



Construindo Aplicações na Web Semântica Buscas Semânticas

Renato Fileto
fileto@inf.ufsc.br

Programa de Pós-graduação em Ciência da Computação - PPGCC
Departamento de Informática e Estatística - DINE
Centro Tecnológico - CTC



Recuperação da Informação

• Medidas da eficácia das soluções

$$\text{precisão} = \frac{\text{nro. de docs relevantes recuperados}}{\text{nro. de docs recuperados}}$$

$$\text{cobertura} = \frac{\text{nro. de docs relevantes recuperados}}{\text{nro. de docs relevantes}}$$

Conteúdo

- Motivação
- Definição do problema
- Abordagens e ferramentas
- Conclusões

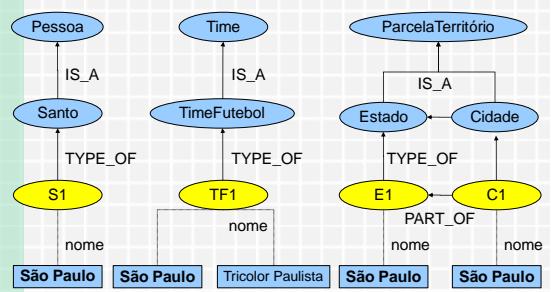
Buscas semânticas

- Consideram os significados (das consultas e do conteúdo ou metadados), procurando:
 - evitar, identificar e/ou tratar ambigüidades,
 - expandir semanticamente as buscas para significados relacionados,
 - orientar o usuário no refinamento das suas consultas,
 - apresentar os resultados classificados segundo os significados a eles atribuídos
 - ordenar os resultados segundo a relevância estimada dos diferentes significados para a consulta e o usuário que a estipulou,
 - Tornar mais rica, agradável e efetiva a interação do usuário com o sistema de busca.

Problemas da RI tradicional

- Usa somente informação léxica e sintática
- Traz todos os resultados relacionados sintaticamente com uma palavra-chave, sem organização quanto aos possíveis significados
- Não recupera resultados cujo conteúdo ou metadados referece outras palavras com significado parecido ou relacionado
- Menospreza as características dos usuários quanto ao significado que atribuem a determinados termos

Trecho de ontologia 1



Classificação de abordagens para buscas semânticas

(Mangold, 2007)

- Arquitetura
- Acoplamento entre ontologias e documentos
- Transparência
- Contexto do usuário
- Modificação de consultas
- Estrutura das ontologias
- Tecnologia para representar as ontologias
- Desempenho, escalabilidade e distribuição
- Adaptabilidade
- Ranking

Arquiteturas para buscas semânticas

- **Máquina de busca stand alone:** armazena metadados sobre os documentos em índices com os quais avalia as consultas
- **Meta-máquina de busca:** distribui as consultas para outras máquinas de busca e combina os resultados

Acoplamento ontologias-documentos

- **Acoplamento forte:** os metadados dos documentos referem-se explicitamente aos conceitos de uma ontologia e vice-versa (requer anotação semântica)
- **Acoplamento fraco:** os documentos são independentes de quaisquer ontologias disponíveis (pode restringir a capacidade de busca)

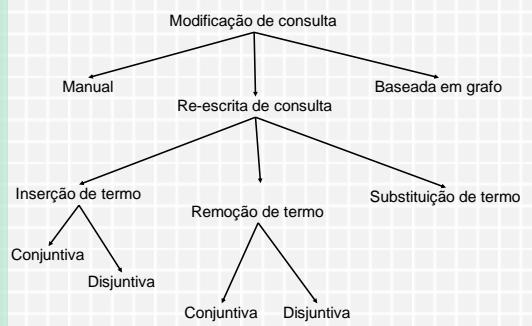
Contexto do usuário

- **Learning:** extraído na interações do usuário com o sistema (histórico das consultas e refinamentos de consultas)
- **Hard-coded:** as consultas são organizadas em categorias que definem a informação solicitada (e.g., "location of", "resources for" ...)

Modificação de consultas

- **Manual:** o sistema retorna um trecho apropriado da ontologia junto com a resposta, permitindo ao usuário navegar na ontologia e reformular a consulta pela adição e remoção de termos
- **Re-escrita automática:** a consulta é otimizada automaticamente pelo sistema
- **Baseada em grafo:** considera conceitos e documentos como nodos de um grafo, efetuando propagação da consulta a partir dos nodos correspondentes a termos de busca

Modificação de consultas



Estrutura da ontologia

- **Propriedades anônimas:** as relações entre conceitos não são nomeadas, indicando apenas que eles compartilham o mesmo contexto
- **Propriedades padronizadas:** sinônima, hipernímia, meronímia, negação, etc.
- **Propriedades específicas de domínio:** além das propriedades padrão, considera propriedades específicas de um domínio (e.g., "tipo de câmera" em sistema para a área de fotografia)

Tecnologia da ontologia

- F-Logic
- RDF
- DAML+OIL
- Frames
- OWL
- :

Comparação de abordagens

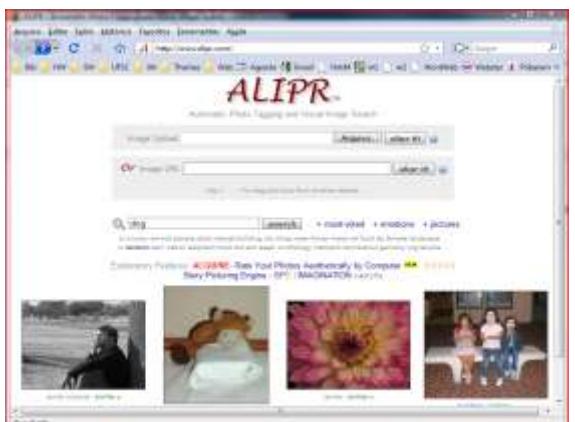
| Projetos/projeto | OWL | JavaML | ZAP | Método gerando arvore | ZSM | Lifframe agent |
|---------------------|------------------------------|------------------------|----------------------------|----------------------------------|----------------------------------|------------------------|
| Auto | Alfons and Hansler (2006) | Gómez et al. (2001) | Dudu et al. (2003) | Ricke et al. (1999) | Bartos-Janczak et al. (2001) | Requena (2005) |
| Type | WWW | WWW | WWW | WWW | WWW | WWW |
| Annotations | Standard | None | Standard | None | Standard | Standard |
| Coupling | Tight | — | Tight | Loose | Tight | Tight |
| Transparency | Enclosure | Transparent | Hybrid | Transparent | Hybrid | Enclosure |
| User control | None | Hard coded | Hard coded and learning | None | None | Learning |
| Query resolution | Massively | Obj. lang. | Graph-based | All sorts of query resolution | Obj. lang. Schemata | Obj. lang. Schemata |
| Ontology instant. | Hypernym Anonymity | — | Anonymity | Domain-specific | Hypernym Anonymity | Anonymity |
| Ontology technology | Frequency | — | EDP | Naive | ZSM, concept + Ford et al. | Proprietary |

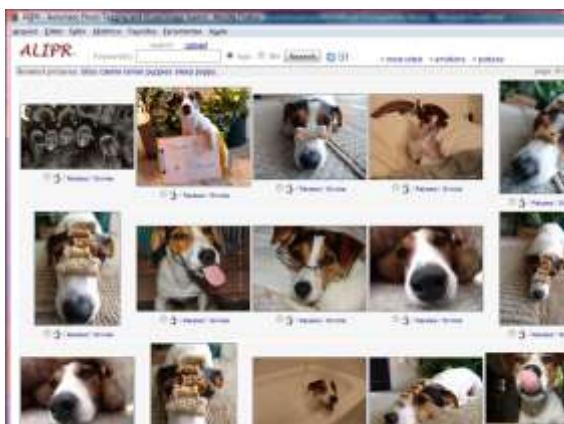
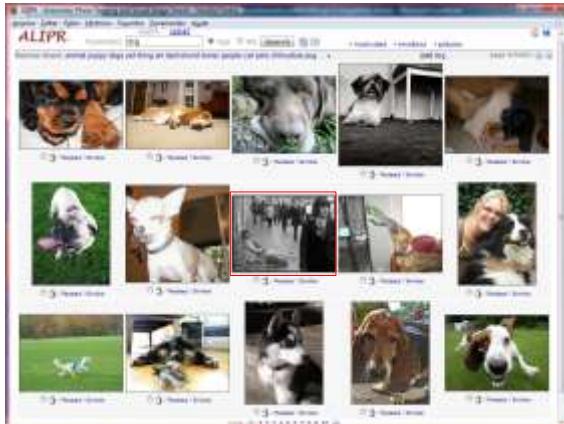
Comparação de abordagens (cont.)

| Projetos/projeto | SCOLE | TRUST | Artefacto orientado | Geoparce | Semantics most effective |
|---------------------|--------------------------------|--------------------------------|----------------------------------------------------|-----------------------------|-----------------------------|
| File | Information system | Information system | To be audio data retrieval, IS for image retrieval | None | — |
| Architecture | Model-driven | Hybrid | Model-driven | Model-driven | Open |
| Coupling | Tight | Hybrid | Tight | Tight | Hybrid |
| Transparency | Enclosure | Object | Transparent | Transparent | Hybrid |
| Unit control | Unclear | Hard-coded | None | Hard-coded | Hard-coded and learning |
| Query resolution | Massify | Obj. lang. | Obj. lang. Schemata | Schemata | Open |
| Ontology instant. | Hypernym Domain specific | Hypernym Schemata Domain | Hypernym Schemata Domain | Hypernym Domain specific | Facilitate use of XML |
| Ontology technology | Unclear | Proprietary | Proprietary | Proprietary | Unclear(?) |

Buscas semânticas na Web hoje

- <http://www.alipr.com> (imagens)
- <http://www.cognition.com>
- <http://www.deepdyeve.com>
- <http://www.bing.com>
- <http://www.cuill.com>
- <http://www.firebaseio.com>
- <http://www.google.com>
- <http://www.kosmix.com>
- <http://www.hakia.com>
- <http://www.powerset.com>
- <http://www.sensebot.net>
- <http://developer.yahoo.com/searchmonkey>
- <http://swoogle.umbc.edu> (ontologias)





A screenshot of a web browser displaying the MEDLINE COGNITION website. The page features a navigation bar with links like Home, About, Help, Log In, and Log Out. Below the navigation is a search bar with the placeholder "Search MEDLINE COGNITION". The main content area is titled "SAMPLE QUESTIONS" and contains several sections of questions and answers. One section is titled "Cognitive Assessments" and includes questions about memory techniques, cognitive behavioral strategies, symptoms of amnesia, geriatric syndromes of dementia, and the MMSE. Another section is titled "Depression" and includes questions about the DSM-IV criteria for depression, the BDI-II, and the BDI-SF. A third section is titled "Anxiolytic Agents" and includes questions about benzodiazepines, selective serotonin reuptake inhibitors, and tricyclic antidepressants. A fourth section is titled "Psychotropic Agents" and includes questions about antipsychotics, mood stabilizers, and drug interactions with psychotropics. A fifth section is titled "Dementia" and includes questions about Lewy bodies, Alzheimer's disease, vascular dementia, and Lewy body dementia.

A screenshot of the DeepDyve website. The header includes the DeepDyve logo, a search bar with placeholder text "Search for Research", and a "Search" button. Below the header is a blue navigation bar with tabs for "Home", "About", "Contact", "Log In", and "Feedback". The main content area features a search bar with the placeholder "Find", a "Search" button, and a "Popular Searches" section. The "Popular Searches" section contains links to "Clinical applications of biodegradable implants in orthopedic surgery", "Knee replacement surgery", and "Cataract extraction and expanded life span: role of resistance to cataract and diabetic retinopathy".

A screenshot of a web browser showing the ebscohost search results for the query "How to increase the efficiency of the D-2". The search bar at the top contains the query. Below it, a search result for a document titled "How to increase the efficiency of the D-2" is displayed. The document is from the journal "Journal of Polymer Science: Part A: Polymer Chemistry" and is dated April 1998. The abstract discusses the effect of various factors on the efficiency of a D-2 polymerization process. On the right side of the screen, there are sidebar sections for "Recent Searches" (including "How to increase the efficiency of the D-2", "ebscohost", and "Journal of Polymer Science: Part A: Polymer Chemistry") and "Help & Information" (with links to "About ebscohost", "Contact Us", "Feedback", "Help", "Help Center", and "Logout").

A screenshot of a web browser showing search results for "Belo-Paulo". The top navigation bar includes links for Home, Search, News, Weather, Entertainment, Apps, and more. The search bar contains the query "Belo-Paulo". Below the search bar, there's a "Did you mean?" suggestion for "Belo Horizonte". The main content area displays several search results, each with a title, a snippet of text, and a green "Ver mais" button. The results include:

- Belo-Paulo** - Belo-Paulo é uma vila no Brasil, no interior do estado de Minas Gerais, pertencente ao município de Itabira. A vila tem cerca de 2.000 habitantes e é conhecida por sua...
- Belo-Paulo** - Belo-Paulo é uma vila no Brasil, no interior do estado de Minas Gerais, pertencente ao...
- Belo-Paulo** - Belo-Paulo é uma vila no Brasil, no interior do estado de Minas Gerais, pertencente ao...
- Belo-Paulo** - Belo-Paulo é uma vila no Brasil, no interior do estado de Minas Gerais, pertencente ao...
- Belo-Paulo** - Belo-Paulo é uma vila no Brasil, no interior do estado de Minas Gerais, pertencente ao...

A screenshot of the Microsoft Bing search engine homepage. The search bar at the top contains the query "New York". Below the search bar, there is a "Did you mean" suggestion for "New York, NY". The main content area displays a grid of 12 image thumbnails, each representing a different view or aspect of New York City, such as the Empire State Building, Central Park, and various landmarks. On the left side of the page, there is a vertical sidebar with a "Search" section containing fields for "Search" and "Search by category", and a "Refine search" section with dropdown menus for "Category", "Type", and "Language". Further down the sidebar, there are sections for "Top 100", "Local", "Hotels", "Flight", "Car", "Local", "Hotels", "Flight", and "Car". At the bottom of the sidebar, there is a "Sign in" link.

The screenshot shows the cuil search engine's homepage. The search bar contains the query "Sao Paulo". Below the search bar, a list of suggested search terms is displayed in a box, including "SAO PAULO-SP", "University of Sao Paulo", "Sao Paulo Fashion", "Sao Paulo Fashion Week", "Sao Paulo State", "Sao Paulo Hotel", "Sao Paulo Hotels", "Sao Paulo Brasiliense", and "Sao Paulo Congonhas". To the right of the search bar is a blue "Frequente" button. At the bottom of the page, there is footer text: "Acesse Cuil no iPhone", "Português © 2008 Cuil, Inc.", and "Sobre Cuil".

The screenshot shows the Facebook homepage with the URL 'www.facebook.com/freebase'. The main header has 'Facebook' and 'Freebase' as the primary links. Below the header, there's a large orange banner with the text 'A wealth of free data.' and a link to 'See how it works'. To the right of this banner is a map of the United States with several location points marked. A box on the left lists 'Search against millions of interconnected objects' with a list of categories like 'Books', 'People', 'Companies', etc., each with a small thumbnail image. On the right side, there are three profile pictures with their names: 'Audrey Tautou', 'Tom Cruise', and 'Keanu Reeves'. Below these profiles is a section titled 'Use Freebase on Your Site' with a sub-section 'Learn how we can help +'. At the bottom of the page, there are links for 'The Wall Street Journal', 'Bing', 'Searches', and 'Case Studies'.

A screenshot of a Google search results page for the query "sao paulo". The results are displayed in a standard grid format with blue links and black text. The first result is a link to the São Paulo Wikipedia page. Other results include links to the official website of the São Paulo state government, the São Paulo tourism board, and various travel and news sites. The right sidebar features local information such as flight prices to São Paulo and travel tips.

A screenshot of a Google search results page. The search query 'Lula presidente' is entered in the search bar. Below the search bar, there are links for 'Procurar' and 'Avançar para a versão móvel'. The results show 1-10 de 16 de aproximadamente 33.300.000 páginas. The first result is a link to 'WIKI' with the title 'Lula presidente' and a snippet: 'Neste artigo são tratados os conceitos de Liderança, Líder e Liderado.' Below this, there are several other news articles and links related to Lula's political career, including mentions of his time as President of Brazil and his role in the PT party.

A screenshot of the Kosmix homepage. The top navigation bar includes links for 'Search', 'Email', 'Edit', 'Delete', 'Pageinfo', 'Environment', and 'About'. The address bar shows the URL 'http://www.kosmix.com/'. The main content area features a large green header with the word 'Kosmix'. Below it, a search bar contains the query 'Sao Paulo' with a dropdown menu showing suggestions: 'Sao Paulo Airport', 'Sao Paulo Plane Crash', and 'Sao Paulo Zoo'. To the right of the search bar is a green 'Explore!' button. On the left, there's a 'Top News' section with a 'YAHOO! BUZZ' button. The right side of the page displays a list of search results, each with a thumbnail image, title, and a brief description.

Kosmix São Paulo

[Search](#)

[Home](#) >

São Paulo

Show me listing for São Paulo (1), São Paulo state newsfeed, São Paulo city newsfeed, or São Paulo São Paulo

[Map](#) - [Video](#) +

From YouTube / Webcam / Images / News / Reference / Wikis

[PostTube](#) [EDITOR'S CHOICE](#) [ADDITIONAL VIDEOS](#)

 Cambridge's a学术 City life (11:00)
 O São Paulo city skyline (22:25)
 O São Paulo - Brazil (00:00)

[Search Kosmix](#)

A screenshot of the hakia search engine's homepage. The top navigation bar includes links for Home, Help, About, Support, Feedback, Signin, and Signout. Below the bar is a logo for "Searched. Searched. Searched." followed by a globe icon. The main search area features the word "hakia" in a large blue font next to a magnifying glass icon. Below the search bar is a search input field containing "soccer player" with a "GO" button to its right. A large, stylized green leaf graphic is positioned in the center of the page.

The screenshot shows a search results page from Swoogle. The search query is "semantic web search". The results are displayed in a table format with columns for URL, Title, Description, and PageRank. The first result is a link to the Semantic Web Center at UFRJ, titled "Semantic Web Search Engine - Semantic Web". The second result is a link to the University of São Paulo's Special Export RDF site, titled "Special Export RDF/Semantic Web". The third result is a link to the University of Columbia's SWEBET tool, titled "SWEBET". The fourth result is a link to the University of Columbia's SPARQL query engine, titled "SPARQL". The fifth result is a link to the University of Columbia's RDF2OWL converter, titled "RDF2OWL". The sixth result is a link to the University of Columbia's SWEBET tool again, titled "SWEBET".

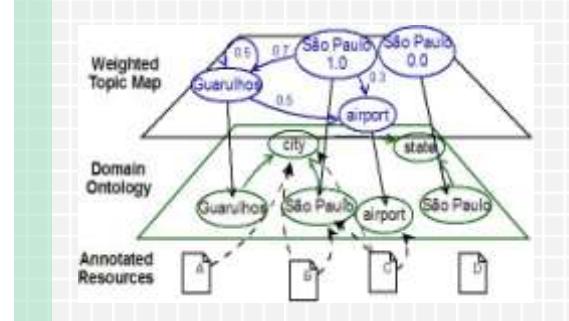
| URL | Title | Description | PageRank |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------|----------|
| http://www.semanticweb.org/search.cgi?query=semantic+web | Semantic Web Search Engine - Semantic Web | Search, 2004, SPARQL, Ontology, Reasoner, Database, Logic | 1.7 |
| http://www.dcc.ufrj.br/~luisse/govontolog/semweb_center.owl | UFRJ - SemWeb, Semantic Ontology, Universidade de São Paulo, Institute, Institute, Semantic Web Division, RDFS/OWL, 2003-11-19, 1.7TCS, generated1300, semantic, cached | University of São Paulo, Institute, Institute, Semantic Web Division, RDFS/OWL, 2003-11-19, 1.7TCS, generated1300, semantic, cached | 1.7 |
| http://www.csie.ntu.edu.tw/~cjlin/paper/SpecialExportRDF/Sem_Web/ | Special Export RDF/Semantic Web | Special Export RDF/Semantic Web, 2004-04-01, 1.7TCS, generated1300, semantic, cached | 1.7 |
| http://swebet.csail.mit.edu/index.php/Special_Export/RDF/University_of_Sao_Paulo.html#UniversityofSaoPaulo | University of São Paulo, 2004-04-26, 1.7TCS, generated1300, semantic, cached | University of São Paulo, 2004-04-26, 1.7TCS, generated1300, semantic, cached | 1.7 |
| http://db.cs.columbia.edu/wikilinks/SWEBET.rdf | SWEBET | SWEBET, 2007-04-06, 1.8TCS, generated1300, semantic, cached | 1.8 |
| http://lt.cs.columbia.edu/~7Ebenstein/dialecticExploit/RDFOWL.owl | DialecticExploit, 2007-04-08, 1.9TCS, generated1300, semantic, cached | DialecticExploit, 2007-04-08, 1.9TCS, generated1300, semantic, cached | 1.9 |
| http://lt.cs.columbia.edu/~7Ebenstein/dialecticExploit/RDF2OWL.owl | RDF2OWL, 2007-04-18, 1.9TCS, generated1300, semantic, cached | RDF2OWL, 2007-04-18, 1.9TCS, generated1300, semantic, cached | 1.9 |
| http://lt.cs.columbia.edu/~7Ebenstein/dialecticExploit/SWEBET.rdf | SWEBET | SWEBET, 2007-09-12, 204K, generated1300, semantic, cached | 204K |

A screenshot of a web browser showing the English Wikipedia search results for the term "Lata". The title "Lata" is at the top, followed by a section titled "Did you mean?" listing various homophones like Lata (person), Lata (pet), Lata (metropolis), and Lata (program). Below this is a section titled "Recent changes" showing recent edits. A large image of a woman's face, identified as Lata Mangeshkar, is prominently displayed. The page also includes sections for "Recent changes", "Random article", and "Recent changes".

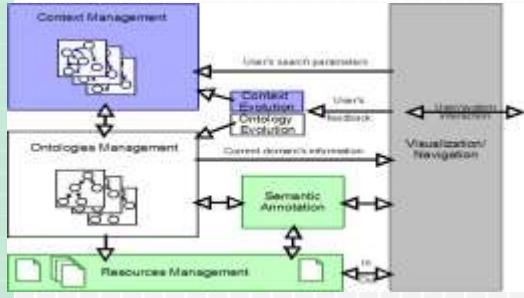
Algumas questões em aberto

- Representação de contextos de usuários e mapeamento desses a ontologias
 - Análise dos efeitos de modificações de consultas
 - Meta-buscas semânticas
 - Análise da aceitação dos usuários
 - Adaptabilidade a diferentes ontologias
 - Ranking dos resultados
 - Integração com sistemas de gerenciamento de documentos/conteúdo
 - Tratamento de dados multimídia
 - Interfaces homem-máquina
 - Desempenho e escalabilidade

Contextual Semantic Search (CSS)
(D'Agostine, Fileto - SBBD2007)



General Architecture for CSS



The Weighted Topic Graph

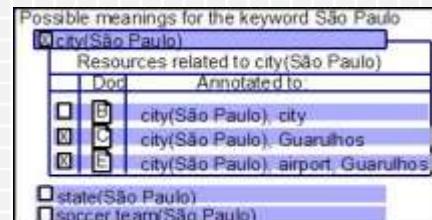
(D'Agostine, Fileto, Dantas, Gauthier - ICEIS 2008)

- A weighted topic graph is a graph $TM(T, A)$
 - T is a set of topics (vertices)
 - A is a set of associations (edges)
- Let be an ontology O
 - Each topic t in T corresponds to a ontology term o in O
- Each topic t has a weight $[0,1]$. The sum from the weights from all topics with same name equals 1
- Each association a has a weight $[0,1]$. The sum from all the weights from all association departing a single topics equals 1

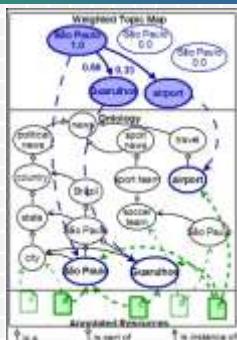
The Weights in the Topic Graph

- **Topic weights:** used to disambiguate
 - E.g., interest in São Paulo city instead of São Paulo state
- **Association weights:** used to semantically expand searches
 - E.g., Actual interest in Guarulhos and airport while searching for São Paulo city

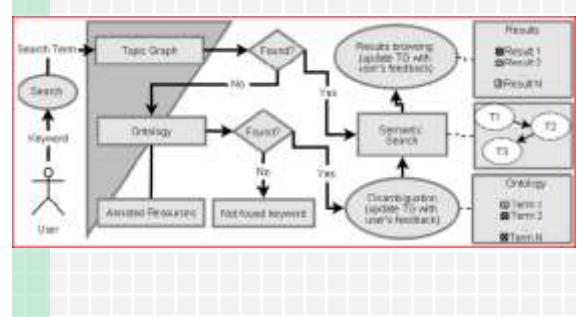
Capturing the User's Context



Generated Topic Graph with the Ontological User's Context



The Contextual Semantic Search Process



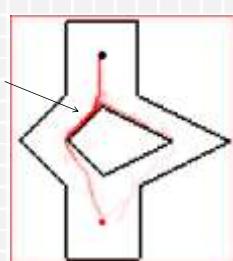
A screenshot of a web-based application interface titled 'New Card'. At the top, there's a navigation bar with links for 'New', 'Open', 'User', 'Edit As', and 'Configure'. Below the navigation is a search bar containing the placeholder text 'Search...'. To the right of the search bar is a dropdown menu labeled 'Results in: Report by date with ("')'. Underneath the search bar, there are two buttons: 'Search' and 'New Transaction'. Further down, there are sections for 'Recent Results' and 'Typical Results', each with a 'Search' link. A message 'None (Create first row to open)' is displayed below the typical results section. The bottom of the page shows a footer with the text 'v1.0.0' on the left and '0.0.0' on the right.

The screenshot shows a 'Disambiguation' dialog box from Semantic MediaWiki. The title bar says 'Disambiguation'. The main area contains the following text:
Choose the intended meaning(s) for the keyword "São Paulo":
São Paulo futebol Clube
SoccerClub
SportTeam
Organisation
Resource

São Paulo
Area
PopulatedPlace
Place
Resource

The screenshot shows the Neo4j browser interface. At the top, there's a navigation bar with 'File', 'New', 'Open', 'Save', 'Import', 'Help' and a 'Logout' button. Below the navigation is a toolbar with icons for 'Create Node', 'Create Relationship', 'Search', 'Import', 'Export', 'Help', and 'About'. The main area features a graph visualization with nodes labeled 'Alice', 'Bob', 'Cesar', 'David', 'Eduardo', 'Fernando', 'Giovana', 'Igor', 'Julia', 'Luis', 'Márcio', 'Natalia', and 'Paulo'. A tooltip for node 'Alice' shows its properties: 'Name: Alice' and 'Age: 30'. Below the graph is a search bar with the placeholder 'Search...'. Underneath the search bar are two buttons: 'Search' and 'Reset Search'. At the bottom left, there's a 'Selected Results' section with a table showing columns for 'Class' and 'Name'. The table lists nodes like 'Alice', 'Bob', 'Cesar', etc., with their names in the 'Name' column.

Fundamento dos algoritmos de busca sobre contexto ontológico



Heurísticas de estigmergia (*Ant Colony Optimization - ACO*)

Algoritmo de busca semântica baseada em ACO

Entrada: palavras-chave[];
Dados: G(T,A), Onto, repositório,
 limite_de_profundidade,
 limite_mínimo_de_peso,
 λ // fator de atenuação do peso;

```
início Algoritmo Busca Contextual 1 Estigmérica
    resultados = busca_G_ou_Onto(palavras-chave[]);
    atenuaG (índice) //  $\lambda \in [0, 1]$ 
    monetização(custo, feedback(resultados));
```

This screenshot shows the Praestro system's user interface. At the top, there is a navigation bar with options like 'New', 'Close', 'Save', and 'Select All'. Below the navigation is a complex ontology diagram with nodes and edges. A search bar at the top right contains the query 'San Paolo'. The main area displays search results under 'Relevant Results / Typed Results' and 'Detailed' sections. The 'Detailed' section lists various entities related to 'San Paolo', such as 'Sao Paulo' (with a population of 10.2M), 'Sao Paulo State', 'Brazil', and 'South America'. There is also a 'Graph' button next to each result entry.

Demo Praestro – Karina Fasolin

This screenshot shows the Praestro system's user interface. It features a search bar at the top with the query 'Sao Paulo'. Below the search bar is a 'Dependency with /' button. The main area displays search results under 'Relevant Results / Typed Results' and 'Detailed' sections. The 'Detailed' section lists entities related to 'Sao Paulo', such as 'Sao Paulo' (with a population of 10.2M), 'Sao Paulo State', 'Brazil', and 'South America'. There is also a 'Graph' button next to each result entry.

Alguns trabalhos relacionados (cont.)

| Name | Article | Description | Image |
|------------------------|---------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------|
| Graupmann et al. 2005 | The SphereSearch engine for unified ranked retrieval of heterogeneous XML and web documents | Searches through around given element to both contexts. Made tests with stress testing report. | |
| Park e Cheyer 2006 | Just For Me: Topic Maps and Ontologies | Manages knowledge in three layers: Topics, Knowledge Structure, Documents. Not implemented (at least at the time of publication). | |
| Michlmayr, et al. 2007 | Adaptive User Profiles for Enterprise Information Access | Represents context as an individual graph. The graph is composed based on tags used to semantically annotate the content which the user considers to be relevant. However, the tags are only that, only labels. They are not tied-down to a formal definition. | |

Alguns trabalhos relacionados (cont.)

| Name | Article | Comments | Image |
|-------------------------|------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------|
| Aleman-Meza et al. 2003 | Context-Aware Semantic Association Ranking | Also defines context as regions. But the regions are static; defined in the ontology which is used by the system. So the contexts are defined the same for every single user. | |
| Mani and Sundaram 2007 | Modeling user context with applications to media retrieval | Creates a graph representing the user's context. Each node corresponds to an instance in the ontology. The edges represent relations between the instances. Built to search through multimedia documents. However, it does not associate the vocabulary the user uses with the instances. | |
| Challam et al. 2007 | Contextual Search Using Ontology-Based User Profiles | Monitors a user's active context coming from open Internet Examples: MS-Office and Mobi. The captured information is stored and used to build a user's contextual profile bases on an ontology. The context is tightly coupled with the ontology. | |
| Vallet et al. 2006 | Personalized Information Retrieval in Context | Stores context as a graph constructed based on the ontology used by the system. Similar to Mani and Sundaram, including its limitations (mapping the vocabulary), however it failed for text documents. | |
| Sieg et al. 2007 | Ontological User Profiles for Personalized Web Search | Similar to Challam et al., but it gets the context information from the ontology terms used to annotate the retrieved relevant context. Also limited by the structure of the ontology. | |

Alguns trabalhos relacionados (cont.)

| Name | Deals with objective knowledge (ontology)? | Deals with subjective knowledge (user's context)? | The knowledge management is transparent? | Idea / Implemented |
|-------------------------|--------------------------------------------|---------------------------------------------------|------------------------------------------|--------------------|
| Graupmann et al. 2005 | X | | | IMPL. |
| Park e Cheyer 2006 | X | X | X | IDEA |
| Michlmayr, et al. 2007 | | X | X | IMPL. |
| Aleman-Meza et al. 2003 | X | | X | IMPL. |
| Mani e Sundaram 2007 | X | X | X | IMPL. |
| Challam et al. 2007 | X | X? | X | IMPL. |
| Vallet et al. 2006 | X | X? | X | IMPL. |
| Sieg et al. 2007 | X | X? | X | IMPL. (incomplete) |
| Praestro | X | X | X | IMPL. (prototype) |

Trabalhos futuros

- Processos parcialmente automatizados para geração e atualização de ontologias e anotações semânticas ???
- Experimentos para comprovar a eficácia das técnicas e ferramentas utilizadas
- Validação de soluções de problemas de busca semântica em diversos domínios de aplicação
- Determinação de níveis adequados de acoplamento entre contextos e ontologias e dessas com os recursos a serem recuperados

Notícias e análises

- <http://news.cnet.com/new-search-engine-cuil-takes-aim-at-google/>
- http://www.readwriteweb.com/archives/is_google_a_semantic_search_engine.php
- http://www.readwriteweb.com/archives/semantic_search_the_myth_and_reality.php
- <http://www.pandia.com/sew/1262-top-5-semantic-search-engines.html>
- <http://mindset.research.yahoo.com/>
- http://news.cnet.com/8301-13953_3-9982015-80.html
- <http://www.searchenginejournal.com/askcom-focuses-on-semantic-search/8252/>
- http://news.bbc.co.uk/2/hi/programmes/click_online/default.stm
- http://news.bbc.co.uk/2/hi/programmes/click_online/8144765.stm

Web semântica e buscas semânticas

- <http://www.w3.org/2001/sw/>
- <http://semanticweb.org/>
- http://en.wikipedia.org/wiki/Semantic_Web
- http://en.wikipedia.org/wiki/Semantic_search

Artigos básicos

- <http://www.aifb.uni-karlsruhe.de/WBS/pha/publications/ontology-ir-ictir07.pdf>
- <http://portal.acm.org/citation.cfm?id=1359840>
- http://espace.library.curtin.edu.au/R/?func=dbin-jump-full&object_id=116029&local_base=GEN01
- <http://www.seco.tkk.fi/publications/2005/makela-semantic-search-2005.pdf>

Referências em buscas semânticas

- Guha, R., McCool, R., and Miller, E. 2003. **Semantic search**. In *Proc. of the 12th international Conference on World Wide Web (WWW)*, Budapest, Hungary. ACM, New York, NY, 2003, 700-709.
- Reeve, L. and Han, H. 2005. **Survey of semantic annotation platforms**. In *Proceedings of the ACM Symposium on Applied Computing*, Santa Fe, New Mexico. ACM, New York, NY, 2005, 1634-1638.
- Eetu Makela. **Survey of Semantic Search Research**
- Mangold, C. **A Survey and classification of semantic search approaches**. Journal of Metadata Semantics and Ontology, 2(1), 2007.
- Stephan Bloehdorn, Philipp Cimiano, Alistair Duke, Peter Haase, Jörg Heizmann, Ian Thurlow, Johanna Völker. **Ontology-Based Question Answering for Digital Libraries**. ECDL 2007: 14-25.
- Hai Dong; Hussain, F.K.; Chang, E. **A survey in semantic search technologies**. In *2nd IEEE International Conference on Digital Ecosystems and Technologies*. 2008, 403-408

Algumas referências grupo UFSC

- D'Agostini, C. S. ; Fileto, R. **Capturing Users' Preferences and Intentions in a Semantic Search System**. In: *21st International Conference on Software Engineering & Knowledge Engineering (SEKE)*, Boston, 2009. p. 587-591.
- D'Agostini, C. S. ; Fileto, R. ; Dantas, M. A. R. ; Gauthier, F. A. O. **Contextual Semantic Search - Capturing, using the User's Context to Direct Semantic Search**. In: *10th International Conference on Enterprise Information Systems (ICEIS)*, Barcelona, Spain, 2008. v. SAIC. p. 154-159.
- D'Agostini, C. S. ; Fileto, R. **Capturing and managing user context for improving information retrieval**. In: *Workshop de Teses e Dissertações do Simpósio Brasileiro de Bancos de Dados (WTDBD/SBBD)*, Campinas: Unicamp, Brazil, 2008.
- Vian, J. ; Silveira, R. A. ; Fileto, R. **Proposal of a Multi-agent System for Indexing and Recovery applied to Learning Objects**. In: *9th IFIP World Conference on Computers in Education (WCCE)*, Bento Gonçalves, Brazil, 2009.

Perguntas?

