

Enterprise Intelligence suportada em Metadados como Ontologia

Francisco Guimarães¹, Carlos Caldeira², Paulo Quaresma³

¹Departamento Informática, Universidade de Évora, Francisco.santana.guimaraes@gmail.com

²Departamento Informática, Universidade de Évora, ccaldeira@di.uevora

³Departamento Informática, Universidade de Évora, pq@di.uevora

Resumo. O contexto da actividade económica actual é caracterizado por um grande volume de dados, estruturados e não-estruturados, com origem em processos internos das organizações e sua extensão para processos externos de acordo com o mercado onde operam. Neste contexto, torna-se fundamental conhecer a localização, estrutura e semântica dos dados, para melhor gerir e tomar decisões. Este conhecimento sobre os próprios dados é assumido sob o conceito de Metadados. Como as organizações são compostas por vários níveis de arquitectura interrelacionados, que se podem resumir em arquitectura de negócio e arquitectura técnica, também os metadados necessitam de manter as 2 perspectivas e devido relacionamento. Na perspectiva técnica, os metadados são amplamente utilizados em soluções de Business Intelligence enquanto suporte ao seu próprio funcionamento. No caso da perspectiva de negócio, onde é necessário manter descritores de relação entre indicadores, objectivos e dados que representam factores que afectam o negócio, o conceito de Metadados de negócio não tem sido tão amplamente utilizado. Por outro lado, a multiplicidade de sistemas operacionais e soluções de Business Intelligence na mesma organização, torna complexo o controlo e centralização destes metadados. Neste domínio de problema, existem várias abordagens de solução para metadados assentes em princípios de interoperabilidade, reutilização e vistos como ontologias com semântica, ou simplesmente enquanto conceitos de Master Data Management e Data Governance com enfoque em qualidade de dados e Data Discovery. Este artigo, explora o conceito de implementação de metadados enquanto ontologia com a semântica de negócio de acordo com metamodelos por sector de actividade, e que possa ser reutilizado por sistemas Business Intelligence enquanto metadados técnicos e de negócio como suporte à inteligência organizacional.

Palavras-Chave: Business Intelligence, DataWareHouse, Metadados, Web Semântica, Arquitecturas Empresariais, Ontologias, Agentes Inteligentes

1 Introdução

As organizações modernas coexistem num mundo dinâmico de informação com integração entre processos internos e externos, criando uma rede virtual de empresas suportada em informação. Por outro lado, estas organizações têm uma realidade ao nível de sistemas de informação que resultam de uma amálgama histórica de implementação de vários tipos de aplicações e bases de dados, assentes em várias tecnologias e com investimentos por vezes elevados em termos de manutenção e evolução.

A proliferação de vários tipos de informação, com estruturas diferentes, em sistemas heterogéneos, com padrões de utilização assentes em diversos processos e com vários tipos de utilizadores, e de necessidades, implica antes de mais conhecer os próprios dados.

A informação sobre as próprias estruturas de dados necessárias para cada organização, isto é Metadados, implica uma necessidade de gerir os dados sobre os dados e utilizar esta informação como base de conhecimento assente em modelos de inferência, manuais ou automáticos, para permitir automatismos e mesmo aprendizagem para acelerar a tomada de decisão.

Daí a importância dos Metadados, não só em termos de dados sobre as aplicações e suas bases de dados integrados numa Arquitectura Técnica, mas igualmente sobre a semântica ou lógica de negócio que advém do que é designado por Arquitectura de Negócio. Este conhecimento deve ser centralizado e dinâmico, quer ao nível de dados técnicos sobre aplicações, bases de dados e processos, mas igualmente sobre dados que representam os conceitos de negócio. Esta identificação, estrutura e relação de dados críticos de negócio, suportados em aplicações para sua gestão, corresponde ao que é designado por Arquitectura Informacional. Na realidade, são metadados de negócio e metadados técnicos.

Numa perspectiva de metadados técnicos, estes sempre existiram nas próprias aplicações, e em particular em soluções de Business Intelligence para suportar o seu funcionamento. Neste domínio existe uma tendência de reutilização deste tipo de informação para melhorar os controlos de segurança, ao nível de Data Discovery e Data Classification para efeito de gestão de risco, integrado no conceito de Data Governance, sendo de destacar tendências de soluções de Master Data Management para controlar redundância de informação entre aplicações [SAS MDM], [ORACLE MDM].

Numa perspectiva de metadados de negócio, os sistemas de Business Intelligence são os que mais consideram a sua necessidade enquanto descritores de lógica de cálculo de indicadores e detalhe de informação associada a relatórios. No entanto, a sua maior utilização, está associada a sistemas de modelação da realidade da organização nas várias perspectivas estratégicas, processual, orgânica, sistemas de informação, e mesmo de contexto onde a organização se insere. A este nível, cria-se normalmente uma linguagem própria que pode ser reutilizada por sectores de actividade. Este tipo de sistemas, denominados de Arquitectura Empresarial, têm sido objecto de investigação para incorporação de ontologia e semântica para se implementar modelos de inteligência empresarial [Kanga], [Rajabi].

Numa perspectiva mais estrutural sobre os próprios metamodelos sob os quais se estrutura os metadados existem tendências de investigação para reutilização ou interoperabilidade, como é o caso das iniciativas Dublin Core Metadata [DCM] ou Common Warehouse Model [CWM], mas que são intrusivas para os sistemas existentes face às alterações que teriam que proceder para reutilizarem estes modelos, ou interoperarem com bases nos mesmos.

Neste contexto, este artigo, apresenta o estado da arte no contexto de Metadados, Arquitecturas Empresariais, Ontologias e Semântica, para enquadrar uma possível solução alternativa, não intrusiva, que permita capturar os metadados técnicos e funcionais, relaciona-los e criar uma ontologia que servirá de base para criar um modelo de inteligência empresarial.

2 Metadados

Metadados corresponde ao conceito de dados sobre dados, visto como descrição de informação sobre as próprias de estruturas de dados numa organização, residentes em activos tecnológicos ou não, como o caso de informação não estruturada em base documental. Pela definição de alguns autores de referência, podemos definir Metadados e a sua gestão da seguinte forma:

- [Marco] define Metadados como qualquer dado físico guardado em aplicações informáticas, e conhecimento existente em aplicações ou em colaboradores, que representam aspectos internos ou externos à organização, incluindo informação sobre processos de negócio, regras e estruturas de dados;
- [Tozer] define o conceito de Metadados como algo relacionado com o comportamento e descrição de outros dados;
- [Chisholm] define Metadados como algo que se traduz nos dados que descrevem todos os aspectos activos de informação de uma empresa, permitindo-lhe utilizar e gerir de forma efectiva esses mesmos activos;
- [Blechar] define o conceito de Metadados como algo que consiste em informação sobre as características de um qualquer artefacto organizacional, tal como o seu nome, localização, importância atribuída, qualidade ou valor para a organização, assim como os seus relacionamentos com outros artefactos da organização.

Apesar de várias definições, existe um conceito comum associado a identificação de objectos, atributos desses objectos e relações entre os mesmos, sendo que os objectos representam dados.

Por outro lado, os metadados são vistos de forma agregada como Metadados de Negócio (definições de negócio, hierarquias de dados, regras de agregação, métricas), Metadados de Processos (origem, destino, frequência, sequência de processos de carregamento), Metadados Técnicos (locais físicos, formatos, estruturas) e Metadados Aplicacionais (Como se acede e utiliza, quer por utilizadores quer por aplicações). Face às características dos vários tipos de metadados, [Marco] agrega os mesmos em 2 grandes grupos de acordo com a figura 1:

- **Metadados Técnicos:** Relatórios, frequências de execução, tempo de execução, mapeamentos operacional/analítico, conversões de dados, modelos lógicos/físicos, identificação campos/tabelas/índices/programas e gestão de versões;
- **Metadados de Negócio:** Estruturas de dados numa perspectiva de utilizador, definições de negócio mapeadas em campos, atributos de negócio mapeados em campos, estatísticas de qualidade de dados, regras de navegação pelos dados, localização/agregação temática dos dados;

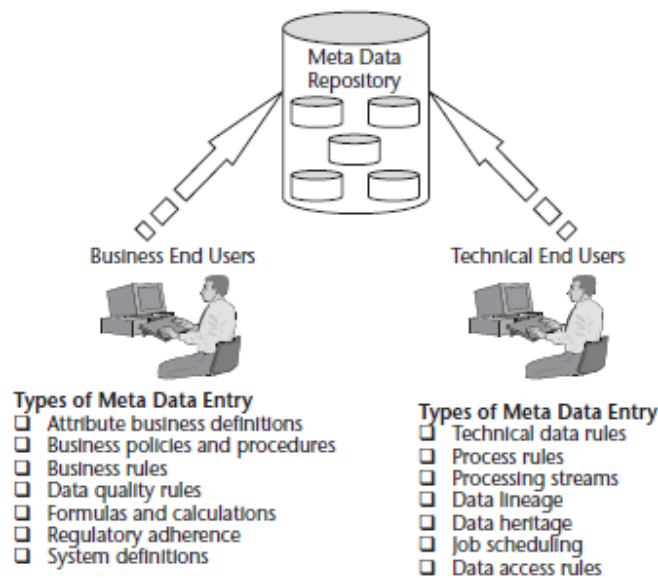


Figure 2.3 Meta data sourcing layer: end-user meta data entry.

Figura 1. Tipos de metadados e utilizadores [Marco]

Para a implementação de conceitos de Metadados, existem vários Standards actuais na indústria, com ênfase nos seguintes:

- MOF (Meta-Object Facility): Definido pela OMG (Object Management Group), este standard [OMG] tinha como objectivo inicial estruturar um metamodelo para o UML (Unified Modelling Language), sendo que foi depois tornado como standard pelo ISO/IEC para a definição, gestão e integração de metadados genéricos;
- CWM Common WareHouse Metamodel: O CWM [CWM] foi definido pela OMG enquanto standard para interoperabilidade tendo como base um metamodelo partilhado, utilizando algumas especificações do [OIM] que foi descontinuado, mas focando em sistemas de Business Intelligence;

- BIDM Basic Interoperability Data Model: O BIDM [R.L.I Group] foi definido pelo Reuse Library Interorability Group e pelo IEEE Software Engineering Standards Commitee tendo gerado o standard IEEE 1420.1 para determinar o conjunto mínimo de activos de informação em qualquer tipo de “reusable libraries”;
- DCMI Dublic Core Metadados Initiative: O DCMI [DCM] é definido por um organismo internacional com o mesmo nome, focando na especificação de elementos de metadados para descrição de recursos para permitir a criação de sistemas inteligentes na descoberta/exploração desses dados, tendo originado o standard ISO 15836-2003 [DCMI];
- ESM Enterprise System Metamodel: O ESM foi definido por empresas de consultoria [Marco] para detalhar todas as estruturas de dados de uma organização e estabelecer a sua relação com metadados;
- ITPMM IT Portfolio Management Metamodel: O ITPMM foi definido por empresas de consultoria [Marco] foi definido tendo por base ideia de que os activos de sistemas de informação como projectos, software, hardware, pessoas e serviços, são fundamentais numa organização e têm estruturas próprias descritivas e de relação entre elas, que podem ser generalizadas.

Considerando os tipos de metadados e standards para modelos de repositório de Metadados, a sua implementação e gestão depende da existência de extractores com baseem várias fontes de dados, para consolidar o repositório e permitir a sua utilização porvários tipos de processos, aplicações e utilizadores. Este modelo referido por [Marco] na figura 2 permite identificar o ecossistema e importância dos metadados vistos como repositório de conhecimento organizacional, em vez da visão de simples repositórios de metadados técnicos aplicacionais.

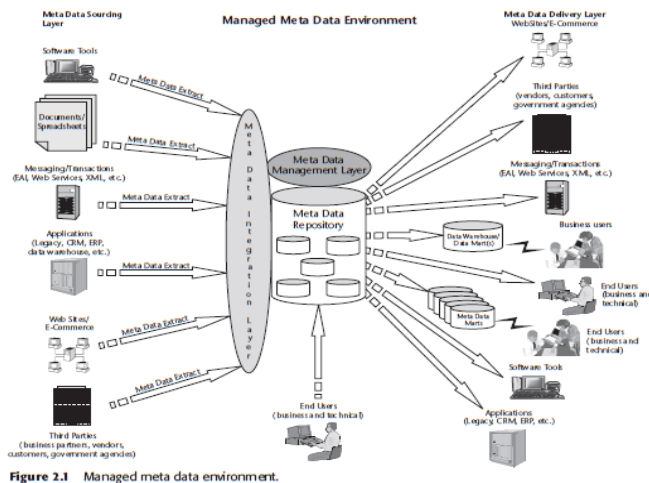


Figura 2. Ecossistema dos Metadados [Marco]

3 Arquitectura Empresarial

No caso de metadados de negócio, numa perspectiva de modelação dos conceitos chave da organização e suas relações, existem implementações ao nível de Metadados integrado em soluções de Business Intelligence. No entanto, para a captura mais abrangente do contexto estratégico, funcional, orgânico e sua relação com aplicações e outros artefactos de sistemas de informação, as soluções de Arquitecturas Empresariais representam a abordagem mais emergente.

De acordo com a norma [IEEE Std1471] as Arquitecturas Empresariais podem ser definidas da seguinte forma:

“The fundamental organization of a system, embodied in its components, their relationships to each other and the environment, and the principles governing its design and evolution.”. Esta definição inclui o meio envolvente, uma vez que, alterações a este nível têm repercussões directas na arquitectura.

Por outro lado, [Towers] refere que a arquitectura empresarial é utilizada com os seguintes objectivos:

- Suporte à implementação da estratégia;
- Clarificação de responsabilidade, tendo por base o conhecimento do posicionamento de cada função, das suas tarefas e da relação entre todos os componentes da organização;
- Definir uma *Framework* que permita identificar onde estamos, e para onde queremos ir, detalhando o caminho e monitorizando o processo de transformação.

Em termos de decomposição por níveis de arquitectura, como referido por [Ritgen] e suportado por uma metodologia e estrutura da arquitectura empresarial em [TOGAF], uma arquitectura empresarial pode ser decomposta pelo seguinte:

- **Arquitectura Organizacional:** Estruturas orgânicas e funcionais em termos de teoria organizacional. Pode-se considerar igualmente a definição da visão, missão, princípios e políticas gerais da organização. No entanto, no entender desta tese, deve-se considerar ainda a definição da oferta de produtos e serviços, e definição clara de canais de distribuição e segmentação de clientes, sendo no entender desta tese, que neste conceito pode ser designada como arquitectura estratégica, ou separando a arquitectura orgânica da arquitectura estratégica;
- **Arquitectura de Negócio:** Modelo de funcionamento detalhado em processos, actividades e tarefas para atingir os objectivos estratégicos e operacionais da organização;
- **Arquitectura Informacional:** Estruturas de informação que são utilizados pela organização, com preocupações de captura da sintaxe e semântica, independente das aplicações e processos que as utilizam e da forma como estão implementados;

- **Arquitetura Aplicacional:** Aplicações informáticas, em termos de serviços de suporte ou que permitem integrar e inovar no modelo de prestação de serviços ou integrados em produtos comercializados;
- **Arquitetura Tecnológica:** Infra-estrutura tecnológica considerando o suporte às aplicações informáticas, mas igualmente as redes, comunicações e serviços base tecnológicos para garantir a fluidez da comunicação em suporte digital.

Existem várias *Frameworks* com vários modelos de representação das dimensões e detalhes das mesmas, como é o caso de várias instituições governamentais Americanas (DOFAF, MODAF, NAF, CEAF, TEAF, FEAF) e a linguagem IDEF0, mas em termos de utilização empresarial, destacam-se as seguintes no contexto desta tese:

- **Zachman:** [Zachman] Utiliza 2 dimensões para definir os Dados (What), as Funções (How), a Localização geográfica (Where), as pessoas (Who), tempo (When) e a motivação (Why) associados aos artefactos, cruzado com nível de detalhe em termos de âmbito (Scope), modelo de negócio (Business Model), modelo de sistemas (System Model), modelo de tecnologias (Technology Model) e representação detalhada (Detailed Representations);
- **MEAF:** Metis Enterprise Architecture Framework, criado pela empresa Troux Technologies a partir de modelos de várias empresas, e que está na base da sua ferramenta de arquitecturas empresariais, IT Portfolio Management, IT Governance Management e Metadados Management [Troux];
- **TOGAF ArchiMate:** A The Open Group Enterprise Framework, apresenta um modelo de arquitecturas de negócio, aplicações e tecnológicas, uma metodologia para a sua implementação e uma linguagem de notação que pode ser reutilizada por várias ferramentas, e que é designada por Archimat [TOGAF].

Apesar da sua relevância, este conceito tem sido descurado nas organizações derivado do esforço de actualização dos modelos. Existem no entanto algumas abordagens de implementação vendo esta arquitectura como ontologia e implementada com conceitos de semântica, nomeadamente ao nível de TOGAF Archimate enquanto linguagem de modelação [Azevedo]

4 Semântica e Ontologia

Se considerarmos alguns sectores de actividade, existem termos que têm um significado estruturante enquanto modelo de negócio e objectos envolvidos. É o caso a título de exemplo do sector financeiro, onde conceitos de cliente, conta, produto bancário, rentabilidade, produto, segmento de cliente, entre outros, têm uma sintaxe e semântica própria, incorporada nos sistemas e no modelo de pensamento do sector, ajustado em significado a cada entidade bancária.

Estes conceitos, fazem parte de modelos de segmentação, modelos de negócio, indicadores de gestão, entre outros modelos de pensamento estratégico e operacional, sendo implementado por modelações de dados em bases de dados, e de processos em vários tipos de sistemas.

A sua consolidação e visão comum pelas organizações enquanto visão de funcionamento e de negócio, são a base dos metadados de negócio, sendo que desalinhamentos de definição têm impacto em problemas de comunicação nas organizações ao nível de processos, pessoas e tecnologia, considerando possíveis automatismos que se queira implementar.

Esta Metadados de negócio, no entanto, tem como raiz filosófica o conceito de ontologia. Isto porque a ontologia corresponde a uma descrição formal de uma área de conhecimento, ou domínio de conhecimento, a partir da qual se montam processos de comunicação com base no significado. Este formalismo permite a criação de automatismos nos processos.

O conceito de ontologia tem origem em palavras gregas pela composição de “On-to” (Ser) e “Logos” (razão), enquanto razão do ser. É como tal um conceito que se encontra no domínio da filosofia, mas igualmente no domínio da ciência da computação.

Em ciência da computação as ontologias são definidas como termos utilizados para descrever e representar áreas de conhecimento utilizando-se normalmente esquemas de acordo com os domínios específicos de representação. Para [Gruber] a Ontologia é a especificação formal ou definição conceptual de ideias, conceitos, relações e outras abstracções no contexto de um domínio ou discurso. Como tal, é um vocabulário para utilização na linguagem nesse domínio, permitindo a comunicação e reutilização. A questão da formalização na definição da Ontologia, torna-se fundamental no domínio da computação, pois permite legibilidade pelas máquinas.

O conceito de ontologias tem sido utilizado [Heflin] para descrever vários tipos de artefactos, incluindo taxonomias, esquemas de metadados do DCMI e mesmo teorias, o que comprova a sua utilizada em diversas finalidades, mas que se reflectem em metodologias, ferramentas e especificações.

Como nomenclatura, pode-se resumir os componentes básicos de acordo com [Corcho]:

- **Classes:** Correspondem a conceitos e utilizado em taxonomias, representando tarefas, acções, estratégias, processos de raciocínio, entre outros conceitos;
- **Relações:** Representam tipos de interacção entre as classes de um domínio, determinando uma classificação dessas relações enquanto “subclasse de”, “conectada a”;
- **Funções:** Tipo especial de relação onde o último elemento da relação, pode ser visto como um elemento determinístico face aos elementos precedentes;
- **Axioma:** Definem o significado e restrições, que permitem modelar expressões sempre verdadeiras;

- **Instâncias:** Representam elementos específicos, ou seja, os próprios dados.

Para tipificar as ontologias, existem várias formas, sendo que ressaltar a definição de [Gruber] que publicaram uma Framework específico no domínio da Ciência da Computação e Informação, agrupada por dimensões:

- **Dimensão Semântica:** Relacionada com a especificação do vocabulário. Pode ser decomposta pelos seguintes níveis:
 - Nível de estrutura: Também visto como nível de formalidade ou nível de dados estruturados, sendo como tal uma ontologia “altamente estruturada” (bem especificada tal como no caso de abstrações matemáticas), “pouco estruturada” (generalista nos conceitos como é o caso de documentos e hiperligações) e “semi-estruturada” (misto de definições formais e informais);
 - Expressividade da Linguagem: Relacionada com o nível de estrutura, a expressividade da linguagem em ontologias altamente estruturadas exige expressividade e formalismo, enquanto que baixo nível de estrutura exige no limite uma linguagem que permita listagens de condições e definições com baixa expressividade formal;
 - Granularidade Representacional: Nível de detalhe dos conceitos.
- **Dimensão Pragmática:** Relacionada com a finalidade e contexto da ontologia. Pode ser decomposta pelos seguintes níveis:
 - Intenção de uso: Qual o grupo de utilização da ontologia, considerando representação de vocabulários em língua natural, partilha de bases de conhecimento, modelo de comunicação entre agentes de *software*, entre outros;
 - Papel da Automação Lógica: Possibilidade de utilização de automação lógica para inferência a partir dos conceitos e relações;
 - Descritiva versus Prescritiva: Ontologias com notações mais descritivas e livres, ou mais prescritivas em termos de rigidez na caracterização e declaração de objectos/classes;
 - Metodologia: as ontologias podem ser construídas com metodologias *bottom-up* a partir dos conceitos base e depois agregando e relacionado ao nível de sub-classes e classes do domínio, ou *top-down* onde se inicia pela caracterização do domínio e depois se efectua a decomposição até às classes mais básicas.

Recentemente, no contexto da World Wide Web, a W3C definiu um novo conceito, designado por Web Semântica [Berners-Lee] enquanto extensão da World Wide Web tendo por base modelos de metadados para suporte de ontologias que as aplicações podem utilizar para interacção no domínio do que é designado por Web 3.0. Para o efeito, padronizou 2 linguagens e frameworks fundamentais:

- **RDF Resource Description Framework:** O RDF foi desenhado como standard pela W3C [RDF] para descrição de recursos ou objectos através de uma categorização de recursos ou objectos que têm várias propriedades e estas têm valores, como numa lógica de predicados utilizada em linguagem natural;
- **OWL Web Ontology Language:** O OWL [OWL] foi criado pela W3C como uma extensão do RDF, acrescentado mais termos de vocabulário para descrever as propriedades, como é o caso de cardinalidade, controlo, tipos de dados mais complexos, classes de dados enumeradas, entre outros conceitos.

Para a implementação das Ontologias, existem assim várias linguagens, com enfoque para as linguagens criadas para explorar as características da World Wide Web e normalmente baseadas em XML, como é o caso das seguintes, além da OWL e RDF:

- **Topic Maps:** É um formalismo de representação de recursos organizados em tópicos, Associações enquanto relações entre tópicos, e Ocorrências enquanto relacionamento de tópicos com outros recursos relevantes [Librelotto];
- **OIL Ontology Interchange Language):** Combina formalismo de Frames com semântica formal e inferência lógica descritiva para classificação e taxonomia de conceitos, tendo evoluído para **DAML+OIL Agent Markup Language**;
- **SKOS Simple Knowledge Organization System:** Aplicação do RDF para representação de Thesaurus e tipos de sistemas de organização de conhecimento.

Pela importância da Web Semântica, enquanto modelo de representação de conhecimento para utilização automática por programas, destaca-se o OWL e RDF, enquanto especificações chave deste conceito, integrado no conceito de ontologias mas com enfoque para a World Wide Web, para permitir que as máquinas naveguem sobre os dados. Este conceito tanto pode ser aplicada para navegação externa, como pela utilização em aplicações enquanto modelo de navegação inteligente pelos dados.

As abordagens de modelação de conhecimento baseadas em ontologias têm uma aplicação imediata a conceitos de arquitectura empresarial como forma de criar um modelo de metadados semântico em determinado domínio de conhecimento. Se considerarmos o domínio de conhecimento como sector de actividade como Banca, Seguros, Corretagem, Infra-estruturas, entre outros, a sua aplicabilidade permitirá consolidar uma linguagem comum dentro da organização e entre organizações do mesmo domínio de conhecimento.

Desta forma ficará facilitada comunicação interna e entre entidades, tendo por base o conhecimento, permitindo a criação de automatismos na comunicação suportada em sistemas de informação, e caminhando para modelos de inteligência empresarial.

5 Proposta de investigação

Considerando a necessidade das organizações terem uma visão integrada de metadados técnicos mas igualmente de negócio, num único repositório e com uma visão

semântica de negócio como parte do modelo com reutilização por toda a organização, quer ao nível de utilizadores, quer ao nível de aplicações informáticas para criação de automatismos, torna-se crítico no domínio da ciência da computação encontrar um modelo de implementação destes metadados assim definidos.

Esta solução tem sido objecto de investigação, sendo que no nosso entender, se podem agrupar as abordagens em 3 modelos a seguir resumidos:

- Modelos de metadados que sejam reutilizáveis ou interoperáveis [CWM], [DCM], mas que se tornam complexos de integrar a linguagem específica de negócio de um sector de actividade e para uma entidade específica desse sector de actividade, além de ser um modelo que implica alterações em sistemas já em funcionamento com custos de adaptação elevados;
- Modelos de metadados de negócio ao nível de ferramentas de Business Intelligence para controlo de qualidade de dados, como é o caso de conceitos de MDM (Master Data Management), Enterprise Information Management e Data Governance, mas com fraca incorporação da visão de semântica de negócio integrada nos metadados, sendo mais focado para qualidade de dados com algum potencial ao nível de definição de regras de negócio, mas sem reutilização por vários tipos de agentes;
- Modelos de metadados com incorporação de semântica e suportada em ontologias para permitir não só a centralização de dados, mas dotar igualmente este repositório de inteligência organizacional, para adaptação dinâmica ao contexto interno e externo onde se insere. Nesta linha, tem-se destacado uma aproximação entre metadados e arquitecturas empresariais, nomeadamente através da linguagem padrão TOGAF ArchiMate. Algumas preocupações a este nível estão relacionadas com encontrar técnicas e linguagens mais adequadas para cada nível de arquitectura (negócio, aplicacional, informacional, tecnológica), respectiva integração e análise dos modelos assim criados [Morais], [Kanga].

No entanto, estas preocupações, não endereçam as seguintes questões:

- Como criar uma ontologia alinhada com modelos por sector de actividades, que se constituam numa linguagem específica de negócio para facilitar a gestão de conhecimento interno e interoperabilidade entre sistemas, ou agentes, comprometidos com esta linguagem?
- Como instanciar uma ontologia automaticamente com base em vários metadados técnicos e funcionais, derivado da inevitabilidade da existência de vários sistemas aplicacionais, como packages ou desenvolvimento à medida, em áreas aplicacionais ou de informação de gestão, criando uma arquitectura informacional, vista como ontologia ou metadados organizacionais?
- Como ajustar os metadados organizacionais, com base em informação de especialistas internos, para garantir que corresponde à realidade específica da cada organização?
- Como reutilizar este modelo pelas aplicações, vistas como agentes comprometidos por estes metadados ou ontologia organizacional, com particular ênfase em sistemas de business intelligence em processos de ETL (Extracção, Transformação e

Carregamento) e na navegação através de sistemas de exploração de dados (reporting, dashboard), usufruindo da semântica assim construída?

Para responder a estas questões, este artigo de acordo com a investigação a decorrer com aplicação em casos de estudo em 3 sectores de actividade (infra-estruturas, banca e corretagem financeira) propõe uma visão de solução por incorporação de semântica em ontologias suportadas em linguagens de arquitecturas empresariais com objectivo de reutilização em contexto de metadados em aplicações de Business Intelligence. A nossa proposta de investigação, de acordo com modelo na figura 3, acrescenta níveis de extensão destes conceitos, já parcialmente aplicados ao repositório de arquitecturas empresariais, para garantir que igualmente os extractores de informação e os utilizadores desta arquitectura de informação sejam agentes com inferência, e que permita uma edição para ajuste específico para cada organização, com base num metamodelo por sector de actividade:

- Para não ser intrusivo para as fontes de dados e face aos diferentes tipos de estruturas, será implementado um ETL Semântico enquanto extractor subordinado às próprias regras de negócio dos Metadados e com inteligência ao nível de captura por eventos, por padrões de produção dos dados e com potencial de acrescentar aprendizagem;
- Os consumidores de metadados estarão disponíveis via WebServices implementados com tecnologias de agentes inteligentes para serem adoptados por utilizadores no contexto de navegação por informação de gestão, transaccional e no contexto de processos de negócio.

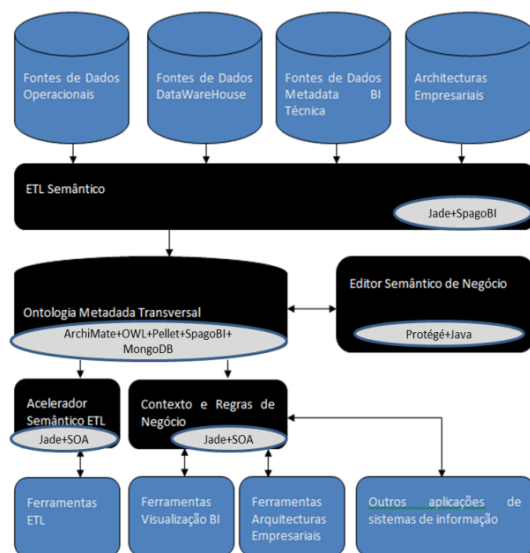


Figura 3. Modelo Proposto

Na base da solução, serão implementados metadados com base em metamodelos por sector de actividade, instanciados por reverse inicial do comportamento da própria organização via extractores semânticos, sendo ajustados com o conhecimento tácito através de um editor semântico, para retroalimentarem o consumo, mas igualmente a aprendizagem da extracção e consolidação de metadados vistos como ontologia organizacional.

Para a implementação do modelo serão utilizadas tecnologias de suporte a ontologias e RDF no repositório de Metadados, a partir do qual se pode desenvolver agentes suportados em JADE [Jade] e Pellet [Pellet] para implementar a inferência, e motores de ETL suportados em SpagoBI [Spago].

6 Conclusão

Este artigo apresenta o estado da arte de metadados, ontologias e arquitecturas empresariais enquanto domínios do problema da implementação de metadados como ontologia transversal numa organização, tendo por base o pressuposto de organizações com proliferação de vários tipos de sistemas com metadados heterogéneos.

O artigo detalha as abordagens de implementação de metadados técnicos e de negócio, com as desvantagens inerentes a alguns modelos, optando por seguir uma abordagem de investigação baseada em metadados como ontologias organizacionais derivados de modelos de arquitecturas empresariais. No entanto, apresenta uma solução de construção dinâmica da ontologia com base em metamodelos específicos por sector de actividade, a partir do qual as organizações podem ajustar o seu próprio domínio de conhecimento.

Com esta abordagem é possível construir dinamicamente os metadados como ontologia, e usufruir deste universo de discurso para retroalimentar os extractores de metadados e os consumidores, em vários processos.

Este modelo será montado em 3 casos de estudo em 3 sectores de actividade diferentes (Banca, Infra-estruturas e Corretagem), testando a construção dos metamodelos, a inferência para captura dos metadados e a inferência como indutor de conhecimento no contexto de navegação em sistemas de Business Intelligence na componente de metadados de negócio.

Referências

1. [Azevedo] Azevedo, Carlos, Almeida, João, Van Sinderen, Marten, Quartel, Dick, Guizzardi, Giancarlo. An Ontology-Based Semantics for the Motivation Extension to ArchiMate. Disponível em http://nemo.inf.ufes.br/files/an_ontology_based_semantics_for_the_motivation_extension_to_archimate.pdf. Último acesso em 12/12/2013.
2. [Blechar] Blechar, M. Metadata Management Technology Integration Strategues, 2H05 to IH06. Gartner Research.

3. [Berners-Lee] Berners-Lee, T, Hendler, J, Lassila, O. The Semantic Web. Scientific American: the Semantic Web. Disponível em www.sciam.com/print_version.cfm?articleID=00048144-10D2-1C70-84A9809EC588EF21. Último acesso em 12/12/2013.
4. [Corcho] Corcho, O, Fernandez-Lopes, M, Gomez-Perez, A. Methodologies, Tools and Languages for Building Ontologies. Where is their meeting point? Data & Knowledge Engineering.
5. [Chisholm] Chisholm, M. Repositories Build or Buy. MPO Newsletter.
6. [CWM] Common Warehouse Metamodel Specification, Volumes 1 & 2. Disponível em <http://www.omg.org/>. Ver também <http://www.cwmforum.org/>. Último acesso em 12/12/2013.
7. [DCM] DCMI. The Dublin Core Metadata Initiative. Disponível em <http://dublincore.org/>. Último acesso em 12/12/2013.
8. [Gruber] Gruber, T. Toward Principles for the Design of Ontologies Used for Knowledge Sharing, In: Guarino, N., Poli, R.(Eds.): Formal Ontology In Conceptual Analysis and Knowledge Representation, Kluwer.
9. [Guimaraes] Guimarães, F. Utilização de sistemas multi-agentes em processos de negócio e sistemas de workflow no domínio da memória organizacional. Tese Mestrado. ISCTE.
10. [Heflin] Heflin, J. OWL Web Ontology Language: Use cases and requirements. W3C Recommendations.
11. [IEEE Std1471] Systems and Software engineering - Recommended practice for architectural description of software-intensive systems: ISO/IEC 42010:2007(E).
12. [Jade] Java Agent Development Platform. Disponível em <http://jade.tilab.com>. Último acesso em 12/12/2013.
13. [Kanga] Kanga, Dongwoo, Leeb, Jeongsoo, Choib, Sungchul, Kim, Kwangsoo. An ontology-based Enterprise Architecture. Disponível em <http://www.sciencedirect.com/science/article/pii/S0957417409006368>. Último acesso em 12/12/2013.
14. [Librelotto] Librelotto, G. Topic Maps: Da Sintaxe à Semântica.
15. [Marco] Marco, D, Jennings, M. Universal Meta Data Models. Wiley Computer Publishing.
16. [Morais] Morais, André. Aplicação de Ontologias à Representação e Análise de Modelos de Arquitetura Empresarial. Dissertação para Obtenção de Grau de Mestre em Engenharia Informática e de Computadores. Instituto Superior Técnico. 2013.
17. [OMG] OMG. Object Management Group – MetaObject Facility (MOF). Disponível em www.omg.org/mof/. Último acesso em 12/12/2013.
18. [OWL] OWL Web Ontology Language. Disponível em www.w3.org/TR/owl-features/. Último acesso em 12/12/2013.
19. [ORACLE MDM] Oracle Master Data Management. Disponível em <http://www.oracle.com/us/products/applications/master-data-management/overview/index.html>. Último acesso em 12/12/2013.
20. [Pellet] Motor de Inferência. Disponível em <http://clarkparsia.com/pellet/features>. Último acesso em 18/12/2013.
21. [Rajabi] Rajabi, Zeinab, Minaei, Behrouz, Seyyedi, Mir Ali. Enterprise Architecture Development Based on Enterprise Ontology. Disponível em <http://www.scielo.cl/pdf/jtaer/v8n2/art07.pdf>. Último acesso em 12/12/2013.
22. [Ritgen] Rittgen, P. Enterprise Modeling and Computing with UML. IDEA Group Publishing.

23. [RDF] RDF Vocabulary Description Language 1.0: RDF Schema. Disponível em www.w3.org/TR/rdf-schema/. Último acesso 12/08/2013.
24. [R.L.I Group] Group R.L.I. The basic interoperability data model. Disponível em <https://kspace.cdvp.dcu.ie/respository/doc/bidm.html>. Último acesso em 12/12/2013.
25. [Spago] Open Source Business Intelligence Suite. Disponível em <http://www.spagoworld.org/xwiki/bin/view/SpagoBI>. Último acesso em 12/12/2013.
26. [Tozer] Tozer, G. Metadata Management for Information Control and Business Success. Archet House, Inc.
27. [Tolbert] Tolbert, D. CWM: A Model-based Architecture for Data Warehouse Interchange. Disponível em www.cwmforum.org/uciwesas2000.htm. Último acesso em 12/12/2013.
28. [Towers] Towers, S, Burlton, R. In Search of BPM Excellence: Straigth from the Thought Leaders.
29. [Togaf] TOGAF v9, The Open Group Architecture Framework. Disponível em www.Togaf.org. Último acesso em 12/12/2013.
30. [Turco] Turco, C. Enterprise Architecture & Metadata Modelling: A Guide to Conceptual Data Model, Metadata Repositoty, Business and System Re-Enginerring. ISBN-0-7414-5304-5.
31. [Trouw] Trouw Enterprise Architecture. Disponível em <http://www.trouw.com/>. Último acesso em 12/10/2013.
32. [SAS MDM] SAS Data Management Software. Disponível em http://www.sas.com/en_us/software/data-management.html#master-data-management. Último acesso em 12/12/2013.
33. [Zachman] Zachman, J. The Zachman Framework™ Evolution. Disponível em <http://www.zachmaninternational.com/index.php/ea-articles/100#maincol>. Último acesso em 12/12/2013.