

Ontology-Based Data Access

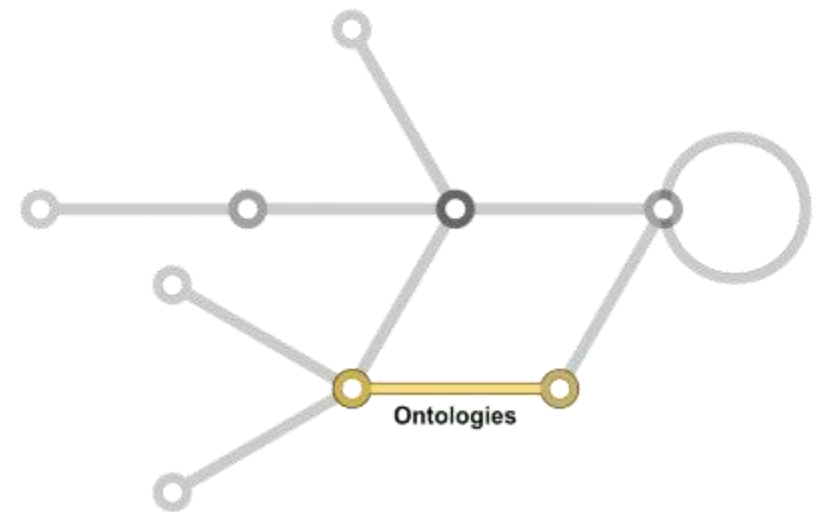
NICOLLE CHAVES CYSNEIROS

Roteiro

- ▶ Introdução
- ▶ Aplicações
- ▶ Arquitetura de um Sistema OBDA
- ▶ Trabalhos Relacionados
 - ▶ Bringing Relational Databases into the SemanticWeb: A Survey
 - ▶ Mastro
 - ▶ OBDA em Repositórios NoSQL
- ▶ Conclusões
- ▶ Referências

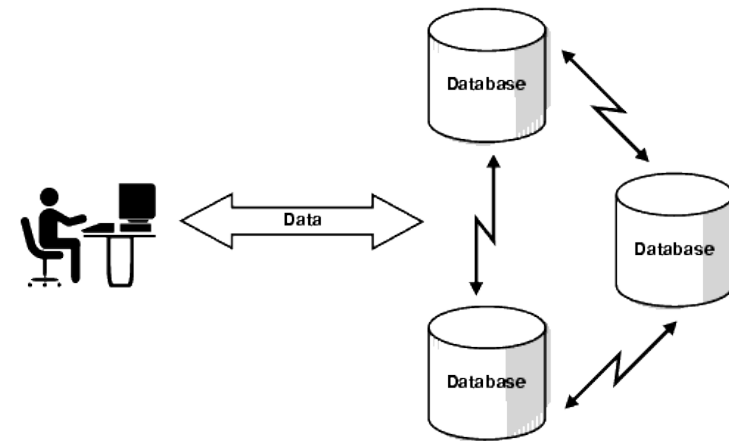
Introdução – O que é um Sistema OBDA

- ▶ Ontologia usada como uma visão conceitual de alto-nível dos repositórios de dados disponíveis.
- ▶ Usuário acessa os dados sem ter o conhecimento específico de como eles estão organizados em suas fontes



Aplicações OBDA

- ▶ Sistemas onde os dados **não** estão organizados de forma uniforme e coerente
- ▶ Banco de Dados manipulados diversas vezes ao longo do tempo
- ▶ Sistemas distribuídos



Aplicações OBDA – Big Data

- ▶ De acordo com Calvanese et al. [5], uma solução OBDA é ideal para uma aplicação Big Data uma vez que:

i.

A ontologia provê **cobertura semântica** para a variedade de dados;

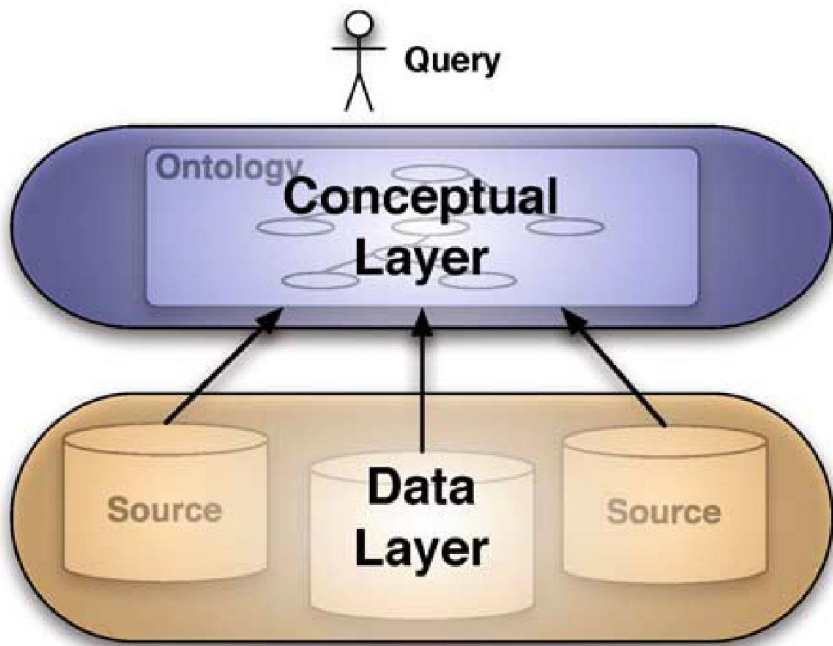
ii.

O acesso a dados no momento em que a consulta é realizada permite a obtenção de **dados recentes** independentemente da velocidade de suas mudanças;

iii.

A integração virtual dos dados permite o gerenciamento de **grandes volumes de dados**

Arquitetura de um Sistema OBDA



- ▶ Camada semântica (ontologia)
- ▶ Mapeamento entre os conceitos da camada semântica e as entidades armazenadas nas camada de dados
- ▶ Camada de dados (fontes)

Arquitetura de um Sistema OBDA

- ▶ A camada semântica é responsável por prover uma visão uniforme dos dados armazenados nas fontes de dados
- ▶ Permite a utilização de raciocinadores para a checagem de inconsistência dos dados e para a inferência de novos conhecimentos
- ▶ Definição da linguagem e estrutura dos mapeamentos

RDF e Sistemas OBDA

- ▶ Resource Description Framework (RDF) é um dos padrões mais utilizados para a descrição do esquema global de Sistemas OBDA.
- ▶ Conhecimento semântico descrito através de classes do RDF Schema
- ▶ Processamento de consultas e inferências utilizando SPARQL

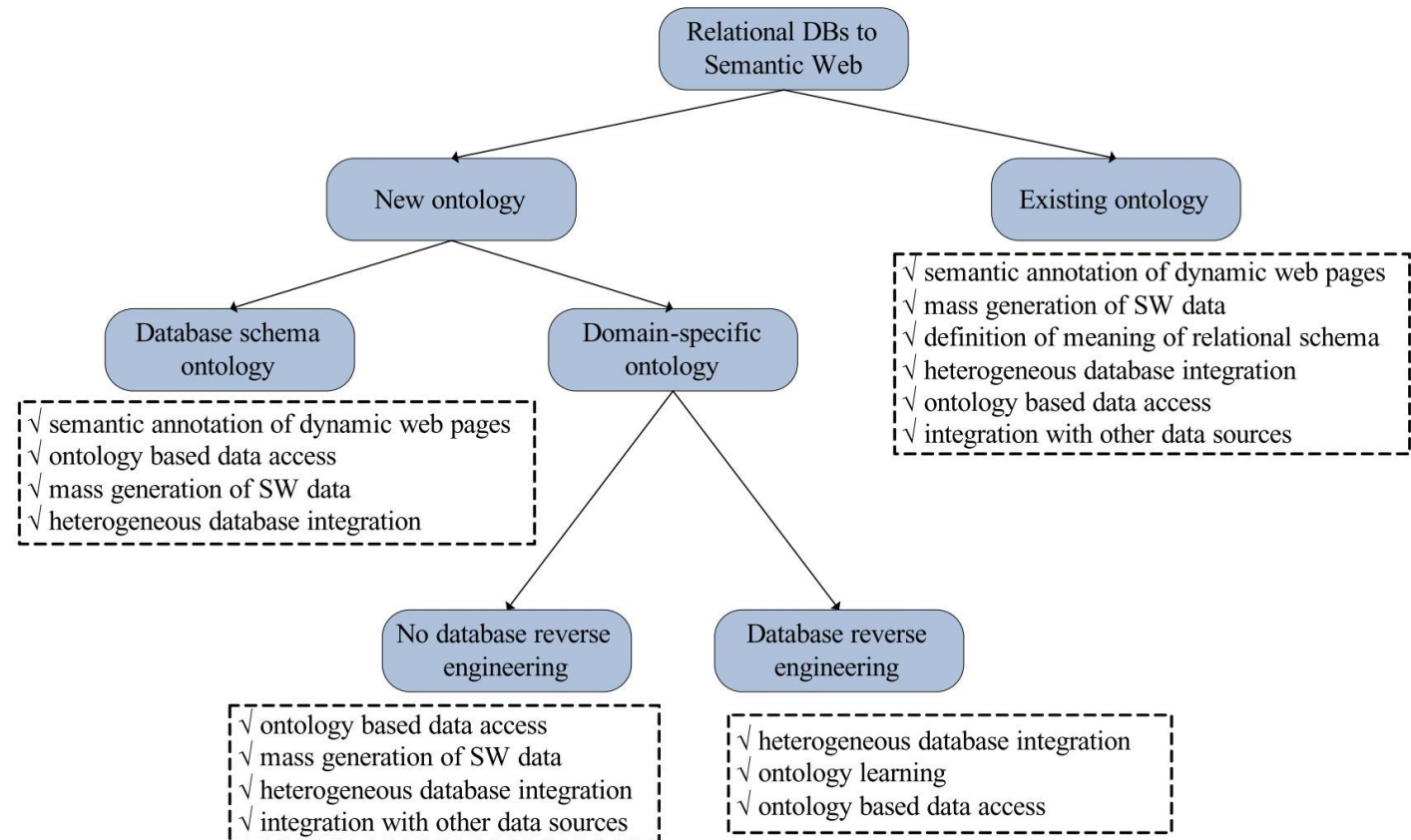
Trabalhos Relacionados

Bringing Relational Databases into the SemanticWeb: A Survey [2]

- ▶ Reúne trabalhos publicados entre 2007 e 2011 que têm como objetivo disponibilizar dados armazenados em bancos de dados relacionais em aplicações de Web Semântica.
- ▶ Propõe uma classificação para as abordagens que visam integrar ontologias e bancos de dados relacionais

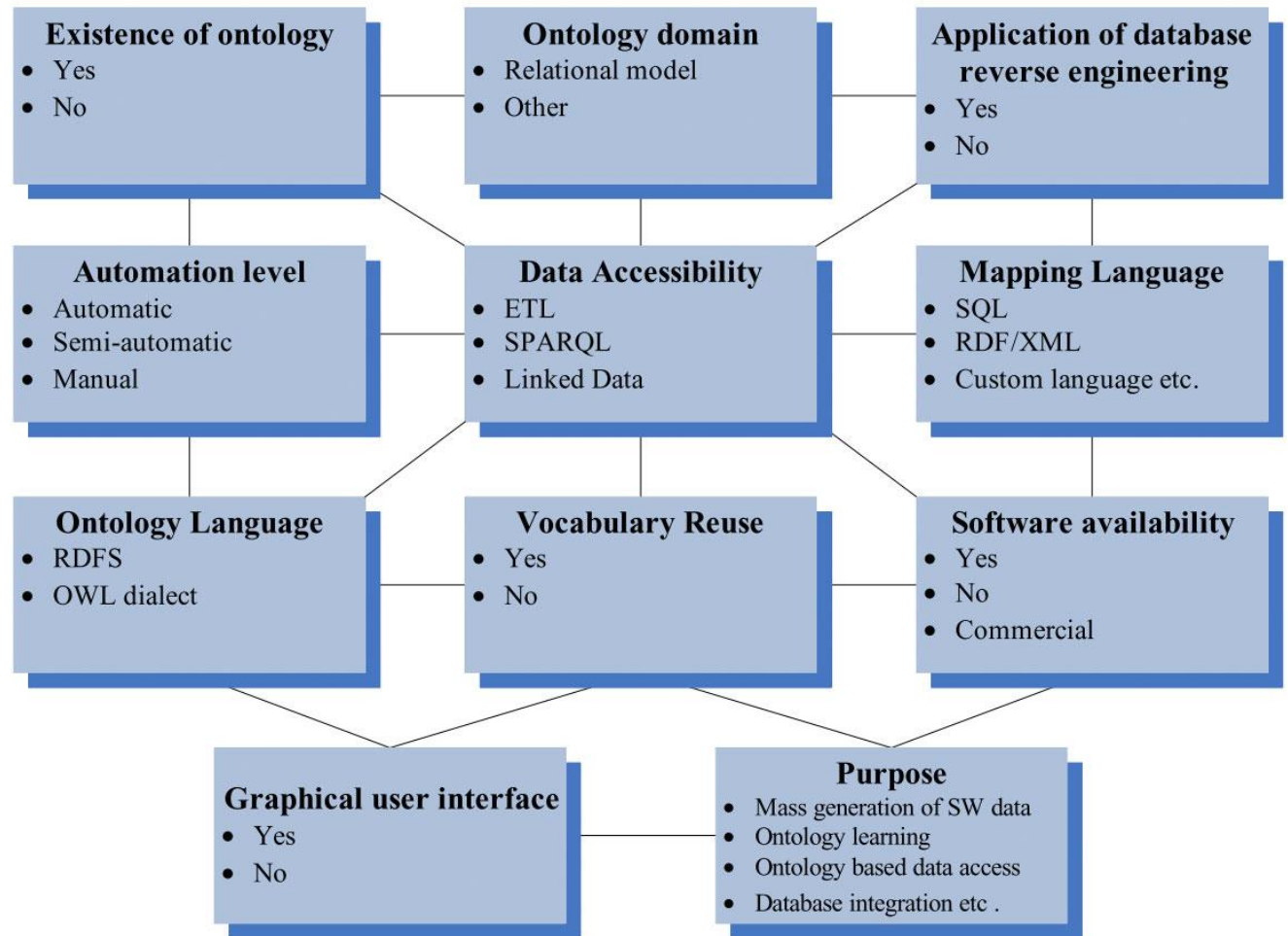
Bringing Relational Databases into the Semantic Web: A Survey

Classificação das soluções de acordo com a técnica utilizada para a definição de mapeamentos



Bringing Relational Databases into the SemanticWeb: A Survey

Características de parâmetros descritivos das abordagens



Trabalhos Relacionados

Mastro [4]

- ▶ Um sistema que possibilita a criação e manutenção de um Sistema OBDA
- ▶ Grupo de pesquisa das universidades *Free University of Bozen-Bolzano* e *Sapienza Universita di Roma*
- ▶ Base teórica bastante sólida

Trabalhos Relacionados

Mastro

- Utiliza a linguagem **DL-Lite**_{A,id}, com expressividade para a modelagem e especificação de restrições de integridade dos dados [1]

$$\text{tempEmp} \sqsubseteq \text{employee} \quad (1)$$

$$\text{manager} \sqsubseteq \text{employee} \quad (2)$$

$$\text{employee} \sqsubseteq \exists \text{WORKS-FOR.project} \quad (3)$$

$$\delta(\text{until}) \sqsubseteq \text{WORKS-FOR} \quad (4)$$

$$(\text{funct until}) \quad (5)$$

$$\text{tempEmp} \sqsubseteq \exists \delta(\text{until}) \quad (6)$$

$$\text{manager} \sqsubseteq \exists \text{MANAGES} \quad (7)$$

$$\text{MANAGES} \sqsubseteq \text{WORKS-FOR} \quad (8)$$

$$\text{manager} \sqsubseteq \neg \exists \delta(\text{until}) \quad (9)$$

$$\rho(\text{until}) \sqsubseteq \text{xsd:date} \quad (10)$$

Trabalhos Relacionados

Mastro

- ▶ O mecanismo de mapeamento adotado permite a resolução de problemas de *impedance mismatch*
- ▶ Consultas são processadas e otimizadas de forma a serem executadas em complexidade LogSpace
- ▶ Desenvolvido em Java, permite conexões JDBC com bancos de dados e possui uma API própria, além de uma interface compatível com OWLAPI e um plugin do Protégè

Trabalhos Relacionados

OBDA em Repositórios NoSQL [3]

- ▶ Devido à sua característica “schemaless”, algumas funcionalidades comuns em bancos de dados tradicionais não são facilmente implementadas em repositórios NoSQL
 - ▶ Otimização de consultas
 - ▶ Detecção de violação a restrições de integridade

Trabalhos Relacionados

OBDA em Repositórios NoSQL

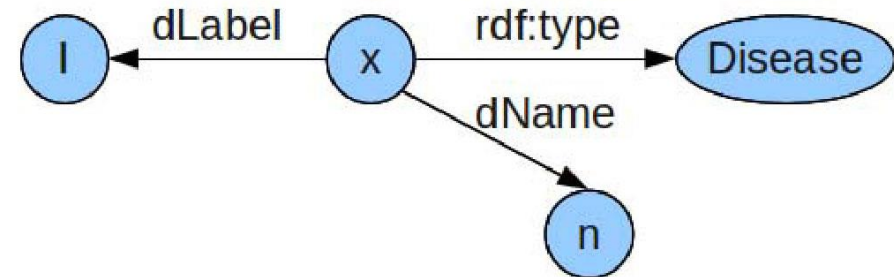
- ▶ Os mapeamentos são descritos em uma linguagem baseada no caminho de acesso aos atributos.
- ▶ Modelo de dados descrito em RDF e consultas feitas em SPARQL
- ▶ Cada consulta SPARQL é reescrita de acordo com os mapeamentos para a linguagem de consulta nativa do repositório NoSQL

Trabalhos Relacionados

OBDA em Repositórios NoSQL

- ▶ Exemplo de mapeamento e representação da consulta em um grafo RDF:

1	$\overleftarrow{k, n, c}$	<i>Disease(k), dName(k, n), dLabel(k, c), Patient(p) sufferFrom(p, k)</i>	Disease(<PKEY AS k; disName AS n, disComment AS c, [p p←Patients]>)
2	$\overleftarrow{k, n, f, g}$	<i>Patient(k), lastName(k, n), firstname(k, f), gender(k, g) Disease(d), sufferFrom(k, d)</i>	Patient(<PKEY AS k; IName AS n fName AS f, gender AS g, [d d ← Diseases]>)



Conclusões

- ▶ Crescente demanda por aplicações que permitam acesso unificado à fontes de dados heterogêneas
- ▶ Variedade de abordagens propostas
- ▶ Ainda faltam padrões e formalizações a serem oficializados e adotados
- ▶ Facilitar integração com sistemas legados

Conclusões

- ▶ Permitir atualização dos dados através da ontologia
- ▶ Ferramentas para facilitar a atualização de mapeamentos
- ▶ Reuso do vocabulário utilizado para a geração de Linked Data.

Referências

- [1] D. Calvanese, G. De Giacomo, D. Lembo, M. Lenzerini, A. Poggi, and R. Rosati, “**Linking data to ontologies: The description logic DL-LiteA**,” *CEUR Workshop Proc.*, vol. 216, 2006.
- [2] D.-E. Spanos, P. Stavrou, and N. Mitrou, “**Bringing relational databases into the Semantic Web: A survey**,” *Semant. Web*, vol. 3, no. 2, pp. 169–209, 2012.
- [3] O. Cure, F. Kerdjoudj, Chan Le Duc, M. Lamolle, and D. Faye, “**On the Potential Integration of an Ontology-Based Data Access Approach in NoSQL Stores**,” *2012 Third Int. Conf. Emerg. Intell. Data Web Technol.*, pp. 166–173, 2012.

Referências

- [4] D. Calvanese, G. De Giacomo, D. Lembo, M. Lenzerini, A. Poggi, M. Rodriguez-Muro, R. Rosati, M. Ruzzi, and D. F. Savo, “**The MASTRO system for ontology-based data access**,” *Semant. Web*, vol. 2, no. 1, pp. 43–53, 2011.
- [5] D. Calvanese, M. Giese, P. Haase, I. Horrocks, T. Hubauer, Y. Ioannidis, E. Jiménez-Ruiz, E. Kharlamov, H. Kllapi, J. Klüwer, M. Koubarakis, S. Lamparter, R. Möller, C. Neuenstadt, T. Nordtveit, Ö. Özcep, M. Rodriguez-Muro, M. Roshchin, F. Savo, M. Schmidt, A. Soylu, A. Waaler, and D. Zheleznyakov, “**Optique: OBDA solution for big data**,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 7955 LNCS, pp. 293–295, 2013.