

Arquitetura da informação para processos de negócio e modelagem de banco de dados: aproximações possíveis

Mariana Baptista Brandt ^I

^I Câmara dos Deputados, Brasília, DF, Brasil;
mariana.brandt@unesp.br; <http://orcid.org/0000-0001-8119-7527>

Silvana Aparecida Borsetti Gregório Vidotti ^{II}

^{II} Universidade Estadual Paulista, Marília, SP, Brasil;
vidotti@marilia.unesp.br; orcid.org/0000-0002-4216-0374

Resumo: Os dados e as informações produzidas e consumidas ininterruptamente nas instituições precisam passar por processos efetivos de gestão da informação e governança de dados, seja por necessidade interna, seja por conformidade legal. Esta pesquisa tem por objetivo verificar a viabilidade de integração de métodos da Ciência da Informação aos já utilizados no processo de modelagem de bancos de dados da Ciência da Computação, para a criação dos bancos de dados nos sistemas de informação. Para tanto, procede-se a pesquisa aplicada, envolvendo a modelagem de um banco de dados para uma aplicação em conjunto com a modelagem da informação desta aplicação, utilizando-se a metodologia de Arquitetura da Informação para Processos de Negócio. O presente trabalho identifica interseções entre as metodologias, sendo observada a correspondência entre o elemento “metadado de negócio” da Arquitetura da Informação para Processos de Negócio e o elemento “atributo da entidade” dos bancos de dados. Tal resultado permite concluir que há aderência entre as metodologias e, ao somar-se a implementação da Arquitetura da Informação para Processos de Negócio à modelagem do banco de dados, é possível traçar a linhagem dos dados e informações, desde o banco de dados até o processo de trabalho.

Palavras-chave: modelagem de banco de dados; arquitetura da informação para processos de negócio; catálogo de dados; governança de dados

1 Introdução

Os dados e informações produzidos e consumidos ininterruptamente nas instituições requerem processos efetivos de gestão da informação e governança de dados. Isso se deve tanto à necessidade interna de uso estratégico e gerencial dos dados quanto à conformidade legal, exemplificada pela Lei Geral de Proteção de Dados Pessoais (LGPD). Dados que não sejam tratados, geridos,

governados e contextualizados não geram informações, não possuem valor, tampouco potencial de apontar caminhos e tendências para a estratégia da instituição. A gestão da informação e a governança de dados também são impactadas por aspectos técnicos como a definição de bancos de dados, de engenharia de dados e de ferramentas tecnológicas empregadas. Para que os dados armazenados nos bancos de dados das várias aplicações da instituição sejam verdadeiramente úteis, é essencial implementar uma gestão e governança eficazes sobre eles:

Uma vez que a maior parte dos ativos informacionais são armazenados e mantidos em ambientes geridos pelas áreas de Tecnologia da Informação (TI), as soluções, ferramentas e serviços de TI são agentes fundamentais para a gestão efetiva da informação corporativa. No entanto, a significativa complexidade dos ambientes e processos relacionados às áreas de TI nas organizações, bem como as deficiências de gestão representam obstáculos para que seus serviços sejam utilizados de forma eficaz para o atendimento às demandas das áreas de negócio; sejam demandas relacionadas aos processos de trabalho típicos do negócio, quanto aquelas relacionadas à gestão da informação (Vilarins, 2013, p. 6).

Os sistemas de gerenciamento de bancos de dados (SGBD) são coleções de dados inter-relacionados que podem ser acessados e modificados pelos usuários, fornecendo uma visão abstrata dos dados e ocultando detalhes de como são armazenados e mantidos (Silberschatz; Korth; Sudarshan, 2020). No desenvolvimento de uma solução tecnológica, é realizada a modelagem dos dados e a seleção do SGBD que irá abrigar potenciais elementos de grande valor para a instituição: os dados. A metodologia de Arquitetura da Informação para Processos de Negócio (AIPN) (Brandt; Vidotti, 2020) oferece meios para que esses dados sejam mapeados, registrados e geridos a partir de identificação, registro e descrição dos metadados de negócio.

O objetivo deste trabalho é verificar a aderência entre as metodologias de modelagem de dados em bancos de dados e a metodologia de AIPN, e identificar a possibilidade de serem implementadas em conjunto, o que deve permitir a gestão e governança dos dados desde sua origem. Para isso, foi realizado um estudo em que a modelagem de dados e a escolha de um banco de dados relacional para uma solução de automatização de um processo de trabalho

foram realizadas concomitantemente à implementação da metodologia AIPN. A seção a seguir trará o referencial teórico da AIPN e da modelagem de dados, utilizados como bases para este estudo.

2 Referencial teórico

Para subsidiar este estudo, são trazidos, neste referencial teórico, os principais aspectos das metodologias aplicadas ao estudo: arquitetura da informação para processos de negócio e a modelagem de dados em bancos de dados a partir do modelo entidade-relacionamento.

2.1 Arquitetura da Informação para Processos de Negócio

A AIPN é uma metodologia para a elaboração da arquitetura da informação de processos de negócio em instituições, empresas e demais tipos de organização de quaisquer áreas de atuação. Essa metodologia parte dos processos de negócio e utiliza como base os métodos, as práticas e os princípios de Biblioteconomia e de Ciência da Informação (CI).

Segundo Choo (2006), a informação é um componente intrínseco de quase tudo que as instituições fazem. Sem uma clara compreensão dos processos organizacionais em que a informação se torna percepção, conhecimento e ação, as empresas não podem compreender a importância de suas fontes e tecnologias da informação.

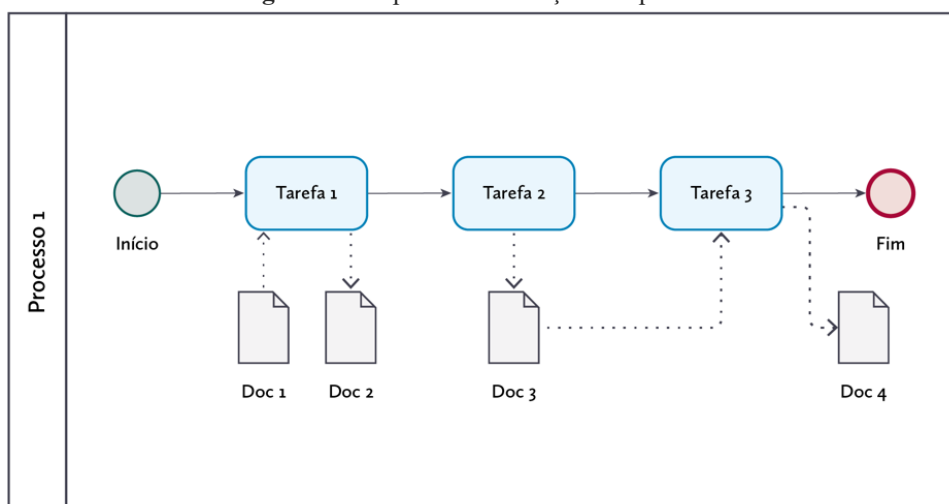
O gestor do negócio deve ser visto como o usuário da informação em projetos de sistemas de informação: “A arquitetura da informação deve reconhecer usuários como agentes do desenvolvimento tecnológico e garantir oportunidades de participação ativa no planejamento dos sistemas de informação, contrariando o determinismo tecnológico”. (Lima-Marques; Lacerda, 2006, p. 253). Apesar das já existentes metodologias de levantamento de requisitos utilizadas pela área de Tecnologia da Informação, acredita-se que a AIPN pode acrescentar aspectos importantes não considerados nessas abordagens, como afirma Victorino:

Esse fluxo [de informação], normalmente, é automatizado através do desenvolvimento de sistemas de informações computadorizados, utilizando metodologias de engenharia de software. No entanto,

essas metodologias raramente utilizam os conceitos, métodos e técnicas da Ciência da Informação de maneira integrada e, normalmente, não levam em conta a influência das estruturas organizacionais (Victorino, 2011, p. 98).

A metodologia AIPN parte do mapeamento de informações relevantes para o negócio e que devem ser gerenciadas. Essas informações são os chamados metadados de negócio, e irão abrigar os dados do negócio (Brandt, 2020). A primeira etapa da metodologia prevê o mapeamento das informações produzidas e consumidas em todas as etapas do processo de negócio. Essas informações estarão representadas em documentos, sistemas, planilhas, anotações ou quaisquer outros suportes de informações. Esta etapa resulta no documento chamado “Mapa das informações do processo”, ilustrado na figura 1 a seguir.

Figura 1 - Mapa das informações do processo



Fonte: Elaborado pelas autoras.

Os documentos identificados no processo de negócio devem ser analisados para a extração dos metadados de negócio, ou seja, as informações chave, essenciais para o negócio. Sherman (2006) afirma que os metadados de negócio, após capturados nos processos da instituição, devem ser geridos para que possam impulsionar os negócios. Marco (2006) compartilha dessa visão. Segundo o autor: “A chave para sua empresa prosperar está em quão bem você reúne, preserva e dissemina conhecimento. Ambientes com gestão de metadados são críticos para reunir, preservar e disseminar conhecimento”. (Marco, 2006, p.

43, tradução nossa).

Assim, na segunda etapa da metodologia AIPN, esses metadados de negócio devem ser registrados e descritos por meio de atributos que os representem: nome, descrição, data de criação, gestor, forma de acesso, entre outros que sejam importantes para a instituição e para o processo. Com isso, obtêm-se os requisitos informacionais do processo de negócio, que permitirão a implementação dos processos de gestão da informação e de governança de dados. Ao final da descrição de todos os metadados de negócio identificados, obtém-se uma matriz, que pode ser armazenada em forma de planilha ou em um repositório. Esse artefato funciona como um Catálogo de Dados, pois descreve como os dados abrigados nos metadados de negócio devem se comportar nos sistemas e como devem ser geridos e governados.

Quadro 1 - Exemplo de matriz de metadados

Nome do metadado	ID	Descrição	Gestor do dado	Regra de formato	Forma de acesso	Alimentação inicial	Restrição de acesso	...
Nome do funcionário	001	Pessoa contratada...	Dpto de Pessoal	Textual	SisRH	Seção de registro de pessoal	Restrito	...

Fonte: Elaborado pelas autoras.

A partir deste artefato é possível realizar a gestão dos metadados de negócio do processo, os quais serão alinhados à estrutura de TI:

[...] os metadados controlados pela Gestão de Metadados e os itens de configuração, com suas características e descrições que compõem o BDGC, devem ficar concentrados em uma mesma estrutura, ou em estruturas diferentes e separadas, mas com ligações que representam os relacionamentos que devem garantir a rastreabilidade dos elementos da infraestrutura de TI até os elementos da Arquitetura de Informação e vice-versa (Vilarins, 2013, p. 50).

Por fim, procede-se à verificação dos requisitos de recuperação da informação, tanto relacionados a critérios de buscas transacionais quanto para identificação de requisitos para a obtenção de informações gerenciais. Os requisitos para recuperação de informações transacionais podem ser quanto a desempenho, funcionalidades, navegação e filtros, ferramentas linguísticas, análise das buscas, entre outros (Brandt; Vidotti, 2020). Enquanto os requisitos

para recuperação de informações gerenciais referem-se às necessidades dos gestores para tomada de decisão, e são informações que devem constar em relatórios do sistema de informação. Para isso, o gestor do processo de negócio deve ser entrevistado e questionado sobre suas necessidades informacionais. Assim, é possível obter os requisitos de informações gerenciais e verificar se os metadados de negócio mapeados anteriormente são suficientes para fornecer dados que atendam às demandas do gestor. Caso haja necessidade de novos metadados para se obter essas informações gerenciais desejadas, deve-se fazer sua inclusão na matriz de metadados nesta etapa.

Quadro 2 - Exemplos de informações gerenciais

Requisito de informação gerencial	Metadados de negócio necessários
Total de vendas no período	Número de pedidos finalizados
	Data inicial
	Data final
Saber a produtividade da equipe	Número de tarefas realizadas
	Nome do funcionário
	Data inicial
	Data final

Fonte: Elaborado pelas autoras.

Com isso, finaliza-se a metodologia de arquitetura da informação para processos de negócio. Os artefatos gerados na metodologia podem ser utilizados para auxiliar no desenvolvimento de sistemas de informações e outros produtos digitais e, posteriormente, na gestão da informação e na governança de dados do processo de trabalho, garantindo assim a efetividade dessas ações. O diferencial desta metodologia é a visão informacional presente nas etapas, com princípios da Biblioteconomia, como os princípios da catalogação e objetivos bibliográficos, e da Ciência da Informação, como a Recuperação da Informação.

2.2 Modelagem de dados em bancos de dados

Os sistemas de informação, em sua maioria, utilizam bancos de dados para armazenar os dados relativos às operações realizadas pelo sistema:

É correto afirmar que os bancos de dados desempenham um papel crítico em quase todas as áreas em que os computadores são usados, incluindo negócios, comércio eletrônico, engenharia, medicina, genética, direito, educação e biblioteconomia. [...] Um banco de dados é projetado, construído e populado com dados com algum significado inerente. Ele possui um grupo específico de usuários e algumas aplicações previamente concebidas nas quais esses usuários estão interessados (Elmarsy; Navathe, 2011, p. 2).

As informações relevantes da empresa são armazenadas nos bancos de dados com base em um modelo de descrição, ou seja, é realizada uma modelagem de dados para representá-los nos sistemas de bancos de dados. Silberschatz, Korth e Sudarshan (2020) afirmam que, no início, os bancos de dados tinham uma estrutura simples de dados, formatados com precisão, mas atualmente essas aplicações podem incluir dados com relacionamentos complexos e estrutura mais variável. Entre tais estruturas, destaca-se o modelo Entidade-Relacionamento (E-R), proposto em 1976 por Peter Chen, bastante utilizado para a modelagem de dados em bancos de dados. Segundo Chen (1976), trata-se de um modelo que propõe uma visão mais natural do mundo real por meio das entidades e relacionamentos entre entidades. O autor explica que:

O banco de dados de uma empresa contém informações relevantes sobre entidades e relacionamentos em que a empresa está interessada. Uma descrição completa de uma entidade ou de um relacionamento pode não ser registrada no banco de dados da empresa. É impossível (e, talvez, desnecessário) registrar todas as peças de informação potencialmente disponíveis sobre entidades e relacionamentos (Chen, 1976, p. 11, tradução nossa).

No contexto de um banco de dados relacional, que ainda constitui a solução mais consagrada em sistemas e mais popularmente utilizada em aplicações de pequeno porte, uma entidade representa um conjunto de objetos da realidade modelada. O modelo E-R busca identificar as entidades necessárias para representar determinado processo da instituição, do qual as informações serão armazenadas em banco de dados, bem como mapear os relacionamentos entre as entidades identificadas. Uma modelagem E-R simples pode ser representada conforme figura a seguir:

Figura 2 - Exemplo de modelo conceitual E-R



Fonte: Elaborado pelas autoras.

Além disso, para cada entidade ou relacionamento identificado, são mapeados os diversos atributos que tais elementos possuam: “[...] um atributo pode ser formalmente definido como uma função mapeada a partir de uma entidade definida ou de um relacionamento definido em um valor definido ou um produto cartesiano de valores definidos” (Chen, 1976, p; 12, tradução nossa).

Para cada item adicionado ao banco de dados é dado o nome de ocorrência ou registro, conforme afirma Heuser (2011, p. 37): “A ocorrência (ou registro) é o item específico da entidade ou do relacionamento (linha da tabela)”. Outro conceito importante no modelo E-R para banco de dados é o de cardinalidade. O autor explica que:

Para fins de projeto de banco de dados relacional, uma propriedade importante de um relacionamento é a de quantas ocorrências de uma entidade podem estar associadas a uma determinada ocorrência através do relacionamento. Esta propriedade é chamada de cardinalidade de uma entidade em um relacionamento (Heuser, 2011, p. 38).

Assim, conforme a metodologia de Chen (1976), a modelagem de dados para um banco de dados utilizando o modelo entidade-relacionamento pode ser resumida em 4 passos:

- a) identificação das entidades e relacionamentos de interesse;
- b) identificação da informação semântica nos relacionamentos;
- c) definição do conjunto de valores e atributos;
- d) organização dos dados em relações de entidades e relacionamentos e decisão das chaves primárias.

Aqui cabe uma observação quanto à chamada “informação semântica” dos relacionamentos nessa proposta de Chen. Neste modelo, a semântica trata-se apenas da identificação da cardinalidade dos relacionamentos, ou seja, de quantas instâncias de cada uma das entidades podem se relacionar com quantas outras instâncias das outras entidades do modelo. No exemplo da figura 4, poderia ser definido que cada instância da entidade “Pessoa” pode se relacionar com (1) instância da entidade “Projeto”, ou seja, seria uma relação em que cada pessoa pode participar de (1) projeto (1:1). Já a relação inversa, entre projetos e pessoas, pode ser definida por “1 projeto poderá ter N participantes (pessoas)”, sendo uma relação de 1:N. Esse tipo de identificação traz um nível de semântica para a modelagem, pois acrescenta algum significado aos dados, por meio da quantificação de seus relacionamentos. Porém, é importante ressaltar que se trata de um nível semântico bastante básico e pouco representativo das relações em termos qualitativos, como o que ocorre, por exemplo, em uma ontologia ou com a utilização de vocabulários da web semântica.

A seção a seguir traz a metodologia utilizada no estudo que buscou aplicar, de forma integrada, as metodologias descritas na presente seção.

3 Metodologia

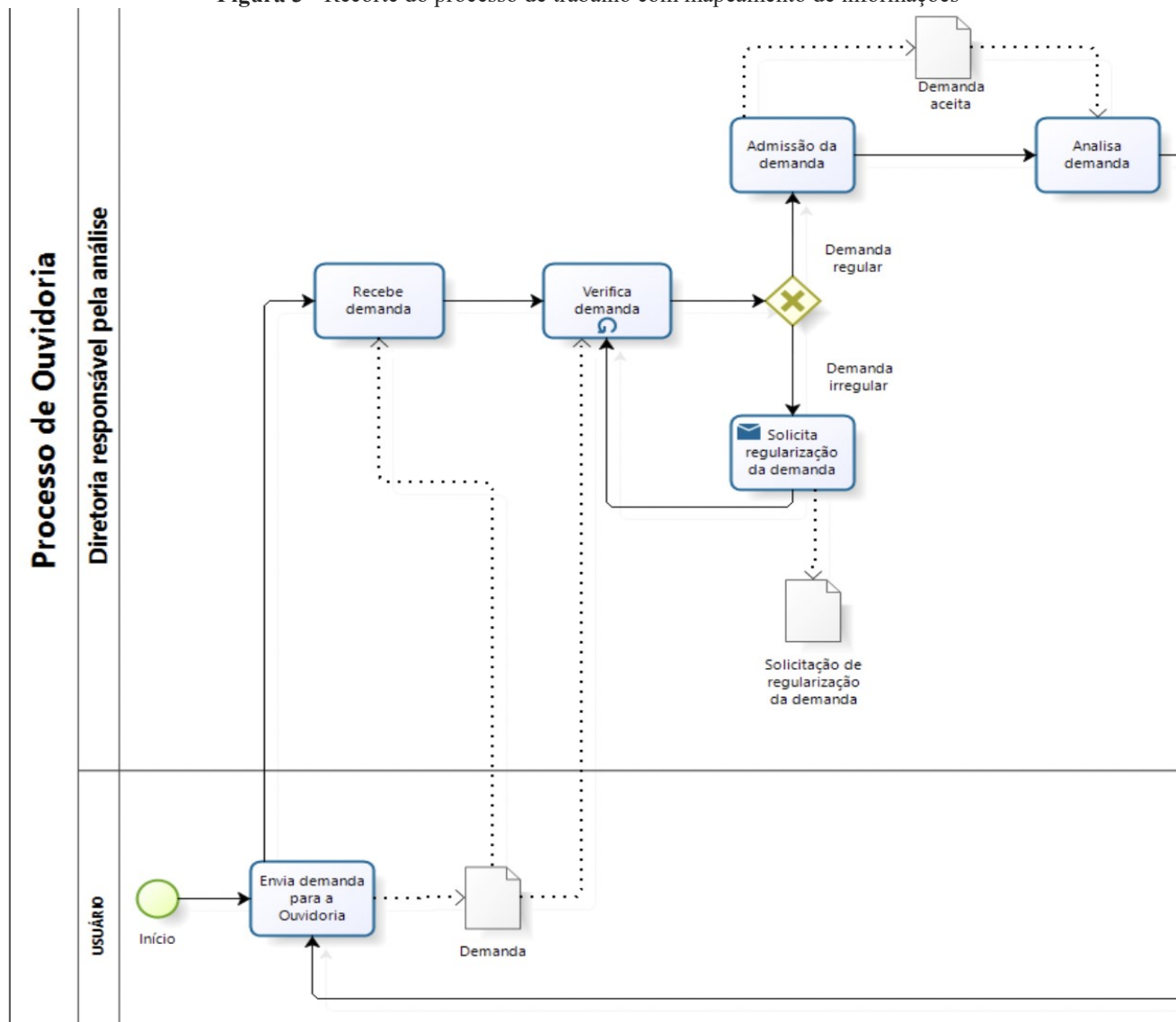
O estudo ora debatido se valeu de pesquisa aplicada, utilizando a metodologia de Arquitetura da Informação para Processos de Negócio em conjunto com a metodologia tradicional de modelagem de dados em banco de dados relacionais, pelo modelo entidade-relacionamento, resultando em um artefato de modelagem de dados enriquecido com informações para sua implementação e posterior gestão e governança dos dados e informações do negócio.

Para isso, partiu-se de um projeto hipotético baseado em uma situação concreta real, que objetivou o desenvolvimento de uma solução web para atendimento de uma ouvidoria de plano de saúde. A primeira etapa elaborada foi o mapeamento do processo de negócio “Atender demandas da ouvidoria”. Foi utilizada a metodologia *Business Process Modelling* e o software *Bizagi*.

3.1 Modelagem com AIPN

Com o processo de trabalho mapeado, foi realizada a primeira etapa da AIPN, com a identificação das informações deste processo de negócio. O diagrama do processo com as informações mapeadas foi validado por um servidor público federal vinculado à ouvidoria em questão.

Figura 3 - Recorte do processo de trabalho com mapeamento de informações



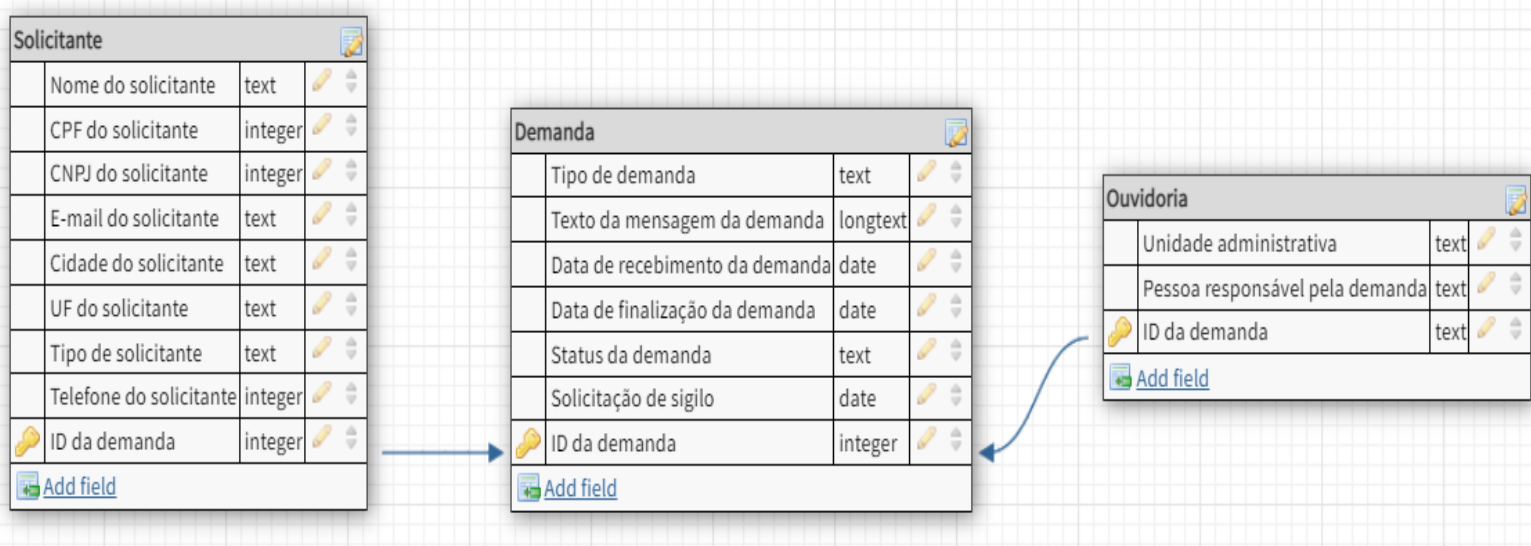
Fonte: Elaborado pelas autoras.

Em seguida, partiu-se para a análise de documentos utilizados no processo. Com base em um modelo exemplificativo de cada um dos documentos identificados, iniciou-se a extração de metadados de negócio, segunda etapa da metodologia de AIPN. Com a lista dos metadados de negócio concluída,

realizou-se a respectiva descrição, gerando a matriz de metadados, fonte para o catálogo de dados do processo.

Em seguida, foi realizada uma modelagem conceitual prévia com os metadados de negócio, em que foram identificadas as entidades do processo, bem como os relacionamentos entre metadados e entidades. Neste ponto, utilizou-se o modelo tradicional de E-R para modelagem de dados. Verificou-se, assim, que os metadados de negócio da metodologia AIPN correspondem aos atributos das entidades da metodologia de modelagem de dados E-R. Com isso, chegou-se ao seguinte resultado:

Figura 4 - Modelagem de dados com metadados de negócio



Fonte: Elaborado pelas autoras.

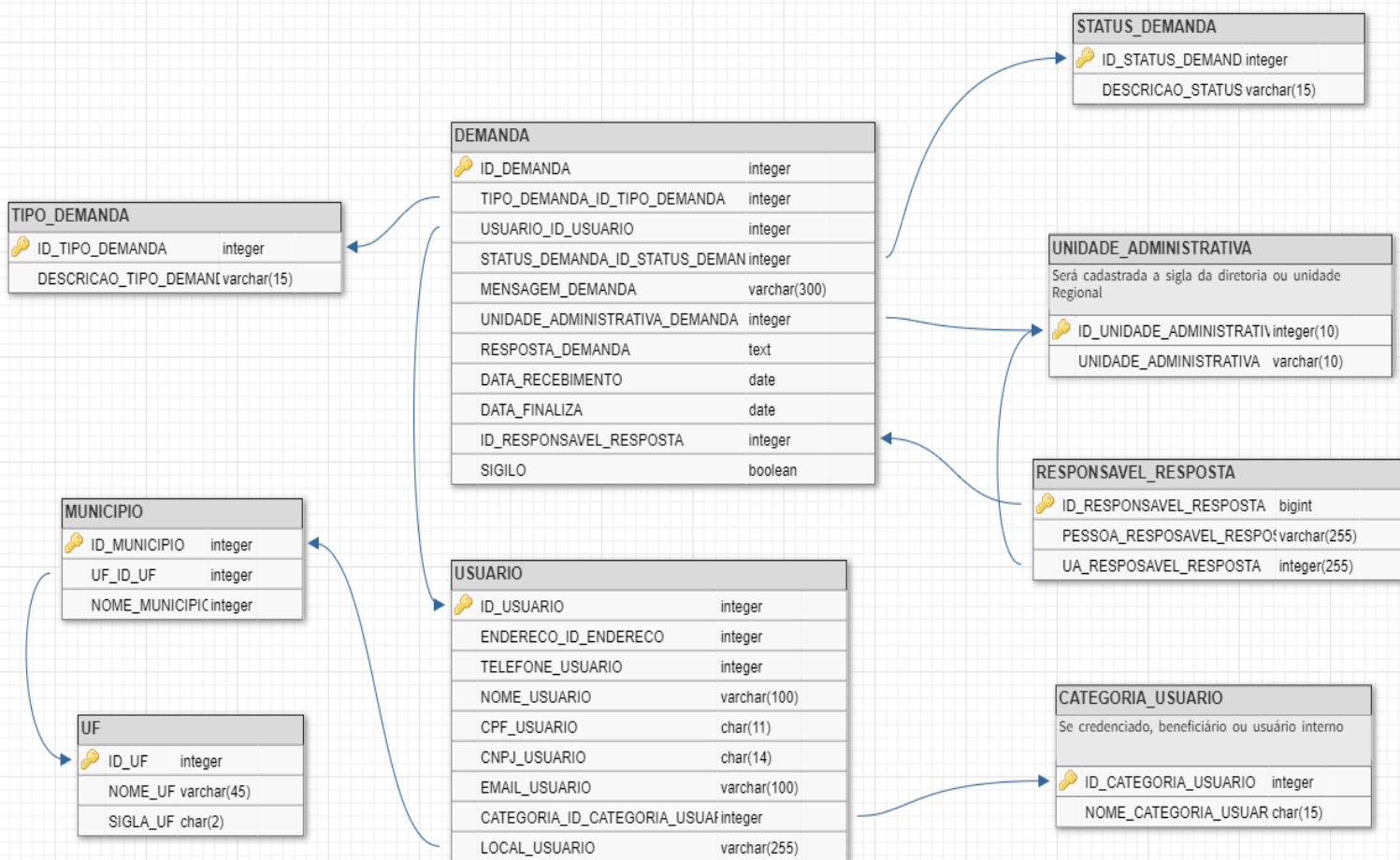
Foram obtidas três tabelas (três entidades), relacionadas por atributos de chaves primárias, além dos demais atributos necessários para a descrição – que haviam sido mapeados como metadados de negócio.

3.1 Modelagem E-R

Enquanto a metodologia AIPN foi implementada por um integrante do projeto, a modelagem dos dados de forma tradicional com o modelo E-R foi realizada por um outro integrante do projeto para que ambas pudessem ser comparadas.

O resultado obtido pela modelagem tradicional está representado na figura 5:

Figura 5 - Modelagem de dados tradicional



Fonte: Elaborado pelas autoras.

O resultado desta modelagem traz nove tabelas com seus atributos e relacionamentos, conforme ilustrado na figura 5.

3.3 AIPN e E-R

Ao comparar o modelo obtido em cada uma das metodologias implementadas para a obtenção do modelo de dados, temos que, pela AIPN, foram mapeadas três entidades, 19 atributos e dois relacionamentos, enquanto na modelagem E-R, nove entidades, 36 atributos e nove relacionamentos foram identificados.

Verificou-se que a modelagem tradicional E-R obteve uma quantidade maior de elementos – entidades, atributos e relacionamentos. Porém, ao comparar os atributos obtidos a partir de cada metodologia, verifica-se que todos os elementos encontrados por meio da AIPN, os seja, os metadados de negócio, foram também mapeados como atributos das entidades no modelo E-R, conforme o quadro 3, a seguir:

Quadro 3 - Relação entre atributos E-R e metadados de negócio AIPN

AIPN (metadados de negócio)	E-R (atributos)
Nome do solicitante	NOME_USUARIO
	ID_USUARIO (FK)
	USUARIO_ID_USUARIO
E-mail do solicitante	EMAIL_USUARIO
Telefone do solicitante	TELEFONE_USUARIO
CPF do solicitante	CPF_USUARIO
CNPJ do solicitante	CNPJ_USUARIO
Cidade do solicitante	NOME_MUNICIPIO
	ID_MUNICIPIO (FK)
	LOCAL_USUARIO
	ENDERECO_ID_ENDERECO
Tipo de solicitante	NOME_CATEGORIA_USUARIO
	ID_CATEGORIA_USUARIO (FK)
	CATEGORIA_ID_CATEGORIA_USUARIO
UF do solicitante	SIGLA_UF
	ID_UF (FK)
	UF_ID_UF
	NOME_UF
Solicitação de sigilo	SIGILO
Tipo de demanda	DESCRICAO_TIPO_DEMANDA
	TIPO_DEMANDA_ID_TIPO_DEMANDA
	ID_TIPO_DEMANDA (FK)
Texto da mensagem da demanda	MENSAGEM_DEMANDA

Data de recebimento da demanda	DATA_RECEBIMENTO
Data de finalização da demanda	DATA_FINALIZA
Texto da resposta da demanda	RESPOSTA_DEMANDA
Status da demanda	DESCRICAO_STATUS
	STATUS_DEMANDA_ID_STATUS_DEMANDA
	ID_STATUS_DEMANDA (FK)
Pessoa responsável pela resposta	PESSOA_RESPONSAVEL_RESPOSTA
	ID_RESPONSAVEL_RESPOSTA (FK)
Unidade administrativa	UA_RESPONSAVEL_RESPOSTA
	UNIDADE_ADMINISTRATIVA
	ID_UNIDADE_ADMINISTRATIVA (FK)
	UNIDADEADMINISTRATIVA_DEMANDA
Id da demanda	ID_DEMANDA (PK)

Fonte: Elaborado pelas autoras.

A partir do quadro 3, é possível verificar que há correspondência entre as modelagens, porém, em diferentes níveis de representação: isso se deve ao fato de a modelagem com os metadados de negócio estar mais próxima de um modelo conceitual do que de um modelo físico, como o da modelagem feita pela metodologia E-R tradicional. Ou seja, os elementos representados são os mesmos, porém, a relação de equivalência não é de 1 para 1, inclusive por existirem atributos do modelo E-R que são equivalentes e apenas se repetem por serem chave estrangeira de outra tabela. Tais atributos estão identificados no quadro 3 com a sigla FK - *foreign key*. A diferença entre as modelagens ocorre, também, devido à modelagem resultante da AIPN não atender aos princípios de normalização do banco de dados exigidos para a implementação do modelo físico no SGBD:

Normalização é o processo de organização de dados em um banco de dados. Isso inclui a criação de tabelas e o estabelecimento de relações entre essas tabelas de acordo com as regras projetadas para proteger os dados e tornar o banco de dados mais flexível, eliminando a redundância e a dependência inconsistente (Microsoft, 2023).

Tais regras são chamadas de formas normais, e há três níveis básicos de normalização: primeira forma normal (1FN), segunda forma normal (2FN) e terceira forma normal (3FN). A 1FN foi criada para impedir atributos multivalorados, ou seja, o domínio de um atributo deve incluir apenas valores

atômicos ou indivisíveis. A segunda forma normal se baseia no conceito de dependência funcional total, ou seja, cada atributo não principal deve ser total e funcionalmente dependente da chave-primária. Já a 3FN é baseada no conceito de dependência transitiva, o que significa que a relação não deve ter um atributo não chave determinado funcionalmente por outro atributo (ou conjunto de atributos) não chave. (Elmarsy; Navathe, 2011). Cumpridas essas três regras, o modelo estará normalizado. Em geral, a normalização requer tabelas adicionais (Microsoft, 2023), o que pôde ser observado na figura 6, com o modelo normalizado.

Ao analisar as duas modelagens (figuras 4 e 5), foi possível verificar a correspondência entre metadados de negócio e atributos das entidades e, assim, conectar a metodologia AIPN com o modelo E-R. O passo seguinte foi retornar ao artefato da matriz de metadados para incluir mais três itens de descrição, que identificam a entidade, a tabela e o atributo correspondente a cada um dos metadados de negócio.

4 Resultados e discussão

Ao final dos passos relatados na seção da metodologia, o artefato obtido continha as seguintes informações:

Quadro 4 - AIPN com modelagem de banco de dados

Finalidade	Descrição	Exemplo
Identificação	Nome do metadado de negócio	UF do solicitante
	Definição	Unidade da Federação ou Unidade Federativa onde se localiza o solicitante da demanda; refere-se aos 26 estados brasileiros, mais o Distrito Federal, em um total de 27 Unidades Federativas (UF)
Gestão da informação e governança de dados	Documentos	Demanda; Solicitação de Regularização da Demanda; Demanda Aceita; Solicitação de informações adicionais; Resposta demais diretorias
	Gestor do metadado	Seção de padronização e controle de vocabulário

Finalidade	Descrição	Exemplo
	Gestor do dado	Ouvidoria PlanAssiste
	Alimentação inicial do dado	Solicitante
	Forma de acesso	SisAssiste
Técnicos (banco de dados)	Entidade (Modelo E-R)	Solicitante
	Atributo	SIGLA_UF
	Tabela	UF
	Entrada padronizada / restrição de domínio	AL; AP; AM; BA; CE; DF; ES; GO; MA; MT; MS; MG; PA; PB; PR; PE; PI; RJ; RN; RS; RO; RR; SC; SP; SE; TO
	Regra de formato / Type	textual / char(2)
Conformidade legal	Restrição de acesso	Restrito
	Dado pessoal?	Não

Fonte: Elaborado pelas autoras.

Cada um dos metadados de negócio identificados no processo de negócio foi descrito com os mesmos parâmetros do exemplo do quadro 4. Verificou-se que alguns itens já utilizados na AIPN para descrição dos metadados de negócio correspondem a itens mapeados na modelagem de dados:

- a) nome do metadado de negócio = Atributo;
- b) entrada padronizada = Restrição de domínio;
- c) regra de formato = *Type* (tipo primitivo do dado).

Com isso, é possível utilizar a modelagem final obtida para implementar o banco de dados e, assim, integrar os metadados de negócio ao banco de dados. Essa implementação possibilita registrar e consultar origem e linhagem dos dados, sua definição, seu gestor, entre outros papéis da governança de dados, por meio de um catálogo de dados. Além disso, os itens de descrição dos metadados de negócio possibilitam a gestão efetiva das informações do processo

de trabalho, bem como auxiliam em processos de melhoria da qualidade dos dados.

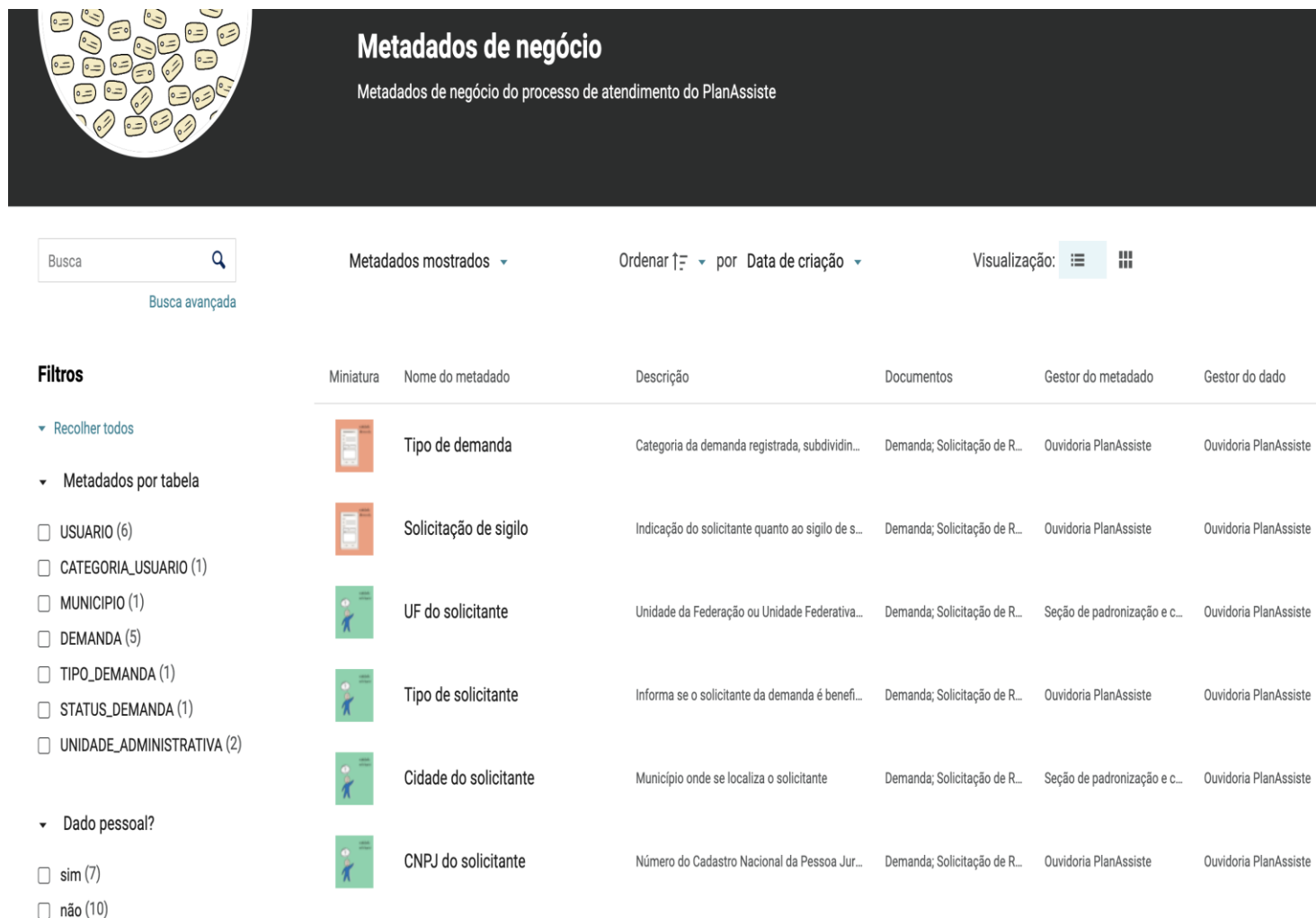
4.1 Repositório de metadados: o catálogo de dados

Para o registro do artefato final dessa modelagem, foi criado um repositório de metadados de negócio:

Tal implementação permite que os metadados de negócio identificados nos processos e registrados nos documentos de AI possuam uma fonte de recuperação única, que poderá estar disponível para a consulta de toda a instituição e com variados níveis de acesso (Brandt; Vidotti, 2020, p.11).

Ou seja, a criação de um repositório permite armazenar as informações estruturadas acerca dos metadados de negócio e possui ferramentas de busca, filtros e outros recursos que facilitam as consultas tanto dos gestores de negócio quanto dos gestores técnicos quanto ao funcionamento dos dados e metadados do processo de negócio. Esse repositório funcionará como um catálogo de dados corporativo, na medida em que forem adicionadas as modelagens de informação dos demais processos de negócio da instituição.

Figura 6 - Repositório de metadados- interface do repositório Tainacan¹



The screenshot shows the 'Metadados de negócio' interface. At the top, there's a search bar with 'Busca' and a magnifying glass icon, and a link to 'Busca avançada'. Below the search bar, there are filters on the left and a table of metadata in the center. The filters include 'Recolher todos', 'Metadados por tabela', and 'Dado pessoal?'. The table has columns for 'Miniatura', 'Nome do metadado', 'Descrição', 'Documentos', 'Gestor do metadado', and 'Gestor do dado'.

Miniatura	Nome do metadado	Descrição	Documentos	Gestor do metadado	Gestor do dado
	Tipo de demanda	Categoria da demanda registrada, subdividin...	Demanda; Solicitação de R...	Ouvidoria PlanAssiste	Ouvidoria PlanAssiste
	Solicitação de sigilo	Indicação do solicitante quanto ao sigilo de s...	Demanda; Solicitação de R...	Ouvidoria PlanAssiste	Ouvidoria PlanAssiste
	UF do solicitante	Unidade da Federação ou Unidade Federativa...	Demanda; Solicitação de R...	Seção de padronização e c...	Ouvidoria PlanAssiste
	Tipo de solicitante	Informa se o solicitante da demanda é benefi...	Demanda; Solicitação de R...	Ouvidoria PlanAssiste	Ouvidoria PlanAssiste
	Cidade do solicitante	Município onde se localiza o solicitante	Demanda; Solicitação de R...	Seção de padronização e c...	Ouvidoria PlanAssiste
	CNPJ do solicitante	Número do Cadastro Nacional da Pessoa Jur...	Demanda; Solicitação de R...	Ouvidoria PlanAssiste	Ouvidoria PlanAssiste

Fonte: Elaborado pelas autoras.

5 Considerações finais

O desenvolvimento de sistemas de informação nas instituições tem como objetivo automatizar, facilitar, simplificar e otimizar os processos de trabalho das áreas de negócio. Esses projetos são usualmente realizados pela área de Tecnologia em conjunto com as áreas de negócio, quase sempre sem a participação de especialistas em informação.

A Arquitetura da Informação para Processos de Negócio (AIPN) fornece um método para que profissionais da informação (arquitetos da informação, cientistas da informação, bibliotecários e arquivistas) possam identificar as necessidades de informação do negócio de forma clara e completa, representar essas informações e fornecer insumos para o desenvolvimento tecnológico. A

presença desses profissionais traz a visão da Biblioteconomia e da Ciência da Informação, sendo este o diferencial da AIPN em relação às metodologias tradicionais da Tecnologia da Informação. Segundo pesquisa que investigou as possibilidades de integração entre a Arquitetura da Informação e a gestão de serviços de Tecnologia da Informação:

A governança explicita os ativos de informação e garante que os processos de trabalho incorporem regras de validação que aumentam a qualidade e cooperam para a percepção do valor desses ativos. Esse controle atribuído à governança da informação é consolidado e operacionalizado por meio da Gestão de Metadados (Vilarins, 2013, p. 49).

Assim, acredita-se possível obter resultados importantes a partir da integração desse profissional com as equipes de projetos que envolvem desenvolvimento tecnológico (sistemas de informação em geral e bancos de dados) para atuar como uma ponte entre a área de negócio e a equipe de tecnologia da informação.

A tecnologia é ferramenta essencial, mas a informação é o que traz valor para o negócio, e uma área especialista em informação pode fazer diferença no resultado. A área de tecnologia fornece um ferramental sofisticado para coleta, tratamento, uso e visualização de dados, por intermédio de numerosas linguagens de programação, frameworks, bibliotecas de código, sistemas gerenciadores de banco de dados, *data warehouses*, *data lakes*, painéis de visualização (*dashboards*) e diversos outros recursos. Não obstante, a gestão da informação e a governança de dados são fundamentais para o fornecimento de contexto e de sentido para os dados, com a subsequente formação de conhecimentos e de competências organizacionais e humanas.

Assim, acredita-se que a utilização da metodologia AIPN aplicada neste estudo poderia ser inserida nas instituições como parte da etapa de análise de requisitos da engenharia de software. A criação dos bancos de dados conforme a modelagem obtida pela AIPN, em conjunto com os processos necessários para normalização dos dados, permite que os dados gerados pela instituição sejam gerenciados desde sua origem, passando pelos processos de armazenamento, gestão e governança, até que sejam convertidos, de fato, em inteligência de

negócio.

O uso da AIPN para os vários processos de negócio da instituição permite o estabelecimento de uma ecologia de dados integrada e para toda a empresa. Assim, facilitando a implementação da governança de dados e gestão das informações de forma consistente e unificada e, com isso, promovendo a qualidade dos dados e gerando valor para a instituição.

Referências

- BRANDT, M. B. **Modelagem da informação legislativa: arquitetura da informação para o processo legislativo brasileiro**. 2020. Tese. (Doutorado em Ciência da Informação) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2020.
- BRANDT, M. B.; VIDOTTI, S. A. B. G. Arquitetura da Informação para processos de negócio: um caminho para a governança de dados. **Informação & Sociedade: Estudos**, João Pessoa, v. 30, n. 4, p. 1-16, 2020. Disponível em: <https://doi.org/10.22478/ufpb.1809-4783.2020v30n4.57349>. Acesso em: 29 nov. 2023.
- CHEN, P. P. The entity-relationship model - toward a unified view of data. **ACM Transactions Database Systems**, United States, v. 1, n. 1, p. 9-36, 1976. Disponível em: <https://dl.acm.org/doi/10.1145/320434.320440>. Acesso em: 18 nov. 2023.
- CHOO, C. W. **A organização do conhecimento**. 2. ed. São Paulo: Senac, 2006.
- ELMARSÍ, R.; NAVATHE, S. **Sistemas de banco de dados**. 6. ed. São Paulo: Pearson, 2011.
- HEUSER, C. A. **Projeto de banco de dados**. Porto Alegre: Grupo A, 2011. v. 4.
- LIMA-MARQUES, M.; LACERDA, F. Arquitetura da informação: base para a Gestão do Conhecimento. In: TARAPANOFF, K. O. (ed.). **Inteligência, informação e conhecimento**. Brasília: IBICT, 2006. p. 241-255.
- MARCO, D. Managing metadata for the business, part 1. **DM Review**, New York, v. 16, n. 2, p. 42-43, 2006.
- MICROSOFT. **Descrição das noções básicas de normalização do banco de dados**. [S.l.], 2023.
- SHERMAN, R. Align metadata and business initiatives. **DM Review**, New York, v. 16, n. 1, p. 50, 2006.

SILBERSCHATZ, A.; KORTH, H.; SUDARSHAN, S. **Sistema de Banco de Dados**. Porto Alegre: Grupo GEN, 2020.

VICTORINO, M. C.; **Organização da informação para dar suporte à arquitetura orientada a serviços: reuso da informação nas organizações**. 2011. Tese (Doutorado em Ciência da Informação) - Faculdade de Ciência da Informação, Universidade de Brasília, Brasília, 2011.

VILARINS, R. **Integrando a arquitetura de informação à gestão de serviços de tecnologia da informação**. 2013. Monografia (Especialização). Curso de Arquitetura e Organização da Informação, Câmara dos Deputados, Centro de Formação, Treinamento e Aperfeiçoamento; Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2013.

Information architecture for business processes and database modeling

Abstract: Data and information produced and consumed uninterruptedly in institutions need to go through effective information management and data governance processes, whether due to internal necessity or legal compliance. This research aims to verify the feasibility of integrating Information Science methods with those already used in the process of modeling databases in Computer Science, for the creation of databases in information systems. For that, applied research is carried out, involving the modeling of a database for an application together with the modeling of the information of this application, using the methodology of Information Architecture for Business Processes. The present work identifies intersections between the methodologies, observing the correspondence between the “business metadata” element of the Architecture for Business Processes and the “entity attribute” element of the databases. This result allows us to conclude that there is adherence between the methodologies and, by adding the Architecture for Business Processes implementation to the database modeling, it is possible to trace the lineage of the data and information, from the database to the business process.

Keywords: database modeling; information architecture for business processes; data catalog; data governance

Recebido: 30/03/2023

Aceito: 14/12/2023

Declaração de autoria

Concepção e elaboração do estudo: Mariana Baptista Brandt

Coleta de dados: Mariana Baptista Brandt

Análise e interpretação de dados: Mariana Baptista Brandt

Redação: Mariana Baptista Brandt, Silvana Ap. B. Gregório Vidotti

Revisão crítica do manuscrito: Silvana Ap. B. Gregório Vidotti

Como citar

BRANDT, Mariana B.; VIDOTTI, S. A. B. G. Arquitetura da informação para processos de negócio e modelagem de banco de dados: aproximações possíveis. **Em Questão**, Porto Alegre, v. 30, e-131304, 2024. DOI: <https://doi.org/10.1590/1808-5245.30.131304>

Parecer(es) aberto(s):

<https://doi.org/10.1590/1808-5245.30.131304.A>

<https://doi.org/10.1590/1808-5245.30.131304.B>



¹ O repositório com todas as informações deste estudo pode ser acessado em: <https://marianabrandt.com.br/102/index.php/metadados-de-negocio/>