada de relação, que representa um vínculo relacional de um tipo específico entre dois vértices (atores) da rede, e.g., coautoria em produção científica.

A coautoria tem sido usada por diversos pesquisadores como medida da colaboração científica entre grupos de pesquisadores ou instituições. A colaboração científica entre atores ou instituições supõe um compartilhamento de ideias e objetivos de um projeto. Para isso, é necessário um clima de confiança, o estabelecimento de uma divisão de trabalho e a interação entre os pesquisadores, além do compartilhamento de informações (FRANCINA, 2008). A Figura 1 apresenta uma rede em que os atores são pesquisadores e as ligações são coautoria em produção científica.



Fonte: Newman (2000).

Figura 1. Exemplo de uma rede de colaboração científica

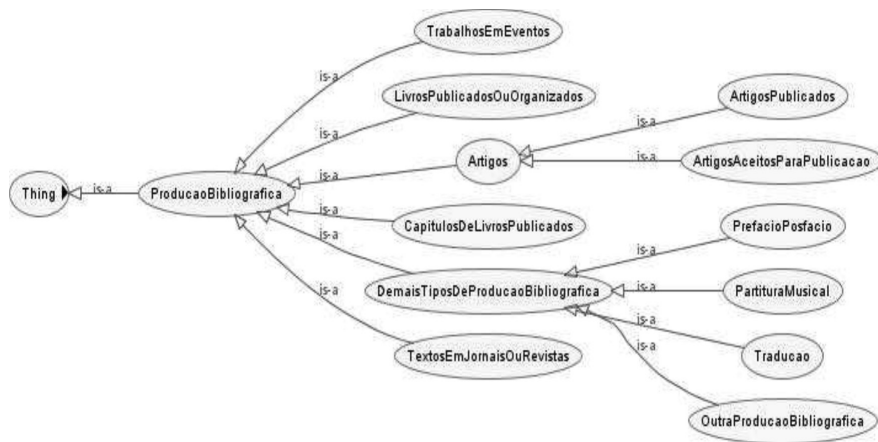
b. Ontologia

Entre as várias referências ao termo ontologia existentes na literatura, citamos neste trabalho de Gruber (1996), que define ontologia como uma especificação formal e explícita de uma conceitualização compartilhada. Para melhor compreender essa definição, vamos detalhar cada termo que a compõe. Especificação formal diz respeito à representação formal passível de ser processada por computadores. Os conceitos, as propriedades e relações são definidos de forma explícita. A conceitualização compartilhada pode ser entendida como consensual entre um grupo de especialistas em determinado domínio de conhecimento. Em resumo, seria o conhecimento de um domínio aceito por um grupo de pessoas.

Outra definição comumente citada na literatura é a de Guarino (1998), que define ontologia como sendo um artefato computacional, baseado em um vocabulário formal, utilizado para descrever uma conceitualização particular comumente aceita do mundo. De forma prática, diz-se que uma ontologia engloba um conjunto de termos, conceitos e seus relacionamentos semânticos. Esses relacionamentos

são criados por especialistas no domínio do conhecimento abordado pela ontologia, em que a estrutura ontológica definida será utilizada, por exemplo, pelos usuários durante a formulação de uma consulta ou por linguagens computacionais dentro da web semântica.

Graficamente, as ontologias se assemelham a árvores de conceitos, em que cada folha é um termo específico relacionado a um galho, que é um conceito mais geral. Sendo assim, dependendo do domínio de conhecimento, uma mesma entidade pode ser caracterizada de diferentes maneiras (ROBREDO; BRÄSCHER, 2010). Na Figura 2, é apresentada uma pequena amostra das classes presentes na ontologia do Grupo *Conscientias* utilizada neste trabalho (CONSCIENTIAS, 2011). Note que a classe produção bibliográfica e suas subclasses foram criadas para contextualizar os possíveis itens de produção, e.g., trabalhos em eventos científicos (TrabalhosEmEventos), livros publicados ou organizados (LivrosPublicadosOuOrganizados) e artigos (Artigos). Note que a classe produção bibliográfica e suas subclasses foram criadas para contextualizar os possíveis itens de produção, e.g., trabalhos em eventos científicos (TrabalhosEmEventos), livros publicados ou organizados (LivrosPublicadosOuOrganizados) e artigos (Artigos).



Fonte: Conscientias (2011).

Figura 2. Exemplos de classes da ontologia do Grupo Conscientias

c. Reputação

Sabater e Sierra (2001) definem reputação como uma opinião ou visão de alguém sobre algo. Os autores apresentam três noções de reputação: individual, social e ontológica. Reputação individual

refere-se a como as impressões de um único indivíduo são julgadas por outros. Reputação social refere-se à impressão sobre indivíduos baseada na reputação do grupo social a que pertencem. Reputação ontológica depende do contexto em que se encontra a reputação; dessa forma, a reputação de um indivíduo pode ter diferentes classificações em comunidades distintas.

Sistemas com modelo de reputação têm se tornado um componente principal em diversos ambientes eletrônicos. Seja em sites de comércio eletrônico, sites de redes sociais ou fóruns de mensagens, a reputação desempenha um papel importante na garantia de qualidade das interações entre os usuários. Pujol, Sangüesa e Delgado (2002) declaram que a medida de reputação é bastante aplicada em comunidades criadas com o propósito do gerenciamento do conhecimento, uma vez que envolve o intercâmbio de informação e conhecimento entre membros da comunidade. O cálculo da reputação dá uma ideia da confiança que um membro pode ter em outro e serve para guiar a procura por pares.

Diferentes técnicas de cálculo de reputação têm sido implementadas, com resultados diversos. Talvez a técnica mais simples seja calcular a média aritmética ponderada dos pesos entre os nós de uma rede, como é usado pelas comunidades eletrônicas, e.g, *eBay*², *Amazon*³ e outras grandes empresas on-line.

Segundo Pujol, Sangüesa e Delgado (2002), o mecanismo usual de medida de reputação deve basear-se no fornecimento de informações pelos usuários ou membros da própria rede de colaboração. Essas informações indicam a autoridade ou reputação que um usuário pode ter, podendo ser propagada de um membro para outro, dependendo da autoridade dos nós relacionados a ele. A desvantagem desse mecanismo de mensuração é a necessidade do envolvimento explícito e frequente dos membros da rede.

Um método alternativo de medir a reputação baseia-se na localização de cada membro na sua comunidade de rede social, que poderá ser usada para inferir alguma propriedade sobre seu grau de

² <http://www.ebay.com/>.

³ <http://www.amazon.com/>.

experiência, por meio da respectiva reputação. Membros que são bem conhecidos e respeitados pela maioria dos membros de sua comunidade (seus pares) tendem a ser facilmente identificados como os nós mais conectados no grafo de rede social da comunidade. Alguns algoritmos usados na classificação de páginas na web baseiam-se nesse método de reputação. Como exemplo, podemos citar o *HITS* (KLEINBERG, 1999), o *PageRanking* (PAGE et al., 1999) e o *NodeRanking* (PUJOL; SANGÜESA; DELGADO, 2002).

Trabalhos correlatos

Existem vários trabalhos de pesquisa no tema de redes científicas e tecnológicas que apresentam soluções integrando pesquisadores e suas respectivas produções científicas disponibilizadas na web. No entanto, não temos conhecimento de outra iniciativa de construção de rede de colaboração científica que utilize associadamente um modelo de reputação e recursos semânticos para auxiliar a gestão de informações científicas no âmbito de comunidades acadêmicas.

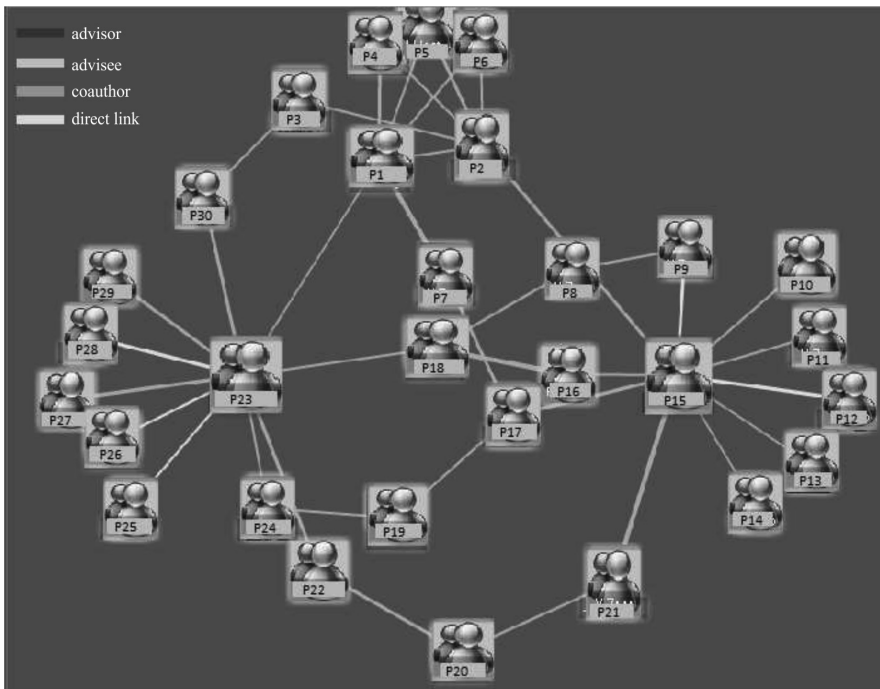
Atualmente, o *Google Scholar*⁴ é uma das mais poderosas ferramentas de colaboração científica, uma vez que permite que pesquisadores compartilhem publicações científicas de forma rápida e segura, além de ser utilizada para medir a produção individual de pesquisadores. Algumas ferramentas automatizadas utilizam as informações disponíveis no *Google Scholar* para calcular a produção científica de pesquisadores. Um aspecto que faz com que essa ferramenta tenha uma grande popularidade é o crescimento exponencial de sua base de referências, que vem sendo mantido pela inserção de novas publicações na base – o critério para essas inserções é ter sido referenciado por outro artigo já indexado (LABBÉ, 2009).

A ferramenta *ArnetMiner*⁵, desenvolvida por Tang et al. (2008), fornece serviços para busca e extração de informações semânticas baseadas na comunidade de pesquisadores na web. Para isso, a ferramenta estende a ontologia FOAF (BRICKLEY; MILLER, 2011) para unificar os perfis dos pesquisadores e permitir a pesquisa de

⁴ <http://scholar.google.com/>.

⁵ <http://arnetminer.org/>.

especialistas por tópicos ou palavras-chave. Esse trabalho utiliza informações de páginas pessoais dos pesquisadores e de diferentes bibliotecas digitais, como DBLP⁶, *CiteSeer*⁷ e ACM⁸. A Figura 3 exibe a rede de colaboração formada pelos pesquisadores fictícios P23 e P15, em que os tipos de relacionamento se distinguem pelas cores, podendo ser do tipo orientador, orientando ou coautor. Vale ressaltar que o ArnetMiner tem recursos semânticos pelo uso da ontologia FOAF, bem como pontua os atores da rede conforme produção disponível na Web; no entanto, não se trata de uma iniciativa no âmbito de uma comunidade acadêmica que auxilie como instrumento no processo de gestão da informação acadêmica e científica institucional.



Fonte: Adaptado de ArnetMiner.

Figura 3. Rede de colaboração de dois pesquisadores

Outro sistema de colaboração científica disponível é o *Microsoft Academic Search* (MAS), uma ferramenta *open-source*, desenvolvida pela *Microsoft Research*⁹, que consiste em uma organização destinada à realização de pesquisas básica e aplicada em Ciência da Computação e Engenharia de Software. O MAS disponibiliza consultas sobre

⁶ <http://dblp.uni-trier.de>.

⁷ <http://citeseer.ist.psu.edu>.

⁸ <http://dl.acm.org/>.

⁹ <http://academic.research.microsoft.com>.

publicações, autores, conferências, periódicos, instituições e palavras-chave mediante um mecanismo de busca que coleta informações na web de instituições de ensino, estudantes e bibliotecas e as armazena em um banco de dados próprio. Esses dados podem ser visualizados em diferentes formatos, como rede de relacionamento, gráficos de publicações de pesquisadores e mapas acadêmicos georeferenciados. A Figura 4 apresenta um mapa de parte da América do Sul, destacando o Brasil. Note que o tamanho da circunferência no mapa é proporcional à quantidade de publicações; nesse caso, a maior circunferência representa a Universidade de São Paulo, com 119.572 publicações até o ano de 2011.



Fonte: Microsoft Research.

Figura 4. Mapa acadêmico do Brasil no Microsoft Academic Search

ObserveUnB

O ObserveUnB é um projeto idealizado e coordenado pelo Decanato de Pesquisa de Pós-Graduação (DPP) da UnB. Foi implementado por equipe técnica vinculada ao Programa de Pós-

Graduação de Informática e ao Curso de Graduação em Ciência da Computação, ambos do Departamento de Ciência da Computação (CIC/UnB), bem como ao Laboratório do Ambiente Construído, Inclusão e Sustentabilidade (Lacis) da Faculdade de Arquitetura e Urbanismo (FAU/UnB) (RALHA et al., 2012). O projeto foi iniciado em março de 2011 e teve sua primeira versão disponibilizada na web com o lançamento institucional em 28/11/2011 (OBSERVEUNB, 2011b).

A proposta da ferramenta é a de um observatório da rede de colaboração científica e tecnológica da UnB, com foco nos relacionamentos de vinculação institucional, programas de pós-graduação e na produção científica dos membros vinculados à instituição. O ObserveUnB permite que atores encontrem potenciais parcerias para colaboração por meio dos relacionamentos visualizados na rede e da reputação dos membros. Dessa forma, o ObserveUnB tem como objetivo principal responder quatro perguntas de pesquisa relacionadas aos membros institucionais:

1. Quem são os pesquisadores?
2. O que os pesquisadores estão produzindo?
3. Onde os pesquisadores estão localizados?
4. Com quem os pesquisadores estão pesquisando?

a. Arquitetura

A arquitetura do ObserveUnB é apresentada na Figura 5. Note que os componentes arquiteturais estão numerados conforme os quatro estágios básicos que formam a estrutura de tratamento da informação: extração, integração, repositório e interface de visualização, que serão detalhados nas respectivas seções a.1 a a3.

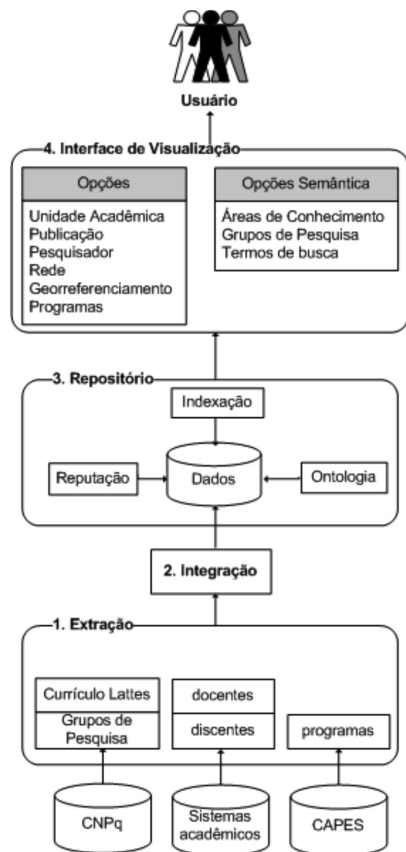


Figura 5. Arquitetura do ObserveUnB

a.1 Extração

No primeiro estágio de tratamento da informação, é realizado o processo de extração de dados. Os dados dos docentes utilizados e armazenados no sistema foram extraídos das bases de dados do Sistema de Pessoal (Sipes), sendo que os respectivos Currículos Lattes foram baixados da base do CNPq. Os dados dos programas de pós-graduação utilizados foram fornecidos pela Capes via seu aplicativo de coleta de dados. No futuro, está prevista a utilização das informações dos discentes regulares de graduação e pós-graduação disponíveis respectivamente nas bases de dados do Sistema de Informação Acadêmica da Graduação (Sigra) e do Sistema de Informação Acadêmica da Pós-Graduação (Sippos). Também está prevista a integração dos grupos de pesquisa certificados pela UnB disponíveis na base de dados do Diretório de Grupos de Pesquisa do CNPq.

Para executar a baixa dos Currículos Lattes do CNPq foi implementado um robô denominado Robo-CV (número registro INPI: 13256-0). O Robo-CV foi desenvolvido na linguagem Java integrado à linguagem XML (*Extensible Markup Language*) dos arquivos de currículo na base do CNPq. Em relação aos dados dos programas de pós-graduação, foram utilizados a listagem de docentes por programa informado via aplicativo de coleta de dados da Capes e os Currículos Lattes dos docentes já baixados pelo Robo-CV.

Para uma próxima versão, o mesmo procedimento de baixa de currículos com o Robo-CV será realizado com os Currículos Lattes dos discentes. A baixa dos grupos de pesquisa certificados pela UnB será operacionalizada por meio do desenvolvimento do Robo-grupos, que fará a baixa dos arquivos dos grupos certificados pela UnB e disponíveis no repositório do CNPq.

a.2. Integração e repositório

O conteúdo dos arquivos gerados durante o estágio de extração é filtrado por um módulo de integração desenvolvido na linguagem PHP e convertido em dados armazenados em um banco de dados MySQL. Também nesse estágio é integrado o banco de dados ao conteúdo da ontologia. Para viabilizar o tratamento semântico, o repositório faz uso de uma ontologia estendida de Diretório de Grupos de Pesquisa e de Currículo Lattes do Grupo *Conscientias* (CONSCIENTIAS, 2011).

O repositório único do ObserveUnB armazena também os dados de reputação dos integrantes da rede. O objetivo do modelo de reputação proposto é auxiliar na busca por potenciais parcerias na ampliação das redes de colaboração de pesquisa e pós-graduação. O modelo proposto foi adaptado de Cervil, Galante e Oliveira (2011) e considera toda a trajetória acadêmica do pesquisador de acordo com as informações extraídas do Currículo Lattes. A Tabela 1 apresenta todos os elementos utilizados pelo modelo de reputação com seus respectivos pesos, os quais são distribuídos em categorias específicas. Note que cada categoria apresenta um peso máximo, em que o somatório é igual a 100. Note ainda que foi atribuído à categoria de

produção científica, tecnológica e artística o maior peso, por ser esse o desafio dos programas de pós-graduação conforme a classificação das Áreas de Avaliação da Capes.

Tabela 1. Categorias, elementos e pesos

Categoria	Elemento	Peso	Peso Máximo
1. Identificação	Pós-doutorado	1	13
	Doutorado	12	
	Mestrado	6	
	Graduação	3	
2. Orientação	Pós-doutorado	5	14
	Doutorado	4	
	Mestrado	3	
	Iniciação Científica	2	
3. Banca	Doutorado	4	7
	Mestrado	3	
4. Membro	Editorial	4	9
	Comitê	2	
	Revisor de Periódico	2	
	Projeto de pesquisa	1	
5. Bolsista	Bolsista do CNPq	12	12
6. Prêmios e Títulos	Prêmios e títulos	3	3
7. Produção	Artigos publicados	10	40
	Livros publicados ou organizados	8	
	Capítulos de livros	6	
	Trabalhos em eventos	3	
	Texto em jornal ou revista	2	
	Demais tipos de produção bibliográfica	1	
	Técnica	5	
	Artística	5	
8.Coautoria	Pesquisador	2	2

Fonte: Adaptado de Cervil, Galante e Oliveira (2011).

As oito categorias apresentadas na Tabela 1 englobam os elementos constantes dos Currículos Lattes e estão descritas a seguir:

1. Identificação: corresponde à titulação do pesquisador (graduação, mestrado, doutorado), bem como à realização de estágio pós-doutoral;
2. Orientação: corresponde às orientações concluídas (iniciação científica, mestrado, doutorado, pós-doutorado);
3. Banca: corresponde às participações em bancas de mestrado e doutorado;
4. Membro: corresponde às conferências em que o pesquisador foi coordenador, membro de comitê de programa, revisor de periódico, membro de corpo editorial ou membro de projeto de pesquisa;
5. Bolsista: indica se o pesquisador possui bolsa de produtividade em pesquisa do CNPq;
6. Prêmios e Títulos: corresponde à premiação ou titulação recebida pelo pesquisador;
7. Produção: especifica a produção científica e técnica do pesquisador por meio da publicação de artigos, livros publicados ou organizados, capítulos de livros, trabalhos em eventos, texto em jornal ou revista, demais tipos de produção bibliográfica; produção técnica; e produção artística, partituras, notas musicais; e
8. Coautoria: corresponde aos outros pesquisadores que possuem relacionamento de produção científica com o pesquisador.

Comparando esse modelo ao de Cervil, Galante e Oliveira (2011), podemos perceber que a quantidade de categorias e alguns elementos variam; por exemplo, o modelo de Cervil, Galante e Oliveira (2011) não contempla as categorias *Prêmios* e *Títulos* e *Bolsista* e o elemento *Membro de Projeto de Pesquisa*. No entanto, julgamos importante o elemento de membro de projeto, uma vez que será associado às informações dos grupos de pesquisa certificados pela instituição. Também vale ressaltar que no modelo de Cervil, Galante e Oliveira (2011) os pesos dos elementos foram definidos levando-se em consideração as particularidades da área de Ciência da Computação, em que, por exemplo, existe uma pontuação para os trabalhos publicados em eventos científicos, diferentemente das demais áreas de ciências.

Em relação à fórmula de cálculo utilizada no modelo proposto por Cervil, Galante e Oliveira (2011), ela também difere da proposta deste trabalho, uma vez que considera o maior valor da reputação entre os membros da rede para ajustar a pontuação dos demais membros, o que se justifica sendo todos da mesma área. No entanto, essa abordagem torna o cálculo mais lento, uma vez que se faz necessário conhecer esse valor antes de proceder ao cálculo de todos os membros da rede.

Neste trabalho, a métrica proposta para o cálculo da reputação consiste no somatório dos pesos de cada elemento multiplicado pelo respectivo peso, totalizando os pesos máximos por categoria e dividindo pela quantidade total de elementos das categorias, conforme a Equação 1. Como na primeira categoria apenas um dos elementos é considerado, também será considerado apenas um elemento no quantitativo.

$$\text{Reputação}_{\text{Pesquisador}} = \sum (\text{Elemento} * \text{Peso})_{\text{Categoria}} / \text{Total}_{\text{Elemento}} \quad (1)$$

Para validar a métrica de cálculo de reputação proposta neste trabalho, foi realizado um estudo com a definição de duas hipóteses distintas:

1. Hipótese Global (H1): são considerados no cálculo os elementos de todas as categorias da Tabela 1; e
2. Hipótese Parcial (H2): são considerados no cálculo apenas os elementos da categoria *Identificação, Produção e Coautoria* da Tabela 1.

Para validar as hipóteses definidas, foi realizado um estudo de caso que comparou oito pesquisadores (P1, P2, P3, P4, P5, P6, P7 e P8) de quatro subáreas de conhecimento distintas: Ciência da Computação, Ciências Biológicas, Antropologia e Psicologia. Os dados dos pesquisadores foram coletados do Currículo Lattes dos pesquisadores, e a amostra inclui bolsistas de produtividade de pesquisa do CNPq. A reputação foi calculada conforme a Equação 1 e os resultados foram classificados em níveis de 1 a 5, conforme apresentado na Tabela 2. Esses níveis são apenas ilustrativos para facilitar a interpretação dos resultados.

Tabela 2. Classificação da reputação

Nível	Reputação	Classificação
1	≥ 0 e < 20	Muito baixa
2	≥ 20 e < 40	Baixa
3	≥ 40 e < 60	Média
4	≥ 60 e < 80	Alta
5	≥ 80	Muito alta

A Tabela 3 apresenta os resultados conforme as hipóteses H1 e H2, considerando a classificação da Tabela 2. Como pode ser observado, houve uma sensível diferença na classificação dos pesquisadores P1, P2, P3 e P8, que recebem bolsas PQ níveis 1D, 2, 1B e 2, respectivamente. Entretanto, não houve diferença na classificação dos pesquisadores P4, P5, P6 e P7, que representam pesquisadores sem bolsa PQ ou bolsistas PQ 1A. Esses resultados indicam que as categorias *Orientação*, *Banca*, *Membro* e *Prêmios e Títulos*, utilizadas no modelo de reputação, influenciam os bolsistas PQ dos níveis 2 a 1B, mas não influenciam os pesquisadores com bolsa PQ 1A ou os não bolsistas.

Como no modelo de Cervil, Galante e Oliveira (2011), nossa proposta considera toda a trajetória acadêmica do pesquisador segundo uma grande variedade de elementos dos Currículos Lattes dos membros. Sendo assim, podemos concluir que, aparentemente, se os pesquisadores apresentam uma baixa produção científica, os elementos utilizados na hipótese global H1 não irão auxiliar para aumentar o valor de reputação do pesquisador. O mesmo raciocínio se aplica no caso oposto, em que o pesquisador já assumiu um patamar elevado de produção científica e, conseqüentemente, recebe uma bolsa PQ 1A. Dessa forma, o modelo definido na hipótese parcial H2 foi implementado na rede do ObserveUnB como um primeiro ensaio em que adotamos o termo Índice de Produção Científica (IPC) para apresentar o total da reputação calculada pela Equação 1. Vale ressaltar que o IPC não está associado a nenhum julgamento de qualidade do pesquisador, não utilizando, por exemplo, a ideia da Tabela 2.

Tabela 3. Resultado da métrica de reputação aplicada

Pesquisador	Subárea	Hipótese Global (H1)		Hipótese Parcial (H2)		Bolsista CNPq
		Reputação	Classificação	Reputação	Classificação	
P1	Ciência da Computação	43	Média	67	Alta	PQ1D
P2	Ciência da Computação	25	Baixa	41	Média	PQ2
P3	Ciências Biológicas	79	Alta	138	Muito alta	PQ1B
P4	Ciências Biológicas	8,4	Muito baixa	16	Muito baixa	Não
P5	Antropologia	107	Muito alta	211	Muito alta	PQ1A
P6	Antropologia	12	Muito baixa	16	Muito baixa	Não
P7	Psicologia	170	Muito alta	162	Muito alta	PQ1A
P8	Psicologia	24	Baixa	38	Média	PQ2

a.3. Interface de visualização

No ObserveUnB, algumas opções de interface de visualização são oferecidas ao usuário com a finalidade de facilitar o acesso à informação. Dessa forma, pode-se visualizar a rede de colaboração da UnB pela vinculação dos membros às respectivas unidades acadêmicas e aos programas de pós-graduação ou por meio de georeferenciamento da localização física nos quatro *campi* da UnB no Distrito Federal: Plano Piloto de Brasília, Planaltina, Ceilândia e Gama. Para viabilizar a localização georeferenciada, foram utilizados mapas do *Google Maps*, mas cabe ressaltar que essa opção de consulta pode ter utilidade, mesmo em se tratando de membros de uma mesma instituição, tendo em vista o cenário de expansão acentuado da UnB viabilizado pelo programa Reuni do governo federal.

A Figura 8 exibe a tela com as opções de pesquisa visualizadas na página principal do ObserveUnB. Como exemplo de consulta, imagine que foi informada pelo usuário a palavra-chave “Sistema de Informação”, que pode estar associada a uma busca das publicações relacionadas ao termo. Como resultado da busca, é apresentada uma lista com o título da publicação e o ano, podendo ainda, por meio da opção mais detalhes, trazer as demais informações do Currículo Lattes do pesquisador associado a essa produção, como: nome do

autor, palavras-chave relacionadas, setores de atividade e áreas do conhecimento relacionados à produção em questão.

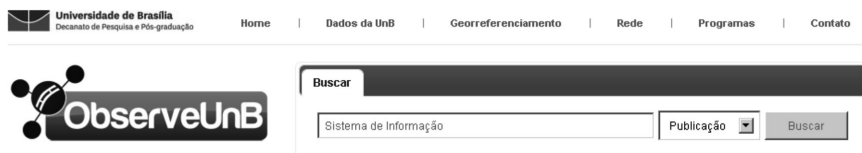


Figura 6. Opções de pesquisa do ObserveUnB

Ao selecionar *Dados da UnB*, conforme opção apresentada na Figura 6, será exibido o total da produção científica da universidade, considerando todos os membros da rede institucional. Podem ser realizadas consultas do total de produção por unidade acadêmica, por pesquisador ou por programa de pós-graduação. As opções de consulta são bem variadas, e.g., caso seja fornecido o nome de um pesquisador, serão retornadas informações de seu Currículo Lattes, incluindo sua produção científica, o relacionamento com outros pesquisadores ou até mesmo os programas de pós-graduação aos que está vinculado no âmbito institucional.

A opção *Programas* exibe os pesquisadores vinculados ao respectivo programa de pós-graduação e o total de produção científica correspondente. A Figura 9 mostra o quantitativo de produção científica dos pesquisadores vinculados ao Programa de Pós-Graduação “A” (conceito 7 na Capes). Note que a Figura 7 apresenta 27 docentes vinculados, um total de produção de 1688 itens, sendo que, no gráfico de barras, há uma maior produção de artigos em periódicos (segunda barra da esquerda para a direita).

Programa A

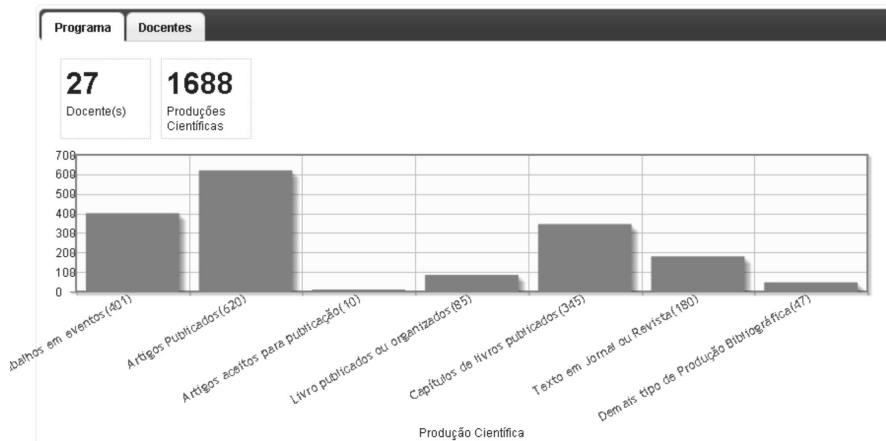


Figura 7. Total de produção científica do Programa de Pós-Graduação “A” da UnB

A opção de consulta pela rede de colaboração pode ser feita por uma determinada *Unidade Acadêmica*, conforme ilustrado na Figura 8, que apresenta a rede das unidades acadêmicas da UnB. Quando realizada a busca por uma unidade acadêmica específica, são apresentados os pesquisadores vinculados; caso se opte por um pesquisador, será apresentada a rede de colaboradores vinculada ao pesquisador.

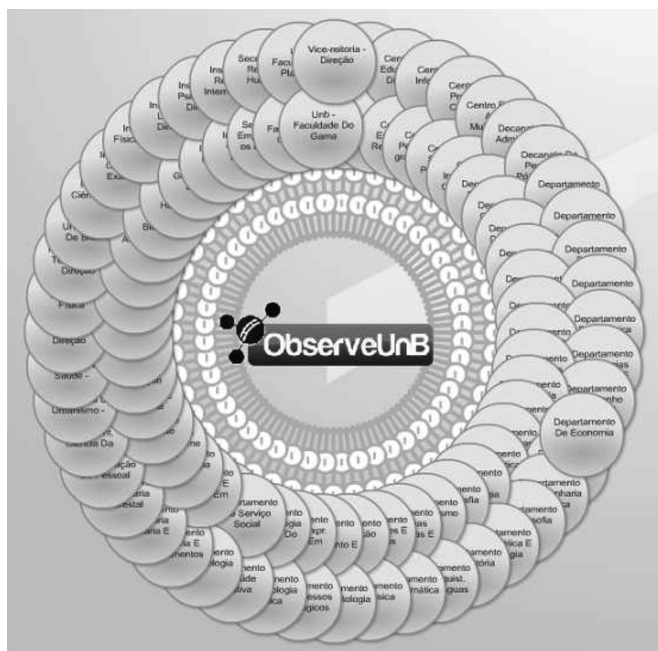


Figura 8. Rede das unidades acadêmicas da UnB

Assim como nas redes sociais e ferramentas apresentadas na seção 3, no ObserveUnB a visualização das conexões entre os pesquisadores possibilita ao usuário identificar as parcerias existentes. Foram definidos três possíveis tipos de relacionamento na rede de colaboração, a saber: “R”, que representa produção científica; “P”, que significa vinculação a um determinado programa de pós-graduação; e “I”, de integrante de unidade acadêmica ou departamento, conforme apresentado na Figura 9. Também na visualização da rede de colaboração do ObserveUnB é possível identificar a reputação de cada membro representada pelo IPC (vide seção a.2).

Análise de resultados

A partir do ObserveUnB, podem-se obter dados estatísticos relacionados à produção científica dos 2.187 pesquisadores vinculados à instituição. Esses dados podem ser visualizados por meio de gráficos quantitativos de cada unidade acadêmica, departamento e programa de pós-graduação. Enfim, podem ser listadas as informações do Currículo Lattes de cada pesquisador. Com a rede de pesquisa, obtêm-se todos os pesquisadores com os quais um determinado docente publicou. Enquanto isso, a rede de colaboração apresenta os relacionamentos em forma de grafo, representando os pesquisadores com os quais o pesquisador produziu (vide Figura 9).

Atualmente, a UnB possui três institutos nacionais de ciência e tecnologia (INCT) aprovados pelo Ministério da Ciência e Tecnologia/CNPq nas áreas de Nanobiotecnologia, Estudos Tectônicos e Inclusão no Ensino Superior e na Pesquisa. Considerando a existência dos INCTs, nota-se, via o ObserveUnB, que a rede de colaboração científica existente entre os coordenadores e pesquisadores vinculados aos INCTs é grande quando comparada a outros pesquisadores não vinculados a esses institutos. A Figura 9 apresenta a rede de um pesquisador “P31” coordenador de INCT da UnB, que integra (I) o Departamento “A”, pertence (P) aos Programas de Pós-Graduação “C” e “D” e produz com (R) 28 pesquisadores (P1 a P28).

O ObserveUnB possibilita respostas às quatro perguntas de pesquisa apresentadas na seção 4, pois, por meio da visualização da rede de colaboração científica, pode-se verificar quem são os pesquisadores da UnB, o que estão produzindo cientificamente, onde esses pesquisadores estão localizados e com quem estão produzindo. Além dessas questões particulares aos membros da rede de colaboração, o observatório permite, por meio da representação em gráficos quantitativos, apresentar a produção científica das unidades acadêmicas e dos programas de pós-graduação, viabilizando análises aos gestores do conhecimento institucional. Diante da variedade de informação disponibilizada pela ferramenta em diferentes formatos, cabe ao usuário filtrá-las e relacioná-las de forma que atenda às suas particulares necessidades de consulta.

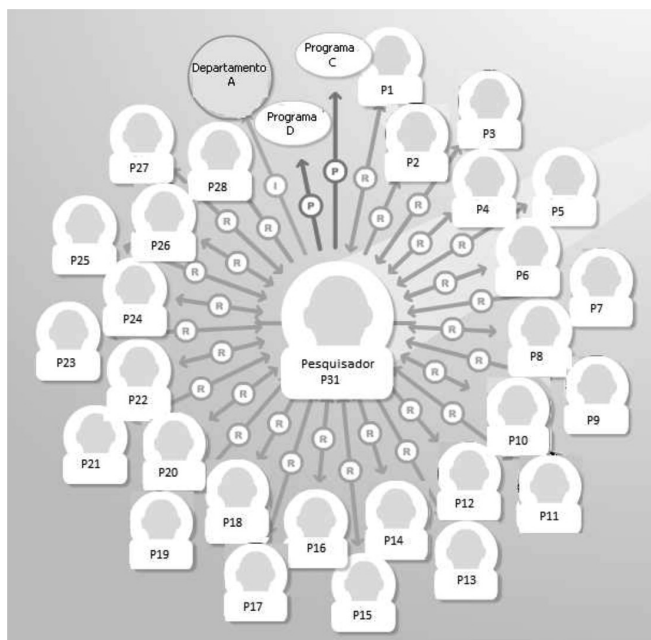


Figura 9. Rede de relacionamentos de um pesquisador membro de INCT vinculado à UnB

Conclusões e trabalhos futuros

Neste trabalho, foi apresentado o observatório da rede de colaboração científica da UnB, denominado ObserveUnB, que tem a

finalidade de servir como um instrumento útil para a identificação de parcerias em redes de colaboração, bem como para auxiliar no processo de gestão institucional da informação. Conforme apresentado na seção 3, existem outros trabalhos e ferramentas que disponibilizam redes de colaboração científica, as quais respondem às quatro perguntas de pesquisa do ObserveUnB com uso de bases de dados internacionais e disponíveis na web (ver seção 4). No entanto, como uma iniciativa no âmbito de uma instituição federal de ensino superior, essa proposta é precursora, além de ser fidedigna à realidade da comunidade acadêmica nacional, uma vez que integra informações dos repositórios do CNPq e da Capes.

Em relação ao modelo de reputação proposto no ObserveUnB, sabe-se que são necessários novos ensaios para melhor equalizar as diferenças existentes entre as diversas áreas de conhecimento, uma vez que cada área dá prioridade a determinados veículos de produção científica, necessitando modelos mais ajustáveis. Nesse sentido, viabilizar a participação de membros mais experientes na reputação de outros menos experientes possibilitaria um modelo de cálculo realizado por meio da colaborações entre os pares. Ainda com relação à reputação, pode-se permitir a visualização da evolução temporal da reputação dos pesquisadores ou de grupos de pesquisadores vinculados aos programas de pós-graduação, o que serviria para fins de análise e avaliação. O repositório do ObserveUnB pode também ser integrado a outras universidades e com isso permitir análises comparativas entre a produção científica das diversas instituições de ensino e pesquisa do País.

Outra questão interessante de pesquisa é a transformação da rede de colaboração científica do ObserveUnB em uma rede social dinâmica, permitindo a interação entre os membros e também com a comunidade de forma integrada e participativa, utilizando diversos veículos de comunicação, e.g., *twitter* e redes sociais.

Conforme apresentado, existem diversas sugestões de trabalhos futuros, uma vez que o projeto do ObserveUnB oferece um grande potencial de evolução pelo seu caráter eminentemente

multidisciplinar. Como áreas relacionadas a esta pesquisa, podemos citar as Ciências Exatas, que incluem aspectos da Matemática, Estatística e Ciência da Computação, no que tange aos modelos de cálculo e à teoria de grafos, às séries estatísticas e à manipulação automática da informação e do conhecimento. As Ciências Humanas e Sociais, por intermédio da Sociologia, Psicologia e Administração, podem auxiliar na avaliação dos atores e de seus relacionamentos nas redes de colaboração científicas, tecnológicas e artísticas.

Recebido em 18/03/2012

Aprovado em 12/04/2013

Referências bibliográficas

BAADER, F.; NUTT, W. Basic description logics. In: BAADER, F. et al. (Eds.). **The Description Logic Handbook**. Cambridge: Cambridge University Press, 2010. p. 47-100.

BALANCIERI, R.; et al. A análise de redes de colaboração científica sob as novas tecnologias de informação e comunicação: um estudo na plataforma Lattes. **Ciência da Informação**, v. 34, p. 64-77, 2005.

BRICKLEY, D.; MILLER, L. **Friend of a friend**. Disponível em: <<http://www.foaf-project.org/>>. Acesso em: 10 out. 2011.

CASTELLS, M. **A sociedade em rede**: A era da informação: economia, sociedade e cultura. 3. ed. São Paulo: Editora Paz e Terra, 1999.

CERVIL, R. C.; GALANTE, R.; OLIVEIRA, J. P. Identificando a Reputação de Pesquisadores Usando um Modelo de Perfil Adaptativo. In: CONGRESSO DA SOCIEDADE BRASILEIRA DE COMPUTAÇÃO, 31., 2011, Rio Grande do Norte. **Anais...**

CONSCIENTIAS, C. **Padronização xml**: Grupos de pesquisa e Curriculum Lattes. Disponível em: <<http://Impl.cnpq.br/Impl/>>. Acesso em: 10 ago. 2011.

CORMEN, T. H.; et al. **Algoritmos: Teoria e Prática**. 4. ed. Rio de Janeiro: Editora Elsevier, 2002.

FRANCINA, E. T. O. Rede de colaboração científica no tema “Estudos Métricos”: um estudo de co-autorias através dos periódicos do SciELO da área de Ciência da Informação. **Brazilian Journal of Information Science (BJIS)**, v. 2, n. 2, p. 35-49, 2008.

GRUBER, T. R. **A translation approach to portable ontology specifications**. Technical Report KSL 92-71. Stanford: Stanford University, 1996.

GUARINO, N. Formal ontology and information systems. In: FOIS, 1., 1998, Trento, Itália. **Proceedings...** Amsterdam: IOS Press, 1998. p. 3-15.

KATZ, J. S.; MARTIN, B. R. What is research collaboration? **Research Policy**, Amsterdam, n. 26, p. 1-18, 1997.

KLEINBERG, J. M. Authoritative sources in a hyperlinked environment. **Journal of the ACM**, v. 46, p. 604-632, 1999.

LABBÉ, C. One of The Great Stars in The Scientific Firmament. **ISSI Newsletter**, v. 6, n. 2, p. 48-52, 2009.

NEWMAN, M. E. J. **Who is the best connected scientist?** A study of scientific coauthorship networks. Paper 00-12-064. 2000. Disponível em: <http://arxiv.org/abs/cond-mat/0011144v1>. Acesso em: 15 set. 2011.

OBSERVEUNB. **Lançamento do ObserveUnB**. 2011b. Disponível em: <http://www.unb.br/administracao/decanatos/dpp/diadi/novinformes/Apresentacaoobserve.pdf>. Acesso em: 28 nov. 2011.

PAGE, L.; et al. **The PageRank Citation Ranking: Bringing Order to the Web**. Technical Report 1999-66, Stanford InfoLab. 1999. Disponível em: <http://ilpubs.stanford.edu:8090/422/>. Acesso em: 10 nov. 2011.

JARDIM, A. D.; PALAZZO, L. A. M. Aplicações da Web Semântica nas Redes Sociais. In: CONGRESSO NACIONAL DE AMBIENTES HIPERMÍDIA PARA A APRENDIZAGEM, 4., 2009, Florianópolis. **Anais...** p. 1-14.

PUJOL, J. M.; SANGÜESA, R.; DELGADO, J. Extracting reputation in multiagent systems by means of social network topology. In: INTERNATIONAL JOINT CONFERENCE ON AUTONOMOUS AGENTS AND MULTIAGENT SYSTEMS: PART 1, 1., 2002, New York. **Proceedings...** p. 467-474.

RALHA, C. G.; et al. ObserveUnB - Um portal Web de rede social científica. In: SEMINÁRIO INTEGRADO DE SOFTWARE E HARDWARE, 39., 2012, Curitiba. **Anais...**

ROBREDO, J.; BRÄSCHER, M. (Orgs). Passeios pelo Bosque da Informação: Estudos sobre Representação e Organização da Informação e do Conhecimento. In: BRÄSCHER, M.; CARLAN, E. **Sistemas de organização do conhecimento: antigas e novas linguagens**. Brasília: IBICT, 2010. p. 147-176.

SABATER, J.; SIERRA, C. Social regret, a reputation model based on social relations. **Newsletter ACM SIGecom Exchanges – Chains of Commitment**, v. 3, n. 1, p. 44-56, 2001.

TANG, J.; et al. Arnetminer: extraction and mining of academic social networks. In: ACM SIGKDD INTERNATIONAL CONFERENCE ON KNOWLEDGE DISCOVERY AND DATA MINING, 14., 2008, New York. **Proceedings...** p. 990-998.