

Interoperabilidade Semântica entre Sistemas de Resposta à Emergências

Rebeca Barros^{1,2}, Vaninha Vieira^{1,2},
Laís Salvador^{1,2}, Reinaldo Almeida¹

¹Departamento de Ciência da Computação
Universidade Federal da Bahia (UFBA)

²Fraunhofer Project Center for Software
and Systems Engineering at UFBA

{rebecasbarros, vaninha, laisns}@dcc.ufba.br, reifa28@gmail.com

Abstract. *Emergency Management is a complex task which involves the communication and collaboration among different organizations and its systems. Data Integration and Interoperability within systems are among the biggest challenges in this area. As a solution, this paper presents a proposal for data integration for Emergency Response Systems (ERS) using a suite of ontologies created based on the EDXL (Emergency Data Exchange Language) standard. The results indicated that EDXL is suitable in local scenario and that the ontologies can be applied as interlingua among ERS.*

Resumo. *Gerenciamento de emergência é uma tarefa complexa pois envolve comunicação e colaboração entre variadas organizações e seus sistemas. Integração de dados e interoperabilidade de sistemas estão entre os maiores desafios nesta área. Como solução, este artigo apresenta uma proposta de integração de dados para Sistemas de Resposta à Emergências (SRE) baseado em uma suíte de ontologias criadas a partir do padrão EDXL (Emergency Data Exchange Language). Os resultados indicam que EDXL é adequado ao cenário local e que as ontologias podem ser aplicadas como interlíngua entre os SREs.*

1. Introdução

Emergência é um evento imprevisto e repentino que exige medidas imediatas para minimizar suas consequências e riscos à vida e à propriedade [DHA 1992]. O Gerenciamento de Emergências constitui a coordenação de atividades envolvidas na preparação, suporte e reconstrução da sociedade quando desastres naturais ou provocados pelo homem ocorrem [Catarci et al. 2008]. Um dos maiores desafios é coordenar ações e permitir a comunicação considerando que a troca de informações e dados envolve múltiplas organizações, especialistas de diversos domínios, além de diferentes sistemas. Por serem criados e mantidos individualmente pelas diversas organizações e/ou sistemas, esses dados são heterogêneos, o que traz problemas de interoperabilidade quando o objetivo é integrá-los.

A Interoperabilidade tem como objetivo superar o problema da heterogeneidade (em todas as formas) e permitir a manipulação de dados de forma transparente entre as diversas fontes. A integração semântica de dados é o processo de usar uma

representação conceitual dos dados e suas relações (geralmente uma ontologia) para eliminar as possíveis heterogeneidades. Nesse cenário, busca-se integrar semanticamente dados de sistemas de resposta à emergência considerando a interoperabilidade entre eles, que a integração ocorra em tempo de execução e que as ontologias representem o domínio de emergência e não tão somente as fontes de dados.

O objetivo deste artigo é apresentar um ambiente de integração de dados semântico para SREs baseado em uma suíte de ontologias criadas a partir do padrão EDXL, um padrão bem aceito e utilizado na comunidade do domínio de emergência. A solução proposta foi validada no escopo do projeto RESCUER¹, que visa o desenvolvimento de uma plataforma que combine informações vindas da *crowdsourcing* com dados abertos para oferecer suporte ao gerenciamento de emergências. O restante deste artigo é organizado da seguinte forma: a Seção 2 apresenta o conceito de Acesso a Dados baseado em Ontologias; a Seção 3 contém os trabalhos relacionados; a Seção 4 contém os detalhes para a implementação do ambiente de integração semântica; a avaliação é discutida na Seção 5; e a Seção 6 faz as considerações finais.

2. Acesso a Dados Baseado em Ontologias

O Acesso a Dados Baseado em Ontologias, do inglês *Ontology Based Data Access* (OBDA), refere-se ao acesso a uma base de dados (ou uma coleção deles) através de uma camada conceitual de alto nível, especificada como uma ontologia, permitindo o acesso sem a necessidade de saber como os dados estão organizados e como são armazenados em cada uma das bases. OBDA pode ser visto como um sistema de integração de dados centralizado que utiliza uma abordagem de ontologia única (expressa usando lógica de descrição, OWL DL) e mapeamentos GaV (Global-as-View), onde cada entidade no esquema global é definida por consultas sob as fontes de dados [Calvanese et al. 2011].

Assim como outros sistemas de integração de dados, OBDA usa os mapeamentos para especificar a correspondência semântica entre uma versão unificada do domínio (esquema global) e os dados armazenados nas fontes. Uma particularidade, é que essa visão é construída a partir da conceitualização do domínio de interesse, independente da representação adotada para os dados armazenados nas fontes. Um outro fator é que os raciocinadores atuais que agem sobre lógica de descrição ainda não conseguem lidar com uma quantidade grande de dados (ABox² extenso). A fim de contornar esse problema, em OBDA, o ABox não é representado diretamente, ou seja, os dados não são carregados na ontologia, consistindo em um banco de dados conectado ao TBox² através de mapeamentos e que é acessado em tempo de execução [Calvanese et al. 2011].

3. Trabalhos Relacionados

Ontologias tornaram-se um campo de aplicação recorrente nas pesquisas relacionadas ao gerenciamento de emergência. Diversos trabalhos discutem o uso de ontologias em problemas como interoperabilidade e integração de dados, descoberta de serviços, recuperação eficiente de informação e apoio a tomada de decisões. O sistema de suporte SoKNOS [Babitski et al. 2011] é um protótipo para sistemas de gerenciamento a

¹www.rescuer-project.org

² Em ontologias do tipo DL um conjunto Tbox (*Terminological Box*), são os conceitos e propriedades de um domínio e um Abox (*Assertional Box*), os indivíduos que pertencem aos conceitos do Tbox.

emergência que faz uso de tecnologias semânticas para vários propósitos. O projeto desenvolveu e faz uso de sete ontologias distintas, sendo um dos casos de uso na Integração de Dados. Em [Zlatanova et al. 2010] é proposto o uso de ontologias no cenário de resposta a emergência para realização de consultas em coleções distintas de dados geográficos. Para tal, uma ontologia local é construída para cada base de dados existente e uma ontologia de alto nível é construída baseada nessas ontologias locais.

Embora semelhante à proposta deste trabalho, a abordagem de [Babitski et al. 2011], envolve uma camada extra, já que ocorrem dois mapeamentos: um entre a fonte de dados e a ontologia criada para essa fonte de dados e um entre a ontologia da fonte de dados e a ontologia base. Além disso, a proposta aqui apresentada visa a utilização de tecnologias da Web Semântica, o que aumenta a possibilidade de reuso dos artefatos produzidos. Na abordagem de [Zlatanova et al. 2010] as classes e propriedades na ontologia alto nível são criadas com prefixos que remetem a qual coleção de dados elas pertencem, o que não torna essa representação relevante a nível de conhecimento de domínio. Diferentemente, as ontologias deste trabalho, ao tomarem como base o padrão EDXL, refletem um conhecimento de domínio específico e ainda foram construídas de forma modular, tal qual o padrão, onde cada uma das ontologias reflete um aspecto particular dentro do domínio do gerenciamento de emergências.

4. Implementação

O ambiente de integração de dados semântico foi implementado na forma de um repositório Sesame³, que funciona como um SPARQL endpoint e pode ser acessado via HTTP por qualquer aplicação para realizar consultas SPARQL. Dessa forma, qualquer SRE pode recuperar dados, que estão armazenados em fontes heterogêneas, sobre uma situação de emergência usando apenas o vocabulário da suíte de ontologias.

A Figura 1 destaca os componentes do ambiente de integração. Três bases de dados distintas foram utilizadas. Duas são conjuntos de dados (*datasets*) de duas organizações envolvidas no gerenciamento de emergência: o conjunto de dados de Desastres do Governo Canadense⁴ e o conjunto de dados do Cecoco, um dos sistemas que registra novas ocorrências utilizado pela Superintendência de Telecomunicações - SSP/BA⁵. A outra é um banco relacional MySQL cujo modelo Entidade-Relacionamento baseia-se na aplicação do projeto RESCUER. Para apoiar a implementação duas ferramentas foram utilizadas. A primeira delas é o Teiid⁶, um sistema de virtualização de dados que permite aplicações usarem dados de múltiplas fontes através de uma única interface JDBC. A segunda é o Ontop⁷, que apoia a construção de um ambiente OBDA.

Como entrada do sistema estão:

- As ontologias: a implementação fez uso principal da Ontologia EDXL-SitRep, cujo domínio é descrever as situações de emergência, mas, como essa ontologia importa conceitos da EDXL-CAP e da EDXL-RM, essas também foram utilizadas.

³ <http://rdf4j.org/>

⁴ <http://open.canada.ca/data/en/dataset/1c3d15f9-9cfa-4010-8462-0d67e493d9b9>

⁵ <http://www.stelecom.ba.gov.br/stelecom/index.php>

⁶ <http://teiid.jboss.org/>

⁷ <http://ontop.inf.unibz.it/>

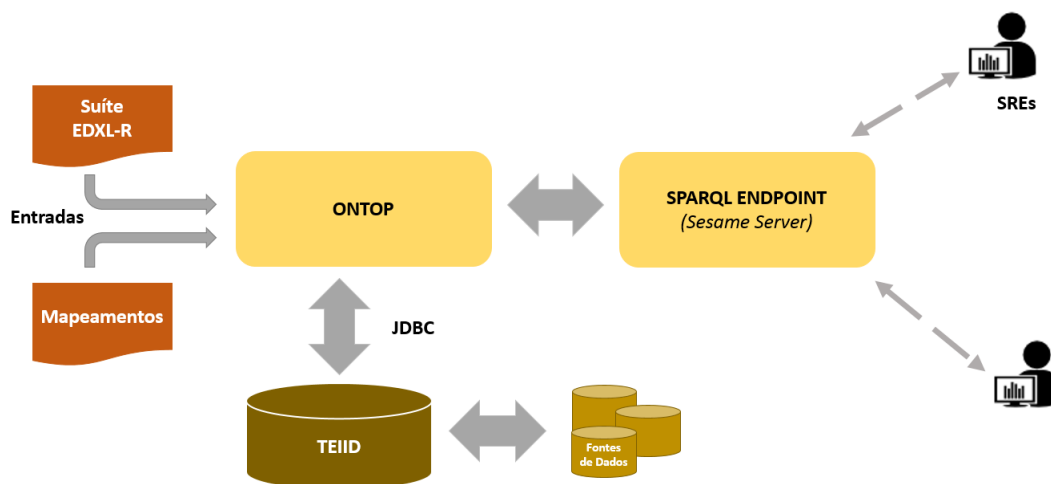


Figura 1. Ambiente de Integração

- Arquivo de Mapeamentos: arquivos que definem a associação entre os dados das bases e o vocabulário da ontologia. São descritos na linguagem de mapeamento nativa do Ontop e possuem a extensão **.obda**.
- Banco de Dados: representam as bases de dados e são conectadas ao sistema através da conexão JDBC com o Teiid.

Código 1. Exemplo de Mapeamento

```
mappingId Incident-MySQL
target  sitrep:incident/{idIncident}/ a sitrep:Incident ; sitrep:incidentID
{idIncident} ; sitrep:fromDateTime {timest} .
source  SELECT idIncident, timest FROM incident;
```

O Código 1 é um exemplo de declaração de mapeamento na linguagem nativa do Ontop, onde `mappingId`, é o nome do mapeamento; `target`, é a declaração em OWL; e `source`, a consulta SQL para recuperar as informações no banco de dados relacional. Nesse exemplo, a ideia é mapear indivíduos da classe `Incident`, assim, é declarado que os valores das colunas `idIncident` e `timest` da tabela `incident` (de um banco relacional) serão usadas para instanciar um indivíduo cuja a IRI⁸ será composta pela concatenação da IRI da ontologia (nesse caso, `sitrep`), a *string* “`incident`” e o valor da coluna `idIncident`. Esse indivíduo pertencerá a classe `Incident` (a `sitrep:Incident`), além disso, os valores das propriedades de dados `incidentID` e `fromDateTime` desse indivíduo serão oriundos das colunas `idIncident` e `timest`, respectivamente.

O arquivo de mapeamento foi criado considerando as três fontes de dados, ou seja, um conceito na ontologia poderia ter mais de um mapeamento, sendo cada um deles correspondente a uma fonte de dados. O exemplo mais claro disso é o conceito de *Incident*, presente em todas as bases utilizadas. O mapeamento da base MySQL é o destacado no Código 1.

⁸Internationalized Resource Identifier

4.1. Suíte de Ontologias EDXL-RESCUER

EDXL consiste em um conjunto de padrões de mensagens baseados na linguagem XML que tem como objetivo facilitar o compartilhamento de informações entre entidades, organizações e pessoas envolvidas em um cenário emergencial [OASIS 2014]. Embora resolva o problema da heterogeneidade sintática e estrutural, visto que é uma suíte de XML-Schemas que padronizam o formato XML a ser trocado pelos sistemas, o EDXL não garante a interoperabilidade semântica. Dessa maneira, foram criadas ontologias, baseadas nesse padrão, para apoiar a integração semântica em sistemas de resposta à emergências.

A suíte denominada EDXL-RESCUER é composta por quatro ontologias baseadas em quatro dos padrões definidos no EDXL. São eles: EDXL-CAP - Protocolo de Alerta Comum; EDXL-DE - Elemento de Distribuição; EDXL-RM - Mensagens de Recurso e EDXL-Sitrep - Relatórios da Situação. As ontologias propostas foram formalizadas em OWL 2 QL⁹ e foram construídas com o apoio do editor Protégé¹⁰. A metodologia para construção das ontologias e descrição são detalhadas nos trabalhos anteriores [Barros et al. 2015a, Barros et al. 2015b].

5. Avaliação

Uma das formas de avaliação para ontologias é através da validação dos conceitos e propriedades por especialistas de domínio. Um questionário foi aplicado com especialistas no gerenciamento de emergência onde foram apresentados os conceitos da ontologia e suas classificações e onde eles receberam a tarefa de avaliar o quanto estavam de acordo com a existência de cada conceito. Ao total, cada um dos 73 conceitos foi avaliado por 6 especialistas, resultando em 438 respostas. Os especialistas possuem uma média de 16 anos de experiência e atuaram em funções como, Comando da Polícia Militar e do Subgruposamento de Bombeiros. Dentro da escala Likert adotada, uma média de aproximadamente 88% das respostas (51% para Concordo Totalmente e 37% Concordo) indicaram que os especialistas estavam de acordo com os conceitos apresentados.

No que diz respeito à integração semântica de dados, a EDXL-RESCUER mostrou-se viável como interlíngua entre os SREs. Das três bases utilizadas foi possível encontrar uma média de 69% de correspondência entre os conceitos existentes nas ontologias e os presentes nas fontes de dados. É importante ressaltar que o sistema Cecoco, por administrar as chamadas do 190, contém também informações a respeito do atendimento, que não estão relacionadas à emergência. Por isso uma filtragem foi realizada previamente nessa base para exclusão dos dados referentes ao atendimento.

Nessa fase, foram realizadas consultas utilizando a linguagem SPARQL a fim de responder as questões de competência determinantes à existência da ontologia EDXL-RESCUER. Consultas sobre detalhes do incidente, seu tipo e nível de confiabilidade da informação; a quantidade de vítimas afetadas por ele e em que estados elas se encontram; e até sobre a estimativa de gastos. Todas essas consultas foram realizadas no repositório criado e retornaram dados que estavam presentes nas bases integradas. Por fim, a EDXL-RESCUER foi validada pelos *reasoners* Hermit 1.3.8 e Quest 1.15.0 (Ontop).

⁹Perfil baseado na família DL-Lite, garante que consultas feitas na ontologia possam ser reescritas em consultas equivalentes sobre as bases de dados.

¹⁰<http://protege.stanford.edu/>

6. Conclusões e Trabalhos Futuros

Esse artigo propôs um ambiente de integração para SREs baseado em uma suíte de ontologias criadas a partir do padrão EDXL. OBDA foi a abordagem escolhida para o ambiente que foi exemplificado pela construção de um repositório disponível na forma de um SPARQL endpoint. Essa abordagem apresenta características apropriadas para o cenário proposto, como o uso de apenas uma ontologia global que representa o conhecimento de um domínio e o fato de recuperar dados nas fontes de dados em tempo de execução.

As principais contribuições são: os resultados sobre o uso dos protocolos da família EDXL como alternativa para a padronização de vocabulário no que se trata a sistemas interoperáveis de gerenciamento de emergência; e o desenvolvimento da integração de dados no estudo de caso com o projeto RESCUER. Adicionalmente, a integração de dados utilizando OBDA demonstra o uso de tecnologias padrões da Web Semântica no contexto dos SREs, permitindo o acesso a dados através de uma camada conceitual composta por uma ontologia. Como trabalhos futuros, pretende-se: (i) alinhar as ontologias criadas com ontologias de alto nível consolidadas (como a DOLCE); (ii) investigar OBDA no cenário de bancos de dados NoSQL; (iii) investigar LOD (Linked Open Data) e a possibilidade de publicação do SPARQL endpoint na web;

Referências

- Babitski, G., Bergweiler, S., Grebner, O., Oberle, D., Paulheim, H., and Probst, F. (2011). Soknos—using semantic technologies in disaster management software. In *The Semantic Web: Research and Applications*, pages 183–197. Springer.
- Barros, R., Kislansky, P., Salvador, L., Almeida, R., Breyer, M., and Gasparin, L. (2015a). Edxl-rescuer ontology: Conceptual model for semantic integration. In *Proceedings of the 12th International ISCRAM Conference*.
- Barros, R., Kislansky, P., Salvador, L., Almeida, R., Breyer, M., Pedraza, L. G., and Vieira, V. (2015b). Edxl-rescuer ontology: an update based on faceted taxonomy approach. In *Proceedings of the Brazilian Seminar on Ontologies (ONTOBRAS 2015)*.
- Calvanese, D., De Giacomo, G., Lembo, D., Lenzerini, M., Poggi, A., Rodriguez-Muro, M., Rosati, R., Ruzzi, M., and Savo, D. F. (2011). The mastro system for ontology-based data access. *Semantic Web*, 2(1):43–53.
- Catarci, T., de Leoni, M., Marrella, A., Mecella, M., Salvatore, B., Vetere, G., Dustdar, S., Juszczak, L., Manzoor, A., and Truong, H.-L. (2008). Pervasive software environments for supporting disaster responses. *Internet Computing, IEEE*, 12(1):26–37.
- DHA, U. (1992). Internationally agreed glossary of basic terms related to disaster management. *UN DHA (United Nations Department of Humanitarian Affairs), Geneva*.
- OASIS (2014). Oasis emergency management tc. https://www.oasis-open.org/committees/tc_home.php?wg_abbrev=emergency.
- Zlatanova, S., De Vries, M., and Van Oosterom, P. (2010). Ontology-based query of two dutch topographic data sets: an emergency response case. In *Proceedings Core Spatial Databases-Updating, Maintenance and Services—from Theory to Practice, Haifa, Israel, 15-17 March 2010; IAPRS, XXXVIII (4-8-2/W9), 2010*. International Society of Photogrammetry and Remote Sensing (ISPRS).