

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE INFORMÁTICA
CIÊNCIA DA COMPUTAÇÃO

MARCOS VINICIUS CAVINATO

**Estudo de padrões de troca de informações
digitais para exploração e produção de
petróleo**

Monografia apresentada para obtenção do Grau
de Bacharel em Ciência da Computação pela
Universidade Federal do Rio Grande do Sul

Mara Abel
Orientador

Joel Carbonera
Co-orientador

Porto Alegre, Julho de 2012

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

Reitor: Prof. Carlos Alexandre Netto

Diretor do Instituto de Informática: Prof. Luís C. Lamb

Coordenador da CIC: Prof. Raul Fernando Weber

Bibliotecária-chefe do Instituto de Informática: Beatriz Regina Bastos Haro

*“Não deixe para amanhã o que pode fazer hoje, pois,
se gostar do que fez hoje, pode repetir amanhã”*

— SIR JAMES A. MICHENER

AGRADECIMENTOS

Ao meu pai, mãe e irmão, por todo o suporte durante a longa caminhada da graduação e por todo o carinho e motivação fornecidos durante este caminho.

À minha namorada, geóloga, por toda a contribuição de conhecimento desta área de conhecimento, nova para mim, que desbravei, além de todo o carinho e amor recebido por ela.

Aos orientadores Mara e Joel, pela paciência e todo o apoio incansável para a realização do trabalho, me ajudando a entender melhor computação e geologia, e como essas duas áreas interagem.

A todos os amigos da faculdade, principalmente aos do time QCB, por me permitirem encontrar uma nova família longe de minha cidade natal.

A todos meus familiares, principalmente meus primos, formando um grupo de amigos com apoio mútuo.

Não menos importante, à Universidade Federal do Rio Grande do Sul, com ênfase ao Instituto de Informática, pelo ensino de excelência, de qualidade e gratuito.

A todos estes citados, meu agradecimento. Todos tiveram papel fundamental para a conclusão de mais uma etapa de minha vida e, por isso, sou profundamente grato.

SUMÁRIO

LISTA DE ABREVIATURAS E SIGLAS	9
LISTA DE FIGURAS	11
LISTA DE TABELAS	13
RESUMO	15
ABSTRACT	17
1 INTRODUÇÃO	19
1.1 Definição de Padrão	20
1.2 Contexto Histórico	21
1.3 Objetivo	22
1.4 Estrutura do trabalho	23
2 CONCEITOS BÁSICOS DE E&P DE PETRÓLEO	25
2.1 A origem do petróleo	25
2.1.1 Rocha Geradora	26
2.1.2 Rocha reservatório	26
2.1.3 Armadilhas	26
2.2 Etapa de Exploração	27
2.2.1 Técnicas geológicas	28
2.2.2 Técnicas geofísicas	28
2.3 Etapa de Produção	29
2.3.1 Log Litológico ou de Amostra	31
2.3.2 Log do Tempo de Perfuração	31
2.3.3 Log de Raios Gama	32
2.3.4 Outros tipos de logs	32
3 PADRÕES PROPOSTOS, CONSÓRCIOS E FORMATOS PROPRIETÁRIOS	35
3.1 Tentativas de padronização	35
3.1.1 PXML	35
3.1.2 RESCUE	36
3.2 Consórcios e Entidades relacionadas	37
3.2.1 OGC - Open Geospatial Consortium	37
3.2.2 INDE - Infraestrutura Nacional de Dados Espaciais	38
3.2.3 Energistics	39

3.2.4	ANP - Agência Nacional de Petróleo, Gás Natural e Biocombustíveis . . .	39
4	CONCEITOS E PADRÕES GENÉRICOS PARA TROCA DE INFORMAÇÕES	41
4.1	Metadados	41
4.2	Linguagem de Marcação	41
4.2.1	XML	42
4.2.2	XSL	43
4.2.3	XML Schema	45
4.3	SOA	46
4.3.1	Web Services	46
4.4	Arquivos ASCII	47
5	PADRÕES PARA TROCA DE INFORMAÇÕES EM GEOCIÊNCIAS	49
5.1	LAS	51
5.1.1	Estrutura LAS	51
5.2	RESQML	53
5.2.1	HDF5	55
5.2.2	EPGS Codes	55
5.2.3	GML	56
5.2.4	Dublin Core Elements	56
5.2.5	GUIDs	56
5.2.6	Objetos de Dados e Conceitos Chave	57
5.3	WITSML	58
5.3.1	Litologias em WITSML	59
5.4	PRODML	60
6	IMPLEMENTAÇÃO DOS PADRÕES RESQML E WITSML	65
6.1	Visão geral do RESQML	65
6.2	Visão geral do WITSML	66
7	RESULTADOS DA UTILIZAÇÃO DOS PADRÕES NA INDÚSTRIA E DA PESQUISA	69
8	EVOLUÇÃO DOS PADRÕES DE TROCA DE DADOS	73
8.1	RESQML	73
8.2	PRODML	74
8.3	WITSML	74
9	CONCLUSÃO	75
	REFERÊNCIAS	77
	ANEXO:	
10	ANEXO A - DESCRIÇÃO DE DOMÍNIOS SUPORTADOS POR RESQML	79

LISTA DE ABREVIATURAS E SIGLAS

ANP	Agência Nacional de Petróleo, Gás Natural e Biocombustíveis
ASCII	American Standard for Computer Information Interchange
BDEP	Banco de Dados de Exploração e Produção
CGI	Commission for the Management and Application of Geoscience Information
CRS	Coordinate Reference System
CSS	Cascading Style Sheets
DLIS	Digital Log Interchange Standard
DTD	Document Type Definition
EPGS	European Petroleum Survey Group
E&P	Exploração e Produção
GUIDs	Global Unique Identifiers
HDF5	Hierarchy Data Format 5
HTTP	Hypertext Transfer Protocol
INDE	Infraestrutura Nacional de Dados Espaciais
LAS	Log ASCII Standard
LIS	Log Interchange Standard
LWD	Logging While Drilling
OGC	Open Geospatial Consortium
OGP	International Association of Oil & Gas Producers
POSC	Petrotechnical Open Software Corporation
PXML	Extensible Markup Language for Petroleum Geologic Information
RESCUE	REServoir Characterization Using Epicentre
SIG	Special Interest Group
SIP	Software Integration Platform
SOA	Service Oriented Architecture
SOAP	Simple data-Object Access Protocol

UDDI	Universal Description, Discovery and Integration
XML	Extensible Markup Language
XML-FO	XML Formatting Objects
XPATH	XML Path Language
XSD	Xml Schema Definition
XSL	Extensible Stylesheet Language
XSLT	Extensible Stylesheet Language for Transformation
WITS	Wellsite Information Transfer Specification
WITSML	Wellsite Information Transfer Standard Markup Language
WSDL	Web Service Definition Language

LISTA DE FIGURAS

Figura 1.1:	Cadeia de exploração e perfis profissionais associados. Extraído de (RAINAUD, 2011)	19
Figura 1.2:	Trecho de um documento WITSML. Adaptado de (ENERGISTICS; WITSML SIG, 2012)	21
Figura 2.1:	Exemplos de armadilhas. Extraído de (INSTITUTE, 2012)	27
Figura 2.2:	Cadeia de exploração. Extraído de (PERRIN et al., 2007)	28
Figura 2.3:	Exemplo de um mapa estrutural com duas candidatas a armadilhas. Extraído de (HYNE, 2001)	29
Figura 2.4:	Exemplo de registro sísmico mostrando estruturas de acamamento da rocha. Extraído de (HYNE, 2001)	30
Figura 2.5:	Exemplo de modelo sísmico 3D - união de diferentes registros sísmicos. Extraído de (GEOSCIENCES, 2012)	31
Figura 2.6:	Exemplos de logs de perfuração. Extraído de (HYNE, 2001)	33
Figura 3.1:	Exemplo de PXML para descrição litológica. Extraído de (XUE; XUE, 2010)	36
Figura 4.1:	Exemplo de documento XML	43
Figura 4.2:	Exemplo de documento XSLT	44
Figura 4.3:	Resultado da transformação mostrada na Figura 4.2 aplicada sobre o documento descrito na Figura 4.1	45
Figura 5.1:	Exemplo do uso de <i>~Version</i> e <i>~Well</i> em um arquivo LAS. Extraído de (HESLOP et al., 2000)	53
Figura 5.2:	Exemplo de tipos de seções e de linhas de seções. Extraído de (HESLOP et al., 2000)	54
Figura 5.3:	Exemplo de código GML	56
Figura 5.4:	Objetos de dados suportados pela versão 1.1 do RESQML. Extraído de (ENERGISTICS; RESQML SIG, 2011a)	57
Figura 5.5:	Exemplo de descrição litológica WITSML. Adaptado de (ENERGISTICS; WITSML SIG, 2012)	60
Figura 5.6:	RESQML, PRODML e WITSML em um processo de negócios E&P. Adaptado de (ENERGISTICS; PRODML SIG, 2010)	61
Figura 5.7:	Exemplo de relatório PRODML	62
Figura 6.1:	Visão geral dos componentes e interações chaves entre clientes e servidores WITSML. Adaptado de (ENERGISTICS; WITSML SIG, 2011)	67

Figura 6.2:	Relações entre objetos WITSML. Adaptado de (ENERGISTICS; WITSML SIG, 2011)	68
Figura 7.1:	Fluxo de validação da caracterização de um reservatório. Adaptado de (MICHEL; VERNEY, 2011)	69
Figura 7.2:	Validação através do uso de ferramentas proprietárias. Adaptado de (RAINAUD, 2011)	70
Figura 7.3:	Resumo das relações entre os esquemas. Adaptado de (RAINAUD, 2011)	71

LISTA DE TABELAS

Tabela 2.1:	Comparativo de logs de perfuração	32
Tabela 5.1:	Comparação entre padrões de log de poços. Adaptado de (GEO-MIND, 2007)	50
Tabela 5.2:	Descrição das características da RESQML. Adaptado de (ENERGISTICS; RESQML SIG, 2011a)	55
Tabela 5.3:	Comparação entre os principais padrões ligados à E&P	63
Tabela 6.1:	Tipos de arquivos e nomeações RESQML. Extraído de (ENERGISTICS; RESQML SIG, 2011b)	65
Tabela 8.1:	Planos futuros para extensão futura do RESQML	73
Tabela 10.1:	Descrição de domínios suportados por RESQML. Adaptado de (ENERGISTICS; RESQML SIG, 2011b)	80

RESUMO

Interoperabilidade é a habilidade de um sistema ou componentes de um sistema de oferecer portabilidade de informação e controle cooperativo de processos entre aplicações de software.

O processo de exploração e produção (E&P) de petróleo é constituído de diferentes etapas: exploração, produção, transporte, refino e comercialização. Cada uma das etapas envolve diferentes perfis profissionais (geólogos, geofísicos, engenheiros, administradores) e trata informações de diferentes fontes. A informação processada em uma das etapas pode consistir em dados de entrada para a próxima etapa. Por essa razão, a interoperabilidade se torna um fator importante para estabelecer conexões entre os sistemas e entre domínios. Nesse contexto, o uso de padrões para comunicação entre um sistema e outro que integre as visões dos diferentes profissionais que usam os dados é fundamental.

Um padrão é um modelo ou paradigma que serve como referência e que são definidos pelos metadados que dão significado e contexto às informações representadas. Essencialmente, um padrão será adequado para troca de informação quando tiver previsto nos seu modelo o tipo de informação que deve ser trocada entre os sistemas. São expressos pela estrutura interna que um documento deve ter (sequenciamento e tamanho de seções) e pelos tipos de dados representados em cada seção. Quando definido para sistemas abertos, o paradigma que define o padrão é construído de forma independente da plataforma onde vão executar os sistemas a serem integrados. Esses são os padrões mais desejados atualmente.

Esse trabalho estuda os padrões abertos para troca de informações digitais nas etapas de exploração e produção da cadeia de petróleo. São analisados os padrões independentes de domínio, como as linguagens de marcação - XML, XSL, XML Schema, SOA, WEB Services - e também os padrões definidos para aplicação na área de petróleo. Os padrões para exportação de dados de perfilagem geofísica - LAS e WITSML - são examinados e comparados, e também o estado da arte na definições de padrões ainda em desenvolvimento como o RESQML, para modelar dados de reservatórios, e PRODML, para tratar informações de acompanhamento de poços de produção são apresentados. Os três últimos são complementares, cobrindo diferentes etapas da E&P, e estão sendo definidos por um consórcio de empresas e institutos de pesquisa, o que garante uma aceitação e adoção mais ampla na indústria. O padrão WITSML tem se mostrado como o padrão de maior potencial de troca de informações na indústria, com as vantagens da independência de plataforma, completeza dos metadados e maturidade do seu desenvolvimento.

A definição de padrões digitais para troca de informações é o passo inicial para real integração de sistemas de diferentes fabricantes desejada ao longo da cadeia de E&P.

Palavras-chave: Exploração, Perfuração, Produção, Petróleo, Padrões, Interoperabilidade, RESQML, WITSML, PRODML, LAS.

A study of digital transfer standards for petroleum E&P

ABSTRACT

Interoperability is the ability of a system or the system components, to provide information portability and cooperative control of processes between software applications.

Petroleum exploration and production (E&P) process is composed of different steps: exploration, production, transportation, refining and commercialization. Each one of those steps involves different professionals (geologists, geophysicists, engineers, administrators) and deals with information from several sources. The processed information from one step may provide input data to the next step. For this reason, interoperability becomes an important issue to establish connection between systems and domains. Thus, standards for systems communication and for integration among different professional who use the data is essential.

A standard is a model or paradigm that serves as reference and is defined by meta-data that provides meaning and context to the represented information. Essentially, a standard will be appropriated for information exchanging when it has on its model the definition of the type of the information exchanged between the systems. They are expressed through the internal structure that a document may have (sequencing and sections size) and the data types represented in each section. When defined to be implemented for open systems, the paradigm that defines the standard is built in a platform-independent way that runs, the systems intended to be integrated. Those are the most wanted standards nowadays.

This work studies the open standards for digital information transfer in petroleum E&P chain. Domain-independents standards are analyzed, like the markup languages - XML, XSL, XML Schema, SOA, WEB Services - likewise standards defined for petroleum area applications. The standards for geophysical logging exporting data - LAS and WIT-SML - are examined and compared, and also the state of art in standard definitions that are still under development, such as RESQML for reservoir data modeling and PRODML for well production data monitoring. The last three are complementary, covering different E&P stages, and are being defined by a consortium composed by companies and research institutes, which are expecting to lead to a higher industry acceptance. WITSML has been proving to be the higher-potential exchanging data standard in industry, with advantages in platform-independence, meta-data completeness and development maturity.

The definition for digital standards to transfer information is the initial step for the required real integrated systems from different manufacturers.

Keywords: Exploration, Drilling, Production, Petroleum, Standards, Interoperability, RESQML, WITSML, PRODML, LAS.

1 INTRODUÇÃO

De acordo com (BISHR, 1998), interoperabilidade é a habilidade de um sistema ou componentes de um sistema de oferecer portabilidade de informação e controle cooperativo de processos entre aplicações. Adicionalmente, interoperabilidade pode ser caracterizada como uma forma de sistema inteligente que atinge cooperação entre componentes de um sistema de informação.

O processo de exploração e produção de petróleo (E&P) é constituído de diferentes etapas. Esse processo é multidisciplinar, envolvendo geólogos, geofísicos, engenheiros, etc. Por essa razão, a interoperabilidade se torna um fator importante para estabelecer ligações entre os limites de diferentes domínios. Nesse contexto, o uso de padrões para comunicação entre os domínios é fundamental. Esse cenário é ilustrado na Figura 1.1, que mostra como esses diferentes perfis de profissionais estão inseridos na cadeia de exploração do petróleo. Ainda é notável que mesmo dentro de uma única etapa, diferentes processos interagem, como interpretação sísmica, modelagem estruturas, etc.

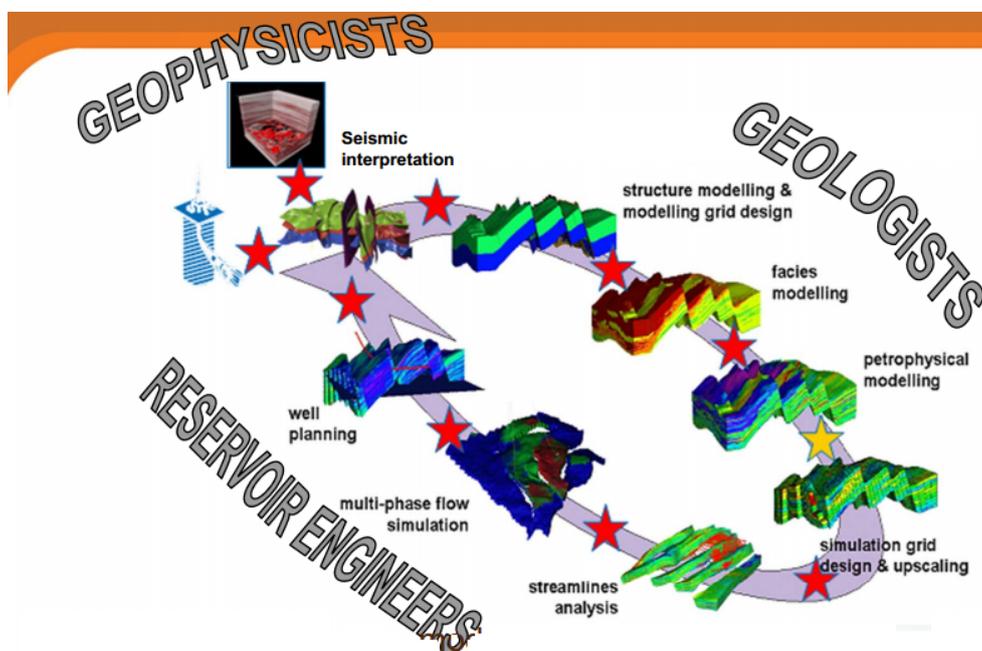


Figura 1.1: Cadeia de exploração e perfis profissionais associados. Extraído de (RAINAUD, 2011)

No passado, não havia preocupação com integração de sistemas e principalmente com interoperabilidade. Com essa liberdade de desenvolvimento durante anos, criou-se um

cenário em que os padrões variam de empresa para empresa e, mais que isso, a própria terminologia referente a eles e aos modelos tornou-se diferente, ocasionando incompatibilidade dos tipos de dados. Visando sanar estes problemas, o estudo dos padrões se faz necessário, de forma que integração e interoperabilidade sejam premissas do desenvolvimento.

Por fatores econômicos, as ferramentas de software que manipulam os modelos são, quase que em totalidade, proprietárias e não se comunicam com outras ferramentas que não do mesmo fabricante ou com interesses relacionados a esse (TUDORACHE, 2009). Desse modo, a troca de informações entre programas dá-se por meio de exportação de dados em formatos que, mesmo que sejam populares, tem sua semântica não revelada pelos fabricantes. Assim, apenas os próprios desenvolvedores das ferramentas conhecem a semântica por trás dos dados.

Na década de 70, eram poucos os prestadores de serviço que desenvolviam software para as etapas de E&P. Naturalmente, cada um criava seus próprios formatos de dados e a comunicação acontecia apenas entre seus diferentes softwares, que abordavam os diferentes processos dentro das diferentes etapas de E&P. Com o passar do tempo, as operadoras, mais importantes usuárias destes pacotes de software, passaram a exigir maior flexibilidade no uso dos softwares, almejando o uso de pacotes de diferentes prestadoras conjuntamente, de modo que houvesse comunicação entre eles. Começa, assim, a buscar-se alternativas para a questão de interoperabilidade.

Uma solução proposta, mas ainda falha, para o problema de integração, é a criação de interfaces uniformes que tratam diferentes formatos e tentam englobar a maior quantidade possível de tipos de dados existentes. Essa alternativa funciona bem quando se lida com fontes estáticas, o que não é o caso da exploração do petróleo (OLSEN et al., 1995). Novos modelos podem surgir e o fato de só serem integrados a essas ferramentas muito tempo depois que os usuários já necessitem de troca de informações continua sendo um problema gerado pelos diferentes padrões.

Para que haja uma unificação dos padrões, garantindo interoperabilidade, é mais importante preservar o uso de um vocabulário comum, que dê significado aos dados, do que garantir a padronização em relação a aspectos que estejam em constante evolução, tais como formatos de arquivos e ferramentas (MASTELLA, 2010). Foi com este pensamento que foram criadas diversas propostas de padrões, a maioria baseadas em XML, e que serão objeto de estudo desse trabalho.

Outro motivador para o desenvolvimento deste trabalho, é o aumento da dependência da economia mundial às áreas energéticas, principalmente no que tange o petróleo, por ser um recurso finito e sustentar economias de nações inteiras. O Brasil, por exemplo, passou a ocupar um papel ainda mais importante nesse contexto, com a recente descoberta do pré-sal. Com isso, este trabalho serve como atualização acadêmica, no sentido de que outros trabalhos possam usar este como base para aprofundar ainda mais os conhecimentos aqui apresentados.

1.1 Definição de Padrão

De acordo com o dicionário Michaelis¹ da língua portuguesa, padrão é um modelo ou paradigma que serve como referência. No contexto deste trabalho, padrão estará diretamente associado ao conceito de modelo, ou seja, são definidos pelos metadados que dão significado e contexto às informações representadas. Essencialmente, um padrão será

¹<http://michaelis.uol.com.br>

adequado para troca de informação quando tiver previsto no seu modelo o tipo de informação que deve ser trocada entre os sistemas.

Padrões são expressos pela estrutura interna que um documento deve ter (sequenciamento e tamanho de seções) e pelos tipos de dados representados em cada seção. Quando previstos para sistemas abertos, esses padrões são aplicados sobre arquivos texto, mas podem ser ocultos na formatação de arquivos binários associados a sistemas proprietários.

```
<documentInfo>
  <documentName>Drill Report</documentName>
  <owner>IIP</owner>
</documentInfo>
<drillReport uidWell="W-12" uidWellbore="B-01" uid="243ef">
  <nameWell>6507/7-A-42</nameWell>
  <nameWellbore>A-42</nameWellbore>
  <name>2006-06-06 prelim</name>
  <dTimStart>2006-06-06T00:00:00.000Z</dTimStart>
</drillReport>
```

Figura 1.2: Trecho de um documento WITSML. Adaptado de (ENERGISTICS; WITSML SIG, 2012)

A Figura 1.2 mostra um trecho de um documento WITSML, um dos principais padrões ligados à perfuração de poços de petróleo. Este exemplo ilustra características do uso de padrões para oferecer validação sintática. Ou seja, documentos que não sigam regras determinadas não se encaixam no padrão.

Os padrões aqui descritos estarão relacionados às formas que diferentes empresas geram informações no processo de extração do petróleo.

1.2 Contexto Histórico

A primeira perfuração de um poço de petróleo para fins comerciais foi realizada em 1859 a uma profundidade de 21 metros (HYNE, 2001). Foi realizada pelo perfurador William Smith e pelo operador Edwin L. Drake, trabalhadores da Seneca Oil Company, dos Estados Unidos. O local foi escolhido devido a uma afloração natural de petróleo, objetivando aumentar o fluxo de óleo que escoava para a superfície. A quantidade foi suficiente para encher 3 metros cúbicos diários de óleo.

Nesse tempo, a forma como o petróleo se formava e migrava através das rochas de subsuperfície, acumulando-se nas armadilhas (ver seção 2.1), ainda não era compreendida. Nos mais de 50 anos seguintes, os perfuradores escolhiam os locais de perfuração baseados na busca por rios e cavernas de subsuperfície, sendo os geólogos raramente utilizados na escolha do local de perfuração. Após encontrado algum sítio, era aplicado o princípio da proximidade (HYNE, 2001): quando mais perto fosse uma proposta de poço de um outro já produtor, melhor era a proposta.

O fato é que todo esse processo era empírico. Não existia nenhuma fundamentação teórica por trás dessas escolhas. A partir do século XX passou a ser aceito que o óleo se acumula em grandes áreas de rochas reservatórias (teoria das anticlinais) e a contribuição (fundamental) dos geólogos passou a ser aceita. A geologia passa então a fazer parte de um processo metodológico, inicialmente, mapeando camadas de rochas sedimentares afloradas na superfície.

Com o surgimento dos aviões e, conseqüentemente, das topografias aéreas, o mapeamento de superfícies se tornou mais eficiente. Desde a década de 70, satélites são utilizados fornecendo imagens importantes das topografias. Através de correspondência entre rochas de áreas candidatas com outras de regiões onde há incidência de petróleo (correlação), métodos mais específicos começam a surgir.

Além da geologia, diversas outras áreas começam a se aprofundar no estudo do solo, procurando métricas que indiquem probabilidades de se encontrar reservatórios, acima de tudo, rentáveis para a indústria. A geoquímica (química aplicada ao estudo da terra), por exemplo, começa a procurar traços de hidrocarbono em amostras da água de solo das áreas de exploração, bons indícios de armadilhas de petróleo nas proximidades.

No entanto, os maiores avanços tecnológicos têm sido encontrados na área de exploração sísmica, parte da geofísica (aplicação de física e matemática para o estudo da terra). Dentre as três principais técnicas de análise da terra (gravimetria, magnetometria e sísmica) a sísmica é a que, devido à evolução tecnológica, tem sido mais usada. O primeiro uso dessa técnica data de 1928, ainda muito ruidosa e pouco precisa.

Essa técnica é baseada no princípio da reflexão do som. Um emissor é usado para gerar ondas. Parte é refletida e parte segue seu caminho. Outro aparelho capta essas ondas. Através do tempo entre a emissão e a recepção, é possível gerar gráficos para análise do solo.

Com a evolução do processamento dos microcomputadores, tornou-se mais fácil o registro, interpretação, processamento e geração de modelos 2D, 3D e até 4D dos dados obtidos da exploração sísmica, que, além do som, aplica uma grande variedade de medidas das emissões das rochas. Na década de 90 (HYNE, 2001) diversos programas de diferentes empresas começam a ser desenvolvidos para armazenamento e estudo dos dados obtidos a partir de explorações sísmicas. Com os modelos prontos, diversos estudos geológicos podem ser feitos sobre os mesmos para estimar a probabilidade de se encontrar reservas de petróleo rentáveis em armadilhas de subsuperfície.

O principal problema é que após essa etapa de exploração ainda existem outras etapas como de perfuração do poço e produção do petróleo. Estas etapas podem ser consideradas independentes e diferentes empresas atuam em cada um destes ramos. Porém, cada etapa é dependente de dados obtidos em etapas anteriores. Com várias empresas criando softwares diferentes e trabalhando em áreas distintas, diferentes padrões de dados acabam surgindo, pois não há preocupação em interoperabilizar esse processo. Em suma: vários softwares salvam os mesmos (ou, pelo menos, muito semelhantes) tipos de dados, mas de diferentes formas e formatos.

Esse cenário começa a mudar a partir do século XXI com novas propostas de formatos abertos para estabelecer um padrão geral para todo o processo de E&P.

1.3 Objetivo

O objetivo deste trabalho é estudar os padrões existentes para transferência de informações entre sistemas, evidenciando as necessidades de transferência de dados e compreendendo a semântica dos dados envolvidos em diferentes etapas da exploração e produção de petróleo. A contribuição esperada é oferecer um guia inicial para desenvolvedores de sistemas em relação ao significado dos dados e especificação dos padrões que permita a criação de sistemas interoperáveis.

No decorrer do trabalho busca-se compreender o porquê da inexistência de um único padrão na representação de dados no ciclo da E&P do petróleo. Relacionado a este obje-

tivo, são mostradas algumas das tentativas de padronização já criadas e como elas contribuíram para a geração do cenário atual, referente às representações.

Além disso, é objetivo também promover a familiarização das pessoas do domínio da computação com os conceitos e desafios do domínio da E&P, bem como oferecer uma análise dos principais padrões de representação e troca de dados geológicos, do ponto de vista da computação, servindo como um guia que facilite o uso deles em sistemas computacionais para o domínio da geologia.

Esta investigação é importante para aprender com as tentativas do passado e buscar novas alternativas de desenvolvimento. Através de toda a discussão no decorrer do texto, o objetivo final é responder à pergunta: é possível que se chegue a uma padronização total nas representações?

1.4 Estrutura do trabalho

No capítulo 2 deste trabalho serão mostrados alguns conceitos específicos de sistemas petrolíferos necessários para compreender o domínio da E&P de petróleo e suprir as necessidades de recursos computacionais do mesmo. Fala-se, de forma resumida, sobre o processo de acumulação, sobre o processo de exploração e o processo de produção do petróleo.

No capítulo 3, será dado foco inicialmente aos padrões adotados pelas empresas mais importantes na área atualmente e, após, às tentativas de padronização propostas, destacando pontos importantes que foram utilizados para que se chegasse aos principais padrões existentes atualmente.

O capítulo 4 apresenta conceitos e padrões sem domínio específico que servem como base para a compreensão dos quatro principais padrões atuais, explicados com detalhes no capítulo 5. Os detalhes de implementação dos padrões são descritos no capítulo 6 e o capítulo 7 fala sobre resultados de aplicações e validações dos padrões no mundo real.

No capítulo 8 são apresentados os planos futuros dos padrões, de acordo com seus desenvolvedores. Por fim, a conclusão (capítulo 9) discute a proposta e engloba a mesma no cenário real.

2 CONCEITOS BÁSICOS DE E&P DE PETRÓLEO

Para entender muitos dos aspectos relacionados aos padrões adotados pelas empresas e os usados atualmente, é necessário ter uma compreensão mínima sobre a área do petróleo. Existem 5 atividades centrais relacionadas ao petróleo (FEIJO, 2010). As etapas, bem como seus objetivos principais, são apresentadas a seguir:

- Exploração: encontrar petróleo em quantidades comerciais;
- Produção: extrair o petróleo com rentabilidade, segurança e respeito ao meio ambiente;
- Transporte: conduzir petróleo bruto e derivados de forma eficaz e segura;
- Refino: obter do petróleo, com eficácia e segurança, os produtos necessários à sociedade;
- Comercialização: fornecer os derivados onde eles são necessários, a preços competitivos.

Convenciona-se chamar as atividades envolvidas até a perfuração do poço de produção de *upstream* e envolve a área de exploração e produção. Já *downstream* se refere às outras áreas de transporte, refino e comercialização. Neste trabalho, a atenção é voltada para as atividades de E&P ou *upstream*, por serem as que mais envolvem padrões para suas representações.

Nas próximas seções é apresentada a origem dos sistemas petrolíferos como introdução para que depois seja explicado o processo de exploração e o de produção, detalhando, assim, cada um deles. Este estudo serve como base para o entendimento dos padrões adotados e para analisarmos as diferenças e a necessidade dos mesmos.

2.1 A origem do petróleo

A palavra petróleo tem origem do grego *Petro* que significa pedra e *oleum* que significa óleo. Óleo da pedra, portanto (HYNE, 2001). Nesse contexto, o estudo da geologia é que explica os conceitos envolvidos nesse processo de transformação do petróleo, ou seja, qual a sua origem.

Para que existam depósitos **comercialmente rentáveis** de petróleo é necessário que três condições geológicas sejam satisfeitas. Primeiro, deve haver uma rocha geradora abaixo da superfície, que é aquela que dá origem ao petróleo em algum momento do passado geológico. Segundo, deve haver uma rocha reservatório de subsuperfície que

armazene o petróleo. Por fim, deve haver nesse reservatório uma armadilha que concentre o petróleo em quantidades comerciais. Na sequência, cada uma das três exigências é explicada com mais detalhes.

2.1.1 Rocha Geradora

A crosta superior da terra em áreas de produção de petróleo é composta de rochas sedimentares. Entre o conjunto possível de tipos de rochas (sedimentares, ígneas, metamórficas), as sedimentares são as mais importantes para o estudo do petróleo, pois elas são as geradoras e as reservatórias (TEIXEIRA et al., 2008). Elas recebem esse nome pois são compostas de sedimentos, que podem ser:

- Fragmentos de outras rochas pré-existentes;
- Conchas marinhas;
- Sal precipitado da água

Durante todo o tempo geológico, o nível do mar tem variado por diversas vezes, expondo e cobrindo a terra. Esse processo deposita sedimentos, formando diversas camadas. A fonte do petróleo está na matéria orgânica (restos animais e vegetais) enterrada e preservada nas rochas sedimentares antigas. O fator externo mais importante é a temperatura: é necessário um mínimo de 65°C para a formação do óleo, alcançada à profundidades de pelo menos 2130 metros. Temperaturas maiores que 150°C formam gás natural e ocorrem a profundidades a partir de 5500 metros.

2.1.2 Rocha reservatório

O petróleo é menos denso que a água - que também ocorre em rochas de subsuperfície - e por isso, após formado, inicia seu processo de ascensão através de fraturas nas rochas de subsuperfície. Eventualmente, o petróleo vai interseccionar uma rocha reservatório: uma rocha sedimentar que contém grande quantidade de pequenos espaços chamados poros. Naturalmente o petróleo flui para dentro destes poros da camada de rochas reservatórias. Esta camada oferece um caminho de menor resistência (permeabilidade) para o fluxo de fluidos, pois a maior parte dos poros são interconectados e, com isso, o petróleo segue seu caminho rumo a superfície em um processo denominado de migração. É devido a migração que o petróleo desloca-se horizontalmente e verticalmente afastando-se do seu local de geração.

2.1.3 Armadilhas

Eventualmente, no processo de migração do petróleo, pode ocorrer uma armadilha, onde o petróleo é parado e concentrado. Aqui, os fluidos separam-se de acordo com suas densidades: o gás é o menos denso e fica no topo, seguido do petróleo e da água. A armadilha ocorre quando uma rocha não porosa, que não permite a passagem dos fluidos adiante, se localiza no topo da rocha porosa onde o petróleo está migrando. As Figuras 2.1(a), 2.1(b) e 2.1(c), extraídas de (INSTITUTE, 2012), mostram três possíveis formas de armadilhas, todas com a mesma característica: armazenamento de petróleo que flua rumo à superfície.

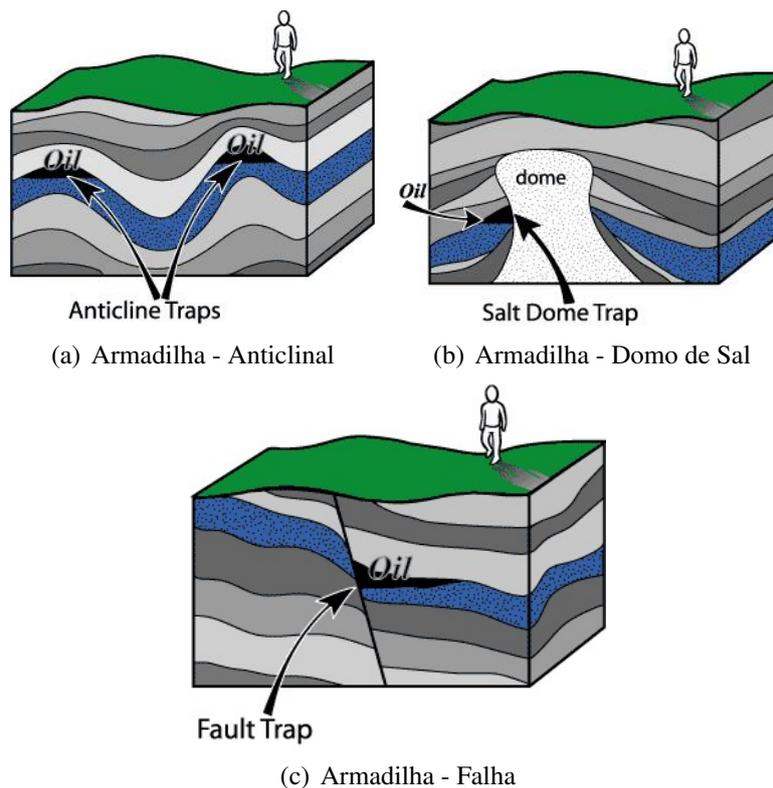


Figura 2.1: Exemplos de armadilhas. Extraído de (INSTITUTE, 2012)

2.2 Etapa de Exploração

Em primeiro lugar, é preciso esclarecer aqui o uso da palavra exploração. A **definição** na língua portuguesa apresenta sinônimos como investigação, estudo e análise. Contudo, o **uso** dessa palavra na língua portuguesa remete a um processo de ação que pode ser confundido com barris de petróleo sendo lacrados e transportados e posteriormente se transformando em gasolina e diversos outros produtos. Embora todas estas etapas sejam pertinentes no estudo do ciclo de vida comercial do petróleo, elas fogem do escopo da área de exploração. Esta palavra deve ser entendida exatamente pelo que ela significa: um processo de investigação e estudo. Portanto, a etapa de exploração está preocupada com a utilização de métodos das áreas de geologia, geofísica, matemática e diversas outras para a localização com maior probabilidade de certeza dos locais onde podem vir a ser realizada a perfuração e produção de petróleo de maneira comercial (MAGOON, 1994). Terminado este estudo, termina também a etapa de exploração.

Na seção 2.1 foi explicado resumidamente o processo de formação do petróleo em subsuperfície, em um processo que ocorre há mais de 2300 metros de profundidade. Também explicou-se que o petróleo, por sua densidade inferior à água, tende a subir para a superfície e eventualmente pára em estruturas denominadas armadilhas. A área de exploração do petróleo está interessada na detecção e estudo do subsolo e demais características que ajudem a encontrar armadilhas com potenciais comerciais, ou seja, com características, como capacidade, características do subsolo, etc, que permitam ações subsequentes que viabilizem o lucro.

A Figura 2.2 mostra a cadeia de exploração, mostrando a variedade de processos que envolvem essa etapa do ciclo de E&P. Durante a cadeia, dados de diferentes dimensões devem ser integrados, como os dados unidimensionais dos logs, as seções sísmicas 2D e

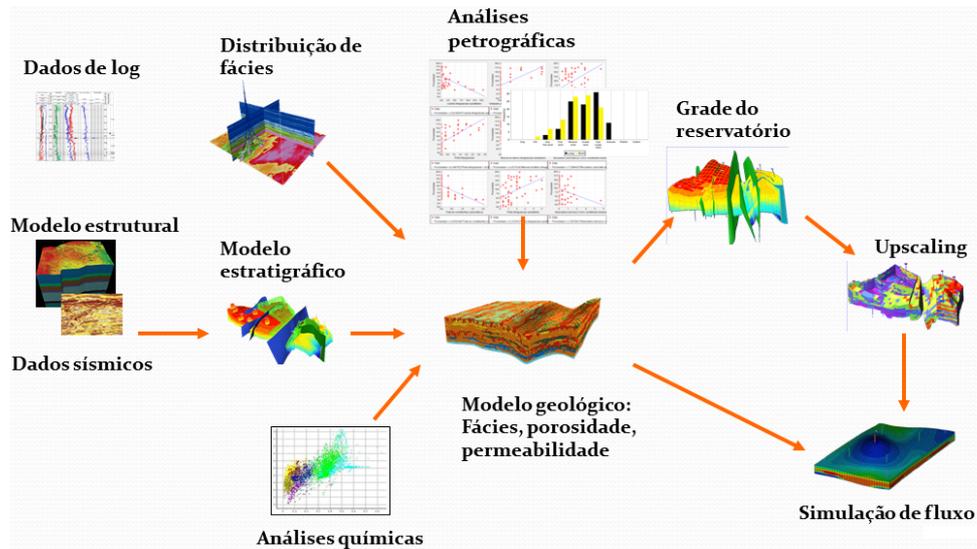


Figura 2.2: Cadeia de exploração. Extraído de (PERRIN et al., 2007)

os modelos 3D, como os estratigráficos.

Diversos processos, técnicas e métodos existem para esse estudo. E todos geram números, modelos e diversos outros artefatos auxiliares. Nesse contexto, sistemas computacionais foram e continuam sendo desenvolvidos para armazenar e realizar operações com base nos dados obtidos (DOVETON, 1986). Abaixo, falamos sobre algumas técnicas de exploração.

2.2.1 Técnicas geológicas

A geologia contribui principalmente com o uso de mapas. Dentre alguns, os mais importantes (HYNE, 2001) são: mapas topográficos, mapas geológicos e mapas de subsuperfície. Esse último pode ser de três tipos, que são os mais importantes para nosso estudo: mapa estrutural, mapa de isópacas e mapa de porcentagem.

Mapas estruturais usam contornos para mostrar a elevação do topo de uma camada de rochas sedimentares de subsuperfície. Com esse tipo de mapa é possível localizar armadilhas onde os topos decrescem rapidamente. A Figura 2.3 mostra um mapa estrutural com duas possíveis armadilhas. Os mapas de isópacas usam contornos para representar a espessura das camadas de subsuperfície e é usada para fornecer uma estatística do tamanho de um possível poço de petróleo. Já os mapas de porcentagem mostram a porcentagem de uma **rocha específica** em uma formação. Porcentagens altas de alguns tipos de rochas aumentam as chances de ser um reservatório grande.

2.2.2 Técnicas geofísicas

As três principais técnicas geofísicas são a gravimetria, magnetometria e a sísmica. Todas objetivam o estudo da subsuperfície. Mesmo que as duas primeiras representem alternativas baratas e portáteis, é a sísmica que tem recebido maiores montantes de investimentos tecnológicos. Pode-se dizer que os maiores avanços em toda a área de exploração do petróleo das últimas décadas foram realizados na área de exploração sísmica, principalmente pelo uso de técnicas de aquisição e registro com processamento digital de dados (DOVETON, 1986).

A etapa de aquisição usa a energia do som que viaja pelas camadas de subsuperfície e

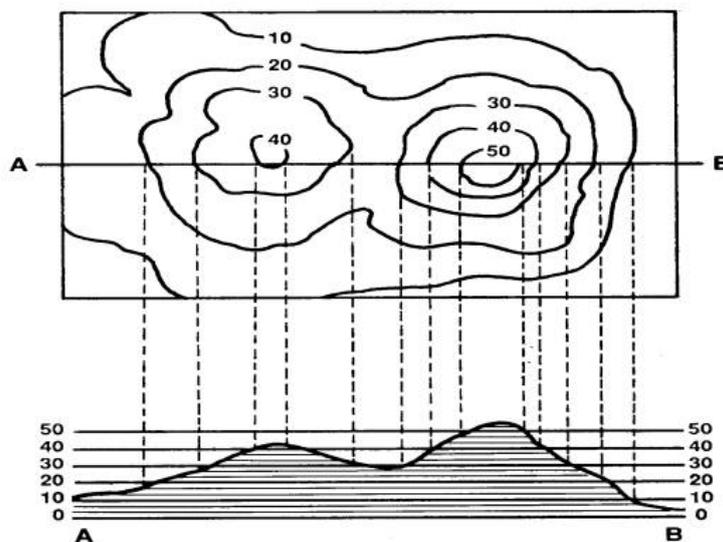


Figura 2.3: Exemplo de um mapa estrutural com duas candidatas a armadilhas. Extraído de (HYNE, 2001)

eventualmente é refletida por camadas diferentes. Nesse método, usa-se um emissor e um receptor. O receptor captura a energia do som que foi refletida pelas camadas de rochas. O tempo de viagem das ondas sísmicas recebidas pelo receptor é o dado mais importante obtido desse processo.

Na Figura 2.4 mostra-se um registro sísmico. A análise destes dados e valores, integrados com outras informações adquiridas, ajudam na construção da estrutura (horizontes e falhas) e da análise estratigráfica (tipos de rochas, ambiente de deposição e conteúdo dos fluidos) das formações de subsuperfície (PENDLETON, 2008).

Após a década de 80, investimentos foram realizados para a construção de modelos sísmicos 3D. Isso foi possível com o aumento do processamento computacional. Algumas técnicas usam vários receptores, em posições previamente calculadas, para receber as ondas sísmicas. Usando cálculos matemáticos complexos, feitos massivamente por computadores, é possível a construção de modelos 3D como o da Figura 2.5, extraída de (GEOSCIENCES, 2012).

A grande maioria dos softwares da área de exploração de petróleo permite a manipulação de modelos sísmicos 3D. A entrada é um conjunto de registros sísmicos. Internamente são feitos cálculos complexos com os dados de entrada, e a saída é o modelo representado.

2.3 Etapa de Produção

A fase de produção se inicia com uma atividade chamada perfuração. Um geólogo, geofísico e/ou engenheiro vai localizar um local de poço ou local de perfuração, que será o local a ser perfurado e, possivelmente, um local onde será encontrado óleo ou gás em quantidades comercialmente viáveis. É identificado um alvo de perfuração, que é um potencial reservatório, e é estimada a distância no subsolo até este alvo.

Entre as etapas *upstream* de exploração e produção, é comum que se diferencie um processo intermediário: a perfuração, que, neste trabalho, é apresentada como parte na etapa de produção. A perfuração pode ser realizada ainda na fase de exploração para que sejam obtidos mais dados para definir a economicidade do campo.

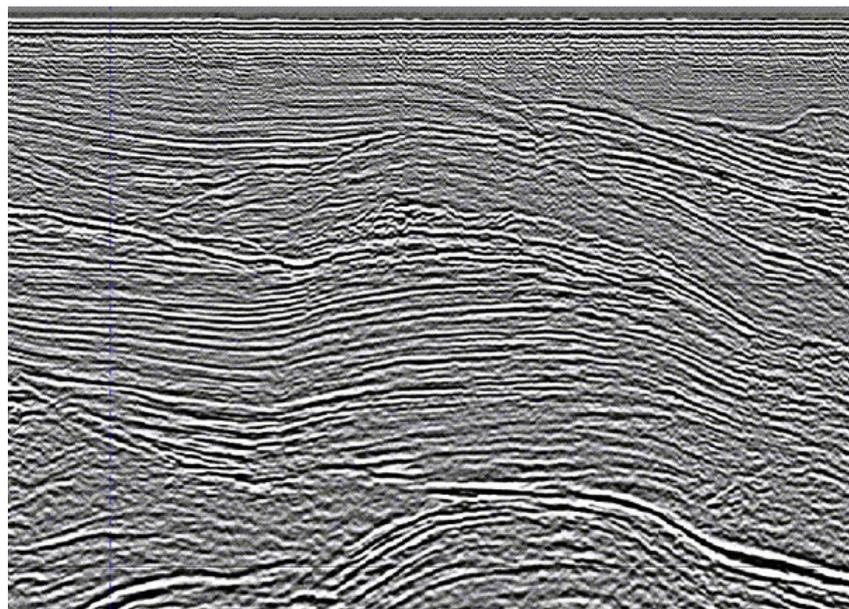


Figura 2.4: Exemplo de registro sísmico mostrando estruturas de acamamento da rocha. Extraído de (HYNE, 2001)

Inicialmente, um poço é perfurado para descobrir uma nova reserva de óleo ou gás e este processo recebe o nome de sondagem. A cada 20 poços perfurados, na média um deles será testemunhado, porém todos serão perfilados. O testemunho geológico é uma amostra retirada da subsuperfície que deve ser analisada com os objetivos de: 1) testar uma armadilha que nunca produziu, 2) testar um reservatório que nunca produziu, ou 3) estender os limites conhecidos de um reservatório já produtor. Se o poço perfurado de fato é uma nova área de atuação, passa a ser chamado de poço descobridor (HYNE, 2001).

Após a perfuração, deve-se completar o poço. Completar um poço é o processo de tornar um poço pronto para produção. Isso envolve preparar o fundo do furo de acordo com especificações necessárias, que podem envolver revestimentos cimentados no furo, dentre outras necessidades. Completar um poço usualmente custa mais caro do que perfurar um poço. Por isso, um poço deve ser cuidadosamente testado após ter sido perfurado. Estes testes são os que indicam a viabilidade comercial do mesmo e são baseados em logs de diversos tipos que envolvem grandes quantidades de dados que são mais rápida e eficientemente manuseados por computadores.

Neste contexto, os logs de poços são um artefato importante obtido durante a etapa de completar o poço, produção e, principalmente, de perfuração. Após 1970, uma técnica chamada Logging While Drilling (LWD) permitiu a obtenção de dados durante a perfuração, em tempo real (antes eram obtidos apenas ao final da perfuração). Diversos dados são fornecidos pelos logs, principalmente informações referentes às rochas da subsuperfície, como porosidade, permeabilidade e conteúdo de fluidos, pois, como mencionado na seção 2.1, estas são características comumente compartilhadas por bons reservatórios.

De forma geral, os logs podem ser vistos como um par onde o primeiro dado é uma profundidade e o segundo dado alguma informação de interesse, uma grandeza numérica absoluta ou vetorial. Para facilitar o entendimento, os logs apresentam um cabeçalho que informa qual é a informação que está sendo adquirida. Por ser de interesse especial deste trabalho pela sua relação com os dados de E&P do petróleo, alguns destes logs serão vistos a seguir.

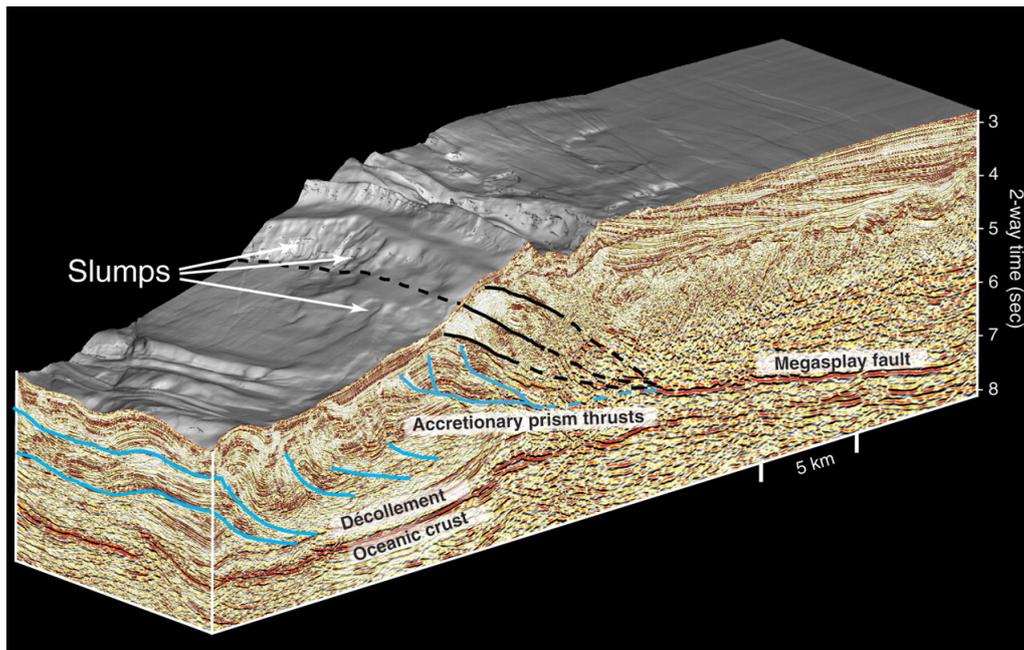


Figura 2.5: Exemplo de modelo sísmico 3D - união de diferentes registros sísmicos. Extraído de (GEOSCIENCES, 2012)

2.3.1 Log Litológico ou de Amostra

Os logs litológicos guardam uma descrição das rochas na qual o poço foi perfurado. No cabeçalho do documento, devem estar descritas informações sobre o poço, tais como: nome do operador, nome do poço e localização do poço. Adicionalmente, há uma lista cujos índices são as profundidades das rochas e cujos valores são símbolos respectivos a tipos de rochas e outras informações como tamanho, porosidade, cor, textura, entre outros.

A Figura 2.6(a) mostra um log de litologia. Na parte de cima encontra-se o cabeçalho com os dados já mencionados. Abaixo, na esquerda, as profundidades, seguidas de um símbolo representativo da rocha e informações adicionais sobre as mesmas.

Este tipo de log é muito importante ainda na etapa de perfuração, pois fornece importantes detalhes geológicos que, analisados, provêm informações sobre a qualidade do poço perfurado. Na seção 8.3, este tipo de log é apresentado como constituinte fundamental de um dos padrões importante atualmente: o WITSML.

2.3.2 Log do Tempo de Perfuração

O log do tempo de perfuração é um registro da taxa de penetração da broca através das rochas. Normalmente, registrado em minutos por metro ou pés. Devido ao fato do log ser gravado enquanto se está perfurando, é considerado um log de tempo real. Mudanças nas rochas de subsuperfície são registradas imediatamente nos logs. Pelo fato de zonas porosas serem menos densas e facilmente perfuráveis, mudanças subitas nos tempos de perfuração podem ser usadas para localizar este tipo de zonas.

A Figura 2.6(b) mostra as diferenças no tempo de perfuração para diferentes tipos de rochas sedimentares. Na parte esquerda da figura, aparecem os nomes da rochas. Pode-se perceber que tanto o xisto (shale) como o arenito (sandstone) apresentam tempos mais rápidos do que os calcários (limestones).

2.3.3 Log de Raios Gama

Os logs de raio gama usam um contador de cintilação para medir a radioatividade natural do potássio, tório e urânio ao longo do poço. Arenito, xisto, calcário, algumas das principais rochas sedimentares, apresentam diferentes respostas a radiação. O xisto é o único radioativo. Estes dados, quando plotados em um log de raio gama, fornecem estimativas da presença de determinadas rochas.

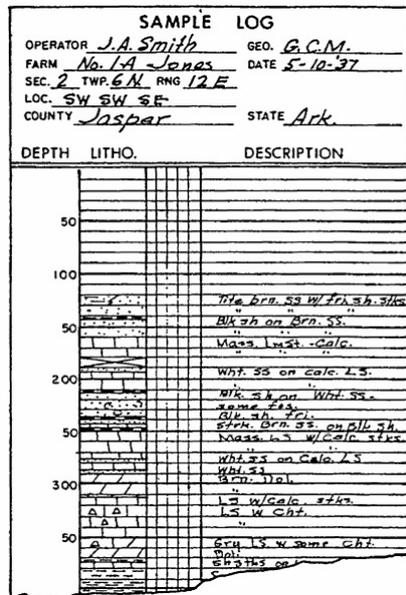
A Figura 2.6(c) mostra que o xisto (shale) apresenta radioatividade alta (H do gráfico) enquanto todos os demais apresentam radioatividade gama baixa (L do gráfico).

2.3.4 Outros tipos de logs

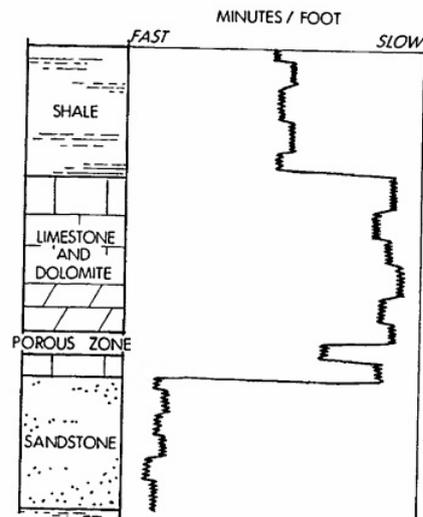
Além dos três citados acima, diversos tipos de logs de poços existem, entre eles: Log de lama, que representa análises químicas sobre os testemunhos, analisando as propriedades químicas das rochas no decorrer da profundidade. Logs elétricos, registram a resistividade elétrica das rochas. Pode-se citar ainda os logs de indução, radioatividade, de sônica, de acústica, dentre diversos outros. A Tabela 2.1 apresenta alguns dos mais importantes logs existentes na etapa de perfuração, identificando o que é medido no log, bem como qual grandeza representa.

Tabela 2.1: Comparativo de logs de perfuração

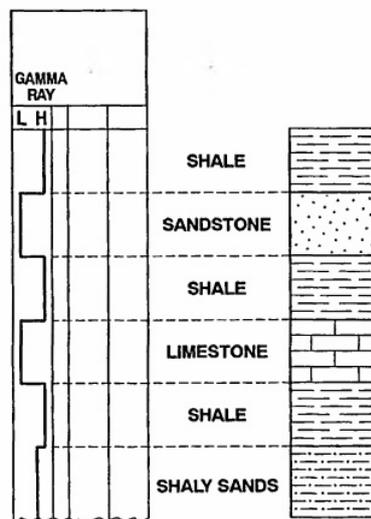
Nome	Medição	Grandeza
Log Litológico	Descrição das rochas pelo processamento combinado de diversos perfis	-
Log do Tempo de Perfuração	Tempo para percorrer certa distância na subsuperfície	Escalar
Log de Raios Gama	Radioatividade das rochas	Escalar
Log de Sônica	Velocidade através de cada camada de rocha	Escalar
Log Elétrico	Resistividade das rochas	Escalar
Log de Inclinação	Orientação das camadas de rocha	Vetorial
Log de Lama (Mud Log)	Descrição das amostras de calha pelo geólogo ou técnico de acompanhamento da perfuração	-



(a) Log de litologia



(b) Log de tempo de perfuração



(c) Log de raios gama

Figura 2.6: Exemplos de logs de perfuração. Extraído de (HYNE, 2001)

3 PADRÕES PROPOSTOS, CONSÓRCIOS E FORMATOS PROPRIETÁRIOS

Este capítulo inicia abordando a comunicação de dados analisando algumas tentativas de padronização, cada uma trazendo alguma melhoria significativa que é incorporada aos padrões mais utilizados atualmente, que serão vistos no capítulo 5. Após, são apresentados alguns consórcios e entidades, estrangeiras e nacionais, que trabalham no sentido de padronizar a representação de dados.

3.1 Tentativas de padronização

A maioria das ferramentas de software de E&P ainda são proprietárias e é comum que lidem com os dados de sua maneira. Dessa forma, a comunicação com outras ferramentas é limitada e, em geral, se dá com as de outros fabricantes que possuam interesses relacionados. A troca de informações é feita, em sua maioria, em formatos não binários, contudo, sem dar acesso aos metadados que permitam ter a semântica revelada. Torna-se, assim, quase impossível realizar qualquer forma de engenharia reversa baseando-se em um conjunto de números sem entender o relacionamento entre eles.

Durante a década de 90 alguns estudos foram feitos e algumas alternativas foram propostas na tentativa da criação de um padrão. Contudo, apenas no final desta década, quando o XML 1.0 se torna uma recomendação da W3C², é que propostas mais concretas e aplicáveis são criadas. A seguir, algumas destas abordagens são apresentadas.

3.1.1 PXML

O Extensible Markup Language for Petroleum Geologic Information (PXML) não tem como objetivo ser uma linguagem de padronização, e sim, expressar informações relacionadas à E&P de uma forma estruturada. Recursos de descrição, plataforma-independente, separação de conteúdo e extensibilidade são herdados do XML (XUE; XUE, 2010).

A linguagem é proposta com base na quantidade enorme de dados armazenados em formatos não estruturados de dados como planilhas e documentos de texto planos. Através de métodos manuais, iterativos ou automáticos, estes dados armazenados podem ser transformados em PXML. Ainda, a própria linguagem seria composta de diversos subconjuntos de linguagens como LXML para registros de dados litológicos e WLXML para descrever registros de dados de poços. As tags representam as propriedades possíveis que os objetos de estudo podem ter. Elas descem até o nível mais básico das propriedades do elemento, onde apresentam os conteúdos, que são os valores reais do elemento.

²<http://www.w3.org/>

Embora a PXML busque uma padronização na representação dos dados ligados à geologia, ela não traz um conjunto de tags pré-definidos. Isto significa que o usuário pode criar seu próprio conjunto de tags. Nesse sentido, há perdas na interoperabilidade sintática, pois perde-se a uniformidade na interpretação dos dados.

O objetivo do PXML é representar dados nos elementos, tanto nas tags quanto no conteúdo e através do uso de DTDs ou XSD (ver subseção 4.2.3), validar os dados, seguindo algum padrão que venha a ser estabelecido. Portanto, a PXML não busca ser um padrão, mas sim, ser um mecanismo para padronização, ou seja, coloca à disposição os conceitos de XML e outras linguagens de marcação, aliados a conceitos da área de geologia para que um possível padrão possa ser gerado.

Um exemplo do uso do PXML para a descrição litológica é mostrado na Figura 3.1.

```

<Lithology>
  <Rock_Association>
    sandstone and mudstone
  </Rock_Association>
  <Sandstone>
    <Color> Gray-gray-white </Color>
    <Granularity> Medium </Granularity>
    <!--...>
  </Sandstone>
  <Mudstone>
    <!--...>
  </Mudstone>
</Lithology>

```

Figura 3.1: Exemplo de PXML para descrição litológica. Extraído de (XUE; XUE, 2010)

Embora o foco do PXML não seja a interoperabilidade, ele pode ser um meio para obtenção desta característica. O capítulo 5 mostra que as abordagens derivadas da Energetics (ver subseção 3.2.3) se assemelham, de uma forma geral, a PXML, apenas adicionando uma maior padronização sintática e, parcialmente, semântica ao modelo.

3.1.2 RESCUE

A REServoir Characterization Using Epicentre (Caracterização de reservatório usando epicentro) é a antecessora do RESQML (ver 5.2), também desenvolvida pela Energetics³, e nasceu para ser um fórum para o desenvolvimento de um padrão aberto de transferência de dados entre desenvolvedores e todas as escalas acima, usando um modelo epicêntrico de dados.

Epicentre, é uma ferramenta criada pela POSC (ver subseção 3.2.3) para exploração e uso produtivo de campos de petróleo. Usa paradigma orientado a objetos, linguagens *Schema* e promove interoperabilidade para prover integração no ciclo de vida de ativos de reservatórios. Adicionalmente, pode-se dizer que o Epicentre é um modelo de dados que define objetos e suas características relevantes para E&P do petróleo, bem como, define relacionamentos entre esses objetos (ALBRECHT, 1999).

A idéia de padrão para o RESCUE se deu pelo desenvolvimento de uma biblioteca de manipulação de dados. Através desta biblioteca, era possível criar modelos de definição

³<http://www.energetics.org/history>

espacial em alguma coordenada do mundo real, seja UTM, *State Plane* ou coordenadas de sistema geográfico. O modelo deveria ter ao menos um bloco/unidade, sendo que os blocos representam divisões verticais da superfície do modelo, enquanto as unidades representam divisões horizontais destas superfícies.

O RESCUE utilizava bibliotecas C++ para seu uso e acabou sendo substituído pelo RESQML a partir de 2009, quando se iniciou o uso das linguagens de marcação para alcançar interoperabilidade independente de plataforma e de linguagem de programação. Esta é uma das principais causas da mudança: por estar atrelado a uma linguagem específica, este padrão exigia que a maioria dos softwares já existentes fossem reescritos utilizando a nova linguagem. Com o uso de XML, o problema se resume a extender, ou seja, adicionar uma nova funcionalidade de leitura e escrita de documentos XML.

3.2 Consórcios e Entidades relacionadas

Existem diversos consórcios e entidades que têm como objetivo padronizar as mais diversas áreas pertencentes à E&P do petróleo. A seguir, são apresentados consórcios e entidades relacionados ao tema, incluindo uma entidade brasileira, com destaque à Ener-gistics.

3.2.1 OGC - Open Geospatial Consortium

O OGC foi criado em 1994 e, anteriormente, era denominado OpenGIS, sendo este termo, atualmente, uma marca registrada que faz referência às diversas especificações do consórcio. Segundo o próprio OGC⁴, a missão do consórcio é:

"Conduzir o desenvolvimento global, a disseminação e a compatibilização de padrões abertos e arquiteturas que viabilizem a integração de dados geográficos e serviços com as mais diversas aplicações e incentivem a geração de negócios na área de geotecnologias."

A OGC possui cinco principais especificações, que são:

- WMS (Web Map Service): especificação voltada para o uso de Web Services através de 4 protocolos que permitem a leitura de múltiplas camadas de informações georreferenciadas;
- WCS (Web Coverage Service): 3 operações que permitem a disponibilização de coberturas através de transporte HTTP em ambiente web;
- SFS (Simple Features Specification): formato SQL padrão para armazenamento, leitura, análise e atualização de dados geológicos;
- WFS (Web Feature Service): uma forma de acesso à dados geológicos que usa o GML para operações de troca entre cliente e servidor e que usa protocolo HTTP para o transporte;
- GML (Geography Markup Language): padrão baseado no XML desenvolvido para permitir o transporte e armazenamento de informações geográficas.

⁴<http://www.opengeo.com.br/?q=node/30>

Os produtos do trabalho do OGC são apresentados sob a forma de especificação de interfaces e padrões de intercâmbio de dados e de serviços. O OGC é uma das primeiras entidades a introduzir o uso de Web Services, que é ainda uma tecnologia emergente que permite disparar aplicações rodando em diferentes máquinas para troca de dados e integração sem necessitar hardware e softwares proprietários adicionais e que é baseada em padrões industriais consolidados e tecnologias já existentes e estabelecidas, como XML e HTTP (PAPAZOGLU, 2008).

Embora largamente utilizado e formado por mais de 250 companhias, agências governamentais e universidades (SANTOS, 2012), as especificações OGC são pouco encontradas na área de E&P do petróleo. Alguns motivos que explicam a razão são:

- **Abrangência:** as especificações OGC abrangem diversos ramos de interesse da geologia. Contudo, são muito genéricas quando se trata de E&P de petróleo, não fornecendo a especificidade necessária para uma padronização entre diversas empresas de porte gigantesco;
- **Detalhamento:** conforme o nome do consórcio indica, a preocupação maior está em características geospaciais dos dados. Quando se tratando de petróleo, características geospaciais constituem apenas uma parcela do volume de dados que devem ser interoperáveis.

O OGC é, portanto, um consórcio bem consolidado no campo da geologia, contudo, sem muito uso na área de E&P do petróleo.

3.2.2 INDE - Infraestrutura Nacional de Dados Espaciais

A INDE foi instituída pelo decreto N° 6.666 em 2008 e é uma entidade brasileira que, segundo a própria entidade, tem foco em⁵:

"conjunto integrado de tecnologias, políticas, mecanismos e procedimentos de coordenação e monitoramento; padrões e acordos, necessário para facilitar e ordenar a geração, o armazenamento, o acesso, o compartilhamento, a disseminação e o uso dos dados geospaciais de origem federal, estadual, distrital e municipal."

A INDE tem o objetivo de ordenar a geração, armazenamento, acesso, compartilhamento, disseminação e uso dos dados geológicos de origem federal, estadual, distrital e municipal, promover a utilização, a produção dos dados geológicos, os padrões e normas homologados, e evitar a duplicidade de ações, através de um conjunto integrado de normas, padrões e especificações homologados através da CONCAR (Comissão Nacional de Cartografia).

No que diz respeito aos padrões, percebe-se que os mesmos são rígidos e só existe interoperabilidade se todos os dados estão completamente compatíveis. Contudo, a preocupação do INDE é meramente sintática, sem preocupação semântica, o que tende a afetar a interoperabilidade de uma forma geral.

Atualmente, as empresas buscam os padrões para compatibilizarem seus dados e metadados aos padrões INDE (SANTOS, 2012). Mas essa afirmativa pouco se aplica às empresas relacionadas a E&P do petróleo principalmente pela maioria não ser nacional e, tal qual o OGC, a sintaxe do INDE não oferece detalhamento suficiente para prover interoperabilidade entre todas as gigantes do petróleo.

⁵<http://www.inde.gov.br>

3.2.3 Energistics

A Energistics foi fundada sob o nome POSC, Petrotechnical Open Software Corporation (Corporação Petrotécnica de Software Aberto), em 1990, patrocinada por 5 das maiores empresas ligadas ao petróleo da época: BP, Chevron, Elf, Mobil e Texaco⁶.

A missão da Energistics, segundo a própria entidade é: desenvolver, dar suporte, evoluir e prover padrões abertos para os aspectos científicos, de engenharia e de operação relacionados à indústria de E&P do petróleo e gás. A Energistics, portanto, dá foco especial aos aspectos ligados ao petróleo e gás.

No decorrer dos anos, diversas outras companhias aderiram à entidade e em 2006 ela é renomeada para Energistics e, atualmente, conta com mais de 100 das maiores empresas e organizações ligadas à milionária indústria do petróleo.

Devido ao fato de os padrões Energistics representarem o estado da arte atual no que se refere à padronização no contexto da E&P de petróleo, o capítulo 5 apresenta a evolução do POSC/Energistics conforme as propostas de padronização também foram evoluindo, e mostra com detalhes os padrões RESQML, WITSML e PRODML.

3.2.4 ANP - Agência Nacional de Petróleo, Gás Natural e Biocombustíveis

A ANP⁷ é o órgão regulador das atividades que integram a indústria do petróleo e gás natural e a dos biocombustíveis no Brasil. Autarquia federal, vinculada ao Ministério de Minas e Energia, a ANP é responsável pela execução da política nacional para o setor energético do petróleo, gás natural e biocombustíveis.

A ANP tem como finalidades:

- Regular: estabelecer regras por meio de portarias, instruções normativas e resoluções para o funcionamento das indústrias e do comércio de óleo, gás e biocombustíveis.
- Contratar: promover licitações e assinar contratos em nome da União com os concessionários em atividades de exploração, desenvolvimento e produção de petróleo e gás natural, e autorizar as atividades das indústrias reguladas.
- Fiscalizar: fazer cumprir as normas nas atividades das indústrias reguladas, diretamente ou mediante convênios com outros órgãos públicos.

A ANP é também um centro de referência em dados e informações sobre a indústria do petróleo e gás natural: mantém o Banco de Dados de Exploração e Produção (BDEP), promove estudos sobre o potencial petrolífero e o desenvolvimento do setor; recebe e torna públicas as notificações de descobertas; divulga as estatísticas oficiais sobre reservas e produção no Brasil; realiza pesquisas periódicas sobre qualidade dos combustíveis e lubrificantes, e sobre preços desses produtos. Na área dos biocombustíveis, mantém e divulga dados sobre autorizações, produção e comercialização de biodiesel e etanol.

Essa agência é uma importante entidade reguladora do país no setor energético do petróleo, gás natural e biocombustíveis. No entanto, suas atribuições fogem do escopo deste trabalho, uma vez que são de natureza jurídica e, até o presente momento, não envolvem esforços de padronização de trocas de informação no setor.

⁶www.energistics.org/history

⁷www.anp.gov.br

4 CONCEITOS E PADRÕES GENÉRICOS PARA TROCA DE INFORMAÇÕES

Neste capítulo, abordam-se os conceitos mais relevantes e os padrões mais usados atualmente no que se refere à interoperabilidade: as linguagens de marcação, com foco em XML e seus derivados, e os Web Services. Os conhecimentos deste capítulo são fundamentais por servirem como base para os padrões RESQML, WITSML, PRODML e LAS, apresentados no capítulo 5. XML, Web Services e arquivos ASCII, apresentados aqui, são considerados padrões sem domínio, pelo fato de não estarem ligados a qualquer campo específico de conhecimento, mas sim, fornecerem estruturas para a construção de padrões mais complexos, como os que serão abordados no capítulo 5.

4.1 Metadados

Metadados são uma maneira de organizar dados para tornar mais fácil a busca por arquivos importantes. Metadados são, na verdade, dados sobre dados ou informações sobre recursos pesquisáveis de dados (DOWNEY, 2007). O objetivo é popular os metadados, torná-los pesquisáveis e manter os elementos metadados para questões de propriedade dos dados, rastreabilidade e retenção de tempo.

Por apresentar essas características, metadados integram alguns dos principais padrões utilizados nas geociências, principalmente os Dublin Core Elements, vistos com mais detalhes na subseção 5.2.4.

4.2 Linguagem de Marcação

Linguagens de marcação são informações adicionadas a um documento que aumenta seu significado de alguma forma, identificando as partes do documento e como estas partes se relacionam entre si (RAY, 2001). Portanto, uma linguagem de marcação é um conjunto de símbolos que podem ser colocados no texto desse documento para demarcar e rotular as partes de um documento. Têm como objetivo incluir no próprio documento informações relativas aos seus metadados: formato interno, tipos de informação e, de forma limitada, seu significado.

Marcações são importantes em documentos eletrônicos, pois elas são processadas por programas de computador. Caso um documento não possua rótulos e limites, um programa não vai saber como tratar um pedaço de texto para diferenciar os diferentes fragmentos de um texto. Este problema não acontece quando é usado XML para a marcação de documentos. A seguir, é apresentada a linguagem XML e algumas outras, baseadas em XML, que são usadas em conjunto para atingir os objetivos de marcação, apresentação e

transformação de documentos.

4.2.1 XML

A linguagem Extensible Markup Language (XML) possui características que incluem (RAY, 2001):

- É uma linguagem de representação de documentos e não uma linguagem de programação;
- É um padrão aberto, independente de plataforma e aplicação;
- É extensível;
- É modular: permite a definição de um novo formato de documento pela combinação e reuso de outros formatos;
- Usa *Unicode* como padrão de caracteres, portanto, suporta diversos tipos de símbolos;
- Possui regras de formatação e diversos outros artefatos para apresentação e transformação de documentos;
- Por ter sintaxe fácil e simples, e estrutura não ambígua, é fácil de ler e analisar por humanos e programas de computador;

Por apresentar essas características, o XML forma a base dos padrões RESQML, WITSML e PRODML, alguns dos padrões do domínio de E&P mais importantes atualmente.

A Figura 4.1 é usada aqui para mostrar algumas das principais características de um documento XML. É fundamental saber que todo documento possui elementos, formados geralmente por uma tag de abertura, conteúdo e a tag de fechamento. Na tag de abertura pode-se encontrar atributos, que devem conter um nome, um símbolo de "=", e um valor, que deve obrigatoriamente estar entre aspas. No exemplo acima, um elemento é a tag `<corredor>`, com o atributo "tipo", o conteúdo, contendo pelo menos uma tag `<produto>`, e a tag de fechamento `</corredor>`.

Algumas regras são obrigatórias na construção de um documento XML e podem ser extraídas do exemplo da Figura 4.1, que mostra uma possível organização de informações em um domínio de supermercados. A primeira característica é que todo documento XML deve ter uma declaração de que o documento é XML bem como de uma versão. Outra característica é que todo documento XML deve possuir um, e somente um, elemento raiz e qualquer outro elemento deve estar dentro deste. As tags de início e fim são obrigatórias e, quando não há conteúdo, as tags podem ter um dos dois formatos: `<exemplo1></exemplo1>` ou `<exemplo1 />`. E, como já mencionado, todo valor de atributo deve estar contido entre aspas.

O XML apresenta algumas outras características, como não possuir tags pré-definidas, possibilitando ao usuário criar seu próprio conjunto de tags, possibilitar o uso de infinitos elementos, bem como de aninhamento de elementos, além de ser uma linguagem *case sensitive*.

Embora o XML seja poderoso, ele costuma ser usado com diversas outras linguagens que acrescentam poder e que são, igualmente, baseadas em XML. Pode-se citar aqui

```

<?xml version="1.0"?>
<mercado>
  <corredor id="1" tipo="Chocolate">
    <produto id="11">
      <descricao>Chocolate Branco</descricao>
      <preco>5.80</preco>
    </produto>
    <produto id="12">
      <descricao>Chocolate Preto</descricao>
      <preco>5.20</preco>
    </produto>
  </corredor>
  <corredor id="2" tipo="Fruta">
    <produto id="21">
      <descricao>Banana</descricao>
      <preco>2.10</preco>
    </produto>
    <produto id="22">
      <descricao>Maçã</descricao>
      <preco>1.80</preco>
    </produto>
  </corredor>
</mercado>

```

Figura 4.1: Exemplo de documento XML

linguagens de formatação de documentos, como o Document Type Definition (DTD) e XML Schema, que definem a formatação da estrutura de um documento, restringindo quais tags podem ser usadas e em que quantidade. Há também as linguagens de estilo e transformação, representadas pelo Cascading Style Sheets (CSS) e Extensible Stylesheet Language (XSL). Este último tipicamente dividido em três partes, que serão vistas com maiores detalhes na próxima seção, juntamente com os XML Schema.

4.2.2 XSL

Conforme visto, XML é um formato baseado em texto que pode ser usado para manter dados em diferentes plataformas e diferentes tipos de aplicações. Mas muitas vezes esta informação em XML pode não ser imediatamente útil e um programa para manipular os dados deve ser escrito. Uma das atividades mais comuns é manipular a forma de apresentação de um XML, que pode ser em formato HTML, documento PDF, texto para e-mail, etc.

O objetivo do XSL em sua concepção era criar uma linguagem de apresentação que pudesse solucionar as limitações do CSS. Em sua criação, o XSL foi dividido em três

linguagens (TENNISON, 2005):

- XSL-FO: A XML Formatting Objects é uma linguagem puramente de apresentação que descreve como a formatação de objetos deve ser mostrada em uma página;
- XSLT: a XSL Transformation define como transformar qualquer linguagem de marcação baseada em XML em outra linguagem de marcação;
- XPath: a XML Path Language é usada junto com XSLT para apontar para pedaços de informações em um documento XML

Um documento XML escrito em XSLT é comumente chamado de XSLT Stylesheet e, normalmente, tem a extensão *.xsl*. Cada documento XSLT descreve como um conjunto de documentos XML (chamados de documentos fonte) deve ser convertido em outros documentos (chamados de documentos destinos ou resultados), que podem ser XSL-FO, XHTML, CVS ou diversos outros formatos. Em geral, um documento XSLT vai usar documentos fonte escritos em uma linguagem de marcação em particular para produzir os resultados.

Para realizar a transformação, o XSLT precisa apontar para as informações no documento resultado, de forma que ele possa ser processado e incluído no resultado. Este é o papel do XPath e explica o porquê de XSLT e XPath trabalharem juntos na transformação de documentos. O papel mais importante do XPath é coletar informações de um documento XML navegando através do documento. Também é função do XPath permitir cálculos, como uma linguagem expressiva.

```
<?xml version="1.0"?>
<xsl:stylesheet version="1.0"
  xmlns:xsl="http://www.w3.org/1999/XSL/Transform">
  <xsl:template match="produto">
    Produto:
    <xsl:apply-templates select="descricao" />
    Preço:
    <xsl:apply-templatesselect="preco" />
  </xsl:template>
</xsl:stylesheet>
```

Figura 4.2: Exemplo de documento XSLT

A Figura 4.2 mostra um pequeno exemplo do uso de XSLT sobre o arquivo XML mostrado na Figura 4.1. Em primeiro lugar, todo documento XSLT é um XML válido, tendo como elemento raiz um elemento *stylesheet* ou *transform*, sendo o *stylesheet* o mais usado e definido pela tag *xsl:stylesheet*. Através da tag *xsl:template*, pode-se definir um texto a ser impresso na tela, caso o valor do atributo *Match* seja o mesmo. No exemplo, a cada vez que uma tag *produto* for encontrada, será impressa a string "Produto:" e, a cada vez que a tag *preco* for encontrada, será impressa a string "Preço:". Por fim, temos a tag *xsl:apply-template* que imprime o valor de uma tag, desde que o atributo *select*, da tag *xsl:apply-template* coincida com a tag desejada. No exemplo, *<xsl:apply-template select="descricao"/>* imprime o conteúdo do elemento *descricao*, o qual aparece 4 vezes no arquivo XML. A saída transformada é mostrada na Figura 4.3.

Embora seja um exemplo básico, já dá uma ideia de como o XSLT pode realizar transformações poderosas. Para gerar saídas HTML, por exemplo, pode-se inserir tags HTML a cada vez que determina tag for encontrada durante a navegação no arquivo XML. O XSLT é, de fato, uma poderosa ferramenta na transformação de arquivos XML e torna a linguagem ainda mais interoperável, pois pode criar compatibilidade, em casos onde ela inexistente.

A principal utilidade das transformações para os formatos ligados à E&P é poder manipular os dados da maneira desejada, podendo até extrair relatórios com maior ou menor grau de detalhamento a partir de relatórios de produção no PRODML ou de dados de logs, em WITSML, por exemplo.

```
Produto: Chocolate Branco Preço: 5.80
Produto: Chocolate Preto Preço: 5.20
Produto: Banana Preço: 2.10
Produto: Maçã Preço: 1.80
```

Figura 4.3: Resultado da transformação mostrada na Figura 4.2 aplicada sobre o documento descrito na Figura 4.1

4.2.3 XML Schema

Um XML Schema, aqui chamado de esquema XML, define uma linguagem de marcação através da especificação de quais elementos e atributos são permitidos e que tipo de conteúdo eles podem conter (TENNISON, 2005). Existem diversos tipos diferentes de esquemas, sendo dois os mais comumente usados:

- DTDs: linguagem de esquema que é embutida no XML;
- XML Schema: linguagem oficial W3C (consórcio que criou as definições referentes às linguagens de marcação).

Qualquer esquema apresenta, ao menos conceitualmente, três tipos de informação:

1. Declaração de elementos, que especifica o tipo esperado de um elemento com nome e contexto particular e o grupo de substituição a que o elemento pertence;
2. Declaração de atributos, que especifica o tipo esperado de um atributo com nome e contexto particular;
3. Definição de tipos, que define o conteúdo permitido para um certo elemento ou atributo e que se relaciona com as outras duas declarações através de uma hierarquia de tipos.

Quando se fala sobre padronização no contexto do uso de documentos XML, cabe aos esquemas XML realizar a validação dos documentos. Isto é fundamental para que se crie um padrão, pois, de outra forma, tal qual acontece com a proposta PXML, visto na subseção 3.1.1, cria-se um conjunto de tags que podem ser facilmente compreensíveis para usuários mas que, por não possuírem uma estrutura definida, podem ser escritas incorretamente, tornando a interoperabilidade incompleta e até inalcançável.

A validação checka se um documento XML está de acordo com os esquemas, gerando uma lista de erros de validação, caso haja, ou indicando sucesso, que indica que o documento é válido. Portanto, os esquemas XML lidam com a sintaxe dos documentos XML. De certa forma, pode-se dizer que há uma preocupação com a semântica do documento, fornecida pelos esquemas XML, pois, para certos elementos, pode-se definir no esquema quais são os tipos de dados que podem ser seu conteúdo, como strings, números, datas, etc.

Quanto aos esquemas DTD e XML Schema, a principal diferença entre as duas é que o XML Schema usa uma sintaxe baseada em XML enquanto os DTDs tem uma sintaxe própria, o que a faz ser constantemente criticada. Porém, ainda assim, os DTDs apresentam uma linguagem concisa, enquanto o XML Schema apresenta uma sintaxe verbosa, no entanto, menos intimidante para quem já é familiarizado com XML.

Um XML Schema, também conhecido como XML Schema Definition (XSD), é o padrão substituto aos DTD. Para seu uso, cria-se um documento, com extensão *.xsd* que contém as definições de um documento válido, segundo as regras que se deseja aplicar. No arquivo XML que vai fazer uso de tais regras para ser validado, deve-se inserir uma sequência de caracteres que definem que este documento XML faz referência a um documento XSD.

A utilidade dos esquemas XML para os padrões RESQML, WITSML e PRODML é, portanto, fornecer interoperabilidade sintática, pois, são estes esquemas que definem o que é válido em documentos dos três padrões.

4.3 SOA

O Service Oriented Architecture (SOA) faz uso da orientação a serviços, um paradigma ainda emergente e que tem serviços como o elemento de construção, para desenvolvimentos rápidos, de baixo custo de composição e de aplicações distribuídas.

Os serviços são módulos auto-contidos que podem ser descritos, publicados, localizados, orquestrados e programados usando tecnologias baseadas em XML em uma rede. A premissa da computação orientada a serviços é que uma aplicação não deve mais ser considerada como um processo único rodando dentro de uma organização. O valor de uma aplicação passa a ser medido não mais pelas suas capacidades funcionais, mas sim, por sua habilidade de integração com o ambiente que a cerca (PAPAZOGLU, 2008).

O modelo orientado a serviços permite uma distinção clara entre provedores de serviços, clientes de serviços e agregadores de serviços. Devido ao fato dos serviços serem oferecidos por empresas diferentes e comunicadas sobre a Internet, eles oferecem uma infraestrutura de computação distribuída para uma empresa em si e também para a comunicação entre diferentes empresas, permitindo fácil integração e colaboração.

O modelo mais comum que faz uso do SOA é chamado Web Service e será apresentado a seguir.

4.3.1 Web Services

Web Services são módulos de software auto-descritos, auto-contidos, disponíveis em uma rede, como a Internet, que realiza tarefas, resolve problemas ou conduz transações através de requisições de clientes. Permite rodar aplicações em diferentes máquinas, que trocam dados e se integram, sem a necessidade de software e hardware adicionais ou de terceiros. É um padrão que permite troca dados independentemente de linguagem, plataforma ou protocolos internos usados e é baseado em outros padrões bem estabelecidos

como XML e HTTP (PAPAZOGLU, 2008).

O funcionamento dos Web Services é baseado em troca de informações e serviços fundamentados em 3 padrões: 1) protocolo de comunicação Simple Object Access Protocol (SOAP), 2) Descritor de serviços Web Services Description Language (WSDL) e 3) publicação e descobrimento de serviços com Universal Description, Discovery, and Integration (UDDI).

Um provedor de serviços descreve sua interface através do padrão WSDL em um agregador de serviços. O cliente de serviços utiliza UDDI para encontrar serviço desejado no agregador. Para a comunicação entre provedor e cliente, o padrão SOAP é utilizado e é feita diretamente entre eles, sem mais comunicação com o agregador.

Web Services são uma tecnologia escolhida pelos padrões mais utilizados na área de E&P de petróleo devido à suas características que tornam interoperabilidade uma consequência de sua implementação. Os padrões WITSML e PRODML já são compatíveis com essa tecnologia. O RESQML tem em seus planos futuros a implementação de Web Services. Já LAS pode vir a ser usado junto com Web Services mas, por não ter XML na base de sua construção, torna mais trabalhosa a adaptação. No capítulo 6 é mostrada uma visão geral da implementação de um Web Service com o padrão WITSML.

4.4 Arquivos ASCII

Arquivos ASCII possuem como característica o fato de conterem caracteres letras, números e símbolos legíveis, de acordo com um padrão universalmente aceito. Seu formato binário representa um padrão aberto, o que torna-os facilmente lidos e manipulados. Os arquivos ASCII não têm um formato próprio, eles são um modelo aberto de dados. Os arquivos LAS, vistos na seção 5.1, por exemplo, têm formato *.las* e são arquivos ASCII.

Por serem lidos e manipulados por programas de uso geral, não há a necessidade de aplicativos específicos ou proprietários. Com isso, arquivos ASCII podem ser a base sobre a qual são adicionadas semântica e sintaxe, para a criação de um padrão de troca de informações específico, como o que ocorre com o padrão LAS.

5 PADRÕES PARA TROCA DE INFORMAÇÕES EM GEOCIÊNCIAS

A principal entidade ligada à padronização no que se refere a E&P do petróleo é a Energistics. Como já mencionado, de 1990 a 2006, o nome ainda era POSC, alterado neste último ano para Energistics. Desde os primeiros anos de existência da entidade, sua preocupação era com a integração entre os diferentes aplicativos existentes no mercado que, em geral, realizavam as mesmas operações, mas utilizando padrões diferentes para entrada e saída de dados.

Ainda em 1992, com a criação do SIP (Software Integration Platform), foram lançadas as primeiras especificações que incluíam modelos de acesso a dados, de troca de dados e interfaces com o usuário. Após 3 anos testando e incrementando a plataforma, chegou-se ao dado de que, o uso da plataforma, gerava economia entre 1 e 3 dólares por barril de petróleo, devido aos avanços na qualidade dos dados, acessibilidade de dados e melhorias da informação e avanços no conhecimento (ENERGISTICS, 1996).

Entre 2000 e 2001, os primeiros esquemas de dados XML começam a surgir na plataforma, direcionando a uma troca de especificações de armazenamento de dados e middlewares para especificações de trocas de dados. Em 2002, surge o embrião do WITSML para troca de dados de perfuração, baseado em XML e com uma concepção inicial da tecnologia Web Services. Neste mesmo ano é formado o primeiro SIG - Special Interest Group (Grupo de interesse em especial), grupo de empresas e entidades que realizam investimentos para desenvolver a tecnologia.

Em 2003 é criado um SIG para dar suporte ao uso do padrão WITSML e no mesmo ano é lançado uma especificação para padrões de dados de logs de poço e documentações adicionais de E&P. Em 2005, surge a segunda família de padrões baseados em XML e Web Services, o PRODML, com o apoio de um SIG, e uma nova versão do WITSML é lançada também, trazendo melhorias significativas. Além disso, é lançado um software de conversão open source de LAS para WITSML.

Em 2007, uma versão de WITSML com XSD é publicada e, em 2008, uma nova versão, que incrementa ainda mais questões de interoperabilidade, é lançada e, por fim, surge, em 2009, o RESQML, substituto do RESCUE, para endereçar padrões de caracterização de reservatórios. No mesmo ano, uma nova versão do PRODML com novas especificações de dados e serviços surge.

Com isso, três padrões da Energistics se destacam no que tange os fluxos de E&P do petróleo: RESQML, WITSML e PRODML. Não é por acaso que são três, e não apenas um padrão: cada um deles está relacionado a uma fase diferente do *upstream*. RESQML relaciona-se com a modelagem de reservatórios. WITSML está relacionado com a etapa de perfuração. Por fim, PRODML está relacionado com a produção.

Paralelamente aos padrões Energistics, há o padrão LAS, relacionado diretamente com o padrão WITSML no que se refere aos dados modelados no padrão. O LAS, contudo, não é baseado em XML como os demais, mas sim, em ASCII.

Além do LAS e do WITSML, há diversos outros padrões importantes relacionados com a etapa de perfuração. Entre eles, pode-se citar o Log Interchange Standard (LIS), Digital Log Interchange Standard (DLIS) e WellLogML.

Tabela 5.1: Comparação entre padrões de log de poços. Adaptado de (GEOMIND, 2007)

Propriedades	LAS	LIS	DLIS	WITSML	WellLogML
Formato	ASCII	Binário	Binário	XML	XML
Manuseio	++	-	-	+	+
Distribuição	++	++	++	O	-
Profundidade da definição	+	O	++	++	+
Número de tipos de dados	+	-	++	++	+
Suporte das aplicações	++	++	++	-	-
Aceitação	++	++	++	O	-
Portabilidade	++	-	-	++	+
Flexibilidade	++	-	-	++	+

	Notação
++	Muito Bom
+	Bom
O	Satisfatório
-	Ruim
-	Muito Ruim

A Tabela 5.1, adaptada de (GEOMIND, 2007), faz uma comparação, sob diversos aspectos, dos principais padrões envolvidos com dados de log de poços. Analisando a tabela, percebe-se que os padrões LAS, LIS e DLIS são aqueles com maior aceitação e uso na indústria. Isso se dá pelo fato de serem mais antigos, ainda baseados em ASCII e binários. No caso do LIS e DLIS, por terem formato binário, sua portabilidade é afetada e o objetivo de interoperabilidade se torna quase inalcançável. Já o WellLogML, é baseado em XML, contudo, acabou sendo ofuscado pelo formato WITSML por modelar uma quantidade menor de tipos de dados. Portanto, no que tange os conceitos de interoperabilidade, os padrões LAS e WITSML são os mais importantes.

Embora a tabela mostre que o padrão LAS trás mais vantagens do que o padrão WITSML, este cenário pode ser alterado nos próximos anos. O LAS foi criado muito antes que o WITSML, por isso, apresenta vantagens de distribuição, suporte e aceitação. Entretanto, o WITSML se baseia em tecnologias mais recentes, como o XML, e atinge interoperabilidade de uma maneira mais natural. Por isso, acredita-se que o WITSML, no futuro, pode ser um substituto ao LAS.

Além de LAS e WITSML, os padrões RESQML e PRODML também são analisados e detalhados. Ao final do capítulo é apresentada a Tabela 5.3, que traz um comparativo entre estes padrões, mostrando quais são os dados modelados por cada padrão, bem como informações sobre versão, data de criação, entidade criadora, suporte a Web Services e informações ao programador.

5.1 LAS

O Log ASCII Standard (LAS) foi inicialmente proposto em 1990 pela sociedade Canadense de Log de Poços (Canadian Well Logging Society) para facilitar e simplificar a troca de informações digitais de poços entre diferentes companhias e clientes. Por ser um arquivo em formato ASCII, os dados em LAS podem ser importados e exportados para qualquer plataforma, estação de trabalho, ou mainframe, sem a necessidade de softwares proprietários e/ou caros. Embora construído para complementar os formatos mais robustos já existentes da época, o LAS se tornou desde então o formato mais usado para transferência digital de dados de logs de poços e um verdadeiro padrão da indústria (HESLOP et al., 1999). Nesta última versão, o LAS já traz a preocupação recorrente do WITSML com aquisição de dados em tempo real.

Em comparação com os padrões RESQML, WITSML e PRODML, o LAS cumpre papel semelhante ao WITSML. Embora o objetivo de ambos seja muito semelhante e ambos sejam padrões não proprietários, os dois padrões são completamente diferentes em suas estruturas. Enquanto o WITSML se baseia em XML, LAS se baseia em arquivos ASCII. Alguns problemas podem aparecer devido a essa característica, como o uso de terminações de linhas diferentes para diferentes plataformas. Enquanto em plataformas Windows/DOS e MAC PC's um final de linha é marcado por um caractere ASCII 13 seguido de um ASCII 10, na plataforma Unix apenas o caractere ASCII 10 é usado. O analisador sintático funciona perfeitamente dentro da plataforma na qual o arquivo LAS foi criado, contudo, problemas podem ocorrer quando há transferências entre plataformas. Portanto, cuidados adicionais devem ser tomados com arquivos LAS para que sejam, de fato, independentes de plataformas. O mesmo problema não ocorre em arquivos XML pois, diferentemente de arquivos ASCII, novas linhas não são determinantes para a estrutura dos documentos, baseados em tags. Em ASCII, quebras de linhas adicionam detalhes sintáticos: nova linha representa mudança de informação sendo armazenada na estrutura.

Para fazer a validação de um arquivo LAS existe o software LAS Certify. Desta forma, é possível garantir que a estrutura de um documento está de acordo com a especificação em (HESLOP et al., 2000). A seguir, será explicada a estrutura básica de arquivos LAS.

5.1.1 Estrutura LAS

Um arquivo LAS é composto de diversas seções. Cada uma começa com uma linha de título, marcada pelo uso do caractere "~" no início da linha. Além de conter uma extensa quantidade de seções pré-definidas, é possível criar novas seções definidas pelo usuário. Há algumas seções obrigatórias, como *~Version*, *~Well*, *~Curve* e *~ASCII*, sendo ainda necessário que *~Version* e *~Well* sejam as duas primeiras seções de um documento, nesta ordem (HESLOP et al., 2000).

Na Figura 5.1, extraída de (HESLOP et al., 2000), é mostrado parte de um arquivo LAS. Neste trecho, são apresentadas duas seções, obrigatoriamente as duas primeiras de qualquer arquivo LAS: *~Version* e *~Well*. O caractere "#" é um dos muitos caracteres reservados do LAS e indica comentário. Todos os mnemônicos (como STRT, STOP, STEP, etc) são obrigatório em uma seção de poço. Contudo, apenas os mnemônicos STRT, STOP e STEP têm a obrigatoriedade de ter um valor associado. Os demais podem ser nulos. Este trecho também mostra o uso dos caracteres ":", também reservado, e que atua como delimitador entre um valor e uma descrição, e do caractere ".", usado como delimitador entre um mnemônico e uma de medida.

Cada seção contém linhas onde os dados são descritos e armazenados. Os dados, por

sua vez, são armazenados como arranjos de uma, duas ou três dimensões, usualmente, indexados por profundidade e tempo, mas, também, podendo ser apresentados como uma medida discreta. Para compreender melhor a estrutura, são apresentados os tipos primários de seções, bem como os tipos de linhas das seções.

Quanto aos tipos primários de seções, estes são quatro, descritos a seguir:

- **Parameter Data Sections:** Contêm qualquer quantidade de linhas do tipo **Parameter Data**.
- **Column Definition Sections:** Contêm qualquer quantidade de linhas do tipo **Column Definition**. Relacionada ao armazenamento de dados detalhados de duas ou três dimensões, armazenados em uma seção **Column Data**. Sempre deve haver uma seção do tipo **Column Data** correspondente;
- **Column Data Sections:** Contêm qualquer quantidade de linhas do tipo **Column Data**. Sempre há uma seção do tipo **Column Definition** que faz uma descrição dos elementos. Também pode haver uma seção **Parameter Data** que possui parâmetros relacionados;
- **User Defined Sections:** É aconselhado usar este tipo somente quando estritamente necessário, dando preferência às acima citadas.

Quanto aos tipos de linhas dentro das seções, existem três tipos:

- **Parameter Data Lines:** Presente nas seções do tipo **Parameter Data**. Cada linha contém um dado unidimensional de 1 ou 2 elementos. Cada linha contém, também, uma descrição completa deste dado. Algumas linhas são obrigatórias em algumas seções;
- **Column Definition Lines:** Aparecem somente em seções do tipo **Column Definition**. Embora possua estrutura idêntica às **Parameter Data Lines**, cada linha **Column Definition** é usada para descrever cada canal contido na seção **Column Data** correspondente.
- **Column Data Lines:** Aparecem somente em seções do tipo **Column Data**. Cada linha contém uma série de valores delimitados de dados. O caractere delimitador é definido pelo valor do parâmetro *DLM* na seção *~Version*. Descrições de cada dado são encontradas nas seções **Column Definition** correspondentes.

A Figura 5.2, extraída de (HESLOP et al., 2000), é mostrada aqui para ilustrar os conceitos de tipos de seções e tipos de linhas de seções mencionados anteriormente. Nesta figura, duas seções são descritas: *~Log_Definition* e *~Log_Data*. A seção *~Log_Definition* é do tipo **Column Definition** e possui diversas **Column Definition Lines**. Já a seção *~Log_Data* é do tipo **Column Data** e possui diversas **Column Data Lines**. Não é coincidência que o número de atributos separados por vírgula em cada linha da seção *~Log_Data* coincida com o número de linhas da seção *~Log_Data*: o número é exatamente o mesmo pois cada valor corresponde a um dos canais da definição. Adicionalmente, a Figura 5.1 mostra a seção *~Well* que é uma seção do tipo **Parameter Data** onde cada linha igualmente é do tipo **Parameter Data**.

Nestes dois exemplos, são mostradas grande parte das características da estrutura do padrão LAS. O manual completo da versão 3.0 do LAS está em (HESLOP et al., 2000).

```

~VERSION INFORMATION
VERS.                3.0 : CWLS LOG ASCII STANDARD -VERSION 3.0
WRAP.                NO : ONE LINE PER DEPTH STEP
DLM .                COMMA : DELIMITING CHARACTER BETWEEN DATA COLUMNS
# Acceptable delimiting characters: SPACE (default), TAB, OR COMMA.
~Well Information
#MNMN.UNIT          DATA          DESCRIPTION
#-----
STRT .M            1670.0000        : First Index Value
STOP .M            1669.7500        : Last Index Value
STEP .M            -0.1250         : STEP
NULL .             -999.25         : NULL VALUE
COMP .             ANY OIL COMPANY INC. : COMPANY
WELL .             ANY ET AL 12-34-12-34 : WELL
FLD .              WILDCAT         : FIELD
LOC .              12-34-12-34W5M     : LOCATION
CTRY .             CA              : COUNTRY
PROV .             ALBERTA         : PROVINCE
SRVC .             ANY LOGGING COMPANY INC. : SERVICE COMPANY
DATE .             13/12/1986       : LOG DATE {DD/MM/YYYY}
UWI .              100123401234W500  : UNIQUE WELL ID
LIC .              0123456         : LICENSE NUMBER
API .              12345678        : API NUMBER
LATI.DEG           34.56789        : Latitude {F}
LONG.DEG           -102.34567      : Longitude {F}
GDAT .             NAD83           : Geodetic Datum
# Lat & Long can also be presented as:
# LATI .           45.37° 12' 58"    : X LOCATION
# LONG .           13.22° 30' 09"    : Y LOCATION
UTM .              : UTM LOCATION

```

Figura 5.1: Exemplo do uso de *~Version* e *~Well* em um arquivo LAS. Extraído de (HES-LOP et al., 2000)

5.2 RESQML

Dentre os 3 padrões (RESQML, WITSML, PRODML), o RESQML é certamente o mais complexo, pois envolve todo o fluxo de trabalho de investigação da existência de um reservatório até antes da realização das perfurações. É também o que apresenta maior documentação e literatura disponível.

Conforme visto na seção 2.2, os dados da exploração vêm de diferentes lugares e envolvem pessoas de diferentes áreas de estudo, eventualmente até de diferentes empresas e usando diferentes ferramentas de software. Neste tipo de ambiente, os dados podem ser analisados, repassados e eventualmente voltam para correções e atualização. Com uma configuração multidisciplinar onde diferentes sistemas trabalham com diferentes formatos, é comum a ocorrência de gargalos, perda e inconsistência de dados. Tempo que poderia ser gasto no estudo do domínio acaba sendo gasto na resolução de problemas com dados.

O RESQML aborda esta questão para todas as ações envolvendo a exploração de petróleo através de um padrão de formato de dados independente de fornecedor. Segundo a própria definição do RESQML (ENERGISTICS; RESQML SIG, 2011a) ele é:

"An industry-developed, vendor-neutral format that facilitates data exchange among the many software applications used along the E&P subsur-

```

~Log_Definition
#MNEM.UNIT          LOG CODES          CURVE DESCRIPTION
#-----
# Format of value in data section F=float, E=0.00E00 S=string
DEPT .M              : DEPTH {F}
DT .US/M            123 456 789      : SONIC TRANSIT TIME {F}
RHOB .K/M3          123 456 789      : BULK DENSITY {F}
NPHI .V/V           123 456 789      : NEUTRON POROSITY {F}
SFLU .OHMM          123 456 789      : SHALLOW RESISTIVITY {F}
SFLA .OHMM          123 456 789      : SHALLOW RESISTIVITY {F}
ILM .OHMM           123 456 789      : MEDIUM RESISTIVITY {F}
ILD .OHMM           123 456 789      : DEEP RESISTIVITY {F}
YME .PA             123 456 789      : YOUNGS MODULES {E0.00E+00}
CDES .              123 456 789      : CORE DESCRIPTION {S}

# A 2D array channel begins here. It has 5 elements.
# value after A: is time spacing of this array element from first element.
NMR[1] .ms          123 456 789      : NMR Echo Array {AF;0ms}
NMR[2] .ms          123 456 789      : NMR Echo Array {AF;5ms}
NMR[3] .ms          123 456 789      : NMR Echo Array {AF;10ms}
NMR[4] .ms          123 456 789      : NMR Echo Array {AF;15ms}
NMR[5] .ms          123 456 789      : NMR Echo Array {AF;20ms}

~Log_Data | Log_Definition
1670.000,123.450,2550.000,0.450,123.450,123.450,110.200,105.600,1.45E+12,DOLOMITE WI/VUGS,10,12,14,18,13
1669.875,123.450,2550.000,0.450,123.450,123.450,110.200,105.600,1.47E+12,LIMESTOVE ,12,15,21,35,25
1669.750,123.450,2550.000,0.450,123.450,123.450,110.200,105.600,2.85E+12,LOST INTERVAL ,18,25,10,8,17

```

Figura 5.2: Exemplo de tipos de seções e de linhas de seções. Extraído de (HESLOP et al., 2000)

face workflow, which helps promote interoperability and data integrity among these applications and improve workflow efficiency."

O RESQML é, portanto, um padrão de troca de dados baseado em XML que visa resolver o problema de múltiplos pacotes de software existentes, ou seja, é focado na interoperabilidade de dados. Sua abordagem é feita sobre as etapas de exploração: análise, interpretação, modelagem e simulação das subsuperfícies.

Alguns dos benefícios que este padrão traz são:

- Rastreabilidade: o padrão permite adicionar metadados que provêm informações sobre a alteração dos dados. É possível informar quem fez alterações, quando foram feitas e o que foi feito;
- Requisitos Regulamentais: com propriedades de rastreabilidade disponíveis, é mais fácil gerenciar os requisitos regulamentais e toda a segurança da informação associada ao uso dos dados;
- Eficiência: tempo que poderia ser gasto resolvendo problemas de dados pode ser gasto resolvendo problemas do domínio;
- Integridade dos dados: ajuda a garantir que todos os dados são transferidos e no formato correto, além de possuir ferramentas que permitem manuseio de grande volume de dados de forma rápida e eficiente;
- Rapidez na tomada de decisões: permite atualização rápida de modelos e até em tempo real. Com isso, as decisões podem ser tomadas mais rapidamente.

Quanto ao escopo e a funcionalidade, o RESQML engloba todas as funcionalidades já contidas nas bibliotecas fornecidas pelo projeto RESCUE e adiciona diversas outras

Tabela 5.2: Descrição das características da RESQML. Adaptado de (ENERGISTICS; RESQML SIG, 2011a)

Características	Descrição
Infraestrutura	Um documento RESQML contém um arquivo XML e opcionalmente, um arquivo HDF, que pode ser usado para processar de forma mais rápida e eficiente conjuntos grandes de dados
Informação Semântica	Metadados, rastreabilidade, referência de dados para propriedades permitindo mapeamento correto de nomes entre diferentes pacotes de softwares, unidades de medida, códigos padrões de localidade
Estrutura	Falhas e horizontes como componentes individuais
Grade 3D	Arranjo de propriedades estáticas e dinâmicas, grade local básica de refinamento, falhas e conexões entre não vizinhos
Poços	Poços bloqueados/IJK e infraestrutura de suporte para objetos de dados WITSML

funcionalidades. A Tabela 5.2 mostra as características e as descrições individuais de cada característica.

Além disso, existem diversos domínios de E&P contemplados pelo RESQML na última versão lançada (1.1). Estes domínios são descritos no Anexo A, na Tabela 10.1, mostrando como o RESQML permite que diferentes pacotes de software consigam interoperar, sendo no mesmo nível de resolução do problema ou quando um pacote precisa finalizar um trabalho para que o próximo possa iniciar seu fluxo.

A seguir, serão apresentados diversos padrões que fornecem funcionalidades para o RESQML, como HDF5, EPGS Codes, entre outros. O conhecimento destes elementos é importante para melhor compreender, na subseção 5.2.6, como eles são usados para compor os objetos de dados do padrão RESQML.

5.2.1 HDF5

Por trabalhar com coordenadas geográficas e espaciais referenciadas por pontos, modelos RESQML podem conter milhões de células. Para tratar características geométricas, além de outras propriedades destes elementos, pode-se usar opcionalmente um arquivo Hierarchical Data Format 5 (HDF5) que é um conjunto de formatos de arquivos aberto para armazenar e organizar grandes quantidades de dados numéricos aumentando a velocidade e a eficiência do processamento (ENERGISTICS; RESQML SIG, 2011b).

O HDF5 fornece:

- Formato "binário" de arquivos independente de arquitetura (suporte em Windows, Linux, etc. APIs em C++, Java, dentre outros);
- Decomposição de arranjo de dados, de forma que sub-arranjos possam ser extraídos de um documento RESQML sem ter que ler o documento inteiro.

5.2.2 EPGS Codes

O EPGS (European Petroleum Survey Group) era um grupo que publicava bancos de dados que incluíam códigos para facilitar a referência para localizações globais bem conhecidas, como a elevação do nível médio do mar ou do cume de uma montanha⁸.

⁸<http://www.epsg.org/>

O grupo acabou sendo absorvido pelo OGP (International Association of Oil & Gas Producers) que agora é o proprietário do banco de dados EPGs.

Os EPGs Codes podem ser usados pelos esquemas RESQML para geolocalização da definição de um CRS (Coordinate Reference System) - o sistema de coordenadas de referência.

5.2.3 GML

O GML (ver 3.2.1) é uma gramática XML para expressar características geoespaciais.

```

<Well>
  <volume>10000</span>
  <depth>200</depth>
  <gml:centerLineOf>
    <gml:LineString>
      <gml:pos>100 200</gml:pos>
      <gml:pos>200 200</gml:pos>
    </gml:LineString>
  </gml:centerLineOf>
</Well>

```

Figura 5.3: Exemplo de código GML

A Figura 5.3 apresenta um exemplo de código GML que ilustra como a gramática pode interagir com uma aplicação. O elemento raiz *Well* (Poço, em português) é o elemento de interesse e, supõe-se ser definido por um esquema já definido, assim como suas propriedades de *depth* (profundidade) e *volume*. Essas tags não são GML, mas sim, os elementos que interessam quando deseja-se fazer mapas para análises. As tags GML vêm na sequência, iniciadas pelo prefixo *gml:*, significando que ali está um elemento GML, que deve estar definido em um GML Schema. A sequência GML da Figura, mostra o traçado de algumas retas para fazer um desenho de um poço. Embora incompleto, este exemplo tem por objetivo esclarecer que o GML é usado em conjunto com outros esquemas, desde que ele próprio tenha um esquema que o defina.

5.2.4 Dublin Core Elements

Quando se trata de arquivos geológicos, como linhas sísmicas, logs de poços, colunas estratigráficas, mapas, etc, é importante manter registros de quem é o proprietário dos arquivos, quem fez a última alteração nos dados, qual o sistema de coordenadas, quando os dados foram criados e qual a versão corrente, entre outros. Metadados solucionam este problema e mantêm o rastro dos dados.

Dublin Core é o padrão de metadados que é usado pelo RESQML para oferecer rastreabilidade.

5.2.5 GUIDs

Os Global Unique Identifiers são usados para cada documento RESQML e cada objeto principal dentro de um documento, como horizontes, falhas e grades. O uso de GUIDs torna possível manipular objetos independentemente de uma grade e trabalhar com Dublin Core Elements para oferecer rastreabilidade.

5.2.6 Objetos de Dados e Conceitos Chave

A Figura 5.6 mostra os objetos que um documento RESQML pode conter: horizontes, falhas e grades, e suas respectivas propriedades.

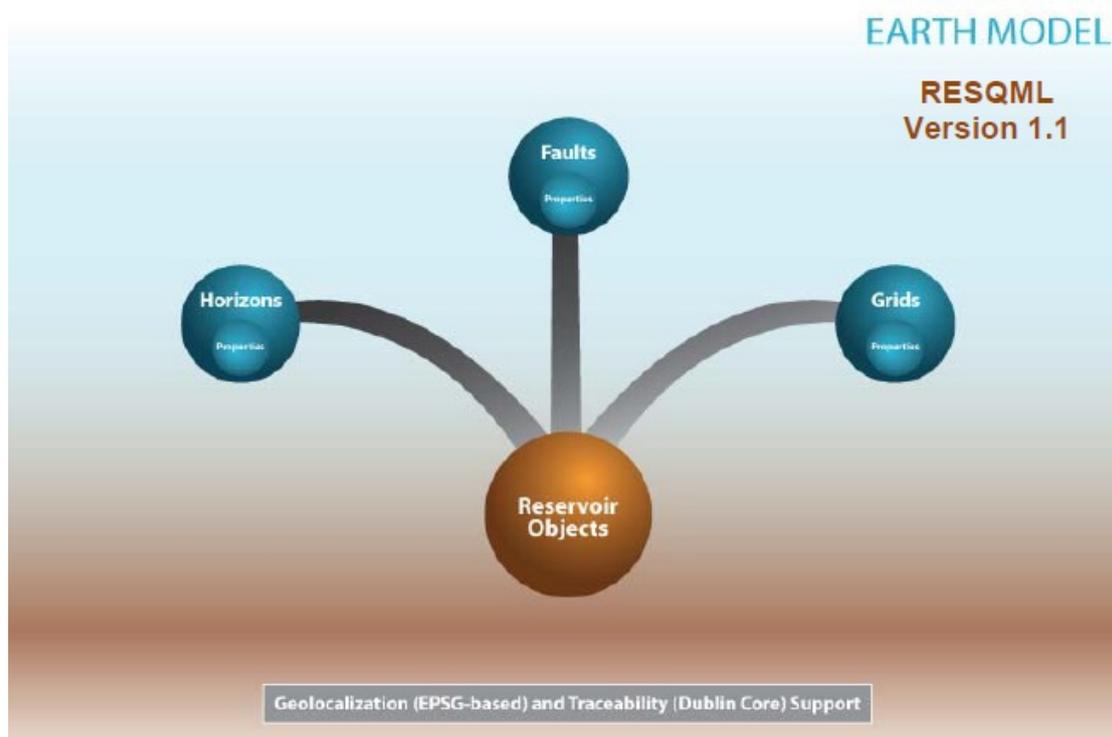


Figura 5.4: Objetos de dados suportados pela versão 1.1 do RESQML. Extraído de (ENERGISTICS; RESQML SIG, 2011a)

Para descrever com maior precisão objetos de dados no processo de modelagem, existem 3 terminologias usadas pelo RESQML:

- **Features (Feições):** Podem ser naturais ou artificiais e representam um conceito associado com objetos geológicos, normalmente de subsuperfície. Um horizonte, por exemplo, é uma feição em um perfil sísmico que pode ser interpretada como um limite entre unidades litológicas, uma falha ou uma discordância entre unidades;
- **Version (Versão):** Interpretação ou opinião. É uma descrição única e consistente de uma feição geológica. Uma feição pode ter uma ou mais versões;
- **Representation (Representação):** descrição digital de uma feição. Um horizonte, por exemplo, é representado por um conjunto de pontos, superfícies triangulares ou grades ortogonais. Em RESQML, representações devem ser em tempo ou profundidade. Um misto não é permitido.

Essas terminologias são importantes para entender os objetos de dados representáveis. Começando pelas falhas e horizontes - que são as feições que definem limites entre unidades litológicas - que podem ter múltiplas versões e múltiplas representações geométricas para cada versão. Estas, podem incluir conjuntos de pontos 3D, grades 2D ortogonais,

malhas triangulares, etc, além de cada horizonte ou falha poderem ter seu conjunto particular de propriedades, que torna possível transferir informação sobre a estrutura central sendo representada.

Os grids, ou grades de pontos, são usados para representar estruturas complexas de reservatórios onde cada célula é definida por seus 8 nós de canto. O RESQML dá suporte às grades através de listas de adjacências, suporte a múltiplas grades, e indicadores topológicos.

O RESQML na versão 1.1 não integra detalhes de objetos de poços (nascentes, trajetórias, etc) mas já apresenta compatibilidade com o WITSML.

Um conceito chave para a compreensão do RESQML é o Coordinate Reference System (CRS). Um CRS define um dado espacial georreferenciado que se relaciona com locais reais, permitindo que modelos sejam precisamente localizados na superfície da terra (HYNE, 2001). O CRS é especificado usando GML, que inclui EPSG Codes para identificar coordenadas para localizações específicas e globais.

Cada representação geométrica é especificada com respeito a coordenadas de sistemas locais, embutidas no sistema CRS global (com opções de rotação, translação e mudança nas unidades de medida).

No capítulo 6 serão apresentados detalhes sobre a implementação do RESQML em um pacote de software.

5.3 WITSML

A perfuração de poços geralmente se insere nas tarefas de *upstream*, sendo comumente vista como um meio termo entre as áreas de exploração e produção. Após a investigação dos locais onde o petróleo ocorre com potencial econômico (etapa de exploração) é necessário realizar perfurações para avaliar a verdadeira qualidade dos reservatórios. Perfurações são muito caras e, por isso, são dependentes de uma etapa de exploração bem conduzida para ter o máximo possível de garantia sobre o preço de sua realização. É comum que as perfurações sejam realizadas por empresas terceirizadas ou por outros setores de uma mesma empresa (FEIJO, 2010) e, com isso, é importante que dados sejam transferidos entre diferentes sistemas de diferentes companhias. Neste contexto, surge o WITSML, como tentativa de padronizar esta comunicação.

Anterior ao WITSML, o WITS (Wellsite Information Transfer Specification) foi o padrão para transferência de dados de perfuração por um tempo. É um formato multi-nível considerado flexível mas que em baixo nível é baseado em formatos ASCII e por isso não muito amigável devido à dificuldade de entendimento por humanos. Com o aumento do volume de dados, ficou claro que este padrão não tinha suporte o suficiente e a Energistics assumiu o desafio de oferecer um novo padrão: o WITSML.

O Wellsite Information Transfer Standard Markup Language (WITSML) nasceu com o objetivo de tratar as funções de perfuração, completação e intervenções da área de *upstream*, visando fornecer um padrão universalmente aceito que facilite o fluxo de dados técnicos entre redes de diferentes empresas, companhias de serviço, contratantes de perfuração, vendedores de aplicações e companhias reguladoras. Segundo a própria definição do WITSML (ENERGISTICS; WITSML SIG, 2011) ele é:

"The Wellsite Information Transfer Standard Mark-up Language (WITSML) consists of XML data-object definitions and a Web services specification developed to promote the right-time, seamless flow of well data

between operators and service companies to speed and enhance decision making."

O WITSML é baseado em padrões existentes como XML, SOAP e WSDL. Assim como o RESQML, a linguagem é definida com o uso de XML Schema que especificam tudo, desde objetos de dados que podem ser representados até unidades de medidas. Isso permite uma validação rápida e precisa dos documentos produzidos e garante que os documentos trocados estejam no formato correto.

Alguns benefícios do WITSML:

- Melhora no conceito Plug&Play de troca de dados entre sistemas;
- Redução de custos de produção de software;
- Maior competitividade entre fornecedores de software;
- Reduz a necessidade de uso de diferentes softwares para diferentes operadores;
- Aumento da eficiência operacional.

Um dos conhecimentos mais importantes no que se refere às perfurações são os dados de logs de perfuração, abordados na seção 2.3. O WITSML tem sua maneira própria de lidar com estes dados. É este o assunto da subseção apresentada a seguir. Questões de implementação do WITSML são abordadas no capítulo 6.

5.3.1 Litologias em WITSML

Uma característica importante do WITSML é lidar com dados litológicos de perfurações. A litologia dos locais de perfuração é um parâmetro geológico baseado em amostras laterais e descrição petrográfica de calhas. Estes dados são essenciais na determinação da profundidade estratigráfica de poços e do contexto geológico de poros e fluidos. Empresas que operam em perfurações têm que relatar os dados litológicos para parceiros e autoridades reguladoras, além de fazer uso desta informação nas decisões diárias. Dados litológicos, obtidos de perfurações, em WITSML podem ser medições numéricas de propriedades (resistividade, gama, etc) ou descrições petrográficas textuais. Frequentemente esses dados são perdidos após 6 meses de completção do poço pois os dados são apresentados como simples imagens de pacotes proprietários e os componentes litológicos, bem como suas interpretações e descrições, não fazem parte destes registros (ENERGISTICS; WITSML SIG, 2012).

Os elementos litológicos WITSML têm por objetivo resolver este problema, aumentando a interoperabilidade entre diferentes softwares que interagem com os mesmos dados. Além disso, a intenção do WITSML é possibilitar interoperabilidade de terminologia e definições com outros padrões também existentes com mesmo objetivo, como o One-Geology e GeoSciML (da comissão CGI).

A Figura 5.5 mostra um exemplo validado de litologia usando WITSML. Diversas informações estratigráficas são encontradas neste documento, incluindo o tempo geológico (eras, períodos, épocas, etc), um importante tópico de estudo para compreensão de subsuperfície, além de informações da litologia e metadados do documento.

```

<?xml version="1.0" encoding="UTF-8"?>
<!--Example of Mud Log data From Lithology Element User Guide -->
<mudLogs xmlns = "http://www.witsml.org/schemas/1series" xmlns:xsi =
"http://www.w3.org/2001/XMLSchema-instance" xsi:schemaLocation =
"http://www.witsml.org/schemas/1series ../xsd_schemas/obj_mudLog.xsd" version = "1.4.1.0">
  <documentInfo>
    <documentName>mudLog</documentName>
    <fileCreationInformation>
      <fileCreationDate>2011-12-15T16:15:31.827Z</fileCreationDate>
      <fileCreator>John Smith</fileCreator>
    </fileCreationInformation>
  </documentInfo>
  <mudLog uidWell = "W-12" uidWellbore = "B-01" uid = "h45a">
    <nameWell>W-12</nameWell>
    <nameWellbore>B-01</nameWellbore>
    <name>Exmouth Marina #1</name>
    <geologyInterval uid = "00001">
      <typeLithology>interpreted</typeLithology>
      <mdTop uom = "ft">495</mdTop>
      <mdBottom uom = "ft">670</mdBottom>
      <chronostratigraphic kind = "era">Mesozoic</chronostratigraphic>
      <chronostratigraphic kind = "period">Triassic</chronostratigraphic>
      <chronostratigraphic kind = "epoch">Middle Triassic</chronostratigraphic>
    </geologyInterval>
    <geologyInterval uid = "00002">
      <typeLithology>interpreted</typeLithology>
      <mdTop uom = "ft">495</mdTop>
      <mdBottom uom = "ft">645</mdBottom>
      <lithostratigraphic kind = "group">Mercia Mudstone Group</lithostratigraphic>
      <lithostratigraphic kind = "formation">Sidmouth
        Mudstone Formation</lithostratigraphic>
      <chronostratigraphic kind = "stage">Ladinian</chronostratigraphic>
    </geologyInterval>
  </mudLog>
</mudLogs>

```

Figura 5.5: Exemplo de descrição litológica WITSML. Adaptado de (ENERGISTICS; WITSML SIG, 2012)

5.4 PRODML

O PRODML está ligado diretamente aos termos: campos petrolíferos do futuro, iFields, eFields, SmartFields. São todos esforços para usar instrumentação e software para otimizar operações em todos os domínios de E&P de óleo e gás. Com o ritmo acelerado dos negócios de óleo e gás, preços voláteis, conformidades regulamentares mais apertadas, encolhimento de preços, terabytes de dados e tempo limitado para análise de informação, as empresas estão continuamente pressionadas a aumentar a eficiência de suas operações de E&P (REDDY, 2010).

Segundo a própria definição do PRODML (ENERGISTICS; PRODML SIG, 2010) ele é:

"An industry initiative to provide open, non-proprietary, standard interfaces between software tools used to monitor, manage and optimize hydrocarbon production."

O escopo do PRODML permite operações, otimização e relatórios da produção, bem como gerência de dados para o domínio de E&P. A versão atual do PRODML aborda, além da otimização da produção, outras questões, como: relatórios de produção, pesquisa de distribuição de fibra-ótica e resultados de produção em laboratórios.

O foco é nos fluxos de trabalho de produção desde os limites dos poços-reservatórios até os pontos de transferência de custódia. São atividades que incluem otimização de

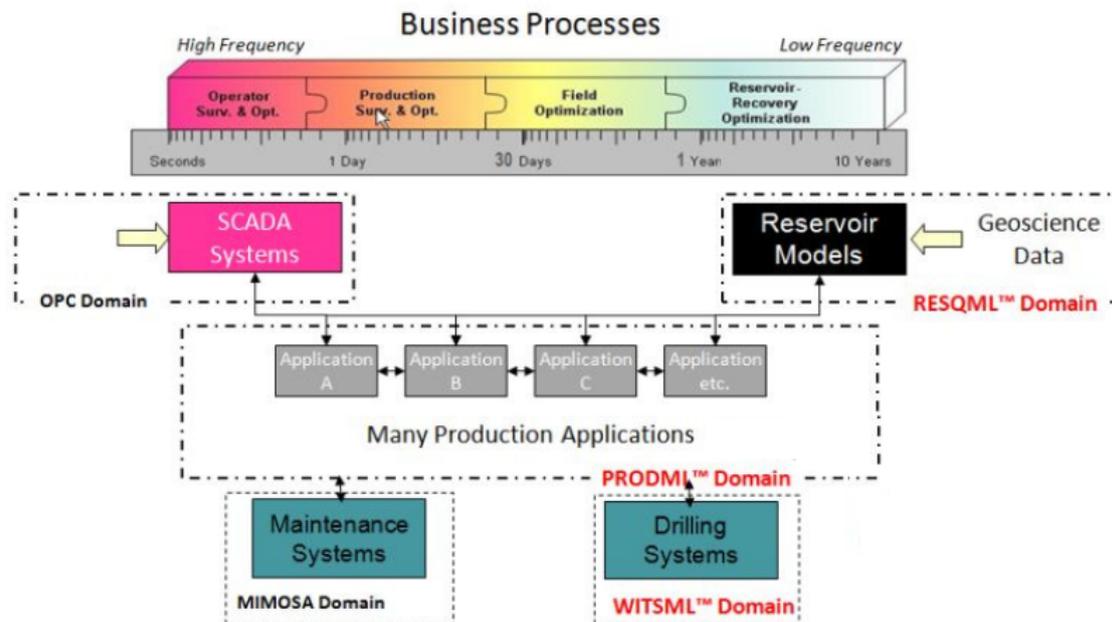


Figura 5.6: RESQML, PRODML e WITSML em um processo de negócios E&P. Adaptado de (ENERGISTICS; PRODML SIG, 2010)

poços e de produção de ativos, monitoração de poços e instalações para produção, integridade e confiabilidade e até transferência de relatórios de produção em formato padrão para parceiros, reguladores e até internamente na companhia.

O PRODML facilita a manipulação e transferência de dados e acelera o processo de trabalho, aumentando e otimizando a produção. Anteriormente, companhias petrolíferas e provedores de serviços dedicavam um tempo muito grande, além de dinheiro e esforço, para tratar dados e criando sistemas para comunicações específicas. Com PRODML, busca-se a padronização para que se possa investir o dinheiro em mais tecnologia e para que a transferência de dados seja natural e a interoperabilidade seja alcançada. Alguns benefícios potenciais do uso de PRODML podem ser listados:

- Aumento da eficiência operacional: Sistemas de otimização de produção mais confiáveis e precisos com custos menores sobre propriedades de dados, pois a informação será usada de forma mais eficiente e efetiva;
- Operações mais seguras: Possibilidade de monitoração remota, colaboração e intervenções oportunas para resolver problemas podem resultar em exposição reduzida para o pessoal, ocasionando um ambiente mais seguro de trabalho;
- Melhoria dos dados - confiabilidade e conformidade: Melhoria da qualidade dos dados facilita a gestão de informações para monitorar e otimizar o desempenho de ativos e relatório. Os operadores podem trocar mais facilmente de dados e colaborar com os parceiros, governos de acolhimento, e prestadores de serviço. além de assegurar o cumprimento das obrigações contratuais, empresariais e regulatórias;
- Aceleração na adoção de práticas de otimização de produção: Devido à padronização, não será necessário reinventar a roda;
- Consciência mais ampla de oportunidades para a otimização da produção.

A Figura 5.7 mostra um exemplo validado pelo esquema PRODML para um arquivamento de relatório. Um pacote de software que implemente leitura e escrita de documentos PRODML deve ser capaz de importar um destes arquivos, através do uso de um *parser* XML ou de um proxy generator (mais detalhes no capítulo 6), para manipular os dados, que nada mais são do que os conteúdos dos elementos da figura.

```
<?xml version="1.0" encoding="UTF-8"?>
<?xml-stylesheet type="text/xsl" href="./stylesheets/generic.xsl"?>
<report
  xmlns="http://www.prodml.org/schemas/1series"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://www.prodml.org/schemas/1series
    ../xsd_schemas/obj_report.xsd">
  <kind>monthly production</kind>
  <title>Rubin Monthly Production for March 2009</title>
  <date>2009-03-12</date>
  <comment>This is an example of a report context for
    a monthly production of a field.</comment>
  <reportVersion>3-internal</reportVersion>
  <reportStatus>draft</reportStatus>
  <contextFacility namingSystem="MyCompany"
    kind="field">RUBIN</contextFacility>
  <operator>
    <name>Norsk Olje AS</name>
  </operator>
  <geographicContext>
    <country>Norway</country>
    <field>RUBIN</field>
  </geographicContext>
  <issueDate>2006-04-19</issueDate>
  <issuedBy>
    <name>Ole Mann</name>
    <role>Driftsleder</role>
    <role>Platformsjef</role>
    <personName>
      <first>Ole</first>
      <last>Mann</last>
    </personName>
  </issuedBy>
</report>
```

Figura 5.7: Exemplo de relatório PRODML

Uma característica do PRODML é que ele foi criado com o objetivo de usar as tecnologias de Web Services, que, como visto em 4.3.1, são baseadas em XML o que torna-as igualmente interoperáveis e independentes de plataforma e aplicação.

Tabela 5.3: Comparação entre os principais padrões ligados à E&P

	RESQML	WITSML	PRODML	LAS
Área de Aplicação	Reservatórios	Perfuração	Produção	Perfuração
Data de Criação	2009	2002	2005	1990
Proposto Por	Energistics	POSC (Energistics)	POSC (Energistics)	Canadian Well Logging Society
Versão Atual	1.1	1.4.1	1.1	3.0
Dados Modelados	Falhas, Horizontes, Grades 3D e Poços (Modelagem)	Dados de Logs	Fluidos e Poços (dados de produção)	Dados de Logs
Modelo da Linguagem	XML	XML	XML	ASCII
Suporta Web Services	Não	Sim	Sim	Não
Possui guia para programadores	Sim	Sim	Não	Não

6 IMPLEMENTAÇÃO DOS PADRÕES RESQML E WITSML

Este capítulo mostra uma visão geral do que envolve a implementação de documentos RESQML e WITSML. Não são detalhados códigos de implementação, pois os mesmos são complexos e poderiam ser um tema completo para novos trabalhos.

A implementação dos formatos RESQML, WITSML e PRODML tem muitos pontos em comum, que envolvem *parsers* XML e clientes e servidores Web Services. Nas seções a seguir, serão analisadas características da implementação RESQML, particulares e detalhes comuns a outras, e WITSML, quanto à Web Services. Não são apresentados detalhes sobre PRODML pois, removidas características particulares dos dados deste padrão, ela segue os mesmos princípios de implementação.

6.1 Visão geral do RESQML

Neste seção, é apresentado um fluxo de trabalho de alto nível para implementar RESQML em um pacote de software. A implementação requer o uso de ferramentas especiais de software para ler os esquemas fornecidos pela Energistics e, opcionalmente, integrar bibliotecas especiais para uso do HDF5 (ENERGISTICS; RESQML SIG, 2011b).

A Tabela 6.1 mostra os possíveis documentos RESQML. Embora seja opcional, a ausência de um arquivo HDF5 é aplicável apenas em projetos muito pequenos.

Tabela 6.1: Tipos de arquivos e nomeações RESQML. Extraído de (ENERGISTICS; RESQML SIG, 2011b)

Tipo de Arquivo	Obrigatório	Convenção de nome	Exemplo de nome
XML	Sim	.resqml	Projeto.resqml
HDF5	Não	nome resqml + .h5	Pojeto.resqml.h5

Para a implementação é necessário obter, através do website da Energistics⁹, os esquemas RESQML, que incluem esquemas Dublin Core, esquemas GML e esquemas da Energistics.

Também, caso deseje-se implementar um arquivo HDF5, é necessário fazer download das bibliotecas HDF5, disponibilizadas pelo grupo HDF¹⁰.

Em todos os padrões (RESQML, WITSML e PRODML), é necessário trabalhar com arquivos XML. Há duas formas principais de extrair os dados destes documentos:

⁹<http://www.energistics.org/current-resqml-standards> - Disponível em 29 de Maio de 2012

¹⁰<http://www.hdfgroup.org/HDF5> - Disponível em 29 de Maio de 2012

- Parser XML - componente de software capaz de ler e escrever documentos XML, como o RESQML;
- Proxy Generator - usado para criar classes de proxy que podem serializar e desserializar objetos RESQML diretamente.

Cada um tem vantagens e desvantagens. *Parsers* são mais flexíveis que Proxy Generators, contudo, podem gerar mais esforço de manutenção e desenvolvimento.

A leitura de documentos XML usando *parsers* envolve 2 passos: 1) validar o documento para garantir que o mesmo está em conformidade com os esquemas RESQML e 2) ler o documento: o *parser* transforma o conteúdo para o formato apropriado. Já a leitura por *proxy generators* envolve: 1) desserializar o documento RESQML usando as classes de *proxy* para criar uma representação em memória do arquivo, 2) criar representações internas dos objetos em memória e 3) se o arquivo RESQML indica que há dados gravados em arquivos HDF5, leia-os para popular os dados internos.

Já a escrita usando *parsers* envolve a transformação da representação interna para o formato de *parser*, como uma hierarquia de objetos, seguida da escrita do conteúdo para o arquivo. A escrita por proxy generators inicia com a criação de uma representação em memória do arquivo XML usando as classes de proxy dos objetos internos, a escrita dos dados no arquivo HDF5 e atualização da memória para indicar a localização nos dados nos arquivos HDF5 e, por fim, a serialização da representação em memória para um arquivo XML.

Outra questão importante é que a Energistics fornece documentos XSLT para realizar transformações entre versões antigas para versões novas, pois, em geral, as novas versões trazem mudanças com relação às antigas.

Detalhes da implementação são encontrados em (ENERGISTICS; RESQML SIG, 2011b).

6.2 Visão geral do WITSML

O WITSML consiste de definições de dados de objetos e de uma especificação de Web Services. Em forma de itens:

- Esquemas de dados que definem os objetos, como logs e poços, e seus relacionamentos. Estes esquemas preocupam-se com a sintaxe correta dos documentos;
- enumValues.xml - um arquivo de carregamento que contém tabelas para definição de valores válidos para os elementos como unidades de medida e tipos litológicos, ou seja, preocupam-se com aspectos semânticos dos documentos;
- STORE API - descreve as funcionalidades e comportamentos dos servidores WITSML e como aplicações clientes podem interagir com os servidores.

A abordagem de validações de documentos, que incluem o uso de *parsers* para conferir se todos os elementos são válidos e a sintaxe que engloba esquemas XML é correta, já foi apresentada na seção anterior. Nesta seção, o foco é na abordagem de Web Services.

A implementação de WITSML pode incluir os seguintes itens (ENERGISTICS; WITSML SIG, 2011)

- Esquemas de dados e arquivos enumValues (obrigatórios);

- Servidor WITSML: processa requisições, adições, deleções, ou atualizações de dados WITSML, conforme definido pelos esquemas de dados WITSML;
- Aplicações clientes: softwares designados para compreender e usar esquemas de dados WITSML e para interagir com servidores WITSML;
- SOAP: usado para transportar pedidos e respostas entre clientes e servidores.

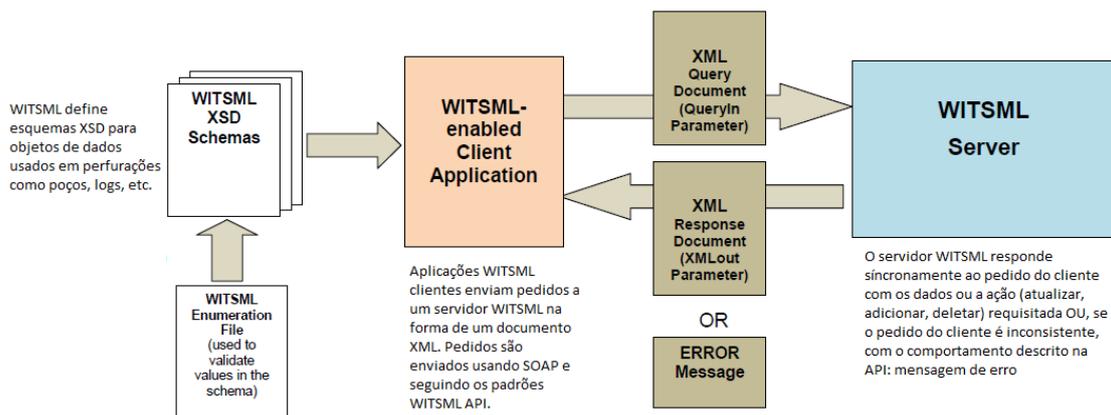


Figura 6.1: Visão geral dos componentes e interações chaves entre clientes e servidores WITSML. Adaptado de (ENERGISTICS; WITSML SIG, 2011)

A Figura 6.1 mostra os principais conceitos envolvidos na implementação de clientes e servidores WITSML. Os *enumeration files* são usados para validar os valores nos esquemas XML (XSD Scheme) que, por sua vez, validam a sintaxe dos documentos XML que são enviados pelas aplicações clientes para o servidor através de protocolo de transporte SOAP. O servidor responde sincronamente aos pedidos dos clientes com os dados ou ações (atualização, remoção, etc) ou, em caso de inconsistência na requisição do cliente, mensagem de erro.

Para definir as funções da interface dos servidores, é usado um arquivo WSDL que faz parte do padrão da linguagem WITSML e, visando garantir a interoperabilidade entre implementações, não deve ser alterado (exceção é o atributo de localização do endereço do elemento SOAP). Quando é feita uma requisição ao servidor, esse deve responder imediatamente e de forma síncrona com uma mensagem de recebimento (ACK) ou mensagem de erro.

Os objetos são as figuras mais importantes dos documentos WITSML. Eles são representações lógicas e organizacionais associados com os componentes e operações envolvidos no processo de perfuração de um poço. Contudo, um objeto WITSML não é um objeto de programação, com métodos, propriedades, eventos, etc. Eles possuem relações entre eles, definidas pelos esquemas de dados WITSML e tabelas de enumeração estáticas associadas, como unidades de medidas, litologia e outras. Cada objeto tem um identificador e a identificação de seu ancestral ou derivação de ancestrais. Eles são organizados em uma árvore de hierarquia com o objeto "poço" sendo o elemento raiz. Este elemento pode ter filhos, que são cada perfuração executada neste poço. Cada perfuração pode ter filhos associados, que são dados da perfuração, como logs, trajetória e outros objetos. Cada objeto (poço, perfurações, logs, etc) tem um ID único associado, que é determinado através da união de seu ID próprio unido aos IDs de seus ancestrais. A Figura 6.2 mostra a hierarquia mencionada.

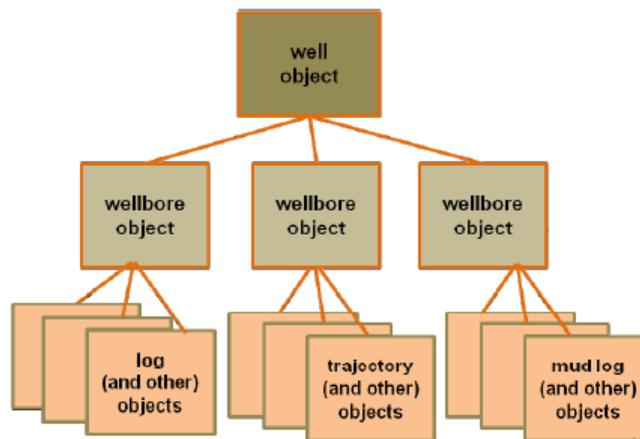


Figura 6.2: Relações entre objetos WITSML. Adaptado de (ENERGISTICS; WITSML SIG, 2011)

É definido uma especificação quanto à representação e transporte dos dados, de modo a garantir a interoperabilidade. Contudo, não é definido um padrão de armazenamento. Quanto à representação, os objetos sendo trocados são sempre documentos XML cujos formatos são definidos pelos esquemas de objetos. Pelo fato de arquivos XML serem simples arquivos texto, não há problemas em usar os protocolos HTTP (não encriptado) e HTTPS (encriptado) como protocolos de transporte. Contudo, embora sejam protocolos recomendados, não são obrigatórios.

Diversos outros aspectos são essenciais para a implementação de WITSML e maiores detalhes de documentação e implementação são encontrados em (ENERGISTICS; WITSML SIG, 2011).

7 RESULTADOS DA UTILIZAÇÃO DOS PADRÕES NA INDÚSTRIA E DA PESQUISA

Os padrões desenvolvidos pela Energistics ainda são muito recentes e há poucos trabalhos que demonstrem resultados quantitativos e qualitativos de seu uso.

Um destes poucos trabalhos de validação sobre o uso do padrão RESQML é descrito em (MICHEL; VERNEY, 2011). O objetivo deste trabalho foi transferir interpretações de resultados internos de horizontes, falhas e marcadores de formações de poços como objetos RESQML para diversas aplicações comerciais de Geologia estrutural, estratigráficas e de reservatórios e, após, importar novamente todos estes resultados para a aplicação interna para garantir, a cada passo, uma consistência geral com as interpretações originais.

Esta sequência de transferência de dados incluiu vários aspectos comuns do fluxo de trabalho de exploração desde interpretações sísmicas de horizontes e falhas, modelagem de frameworks estruturais de reservatórios, grades 3D, até simulações do fluxo de fluidos de reservatório.

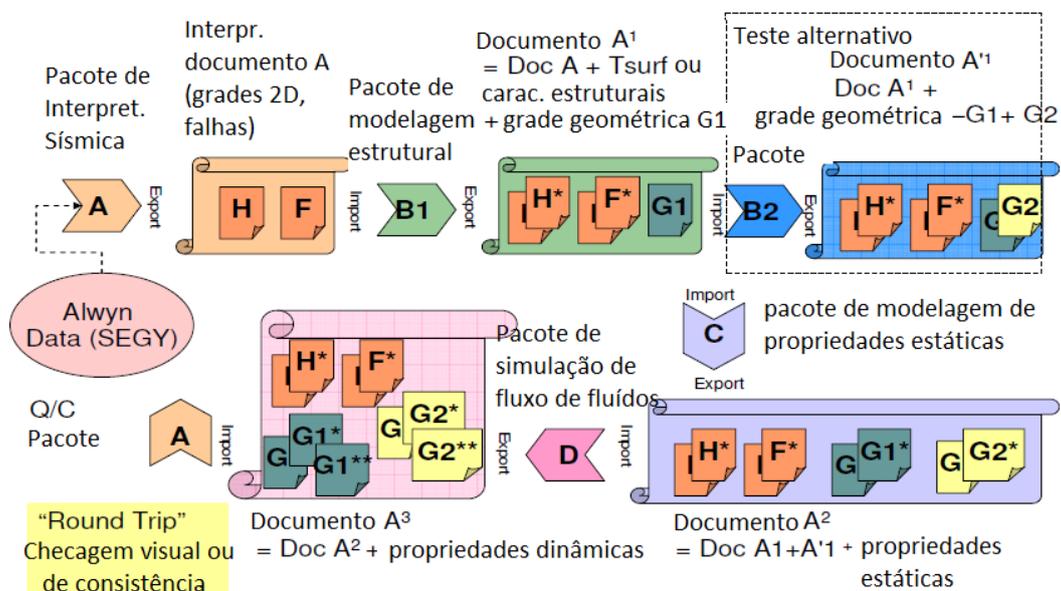


Figura 7.1: Fluxo de validação da caracterização de um reservatório. Adaptado de (MICHEL; VERNEY, 2011)

A Figura 7.1 mostra um caso de uso, onde H representa horizonte, F representam falhas, * indica triangulação e G representa grade. O fluxo mostra a exportação de interpretações sísmicas para pacotes de modelagem estrutural, modelagem de propriedades

estáticas, simulação de fluidos e uma nova importação para validar a consistência de dados.

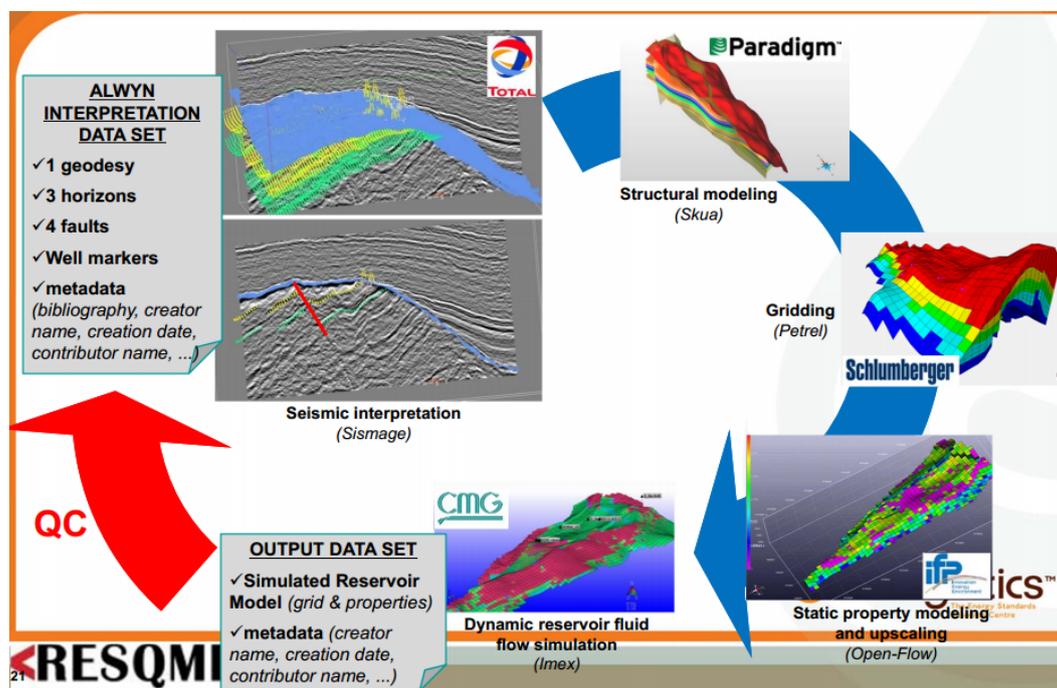


Figura 7.2: Validação através do uso de ferramentas proprietárias. Adaptado de (RAINAUD, 2011)

A Figura 7.2 mostra uma validação do padrão RESQML através do fluxo de dados desde o início da cadeia de exploração, na interpretação sísmica, até a adição das propriedades dinâmicas de um reservatório. Ao final, os dados deveriam estar íntegros e corretos, o que, de fato, aconteceu. Essa validação tem importância especial, pois, para fazê-la foram usadas algumas das principais ferramentas proprietárias do mercado, como Petrel, para gradeamento, e Open-Flow, para adicionar as propriedades estáticas ao reservatório.

Não é possível saber se essas propriedades de importação e exportação mostradas na validação estão disponíveis em todas as ferramentas, pois entre os testes e a divulgação no mercado há uma grande distância. Contudo, pelo fato da demonstração envolver diversos pacotes de software de modelagem diferentes, este documento acaba sendo uma importante validação do padrão RESQML.

Quanto ao padrão PRODML, a Energistics divulgou em (ENERGISTICS; PRODML SIG, 2010) uma lista que apresenta o uso do padrão PRODML por companhias ligadas à área de energia. Na lista, são mostradas empresas que já fazem uso constante do padrão e uma lista de testes piloto em diversas outras empresas, sendo algumas das maiores do mercado petrolífero, como Chevron e BP. Esses dados são importantes para mostrar que, embora incipiente, o padrão começa a ser testado em ambientes reais.

Por fim, quanto ao resultado da análise dos diferentes padrões, a Figura 7.3 trás um resumo importante. A figura mostra que, na construção dos esquemas RESQML, os esquemas GML são importantes para localizar geograficamente os elementos RESQML, bem como os esquemas Dublin Core são importantes para adicionar rastreabilidade através de metadados aos elementos. A figura mostra ainda, dentro do contorno vermelho, os elementos que devem ser tratados através das ferramentas proprietárias, que devem ser capazes de importar e exportar esses elementos. Todos os elementos fora do contorno são

fornecidos pela Energistics e são estes que garante a padronização.

Contudo, o mais importante desta figura é a compreensão do significado das setas tracejadas que partem do elemento WITSMML e PRODML, e chegam até o elemento RESQML. Essas setas significam que essa funcionalidade ainda não é existente e devem ser um lançamento futuro. Ou seja, atualmente, ainda não há suporte entre os diferentes padrões. Isso é importante, pois significa que os padrões da Energistics ainda não fornecem interoperabilidade entre as diferentes etapas do ciclo de E&P do petróleo. A interoperabilidade existente, no formato atual, é entre os diversos processos que existem dentro de cada uma das etapas, como os que aparecem na Figura 7.2. Portanto, a integração dos dados entre as etapas ainda é feita fora dos padrões, pelos profissionais que manuseiam os dados.

Essa questão de interoperabilidade entre diferentes etapas é uma, dentre diversas outras, das limitações dos padrões da Energistics. Por isso, a entidade faz planos de evolução para os padrões. Essa evolução é vista com mais detalhes no capítulo seguinte.

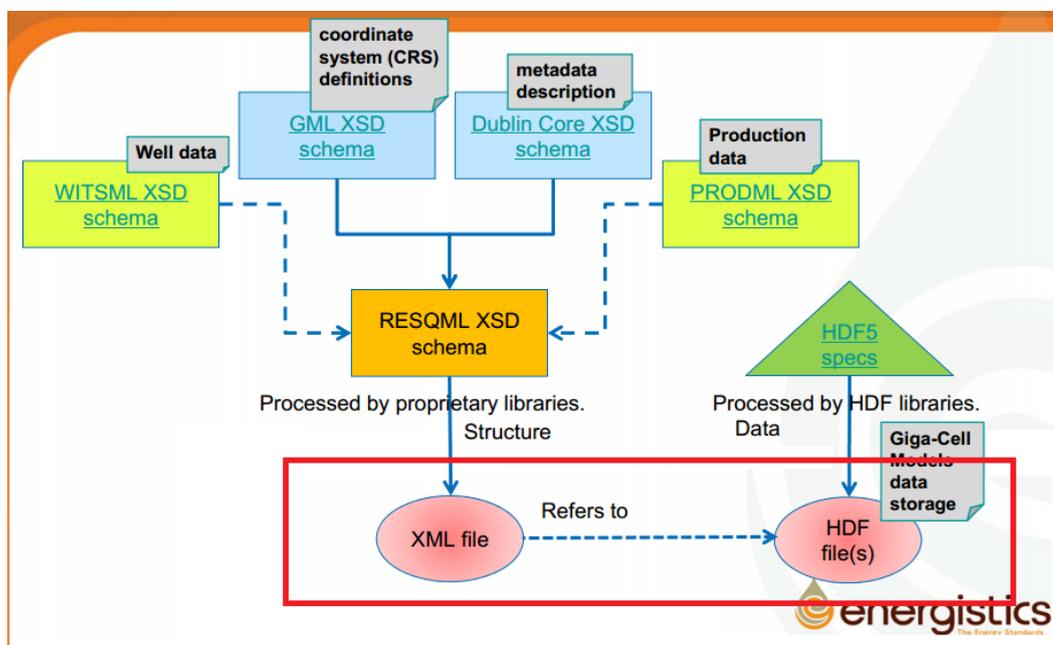


Figura 7.3: Resumo das relações entre os esquemas. Adaptado de (RAINAUD, 2011)

8 EVOLUÇÃO DOS PADRÕES DE TROCA DE DADOS

Para tratar algumas das limitações dos padrões, como a integração semântica, a Energistics disponibiliza planos de lançamentos futuros e melhorias para todos os padrões. O RESQML é o mais detalhado, enquanto PRODML e WITSML ainda não têm detalhes, embora os SIGs correspondentes estejam trabalhando nisto. Há uma iniciativa em planejamento para meados de 2014/2015 para ajudar a garantir que os padrões Energistics são interoperáveis e baseados nas tecnologias arquiteturais mais utilizadas. Quanto ao LAS, não há divulgação, por parte de seus criadores, de versões futuras.

8.1 RESQML

Quanto ao RESQML, há previsão de grandes lançamentos para 2012 e 2014, além de lançamentos incrementais para 2013 e 2015. Há ainda planos futuros, sem datas definidas. A Tabela 8.1 mostra quais características devem ser adicionadas ao padrão em cada lançamento.

Tabela 8.1: Planos futuros para extensão futura do RESQML

	2012	2014	Sem data
Infraestrutura e fluxos de trabalho	Modelos maiores; Documentos modulares; Integração com EnergyML.	Arquivamento de projetos; Modelagem de atividades.	Web Services
Estrutural	Escala estrutural e estratigráfica de reservatórios; Tabela de referência geocronológica.	Escalas sísmicas e de bacias; Organização de frameworks estruturais e estratigráficos; Metodologia de preenchimento	Tipos litológicos
Grades 3D	Grades radiais e desestruturados; Relacionamentos entre estruturas; Melhora nas propriedades de rochas e fluidos.	Otimização na criação de grades; Deque de simulação de dados;	Fraturas discretas
Poços	Trajcetórias, marcas e logs WITSML	Objetos PRODML	Poços complexos

Diferentemente dos padrões WITSML e PRODML que já possuem Web Services como características atuais, o RESQML ainda não suporta o uso desta nova tecnologia.

8.2 PRODML

O PRODML possui pouca documentação em relação a incrementos futuros, informando apenas que versões futuras devem incluir os seguintes itens:

- Perdas e adiamentos de produção;
- Testes transientes e testes de formação;
- Validação de testes de produção de poços;
- Suporte adicional para informações de ativos.

8.3 WITSML

O SIG do WITSML planeja uma nova API para o fim de 2014, e será baseada em mais tecnologias correntes além de diminuir a latência de dados de tempo real. Outros requisitos arquiteturais também devem ser atendidos, para prover um padrão ainda mais interessante às necessidades da indústria. Paralelamente, o SIG está trabalhando em uma API suplementar para permitir que a versão atual trabalhe com troca de dados em tempo real. Melhorias nos objetos existentes e novos objetos de dados também fazem parte dos planos futuros do SIG WITSML.

9 CONCLUSÃO

Neste trabalho, foram apresentados os principais padrões no que tange a interoperabilidade e integração de dados no fluxo de trabalho de E&P de petróleo. Foi mostrado que os principais padrões, RESQML, WITSML e PRODML, são construídos pelo suporte de outros padrões genéricos, independentes de domínio e plataforma, como XML, XML Schema, XSL e Web Services. Já o padrão LAS faz uso de código aberto mas em formato ASCII, que é de difícil compreensão para humanos.

Este estudo foi motivado principalmente por ser um assunto relativamente recente, com tecnologias igualmente recentes, e em uma área que é cada vez mais importante para a economia mundial, dada a importância do petróleo como fonte energética, motivadora de guerras e sustentadora da economia de diversos países. Além disto, o próprio Brasil passa a ocupar papel mais importante nesta área a partir da descoberta do pré-sal na costa do país. Portanto, a atualização acadêmica em relação a assuntos relativos ao tema de E&P do petróleo passa a ser primordial para a formação de mão-de-obra qualificada nacional.

Com relação aos objetivos, de apresentar o assunto de E&P do petróleo e, mais importante, estudar e analisar os principais padrões existentes no mercado, conhecendo melhor suas qualidades e limitações, pode-se dizer que este trabalho realizou uma análise que auxilia na comunicação entre áreas da informática, geologia e geofísica.

Em relação à literatura, pode-se dizer que ainda há muito pouco disponível sobre os principais padrões atuais. Por estar relacionado a uma área de aplicação industrial, pouco é publicado abertamente, principalmente quando algumas particularidades de pacotes de software oferecem vantagens competitivas de mercado.

Ainda assim, há diversos trabalhos que tangem o assunto de padrões de troca de dados na área de E&P, embora poucos resultados quantitativos e qualitativos comprovem a eficácia do emprego dos padrões em ambientes reais. As tecnologias de base, como XML e Web Services também são recentes, mas já possuem diversas aplicações práticas e são, de fato, abordagens que cumprem com seus objetivos de permitir interoperabilidade e integração de dados de uma maneira mais natural. Pelo fato de serem baseados nestas linguagens, os padrões RESQML, WITSML e PRODML têm vasto potencial de sucesso e os SIGs particulares de cada um destes padrões tem trabalhado com esforço para garantir melhorias e incrementos que tornem os padrões a alternativa mais natural contra os pacotes proprietários que não se comunicam.

O resultado esperado deste processo de interoperabilização e integração é que os softwares proprietários continuem existindo, mas, ao mesmo tempo, permitam importação e exportação de dados como uma característica dos pacotes de software. Consequentemente, a competitividade entre os diversos softwares que atuam em toda E&P de petróleo será dada pela qualidade da aplicação em si, em cada área diferente, e não mais como uma escolha de apenas um tipo de pacote, que aborda todas as etapas de traba-

lho, pois utiliza formatos que só são compreendidos internamente. A criação de padrões completos e amplamente adotados para a troca de dados constitui-se em um pré-requisito essencial para a integração das complexas tarefas de exploração e produção de petróleo e vai expandir a fronteira dos reservatórios produtivos que existe atualmente.

REFERÊNCIAS

- ALBRECHT, J. Geospatial information standards. A comparative study of approaches in the standardisation of geospatial information. **Computers & Geosciences**, [S.l.], v.25, p.9 – 24, 1999.
- BISHR, Y. Overcoming the semantic and other barriers to GIS interoperability. **(1998): Overcoming the semantic and other barriers to GIS interoperability, International Journal of Geographical Information Science**, [S.l.], v.12:4, p.299–314, 1998.
- DOVETON, J. H. (Ed.). **Log analysis of subsurface geology: concepts and computer methods**. [S.l.]: John Wiley and Sons, New York, NY, 1986. 288p.
- DOWNEY, D. W. What Do Geologists Need to Know about Metadata? **AAPG Annual Convention**, [S.l.], v.2, p.1–4, 2007.
- ENERGISTICS. **<http://www.energistics.org/history>**. 1996.
- ENERGISTICS; PRODML SIG the. **Business Overview of PRODML**. [S.l.: s.n.], 2010.
- ENERGISTICS; RESQML SIG the. **RESQML Overview Guide**. [S.l.]: Energistics, 2011.
- ENERGISTICS; RESQML SIG the. **RESQML Usage Guide**. [S.l.: s.n.], 2011.
- ENERGISTICS; WITSML SIG the. **WITSML STORE Application Program Interface (API)**. 1.4.1.ed. [S.l.: s.n.], 2011.
- ENERGISTICS; WITSML SIG the. **WITSML Lithology Element Usage Guide**. 1.01.ed. [S.l.: s.n.], 2012.
- FEIJO, F. J. **Noções Básicas de Geologia do Petróleo**. [S.l.]: Universidade Petrobrás, 2010.
- GEOMIND. **Specifications of standards for digital geophysical content**. 2007.
- GEOSCIENCES, J. S. of. **<http://www.jsg.utexas.edu/news/2007/11/tsunami-factory/>**. 2012.
- HESLOP, K.; PRENSKY, S.; KARST, J.; SCHMITT, D. Log Ascii Standard (LAS) Version 3.0. **Society of Petroleum Engineers**, [S.l.], 1999.
- HESLOP, K.; PRENSKY, S.; KARST, J.; SCHMITT, D. **Log ASCII Standard 3.0 File Structures**. [S.l.: s.n.], 2000.

HYNE, N. J. (Ed.). **Nontechnical Guide to Petroleum Geology, Exploration, Drilling, and Production**. 2.ed. [S.l.]: PeenWell, 2001. 598p.

INSTITUTE, P. R. <http://www.priweb.org/ed/pgws/systems/traps/structural/structural.html>. 2012.

MAGOON, L. B.; DOW, W. G. (Ed.). **The Petroleum System**. [S.l.]: American Association Of Petroleum Geologist, 1994. 644p.

MASTELLA, L. S. **Semantic Exploitation of Engineering Models**: application to petroleum reservoir models. 2010. Tese (Doutorado em Ciência da Computação) — Lecole Nationale Superieure Des Mines De Paris.

MICHEL, B.; VERNEY, P. RESQML v1.0 Case Study. **Digital Energy**, [S.l.], p.4, 2011.

OLSEN, G.; CUTKOSKY, M.; TENENBAUM, J.; GRUBER, T. Collaborative engineering based on knowledge sharing agreements. **SAGE Journals**, [S.l.], v.3, n.2, p.145–159, Jan. 1995.

PAPAZOGLU, M. P. (Ed.). **Web Services**: principles and technology. [S.l.]: Pearson Prentice Hall, 2008.

PENDLETON, V. Seismic Exploration. **Encyclopedia of Earth**, [S.l.], 2008.

PERRIN, M.; RAINAUD, J.-F.; MASTELLA, L. S.; ABEL, M. **Knowledge related challenges for efficient data fusion**. Vancouver: AAPG, 2007.

RAINAUD, J. F. **RESQML Update**: reservoir and earth modeling data exchange standard status. 2011.

RAY, E. T. (Ed.). **Learning XML**. [S.l.]: O'Reilly, 2001.

REDDY, B. R. **Use of Semantic models in Integrated Operations for Oil & Gas and New Energy**. 2010. Dissertação (Mestrado em Ciência da Computação) — Telemark University College.

SANTOS, A. M. R. F. dos. **Diretrizes para o resgate do esquema conceitual e seu compromisso ontológico a partir de um banco de dados**: um estudo de caso no domínio da litoestratigrafia. 2012. Dissertação (Mestrado em Ciência da Computação) — Instituto Militar de Engenharia.

TEIXEIRA, W.; TAIOLI, F.; FAIRCHILD, T.; TOLEDO, C. **Decifrando a terra**. [S.l.]: IBEP NACIONAL, 2008. 558p.

MILLS, C. (Ed.). **Beginning XSLT 2.0 From Novice to Professional**. [S.l.]: Apress, 2005. v.5.

TUDORACHE, T. **Ontologies in Engineering: modeling, consistency and use cases**. **VDM Verlag**, [S.l.], 2009.

XUE, F.; XUE, L. Extensible Markup Language for Petroleum Geologic Information - PXML. **IEEE**, [S.l.], October 2010.

10 ANEXO A - DESCRIÇÃO DE DOMÍNIOS SUPORTADOS POR RESQML

Tabela 10.1: Descrição de domínios suportados por RESQML. Adaptado de (ENERGIS-TICS; RESQML SIG, 2011b)

Domínio	Descrição
Interpretação Sísmica	Permite que um geofísico interprete um horizonte sísmico no domínio do tempo em uma interpretação sísmica em um dado pacote de software e, após, realize alguma outra operação, como extrair atributos do volume sísmico no decorrer da superfície em um software diferente;
Interpretação Geológica	Supondo que dois geólogos trabalhem com dados de diferentes poços em diferentes cidades e ambos gerem modelos e mapas estruturais baseados nestes dados. O RESQML fornece mecanismos para que qualquer um deles possa importar dados do outro ao seu modelo, de forma a aumentar o entendimento regional do problema;
Interpretação e Framework Estrutural	Interpretações de horizontes e falhas podem ser importadas e exportadas de forma que os mesmos dados possam ser facilmente usados por geofísicos na geração de grades e por geólogos para geração de frameworks 3D selados;
Modelagem de Reservatórios	Um modelo de reservatório construído baseado em interpretações sísmicas e geológicas pode ser compartilhado, incluindo entradas de horizontes, falhas, dados de poços, grades 3D resultantes e propriedades específicas;
Modelagem de Bacias	A geração de um modelo de bacia é complexo e pode envolver uma variedade de dados de diferentes fontes de software, cobrindo diversas áreas e diversos tipos de dados. Com RESQML, os dados podem ser exportados por estas diferentes fontes e importados pelo pacote de software que gerará o modelo;
Planejamento de Poços	Superfícies, incluindo falhas e horizontes, podem ser importadas, juntamente com metadados de rastreabilidade que incluem informações de quem editou os dados pela última vez e quando. RESQML garante a integridade geoespacial, necessária para planejamento de poços;
Simulação Dinâmica	Informações estáticas de reservatório, como porosidade, permeabilidade e saturação da água, podem ser exportadas juntamente com grades geométricas 3D para que engenheiros de reservatórios possam fazer simulações dinâmicas;
Grades 3D grandes	O RESQML consegue manipular grandes quantidades de dados para a construção de grades de forma rápida e eficiente com o uso de HDF5 (Hierarchy Data Format 5);