

Integração de Padrões para Transferência de Informações Digitais no Fluxo de Trabalho de Modelagem de Reservatórios Baseada em Ontologias

Ricardo Werlang¹, Mara Abel¹

¹Instituto de Informática – Universidade Federal do Rio Grande do Sul (UFRGS)
Porto Alegre – RS – Brasil

{rwerlang,marabel}@inf.ufrgs.br

Abstract. *In the workflow of petroleum reservoir characterization, experts seek an integrated vision of the data issued from the same oil field, generated by the application of different techniques and represented in different standards and formats. Also, the same geological objects analyzed by different professionals assume distinct semantic representations and complementary in supporting decision-making. In this work, we aim to delimit the most used data and its formats in the construction of structural models and also to propose a semantic-based integrated approach using ontologies to capture the real meaning of the information and not just the specific technology used to represent it.*

Resumo. *No fluxo de trabalho de caracterização de reservatórios de petróleo, especialistas buscam uma visão integrada dos dados do mesmo campo petrolífero, originados pela aplicação de diferentes técnicas e representados em padrões e formatos diversos. Além disso, os mesmos objetos geológicos analisados por diferentes profissionais assumem significados semânticos e representações distintas e complementares no suporte à tomada de decisão. Nesse trabalho, pretende-se delimitar os dados e os formatos mais utilizados na construção de modelos estruturais e, também, propor uma abordagem de integração semântica utilizando ontologias que capture o significado real da informação e não apenas a tecnologia específica utilizada para representá-la.*

1. Introdução

O processo de modelagem de reservatórios de petróleo envolve uma complexa sucessão de atividades que dependem do tipo de modelo que é construído e das opções que são levadas em consideração para construí-lo. Estas atividades podem ser classificadas de acordo com o objetivo de estudo do modelo de reservatório em construção, que pode estar relacionado, segundo [Perrin and Rainaud 2013], com a geometria, na construção do modelo estrutural; com as características da rocha, na construção do modelo estratigráfico; e com os fluídos de um reservatório de petróleo, no modelo de reservatório.

Grande parte destas atividades, contudo, envolvem o uso de uma grande quantidade de dados, que são gerados diariamente por técnicas distintas e que necessitam ser analisados e interpretados pelos diferentes profissionais envolvidos na construção de modelos de reservatórios. No entanto, para usufruir do valor da informação contida nesses dados em diferentes formatos, os profissionais necessitam de acesso imediato às informações, de uma visão integrada das informações e de um completo gerenciamento do conhecimento já adquirido sobre as informações disponíveis [Soma et al. 2008]. Necessita-se,

portanto, de uma abordagem que permita a integração real dos objetos modelados e não apenas dos formatos de dados. Além disso, a abordagem de integração deve descrever tanto objetos quanto propriedades em uma linguagem uniforme, permitindo a todos os envolvidos o acesso às informações em qualquer etapa do processo de modelagem. Necessita-se, neste caso, de uma abordagem de integração semântica dos dados.

O uso de ontologias para modelar os diversos elementos do domínio, como rocha e petróleo, e de bases de conhecimento para armazenar as informações das instâncias desses elementos, como uma determina porção de rocha ou de petróleo, vem sendo proposto para resolver os problemas de integração de dados enfrentados nas operações de modelagem de reservatórios de petróleo. Além disso, nas últimas décadas, diversos padrões abertos ou proprietários foram propostos e utilizados para transferência de informações digitais entre as diferentes etapas e atividades envolvidas no fluxo de trabalho de modelagem de reservatórios. Esses padrões foram criados para resolver o problema de dados criados em diferentes sistemas com formatos proprietários. A integração dos formatos evidenciou o problema de que objetos geológicos podem assumir diferentes significados em etapas distintas da exploração de petróleo. Embora sejam muitas vezes referenciados pelo mesmo vocabulário, são definidos por atributos distintos, que buscam atender o papel daquele objeto em uma etapa particular da exploração.

Nesse trabalho, serão definidas as atividades que envolvem a interpretação de diferentes formatos de dados no processo de construção de um modelo estrutural e, também, proposta uma abordagem de integração semântica, utilizando ontologias, que resolva os principais problemas enfrentados no processo de modelagem.

O restante do texto está organizado da seguinte maneira: a Seção 2 apresenta os padrões para troca de informações mais utilizados na indústria de petróleo; a Seção 3 contém uma visão geral da cadeia de exploração de petróleo, destacando as atividades da modelagem estrutural; a Seção 4 comporta o uso de ontologias e a abordagem de integração proposta; e a Seção 5 conclui o trabalho, expondo os trabalhos futuros.

2. Padrões para Trocas de Informações Digitais

A criação de uma plataforma de trabalho comum sempre foi uma das maiores preocupações dos fornecedores de softwares para a indústria do petróleo. No entanto, do ponto de vista do usuário, ainda é muito difícil transferir dados de uma plataforma para outra. O principal motivo é o histórico complicado do mercado de geomodelagem, caracterizado por extensões ou reconstruções de produtos, aquisições de softwares e fusões de empresas. Perrin *et al.*, em [Perrin and Rainaud 2013], apresenta a evolução histórica das ferramentas para modelagem, que começaram a surgir em meados dos anos 80, com o objetivo de representar as superfícies geológicas e as propriedades petrofísicas das rochas.

No início de 1990, as descrições de reservatórios só poderiam ser trocadas através da escrita e leitura de arquivos de dados proprietários, que transportavam informações limitadas. Na tentativa de facilitar a comunicação entre os diversos softwares utilizados no fluxo de trabalho para construção de modelos da Terra, diversos estudos foram feitos e algumas tentativas de padronização foram propostas. O objetivo da criação de padrões, que são definidos pelos metadados que dão significado e contexto às informações representadas, é permitir que a semântica dos dados seja revelada.

Entre os padrões mais utilizados na indústria, destaca-se o *Log ASCII Standard*

(LAS), que foi proposto em 1990, pela *Canadian Well Logging Society*¹. O LAS é um padrão utilizado para facilitar e simplificar a troca de informações digitais de dados de *log* de poços. Por ter sido proposto no formato ASCII, que possibilita a importação e exportação para qualquer plataforma, o LAS teve ampla aceitação e utilização na indústria. Contudo, padrões mais recentes, que utilizam tecnologias que promovem a interoperabilidade de uma maneira mais natural, como o XML (*eXtensible Markup Language*), vêm sendo propostos nas últimas décadas e estão sendo cada vez mais utilizados. A Energistics² tem sido a líder na exploração do XML dentro da indústria do petróleo, através da proposta de diversos padrões: o WITSML, como um padrão para transferência de informações de perfuração de poços; o PRODML, para transferência de dados de operações de produção; e o RESQML, para transferência de dados contendo descrições de modelos.

3. Atividades de Modelagem de Reservatórios para Exploração de Petróleo

Exploração de petróleo é uma atividade na qual a aquisição, a distribuição e o uso do conhecimento dos especialistas são críticos para a tomada de decisão. Modelos geológicos são ferramentas chaves para a identificação e caracterização de potenciais reservatórios de hidrocarbonetos. Modelos geológicos são representações, 3D ou 4D, de dados e interpretações relacionadas com recursos do subsolo e são desenvolvidos por diferentes profissionais, como geólogos, geofísicos e engenheiros, que são responsáveis pela evolução de um potencial reservatório de petróleo, através de várias etapas de modelagem conhecida como fluxo de trabalho de modelagem de reservatórios [Mastella 2010].

O fluxo de trabalho de modelagem de reservatórios começa com a definição da área de interesse a ser modelada, conhecida como prospecto. É realizada, então, a aquisição dos dados sísmicos e de perfuração de poços, assim como dos documentos e mapas da geologia regional. Com estes dados, é realizada a construção do modelo estrutural, que é essencial para determinar a localização de armadilhas de hidrocarbonetos e, conseqüentemente, para a identificação de possíveis campos petrolíferos e para uma possível avaliação do volume de óleo disponível. A Figura 1 ilustra as principais atividades realizadas na construção do modelo estrutural, como a interpretação de log de poços, as interpretações geológica e sísmica, o encaixe das marcações de poços e a modelagem de superfície. A descrição detalhada das atividades pode ser encontrada em [Perrin and Rainaud 2013]. Pelo fato de todas estas atividades envolverem uma grande quantidade de dados, enfatizamos os principais formatos de dados utilizados. Entre estes dados, alguns tendem a seguir um determinado padrão, como o LAS e o WITSML, que são os mais utilizados. Outros dados, contudo, não seguem praticamente nenhuma estrutura, como os tipos DAT, SEG-Y, PLO e documentos, planilhas, anotações e mapas geológicos.

Após a construção do modelo estrutural, são construídas as malhas estratigráficas dentro de cada bloco geológico definido no modelo estrutural. Estas malhas são divididas por diversas células, que devem carregar as propriedades petrofísicas das rochas. Para isso, são utilizados dados de estudos de laboratório. Estas propriedades são propagadas para todo o volume utilizando simulações geo-estatísticas. O resultado deste processo é o modelo estratigráfico. Este modelo, juntamente com o resultado do processo de aprimoramento da malha estratigráfica, resulta no modelo de reservatório, utilizado para simular a quantidade de óleo acumulado no subsolo.

¹<http://www.cwls.org/>

²<http://www.energistics.org/>

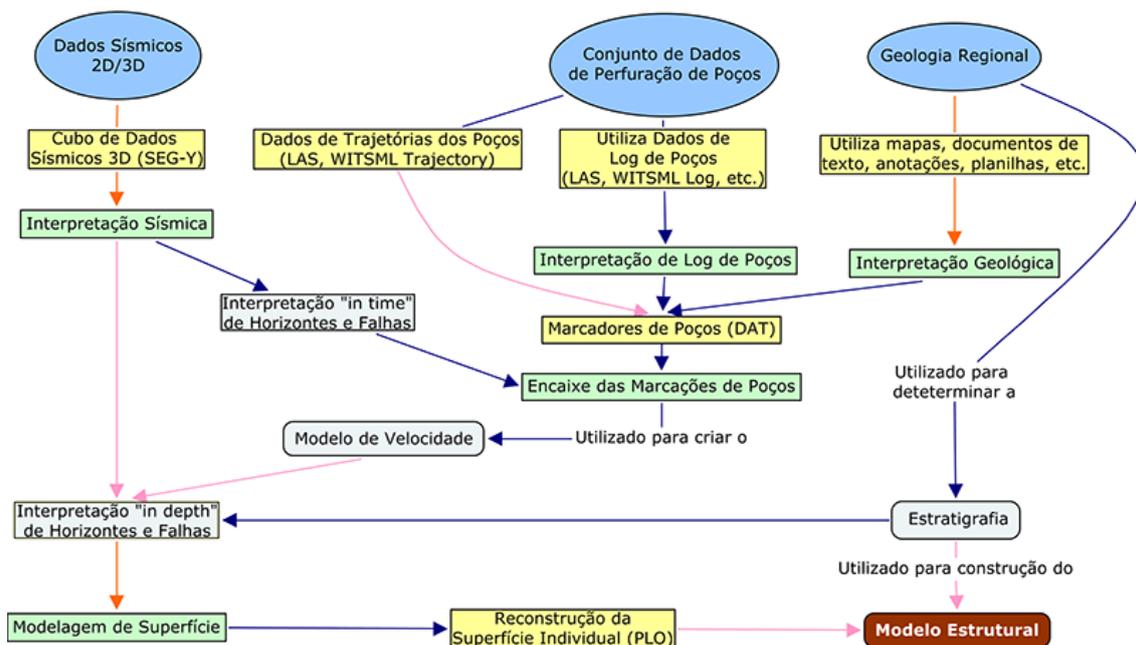


Figura 1. Construção do Modelo Estrutural [Perrin and Rainaud 2013]

4. Integração Baseada em Ontologias

Diversas entidades, companhias de petróleo e consórcios de Geociências vêm emitindo diversas codificações e formalizações do conhecimento geológico durante os últimos anos. Entre os levantamentos geológicos, destacam-se: o modelo GeoSciML, que é uma formalização baseada na normativa GML (*Geography Markup Language*), para a representação de características geográficas e geométricas; o NADM (*North American Geologic Map Data Model*), projetado como uma ontologia para desenvolvimento interoperável de banco de dados geológicos centrados em mapas; e o projeto GEON (*Geosciences Network*)³, para a integração de mapas geológicos, cujos arquivos de origem contêm informações com esquemas e vocabulários diferentes. Entre projetos de companhias de petróleo, mesmo não englobando áreas das Geociências, merece destaque a ontologia proposta no projeto IPP (*Integrated Information Platform*) para formalização de terminologias usadas na etapa de produção de petróleo, baseada no padrão ISO 15926.

Além disso, muitas ontologias foram propostas para domínios específicos, tanto ontologias de domínio quanto ontologias de nível superior. Entre as ontologias de domínio, destacam-se: a ontologia proposta em [Abel et al. 2004], para descrição petrográfica de rochas de reservatório; a ontologia proposta em [Lorenzatti et al. 2009], para modelagem de estruturas sedimentares e atributos textuais de rochas; e a ontologia proposta em [Perrin et al. 2005], que descreve os principais conceitos utilizados na modelagem estrutural. Entre as ontologias de nível superior, destacam-se: a ontologia proposta pelo projeto SWEET (*Semantic Web for Earth and Environmental Terminology*), desenvolvida na NASA, que fornece milhares de termos sobre todo o sistema da Terra [Raskin and Pan 2005]; e a ontologia proposta em [Mastella 2010], a *Basic Geology*, que descreve e interconecta as entidades geológicas consideradas na modelagem de reservatórios. A *Basic Geology* utiliza a *GeoLocation*, uma ontologia com termos geográficos;

³<http://www.geongrid.org/>

ontologias para as disciplinas de *Paleogeography*, *Hydrogeology* e *Lithology*; e ontologias para definição e mapeamento de eras geológicas, *Geological Time* e *Geological Dating*.

Nesse trabalho, propomos a reutilização dessas ontologias em destaque, objetivando a construção de um modelo conceitual único que formalize tanto os conceitos geológicos, que estão envolvidos no processo de modelagem da Terra, quanto os conceitos gerais do domínio. Utilizamos, para isso, o uso da abordagem de ontologia única para integração de informações [Wache et al. 2001]. Contudo, no trabalho apresentando em [Wache et al. 2001], o autor assume que todas as fontes de dados são bancos de dados. Nosso problema, no entanto, é mais complexo, uma vez que as informações que desejamos integrar estão representadas por diferentes formatos de arquivos e armazenadas em diferentes fontes de informações, que devem ser mantidas nos locais e formatos originais [Mastella 2010]. Além disso, em grandes indústrias de petróleo, os profissionais envolvidos na modelagem necessitam de uma maneira eficiente para encontrar os dados desejados, o que não corresponde à realidade atual. A forma de busca ideal é através de consultas relacionadas aos significados reais dos dados, isto é, pela semântica dos dados.

Identificou-se, portanto, a necessidade de resolver dois problemas: (i) localizar os dados desejados e (ii) mapear as informações desses dados com o modelo conceitual proposto. A fim de resolver o primeiro problema, propomos o uso de metadados, que são definidos como dados sobre dados. Desse modo, para cada objeto de dados (arquivos LAS, WITSML, DAT, PLO, Documentos Geológicos, etc.), são definidos dois tipos de metadados: (i) metadados de informações de acesso, que definem como os objetos de dados podem ser acessados, isto é, o local, o nome e o tipo (extensão) do arquivo; (ii) metadados de proveniência, que descrevem como os objetos de dados foram criados, incluindo o autor, a data de criação e a data de última modificação.

Uma vez com os metadados de informações de acesso e proveniência, os objetos poderão ser processados por um analisador sintático, específico para cada padrão. Considerando como gramática a sintaxe do padrão utilizado no arquivo, que declara precisamente quais são os possíveis elementos a serem descritos, este analisador sintático poderá auxiliar no mapeamento entre os objetos de dados e o modelo conceitual proposto. Nesse processo, as informações dos arquivos deverão ser transformadas em instâncias ontológicas, que serão armazenadas em uma base de conhecimento. Dessa maneira, os usuários poderão realizar buscas nessa base de conhecimento através de consultas relacionadas aos dados geológicos representados pelos objetos de dados.

5. Conclusão

Apresentamos, neste artigo, a necessidade de uma abordagem de integração dos objetos de dados utilizados no fluxo de trabalho de modelagem de reservatórios. Esta abordagem, contudo, deve levar em consideração o significado real dos dados e não apenas a tecnologia utilizada para representá-los. Para isso, identificamos os principais objetos de dados e padrões utilizados no processo e realizamos um levantamento das principais ontologias disponíveis para esse domínio. Propomos a reutilização dessas ontologias para a construção de um modelo conceitual único, que é uma ferramenta chave para atender as necessidades dos profissionais envolvidos no processo de modelagem: encontrar, de maneira eficiente, os dados desejados, isto é, através da realização de consultas pelos significados reais dos dados. Identificamos, então, os dois principais problemas enfrentados

nas indústrias de petróleo e propomos soluções: (i) utilizar metadados para auxiliar na localização dos arquivos desejados e (ii) extrair as informações dos dados, descritos por padrões, para auxiliar os profissionais no mapeamento das informações geológicas, que estão representadas nos objetos de dados, com o modelo conceitual proposto.

No estado atual de desenvolvimento do trabalho, focamos no processo de construção do modelo estrutural. Em trabalhos futuros, pretendemos estender esse estudo para todo processo de modelagem de reservatórios. Para isso, iremos analisar a construção dos modelos estratigráfico e de reservatório, a fim de definir os principais padrões utilizados e uma forma de mapeá-los ao modelo conceitual. A validação do modelo de integração proposto será realizada através do desenvolvimento de um protótipo que permita o cadastramento de novos objetos de dados e, posteriormente, a recuperação destes objetos de dados através de consultas que relacionem os objetos geológicos que eles representam.

6. Agradecimentos

Este trabalho foi parcialmente financiado pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES, pelo Programa de Excelência Acadêmica - PROEX, assim como pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq.

Referências

- Abel, M., Silva, L. A., De Ros, L. F., Mastella, L. S., Campbell, J. A., and Novello, T. (2004). Petrographer: managing petrographic data and knowledge using an intelligent database application. *Expert Systems with Applications*, 26(1):9–18.
- Lorenzatti, A., Abel, M., Nunes, B. R., and Scherer, C. M. (2009). Ontology for imagistic domains: Combining textual and pictorial primitives. In *Advances in Conceptual Modeling-Challenging Perspectives*, pages 169–178. Springer.
- Mastella, L. S. (2010). *Semantic Exploitation of Engineering Models: Application to Petroleum Reservoir Models*. PhD thesis, Ecole Nationale Supérieure des Mines De Paris.
- Perrin, M. and Rainaud, J.-F. (2013). *Shared Earth Modeling: Knowledge Driven Solutions for Building and Managing Subsurface 3D Geological Models*. Editions Technip.
- Perrin, M., Zhu, B., Rainaud, J.-F., and Schneider, S. (2005). Knowledge-driven applications for geological modeling. *Journal of Petroleum Science and Engineering*.
- Raskin, R. G. and Pan, M. J. (2005). Knowledge representation in the semantic web for earth and environmental knowledge representation in the semantic web for earth and environmental terminology (sweet). *Computers & Geosciences*.
- Soma, R., Bakshi, A., Prasanna, V., DaSieg, W., and Bourgeois, B. (2008). Semantic web technologies for smart oil field applications. In *Intelligent Energy Conference and Exhibition*.
- Wache, H., Voegele, T., Visser, U., Stuckenschmidt, H., Schuster, G., Neumann, H., and Hübner, S. (2001). Ontology-based integration of information—a survey of existing approaches. In *IJCAI-01 Workshop: Ontologies and Information Sharing*.