

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
INSTITUTO DE INFORMÁTICA  
PROGRAMA DE PÓS-GRADUAÇÃO EM COMPUTAÇÃO

PAULO ALBERTO COUTINHO PRECHT

**Do mapeamento de Dados Abertos à Modelagem Conceitual:  
um *Data Warehouse* para análise de informações multidimensionais**

Dissertação apresentada ao Programa de Pós-Graduação em Computação, Instituto de Informática da Universidade Federal do Rio Grande do Sul, como requisito parcial para obtenção do título de Mestre em Computação.

Orientador: Prof. Dr. Leandro Krug Wives

Porto Alegre  
2019

## CIP – CATALOGAÇÃO NA PUBLICAÇÃO

Coutinho Precht, Paulo Alberto

Do mapeamento de Dados Abertos à Modelagem Conceitual: um Data Warehouse para análise de informações multidimensionais / Paulo Alberto Coutinho Precht. - 2019.

63 f.:il

Orientador: Dr. Leandro Krug Wives.

Dissertação (Mestrado) - Universidade Federal do Rio Grande do Sul, Instituto de Informática, Programa de Pós-Graduação em Computação, Porto Alegre, BR-RS, 2019.

1. Modelagem conceitual. 2. ETL - Extract, Transform, Load. 3. DW - Data Warehouse. 4. Dados abertos. I. Krug Wives, Leandro. Título.

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

Reitor: Prof. Rui Vicente Oppermann

Vice-Reitor: Profa. Jane Fraga Tutikian

Pró-Reitor de Pós-Graduação: Prof. Celso Giannetti Loureiro Chaves

Diretor do Instituto de Informática: Profa. Carla Maria Dal Sasso Freitas

Coordenador do PPGC: Profa. Luciana Salette Buriol

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

## **AGRADECIMENTOS**

Agradeço a todos que de alguma maneira contribuíram e me acompanharam ao longo da realização desta dissertação. Um agradecimento em especial ao meu orientador Dr. Leandro Krug Wives, a Angélica Siqueira, aos meus amigos e principalmente a minha família. A todos meu muito obrigado.

## RESUMO

Com a ampliação do acesso à Internet, os sistemas Web de visualização de dados abertos têm se tornado ferramentas importantes no controle social. Apesar de oferecerem um grande potencial para extrair informações relevantes para a gestão e controle de políticas públicas no país, os dados abertos são usualmente disponibilizados em linguagens diversas, formatos estáticos e de difícil leitura e interpretação, trazendo obstáculos para importações e a integração das informações. Nesse contexto, o presente trabalho pretende demonstrar a relevância da Modelagem Conceitual para o desenvolvimento de sistemas que gerenciem dados abertos multidimensionais de maneira extensível, adequados às demandas das instituições que atuam no monitoramento da Política Pública brasileira. Para tal, foi avaliado um conjunto amplo de dados secundários trabalhados pelo Observatório Socioambiental em Segurança Alimentar do Rio Grande do Sul (OBSSAN). A experiência resultou no desenvolvimento de um *Data Warehouse* modelado em formato estrela, carregando e integrando os dados por meio de processos de extração e carga, o chamado *Extract Transform Load* (ETL). Observou-se que o modelo desenvolvido proporciona consistência ao banco de dados analisado neste trabalho, possuindo a capacidade de validar os dados a serem carregados e viabilizar a inserção de novos dados de maneira compatível com as dimensões e restrições necessárias; garantindo, portanto, a confiabilidade das informações que serão apresentadas na ferramenta de visualização. Cabe ressaltar que o ambiente Web desenvolvido será utilizado por agentes sociais da gestão pública, do controle social, bem como acadêmicos e demais interessados sobre o tema.

**Palavras-chave:** Modelagem conceitual. Extract Transform Load. Data Warehouse. Dados abertos.

## **From Open Data Mapping to Conceptual Modeling: a Data Warehouse for Multidimensional Information Analysis**

### **ABSTRACT**

With internet access expansion, open data visualization Web systems have become essential tools in social control. Despite offering a great potential to extract relevant information to the Public Policies management and control in the country, open data is usually available in a variety of languages and static formats that are difficult to read and interpret, bringing obstacles to import and integrate the information. In this context, the present work intends to demonstrate the relevance of Conceptual Modeling for the development of systems that manage multidimensional open data in an extensible form, suitable to the institutions' demands for the Brazilian Public Policy monitoring. To this end, a broad set of secondary data provided by the Socio-Environmental Observatory in Food Security of Rio Grande do Sul (OBSSAN) was evaluated. The experience resulted in the development of a Data Warehouse (DW) modeled in star format, loading and integrating the data through Extract Transform Load (ETL) processes. It was observed that the developed model provides consistency to the database analyzed in this work, having the ability to validate the loaded data and to enable the new data insertion in a way that is compatible to the necessary dimensions and restrictions; guaranteeing, therefore, the information reliability in the visualization tool. It should be noted that the developed Web environment will be used by public managers and social control agents, academics and other stakeholders.

**Keywords:** Conceptual model. Extract Transform Load. Data Warehouse. Open Data.

## LISTA DE FIGURAS

Figura 2.1 - Portal Brasileiro de Dados Abertos: formato dos dados.....	16
Figura 2.2 - Como cada indivíduo pode visualizar um projeto.....	17
Figura 2.3 - Processo de ETL.....	20
Figura 2.4 - Fato compra e seus elementos.....	21
Figura 2.5 - Estrutura de um Modelo Estrela ( <i>Star schema</i> ).....	22
Figura 2.6 - Estrutura de um Modelo Floco de Neve ( <i>Snowflake</i> ).....	23
Figura 4.1 - Representação das relações entre as características dos dados.....	29
Figura 4.2 - Organograma dos dados do OBSSAN.....	31
Figura 4.3 - DW Obssan - Modelo Estrela.....	33
Figura 5.1 - Área restrita - Cadastro de dimensões.....	38
Figura 5.2 - Área restrita - Importação de dados.....	39
Figura 5.3 - Área restrita - Cadastro de dimensões.....	40
Figura 5.4 - Elementos da tela inicial do sistema.....	41
Figura 5.5 - Filtros da Tela “Dimensões/Indicadores”.....	42
Figura 5.6 - Resultados em “modo planilha” da Tela “Dimensões/Indicadores”.....	43
Figura 5.7 - Resultados em “modo gráfico” da Tela “Dimensões/Indicadores”.....	44
Figura 5.8 - Resultado em formato de cartograma da Tela “Dimensões/Indicadores”.....	45
Figura 5.9 - Cartogramas da tela “Agregações territoriais”.....	46
Figura 5.10 - Detalhamento das “Agregações territoriais”.....	47
Figura 5.11 - Tela “Municípios”.....	47
Figura 5.12 - Tela “Municípios” - Relatório municipal.....	48

## LISTA DE TABELAS

Tabela 4.1 - Dicionário de Dados: Entidades.....	33
Tabela 4.2 - Dicionário de Dados: Atributos.....	57
Tabela 5.1 - Planilha de inserção e atualização de dados no DW.....	37
Tabela 5.2 - Perfil dos participantes da avaliação de usabilidade.....	51
Tabela 5.3 - Resultados da avaliação de usabilidade.....	52

## LISTA DE ABREVIATURAS E SIGLAS

ANS	Agência Nacional de Saúde Suplementar
BI	Business Intelligence
CONSEA	Conselho Estadual de Segurança Alimentar e Nutricional
CRM	Customer Relationship Management
DATASUS	Departamento de Informática do SUS
DW	Data Warehouse
ETL	Extract Transform Load
FEE	Fundação de Economia e Estatística
FUNASA	Fundação Nacional de Saúde
IBGE	Instituto Brasileiro de Geografia e Estatística
OBSSAN	Observatório Socioambiental em Segurança Alimentar e Nutricional
OLAP	On-Line Analytical Processing
PBDA	Portal Brasileiro de Dados Abertos
PGDR	Programa de Pós-Graduação em Desenvolvimento Rural
RS	Rio Grande do Sul
SGBD	Sistema Gerenciador de Banco de Dados
SLTI	Secretaria de Logística e Tecnologia da Informação
UFRGS	Universidade Federal do Rio Grande do Sul



# SUMÁRIO

<b>1. INTRODUÇÃO</b>	<b>10</b>
<b>1.1. OBJETIVOS</b>	<b>12</b>
<b>1.1.1. OBJETIVO GERAL</b>	<b>12</b>
<b>1.1.2. OBJETIVOS ESPECÍFICOS</b>	<b>12</b>
<b>1.2. ESTRUTURA DO TEXTO</b>	<b>12</b>
<b>2. FUNDAMENTAÇÃO</b>	<b>14</b>
<b>2.1. DADOS ABERTOS</b>	<b>14</b>
<b>2.2. MODELAGEM CONCEITUAL</b>	<b>16</b>
<b>2.3. PROCESSOS DE ETL, DW E MODELAGEM DIMENSIONAL</b>	<b>19</b>
<b>3. TRABALHOS RELACIONADOS</b>	<b>24</b>
<b>3.1. MODELOS CONCEITUAIS</b>	<b>24</b>
<b>3.2. PROJETOS PARA O OBSSAN</b>	<b>26</b>
<b>4. MODELO CONCEITUAL DESENVOLVIDO</b>	<b>28</b>
<b>4.1 O MODELO DESENVOLVIDO</b>	<b>29</b>
<b>5. VALIDAÇÃO E AVALIAÇÃO DO MODELO</b>	<b>36</b>
<b>5.1. GERENCIAMENTO DOS DADOS NO DW</b>	<b>36</b>
<b>5.2. FERRAMENTA DE VISUALIZAÇÃO DE DADOS</b>	<b>40</b>
<b>5.3. AVALIAÇÃO DA FERRAMENTA</b>	<b>49</b>
<b>6. CONCLUSÃO</b>	<b>54</b>
<b>REFERÊNCIAS</b>	<b>56</b>
<b>APÊNDICE A - DICIONÁRIO DE DADOS: ATRIBUTOS</b>	<b>59</b>
<b>APÊNDICE B - STORED PROCEDURE PARA INSERÇÃO DE DADOS</b>	<b>62</b>

## 1. INTRODUÇÃO

Desde 2009, através da Secretaria de Logística e Tecnologia da Informação (SLTI) e do Ministério do Planejamento Orçamento e Gestão, o governo brasileiro vem realizando ações para o desenvolvimento de uma política de disseminação de dados e informações governamentais para o livre uso pela sociedade. No Brasil existem diversas organizações públicas que disponibilizam tais dados, algumas delas são: Instituto Brasileiro de Geografia e Estatística (IBGE); Agência Nacional de Saúde Suplementar (ANS); e Fundação Nacional de Saúde (FUNASA). A política de dados abertos tem como principais objetivos o incentivo da transparência, o envolvimento na participação social, o desenvolvimento de novos e melhores serviços governamentais e o aumento da confiança pública (PORTAL BRASILEIRO DE DADOS ABERTOS, 2018). Nesse contexto, com a ampliação do acesso à internet, sistemas Web de visualização de dados abertos têm se tornado ferramentas importantes na gestão pública e no controle social.

Os dados disponibilizados oferecem um grande potencial para extrair informações que podem ser utilizadas pelos mais diversos órgãos públicos e também de toda a população em geral. Entretanto, esses dados são normalmente disponibilizados em formatos estáticos e de difícil leitura e interpretação, além de linguagens diversas que trazem obstáculos para importações e a integração entre diferentes plataformas. Outra questão observada é que os dados brutos não possuem *layouts* e formatos padronizados, tornando a análise de temas multidimensionais um desafio ainda maior (PORTAL BRASILEIRO DE DADOS ABERTOS, 2018). Desse modo, considerando a necessidade de fortalecer a transparência dos governos são percebidas novas demandas para a gestão pública: a simples disponibilização de dados brutos não é suficiente, necessitando de sistemas de informação capazes de apresentar as informações de maneira cada vez mais intuitiva e dinâmica para os usuários.

Avanços na tecnologia da informação disponibilizam recursos que possibilitam armazenar, processar e controlar grandes volumes de dados. Nesse contexto, os sistemas mais utilizados são os de *Business Intelligence* (BI), que consistem em um conjunto de tecnologias e ferramentas que fazem uso de dados que já se encontram disponíveis dentro de um ambiente. Esses sistemas têm como principais objetivos proporcionar o acesso interativo aos dados e permitir a manipulação dos mesmos para atingir os resultados esperados. O processo

de BI permite a transformação de dados em informações, depois em decisões e finalmente em ações (KIMBALL; ROSS, 2013).

Com o intuito de adquirir informações para os mais diversos interesses, instituições podem utilizar esses sistemas para integrar seus grandes volumes de dados. Por meio de processos de extração e carga, o chamado *Extract Transform Load* (ETL), esses dados, que podem ser provenientes de diferentes fontes, são carregados e integrados dentro de um *Data Warehouse* (DW). Dessa forma, passam a ser utilizados por ferramentas, tais como *On-Line Analytical Processing* (OLAP), as quais permitem visualizar e explorar os dados em diversas perspectivas e dimensões, realizando cruzamentos multidimensionais entre as mesmas (KIMBALL; ROSS, 2013).

No Brasil existe um grande repositório de Dados Abertos sobre políticas públicas disponíveis em portais oficiais do Governo. Esses dados encontram-se em formatos e *layouts* diferentes, apresentando potencial para serem explorados através das ferramentas descritas anteriormente (PORTAL BRASILEIRO DE DADOS ABERTOS, 2018). Um exemplo de Instituição que pode se beneficiar através do uso de um ambiente de BI é o Observatório Socioambiental em Segurança Alimentar e Nutricional (OBSSAN) do estado do Rio Grande do Sul (RS), o qual coleta e disponibiliza dados oriundos de fontes secundárias, tais como, Instituto Brasileiro de Geografia e Estatística (IBGE), Ministério do Desenvolvimento Social, Departamento de Informática do SUS (DATASUS), entre outros. O OBSSAN foi criado pelo Núcleo de Estudos e Pesquisas em Segurança Alimentar e Nutricional, vinculado ao Programa de Pós-Graduação em Desenvolvimento Rural (PGDR) da Universidade Federal do Rio Grande do Sul (UFRGS), em parceria com o Conselho Estadual de Segurança Alimentar e Nutricional (CONSEA) do Rio Grande do Sul (RS) e o Departamento de Informática da UFRGS (SIQUEIRA et al., 2015).

Dada a característica do público-alvo atendido pelo OBSSAN-RS, bem como a natureza multidimensional e complexa da temática trabalhada (Segurança Alimentar e Nutricional), é imperativa a construção de um modelo de dados que atenda às necessidades da instituição, o que demanda conhecimento aprofundado sobre os dados em questão, e um grau elevado de personalização do sistema de visualização das informações. Este trabalho propõe um modelo conceitual extensível para tal domínio, sendo fundamental para auxiliar no desenvolvimento do sistema do OBSSAN que o modelo tenha a capacidade de validar os

dados a serem carregados, além de viabilizar a inserção de novos dados de maneira compatível com as dimensões e restrições necessárias.

## **1.1. Objetivos**

Os objetivos deste trabalho foram organizados e divididos em Objetivo Geral e Objetivos específicos, os quais são apresentados a seguir.

### **1.1.1. Objetivo geral**

Desenvolver e validar um modelo conceitual para o Observatório Socioambiental em Segurança Alimentar e Nutricional (OBSSAN) do estado Rio Grande do Sul (RS), proporcionando a concepção e o desenvolvimento de um *Data Warehouse* (DW) multidimensional e extensível, concretizado através de um ambiente Web que serve como ferramenta de análise de dados.

### **1.1.2. Objetivos específicos**

- a) criar um modelo conceitual para o desenvolvimento do *Data Warehouse* do OBSSAN;
- b) permitir que o modelo seja extensível para a inserção de novos dados de maneira compatível com as dimensões e restrições existentes;
- c) avaliar o ambiente Web desenvolvido a partir do modelo conceitual desenvolvido, através de prova de conceito, mensurando sua efetividade e usabilidade, em conjunto com os usuários do OBSSAN.

## **1.2. Estrutura do texto**

O texto está estruturado em seis (06) capítulos. Após a introdução do tema e identificação dos objetivos do trabalho, são apresentados e discutidos os fundamentos teóricos que embasam a temática desenvolvida nesta dissertação, incluindo um tópico contextual sobre Dados abertos, e dois tópicos específicos tratando da Modelagem conceitual e dos Processos de ETL, DW e Modelagem dimensional. Posteriormente alguns trabalhos relacionados são

descritos, abordando a relevância dos Modelos conceituais no desenvolvimento de sistemas, e também os Projetos já desenvolvidos para o OBSSAN. No capítulo seguinte é detalhada a metodologia aplicada no desenvolvimento do Modelo conceitual, bem como o resultado obtido pelo autor. A proposta de validação e avaliação do modelo, bem como o detalhamento sobre o protótipo da ferramenta de visualização desenvolvidos, são apresentados no capítulo 5. Por fim, as conclusões são apresentadas no fechamento do trabalho, e o leitor pode aprofundar seu conhecimento sobre o trabalho através de três Apêndices, contendo o dicionário de dados do modelo, o script de importação e validação de dados do modelo, e ainda o teste aplicado para validação da ferramenta.

## 2. FUNDAMENTAÇÃO

Neste capítulo será apresentada a fundamentação teórica na qual este estudo é baseado. São eles: Dados abertos; Modelagem conceitual; e, Processos de ETL, DW e Modelagem Dimensional.

### 2.1. Dados abertos

O objetivo dessa seção é apresentar o conceito de “Dados Abertos”, o qual segundo a *Open Knowledge Foundation* (2018) é um conjunto de dados que qualquer pessoa pode livremente acessar, modificar e compartilhar conforme seus interesses, desde que preservem sua proveniência e abertura. Além disso, os dados devem estar disponíveis, preferencialmente em algum portal público na Internet e com um formato que aceite modificações, reutilização e redistribuição, inclusive com outros dados.

Algumas das muitas formas de utilizar Dados Abertos são:

- a) supervisionar políticas públicas;
- b) extrair novas informações através da combinação de diferentes fontes;
- c) encontrar padrões em grande volume de dados;
- d) visualizar os custos públicos.

No Brasil existem muitos projetos desenvolvidos a partir de dados abertos para atender aos mais diferentes objetivos. O portal Open Knowledge Brasil (2018), o qual é vinculado a *Open Knowledge Foundation*, reúne alguns deles, tais quais pode-se destacar:

- a) Eu voto: uma ferramenta online onde a população do município de São Paulo/SP pode opinar sobre os projetos de lei em tramitação na Câmara Municipal;
- b) Vai mudar: um projeto colaborativo que visa a participação social. Nele são apresentadas visualizações das relações entre poder econômico e poder público com a possibilidade de participar de discussões sobre os temas;
- c) Orçamento ao seu Alcance: um portal que tem o objetivo de mostrar o orçamento federal de forma mais simples e objetiva. Ele resume informações mensais referentes aos desembolsos financeiros dos ministérios e demais órgãos federais.

Fora do Brasil, projetos como o finlandês "*tax tree*" e o britânico "*where does my money go*" apresentam como o dinheiro dos impostos é utilizado pelos governos. Outros como

o Holandes “vervuilingsalarm.nl” avisa com uma mensagem como será a qualidade do ar na sua vizinhança. Já em no Reino Unido e na Alemanha, serviços como o “mapumental” e o “mapnificent” permitem aos usuários encontrar locais para morar, levando em conta diversos pontos, como por exemplo a distância e o tempo que eles levarão para chegar de suas casas até seus trabalhos. (OPEN KNOWLEDGE FOUNDATION, 2018).

No Brasil, a Lei de Acesso à Informação Nº 12.527, de 18 de novembro de 2011 regula o acesso a informações públicas. Através dela foram criadas ferramentas que possibilitam, a qualquer pessoa, física ou jurídica, sem necessidade de apresentar motivo, o recebimento de informações públicas dos órgãos e entidades. A Lei vale para os três Poderes da União, Estados, Distrito Federal e Municípios, inclusive aos Tribunais de Conta e Ministério Público. Entidades privadas sem fins lucrativos também são obrigadas a dar publicidade a informações referentes ao recebimento e à destinação dos recursos públicos por elas recebidos (BRASIL, 2011).

Através do Portal Brasileiro de Dados Abertos (PBDA) essas organizações disponibilizam diversos conjunto de dados conforme suas áreas de atuação. Além do PBDA, existem diversas organizações públicas que disponibilizam dados públicos em nível nacional, algumas delas são: o Instituto Brasileiro de Geografia e Estatística (IBGE); Agência Nacional de Saúde Suplementar (ANS); e Fundação Nacional de Saúde - (FUNASA). (PORTAL BRASILEIRO DE DADOS ABERTOS, 2018). No estado do Rio Grande do Sul, alguns exemplos seriam: Fundação de Economia e Estatística (FEE), OBSSAN-RS e ObservaPOA (SIQUEIRA et al, 2016).

Entretanto, conforme ilustra a Figura 2.1, normalmente esses dados são disponibilizados em diversos formatos, o que acaba dificultando a realização de importações, integrações, leituras e interpretações. Além disto, seus dados brutos não possuem um *layouts* e formatos padrão, o que dificulta ainda mais realizar integrações e análises.

Figura 2.1 - Portal Brasileiro de Dados Abertos: formato dos dados



Formatos	
CSV (4546)	
HTML (3773)	+
JSON (3372)	
wSDL (2945)	
KML (699)	
ZIP SHP (656)	
PDF (642)	
GeoJSON (606)	
XML (356)	
XLSX (120)	

Fonte: Portal Brasileiro de Dados Abertos, 2018.

Todavia, os dados disponibilizados pelo PBDA oferecem um grande potencial para extrair diversas informações que poderão ser utilizadas pelos mais diversos órgãos públicos e também de toda a população em geral. Para isso, através de processos de ETL é possível desenvolver um *Data Warehouse* de um domínio específico e integrar diversas fontes de dados, resultando em um ambiente capaz de prover informações, tais como: o acompanhamento e desenvolvimento de políticas públicas; informações sobre a transparência da gestão pública; e prover informações para aprimoramento na qualidade dos dados governamentais.

## 2.2. Modelagem conceitual

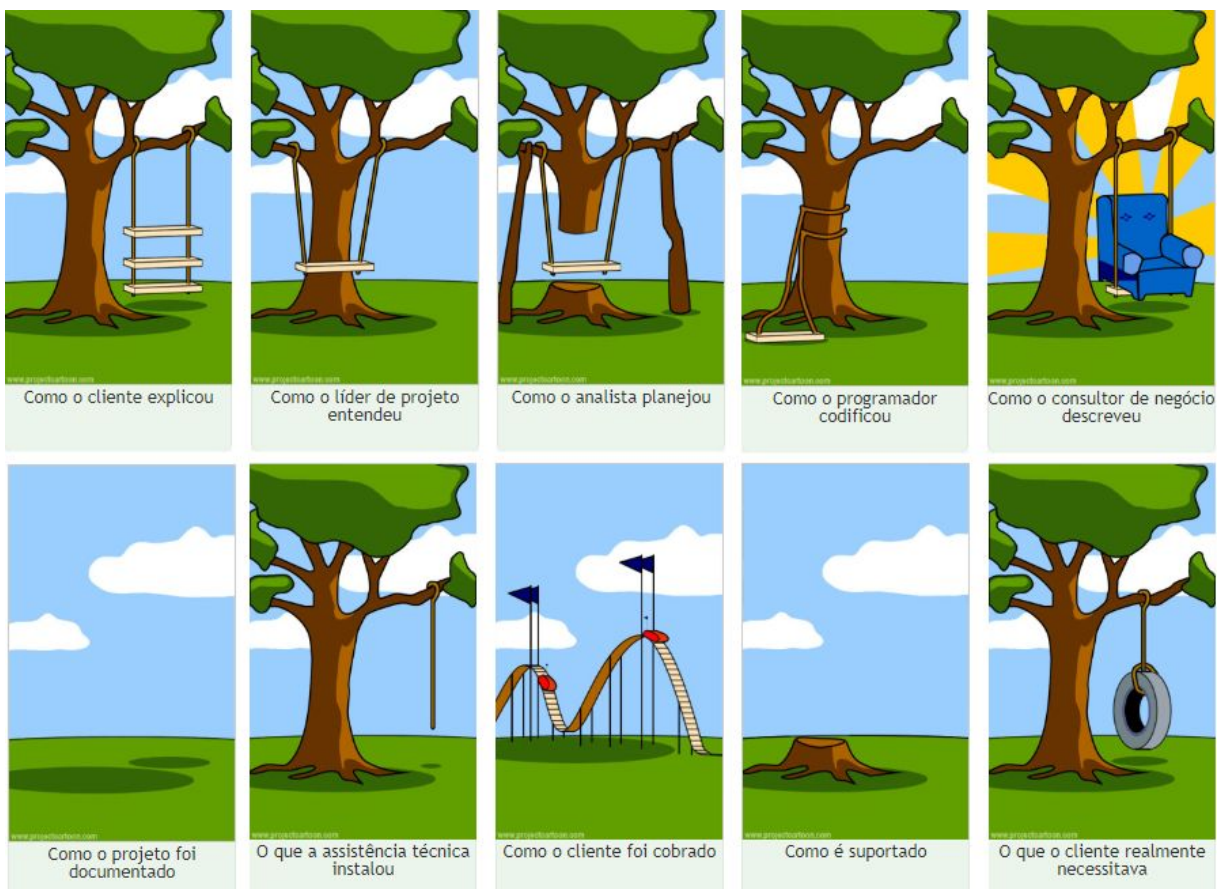
Para Cougo (1997), um modelo conceitual é caracterizado como um conjunto de objetos, suas propriedades e relações, de modo que este seja uma representação clara de um determinado contexto. Ele caracteriza-se por ser representado no mais alto nível de abstração, por exemplo, nele não é levado em conta qual será o Sistema Gerenciador de Banco de Dados



(SGBD) em si, mas a forma como as estruturas serão criadas para armazenar os dados independente do SGBD, ou seja, representar apenas os conceitos formais.

O modelo deve ser desenvolvido independente das linguagens de programação, tecnologias ou meios físicos utilizados. Além disso, deve facilitar o entendimento, compreensão e validação de um determinado domínio por todos os participantes envolvidos em determinado projeto, desde o usuário de um sistema, até seus analistas e programadores (The Project Cartoon, 2019). Um exemplo clássico da falta de entendimento entre todos os envolvidos é ilustrado pela Figura 2.2. Nela se torna evidente que a conceituação através de um modelo se torna cada vez mais importante.

Figura 2.2 - Como cada indivíduo pode visualizar um projeto



Fonte: The Project Cartoon, 2019.

Em um projeto de *Data Warehouse (DW)* onde é armazenado um grande volume de dados - que podem ser originados de diversas fontes de um determinado domínio - é essencial que o DW seja desenvolvido com base em um modelo conceitual capaz de fornecer os

conceitos abstraídos do domínio em específico. O modelo deve conter com precisão os requisitos dos usuários para aprimorar a comunicação entre os analistas, desenvolvedores, especialistas em domínio e os usuários.

Nesse sentido, cabe resgatar a reflexão de Antoni Olivé (2007), indicando que quando assumimos que um domínio consiste em objetos, relacionamentos e conceitos, nos comprometemos com uma maneira específica de ver os domínios. No campo dos sistemas de informação, esse compromisso de visualizar domínios de uma maneira particular é chamado de modelo conceitual. O mesmo modelo conceitual pode ser aplicado a muitos domínios diferentes, e vários modelos conceituais podem ser aplicados ao mesmo domínio, sendo que o conjunto de conceitos usados em um domínio particular constitui uma conceituação desse domínio.

O termo usado na ontologia para denotar esse compromisso é o compromisso ontológico. Note-se que o termo ontologia é usado para denotar duas idéias diferentes: um ramo da filosofia e uma especificação de uma conceituação. Na ciência da computação, o último é o significado usual. Desse modo, pode haver várias conceituações do mesmo domínio e, portanto, várias ontologias. Além disso, uma ontologia é uma visão concreta de um domínio específico, correspondendo também a um compromisso ontológico para as pessoas que observam e atuam em tal domínio (OLIVÉ, 2007).

Na prática, pode-se afirmar que uma ontologia é uma descrição de um vocabulário utilizado em um determinado domínio com relações hierárquicas de classes abstratas. Construídos como uma forma de apresentar aos usuários qual a proposta do sistema em desenvolvimento, esses modelos também tratam tanto das tarefas a serem realizadas pelo sistema, quanto às partes que o compõem e sua relação. Ambas representações são formas de se explicitar o conhecimento para uma determinada tarefa, decompor um problema através da abstração de conceitos úteis e sua formalização, que permite que eles sejam usados em raciocínios confiáveis e computacionais. O domínio do modelo é um aspecto importante, ambas representações são usadas para resolver problemas específicos, mas sistemas de ontologias computacionais devem considerar que tudo o que não definem é indefinido, enquanto modelos conceituais consideram essas pressuposições como falsas (MIZOGUCHI e IKEDA, 1997).

Modelos conceituais de sistemas de informação, portanto, fazem suposições mais complexas do que simplesmente considerar que um domínio consiste em objetos e

relacionamentos. Um modelo conceitual assume que um domínio inclui outras “coisas” e que objetos, relacionamentos e conceitos têm várias propriedades que devem ser distinguidas (OLIVÉ, 2007).

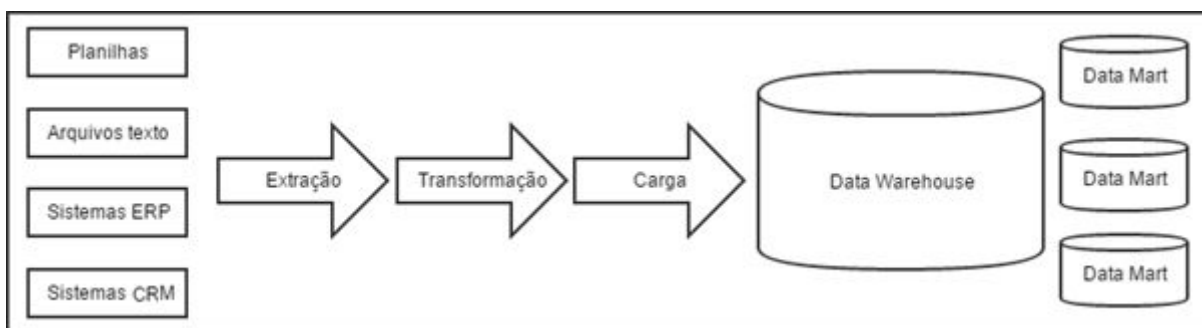
### **2.3. Processos de ETL, DW e Modelagem dimensional**

O objetivo dos processos de *Extract Transform Load* (ETL), conforme sua tradução “Extração, Transformação e Carga”, é fazer a integração de informações de diversas fontes de dados. Esses processos têm a capacidade de se conectarem com diversas fontes de dados, ler os dados, transformá-los e carregá-los em um *Data Warehouse* ou *Data Mart*. As informações carregadas podem ser combinadas e provenientes das mais variadas fontes de dados, como por exemplo: sistemas *Customer Relationship Management* (CRM), planilhas ou arquivos texto (MACHADO, 2000).

Para dar início a este processo, primeiramente é necessário definir as origens das fontes de dados e realizar a sua “Extração”. Nesta etapa, é importante ressaltar que, tendo em vista que os dados podem se originar de diferentes fontes, às vezes é necessário identificar os diferentes formatos encontrados. Por exemplo, um atributo o qual armazena o sexo de uma pessoa cujo o valor pode ser “M” ou “F” pode ser facilmente distorcido, o valor “M” na “Fonte de dados A” pode representar “Masculino”, já na “Fonte de dados B” o “M” pode representar “Mulher”. Dessa forma o próximo passo é realizar a “Transformação” e padronização dos dados para um formato uniforme, normalmente sugerido pelo próprio usuário para que possam ser “Carregados” no repositório de dados (MACHADO, 2000).

A carga dos dados consiste em fisicamente, seguindo um modelo dimensional, estruturar e carregar os dados para dentro do *Data Warehouse* (DW). De forma geral o processo pode ser entendido conforme demonstra a Figura 2.3.

Figura 2.3 - Processo de ETL



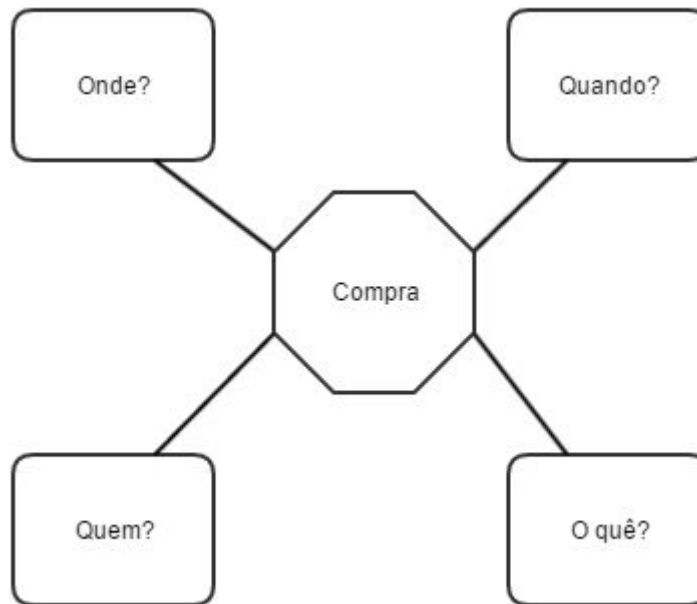
Fonte: Autor, 2018.

Através de processos de ETL é possível desenvolver um DW de um domínio específico e integrar diversas fontes de dados, resultando em um ambiente capaz de prover informações. O DW caracteriza-se como um grande repositório de dados, os quais ficam organizados e acessíveis com as informações necessárias para gestão de negócios. Seu principal objetivo é fornecer informações confiáveis e úteis para ajudar no processo de tomada de decisão. Além disso, seus dados podem ser originados de diversas fontes e sistemas existentes dentro de uma organização. Entretanto é importante ressaltar que esses dados são somente inseridos, não existindo atualização. Seu objetivo é prover dados para ferramentas voltadas para o acesso com diferentes níveis e formas de apresentação (MACHADO, 2000).

Para desenvolver um projeto de DW é fundamental realizar um levantamento e descobrir as necessidades da informação e indicadores que se deseja obter. Com base nesta análise, deve-se projetar, modelar e desenvolver o banco de dados que irá armazenar os dados em formato multidimensional, de modo que os dados possam se relacionar e serem representados como um cubo. Sendo assim é possível dividir este cubo e navegar em cada uma de suas dimensões ou eixos. Basicamente esta modelagem pode ser implementada através de dois modelos, *Star Schema* e *Snowflake Schema*. Em ambos os modelos, existe uma tabela dominante no centro, denominada de tabela fato com múltiplas junções a outras tabelas, denominadas de tabelas dimensões (KIMBALL; ROSS, 2013).

De acordo com Machado (2000), dimensões são os elementos e as entidades que participam de algum fato, conforme pode-se observar na Figura 2.4. Por exemplo, toda ocorrência de uma compra envolve no mínimo quatro elementos: quando foi realizada a compra, onde foi realizada a compra, quem realizou a compra e o que foi comprado.

Figura 2.4 - Fato compra e seus elementos



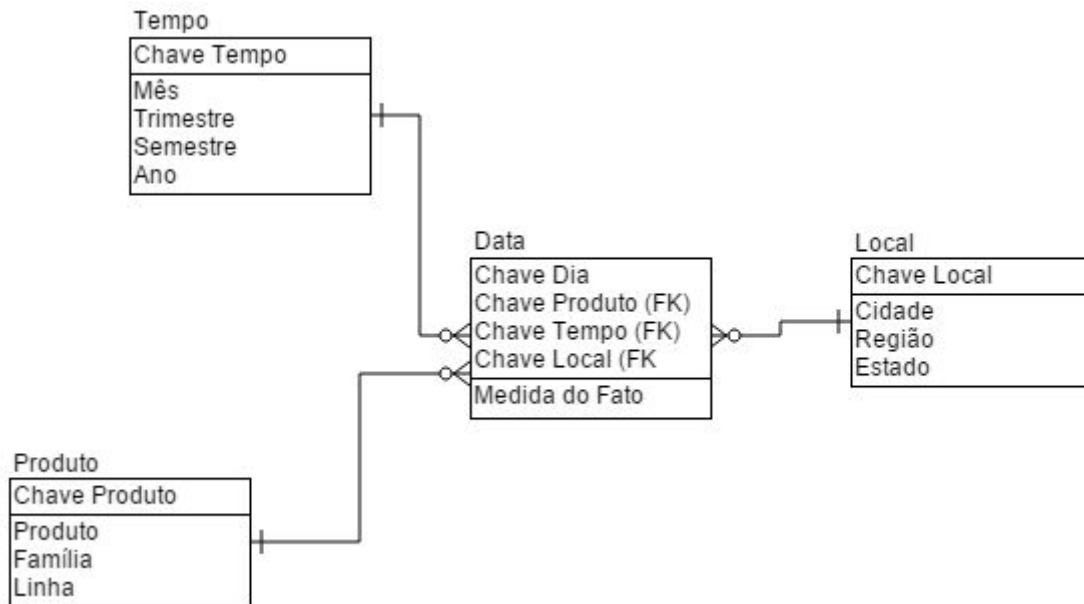
Fonte: Adaptado de Machado (2000).

Independente do modelo utilizado, para realizar esta modelagem existem dois tipos de tabelas, são elas:

- a) Tabela Fato: a tabela dominante, que fica no centro do modelo, é responsável por armazenar as medidas numéricas referentes à situação de negócio. Cada uma das medidas é obtida a partir da união de todas as dimensões. Ela também armazena as chaves de junções para as tabelas de dimensões (Machado, 2000);
- b) Tabela Dimensão: é conectada a tabela fato. Nela ficam armazenadas as descrições textuais das dimensões, fornecendo um maior nível de detalhamento para o negócio. Seu relacionamento é de um para muitos, o que possibilita um melhor aproveitamento do processamento das informações para as tabelas fato. Possuem diversos atributos de todos os tipos de dados, podendo armazenar as mais variadas informações (KIMBALL; ROSS, 2013).

O modelo estrela (*Star Schema*) é a estrutura básica de um modelo de dados multidimensional. Ele é composto por uma tabela central denominada fato e um conjunto de tabelas menores denominadas dimensões, as quais ficam dispostas ao redor desta tabela central formando uma estrela (MACHADO, 2000).

Figura 2.5 - Estrutura de um Modelo Estrela (*Star schema*)

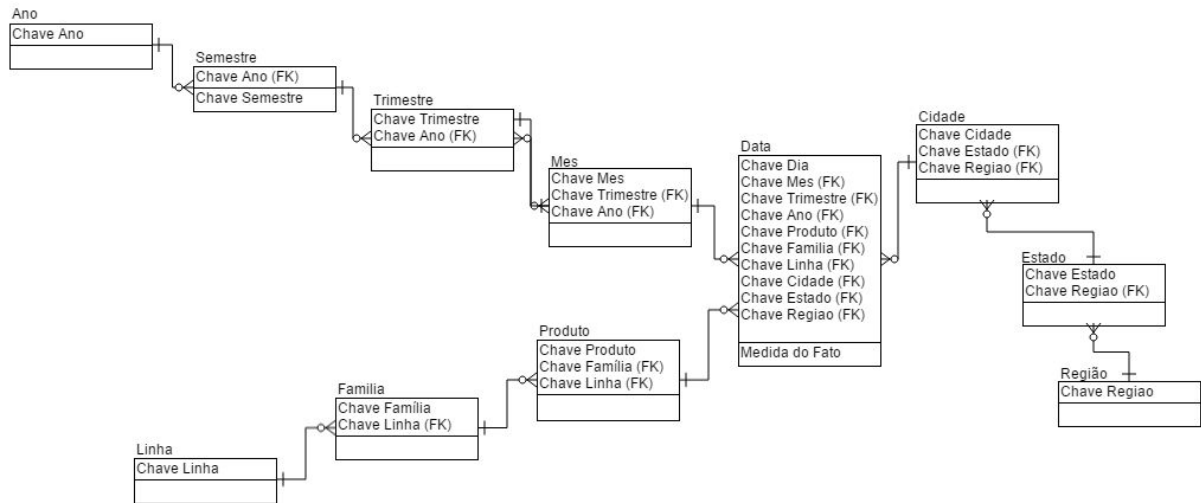


Fonte: Adaptado de Machado (2000).

Conforme pode-se observar na Figura 2.5, o centro da estrela é o fato e ao seu redor estão dispostas as dimensões que participam deste fato. O relacionamento entre a tabela central e as tabelas dimensões é uma simples ligação entre as entidades em um relacionamento de um para muitos.

Conforme Machado (2000), o modelo *Snowflake* (Floco de Neve) tem dimensões separadas em hierarquias, uma tabela para cada nível. A hierarquia é o resultado da decomposição de uma ou mais dimensões que possuem hierarquia entre seus membros, ou seja relacionamentos muitos para um entre os membros de uma dimensão, formando relacionamentos entre tabelas dimensões. Sua utilização preserva os meios de armazenamento, como ocorre em um modelo normalizado, desta forma, ele evita a redundância de valores textuais em uma tabela.

Figura 2.6 - Estrutura de um Modelo Floco de Neve (*Snowflake*)



Fonte: Adaptado de Machado (2000).

Na Figura 2.6 é possível notar que o relacionamento entre a tabela fato e as tabelas dimensões ocorre seguindo uma hierarquia, com várias tabelas de sub dimensões, onde cada tabela é responsável por um nível de detalhe.

Conforme pode-se observar nos dois modelos, *Star Schema* e *Snowflake*, ambos possuem uma tabela central, denominada fato e disposta ao seu redor as tabelas dimensões. É possível notar uma grande semelhança entre ambos, porém no modelo *Star Schema* não ocorre a normalização dos dados contidos nas tabelas dimensões ao contrário do que ocorre no modelo *Snowflake*, onde as tabelas fatos são normalizadas e possuem hierarquia entre suas tabelas.

### 3. TRABALHOS RELACIONADOS

Neste capítulo serão apresentados trabalhos relacionados ao tema desta dissertação. Primeiramente são revisados alguns trabalhos sobre modelos conceituais, visando investigar a relevância do desenvolvimento desta atividade previamente às demais etapas de elaboração do sistema proposto. Além disso, são apresentados ainda trabalhos e projetos desenvolvidos anteriormente para o OBSSAN, buscando compreender as lacunas, aprendizados e indicações dos autores para aprimorar o presente trabalho.

#### 3.1. Modelos conceituais

Bernasconi e colaboradores (2017) apresentam um modelo conceitual para a área da genômica, um ramo da bioquímica. Os autores justificam sua proposta em razão dos conjuntos de dados muitas vezes utilizados por biólogos para validar ou enriquecer seus experimentos não serem padronizados em suas fontes, de modo que se torna relevante uma proposta para o desenvolvimento de um modelo conceitual para a homogeneização desses dados. Primeiramente são identificados os atributos de metadados mais comuns de fontes genômicas e definidas suas propriedades semânticas. Após isso, através de um método *top-down* é criado um esquema integrado abstraindo as propriedades conceituais mais relevantes das fontes. Para validar a proposta, foram mapeadas três fontes de dados: TCGA, ENCODE e Gene Expression Omnibus.

Panos Vassiliadis e colaboradores (2002) abordam a questão de definição das atividades de ETL - responsáveis pela extração de dados de várias fontes, sua limpeza, personalização e inserção no DW -, fornecendo fundamentos formais para a representação conceitual. O trabalho contempla a customização de um modelo conceitual para o rastreamento de relações entre atributos e as respectivas atividades de ETL nos estágios iniciais do projeto de DW, bem como agrega uma 'paleta' de um conjunto de atividades de ETL usadas com frequência (p.ex.: atribuição de chaves substitutas, verificação de valores nulos, etc.) personalizável e extensível, permitindo que o projetista possa enriquecê-lo com seus próprios padrões recorrentes para atividades de ETL.

Charles E. Barr e colaboradores (1988) utilizam a modelagem conceitual como embasamento para uma rede semântica. Os autores argumentam que esta técnica pode auxiliar



na definição de um conjunto completo e consistente de objetos e relacionamentos para incluir na rede semântica. O trabalho busca facilitar o intercâmbio e a utilização de informações de múltiplas fontes em um Sistema Unificado de Linguagem Médica, proposto pela Biblioteca Nacional de Medicina, e as redes semânticas são utilizadas como esquema de representação de conhecimento em um sistema protótipo para explorar como realizar esses objetivos. Ambas abordagens *top-down* e *bottom-up* foram consideradas úteis em um processo de sete etapas de construção da rede semântica. Questões teóricas e práticas são discutidas, bem como futuras extensões do protótipo em vigor.

Tryfona e Jensen (1999) olharam para a questão dos aspectos espaciotemporais da informação, tendo em vista que as técnicas de banco de dados existentes, os idiomas e as ferramentas associadas forneciam pouco suporte interno para o gerenciamento das informações. Os autores desenvolveram um trabalho de aprimoramento de modelos conceituais já existentes agregando novas construções, demonstrando a relevância e os meios para acelerar o processo de modelagem de dados e facilitar a compreensão e manutenção dos diagramas, com base em fundamentos ontológicos. O trabalho apresenta um pequeno conjunto de construções novas e genéricas de modelagem que podem ser introduzidas em diferentes modelos de dados, utilizando uma semântica de ER espaço-temporal resultante (STER), com modelo ER subjacente. O STER foi acompanhado por uma contraparte textual e uma ferramenta CASE baseada no STER, usando a contraparte textual como sua representação interna.

Mizusaki (2016) se debruçou sobre a interação entre as teorias de aprendizagem e as diferentes ferramentas computacionais para educação. Analisando ontologias computacionais frente a metodologias de ensino cognitivistas, propõe um novo modelo computacional de avaliação denominado grafos de avaliação. O autor desenvolveu um modelo conceitual capaz de registrar os trabalhos dos alunos junto com as avaliações dos professores através da indicação de relações entre esses que evidenciam aprendizagens.

Em síntese, observa-se que independente da área de domínio - seja na genômica, processos de ETL em geral, na linguagem médica, nos dados espaciotemporais ou na educação -, os autores identificam a relevância da modelagem conceitual para a organização de dados de quaisquer temáticas. Diferentes ferramentas podem ser utilizadas para o desenvolvimento da ferramenta de visualização, porém, os resultados parecem ser influenciados positivamente pela realização prévia de um trabalho de modelagem conceitual.

Deste modo, podemos observar que para viabilizar a concepção e o desenvolvimento de um Data Warehouse (DW) multidimensional e extensível para monitoramento de Políticas Públicas a partir de dados abertos, conforme proposto neste trabalho, pode ser relevante o desenvolvimento de um Modelo Conceitual prévio.

### **3.2. Projetos para o OBSSAN**

As monografias de Rigo (2012) e Rigon (2013), intituladas de "Sistema de Apoio à Decisão com Apresentação Georreferenciada de Dados no âmbito do Projeto Agroflorestas" e "Sistema Web de Acompanhamento do Projeto Agroflorestas", respectivamente, apresentam o desenvolvimento de um sistema Web para o projeto intitulado "Fortalecimento das Agroflorestas no Rio Grande do Sul". O principal objetivo de ambos os trabalhos foi realizar o processamento analítico dos dados coletados pelos pesquisadores do OBSSAN-RS. Para armazenar estes dados foi modelado um *Data Warehouse*, capaz de gerar visualizações georreferenciadas utilizando as API do Google Maps. Em ambos os trabalhos os objetivos não foram atendidos em sua totalidade. Algumas melhorias e trabalhos futuros foram sugeridos pelos autores, como por exemplo: uma maior interação com os mapas; novos filtros e pesquisas mais dinâmicos; novas maneiras de visualização dos dados; e melhorias no módulo responsável pela carga de dados no sistema (e esse é um ponto abordado nesta dissertação).

Castro (2014) apresenta em sua monografia "Sistema de Apoio ao Projeto OBSSAN", o desenvolvimento de um sistema Web para o OBSSAN. O sistema auxilia na visualização e análise dos dados referentes à produção alimentar por município do Rio Grande do Sul. Os dados do sistema são visualizados através de buscas com a aplicação de filtros, as quais têm como principal forma de apresentação um mapa com o contorno dos municípios demarcados utilizando a API Javascript disponibilizada pelo Google Maps. Seus objetivos iniciais foram alcançados e o sistema teve uma boa aceitação pelo integrantes do OBSSAN, entretanto, a partir do seu uso, surgiram novas ideias e limitações no ambiente, como por exemplo: permitir exportar os análises em formatos planilha e pdf; novos filtros para as consultas; a necessidade uma interface administrativa para inserir, alterar e/ou excluir dados; e novos atributos e informações que somente foram identificadas ao final do trabalho.

Dentre os trabalhos citados notou-se que nenhum realizou um estudo mais aprofundado e amplo de todas as dimensões e indicadores do OBSSAN, além disso, surgiram

novas necessidades de filtros e visualizações mais dinâmicas. Tais questões surgiram após a finalização dos trabalhos citados, o que pode estar relacionado ao fato de que não foram documentadas previsões detalhadas sobre as necessidades do sistema. Nesse sentido, observa-se que os autores desenvolveram seus trabalhos com enfoque maior sobre a ferramenta de visualização - voltada aos usuários finais da plataforma - sem no entanto realizar um estudo sobre os dados das oito dimensões trabalhadas pelo Observatório. Devido aos objetivos dos referidos trabalhos ter outro enfoque, não encontra-se nos mesmos um aprofundamento da descrição de como foi realizado o desenvolvimento do banco de dados. Pode-se concluir, portanto, que a ausência do desenvolvimento de um Modelo Conceitual para o OBSSAN em trabalhos anteriores, resulta em um sistema que enfrenta limitações para sua ampliação, bem como para garantia de sua funcionalidade e atendimento dos objetivos a que se propõe no decorrer de todo o processo.

#### 4. MODELO CONCEITUAL DESENVOLVIDO

Neste trabalho é apresentada a elaboração de um modelo conceitual baseado em dados multidimensionais voltados para o monitoramento de Políticas Públicas de Segurança Alimentar e Nutricional. O seu objetivo é auxiliar os processos de *Extract Transform Load* (ETL) em dados abertos, tendo como resultado um *Data Warehouse* (DW) capaz de fornecer dados para sistemas de visualizações, tais como sistemas *On-Line Analytical Processing* (OLAP), geradores de gráficos e *dashboards*. Essa tarefa poderá também agregar maior consistência ao banco de dados, garantindo a confiabilidade dos dados que serão apresentados na ferramenta de visualização.

A primeira etapa para a concepção do modelo desenvolvido foi a definição de uma fonte e de um domínio de dados. Uma das instituições beneficiada pela proposta aqui apresentada é o Observatório Socioambiental em Segurança Alimentar e Nutricional do Rio Grande do Sul (OBSSAN-RS). Ele foi criado pelo Núcleo de Estudos e Pesquisas em Segurança Alimentar e Nutricional, vinculado ao Programa de Pós-Graduação em Desenvolvimento Rural (PGDR), em parceria com o Conselho Estadual de Segurança Alimentar e Nutricional (CONSEA-RS) e o Departamento de Informática da UFRGS (SIQUEIRA et al., 2015). O Observatório alinha-se ao propósito da Lei nº 11.346/2006 - Lei Orgânica de Segurança Alimentar e Nutricional (LOSAN), que criou o Sistema Nacional de Segurança Alimentar e Nutricional (SISAN) e estabeleceu as bases para a construção da Política Nacional de Segurança Alimentar e Nutricional (PNSAN) e do Plano Nacional de Segurança Alimentar e Nutricional (PLANSAN) (Brasil, 2006). Um de seus objetivos é coletar todos os dados necessários para compor e operacionalizar indicadores de sete dimensões, no estado do RS.

Buscando atender a necessidade do Observatório de desenvolver e disponibilizar um sistema *on-line* para os seus dados, coletados em diversos portais de dados abertos, foi estabelecida uma parceria entre o OBSSAN-RS e o autor. A partir dos dados referentes a Segurança Alimentar e Nutricional do Estado do Rio Grande do Sul coletados pelo OBSSAN, o autor em conjunto com os pesquisadores do projeto levantou os requisitos de dados a fim de construir o modelo conceitual. Esse modelo conceitual, para fins de validação, foi materializado como um banco de dados e um protótipo de uma aplicação Web.

#### 4.1 O modelo desenvolvido

A metodologia de desenvolvimento do Modelo Conceitual foi baseada a partir da proposta apresentada por W. Embley e Wai Yin Mok, intitulada de “Mapeando Modelos Conceituais para Esquemas de Banco de Dados” (EMBLEY, 2011, p.123). A publicação analisa os princípios e práticas de mapeamento algorítmico de modelos conceituais para esquemas de banco de dados, aplicando os princípios de mapeamento aos modelos de entidade-relacionamento estendidos, à linguagem de modelagem unificada e aos hipergrafos genéricos de modelo conceitual.

Realizou-se inicialmente o reconhecimento das informações trabalhadas pelo OBSSAN, identificando suas características e hierarquias entre os dados. A equipe do OBSSAN disponibiliza planilhas de dados, bem como esquemas de representação das relações entre as características dos dados, conforme ilustra a Figura 4.1.

Figura 4.1 - Representação das relações entre as características dos dados



Fonte: Siqueira, 2017.

Segundo o OBSSAN-RS, os itens da Figura 4.1 possuem a seguinte descrição:

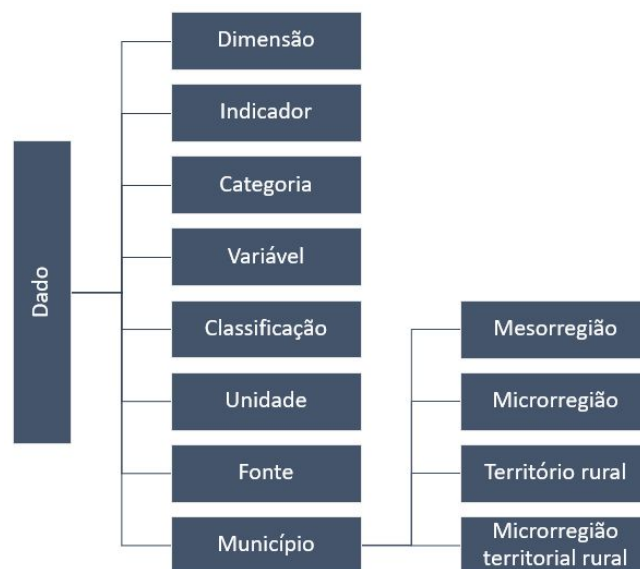
- a) **dimensão** - agrupamento de indicadores sociais que correspondem a uma mesma temática, sendo o primeiro nível na hierarquia dos dados;
- b) **indicador** - medidas utilizadas para traduzir um conceito abstrato de interesse social.
- c) **categoria** - agrupamento de dados de diferentes variáveis, com alguma característica em comum. É o segundo nível na hierarquia de agregação de dados do OBSSAN-RS;
- d) **variável** - agrupamento de dados com alguma característica em comum. É o terceiro nível na hierarquia de agregação de dados do OBSSAN-RS;
- e) **classificação** - Agrupamento de dados de diferentes unidades, com alguma característica em comum. É o último nível na hierarquia de agregação de dados do OBSSAN-RS;
- f) **dado** - informação quantitativa bruta. Suas características serão identificadas através da combinação de dimensões, indicadores, categorias, variáveis, classificações e agregações de municípios.
  - i) **quantidade** - é o dado em si, em formato decimal;
  - ii) **unidade** - identificação que dá singularidade a um dado específico e o diferencia dos demais. Expressam a informação básica necessária para o entendimento de um dado numérico;
  - iii) **ano** - refere-se ao ano em que o dado foi coletado - todas as fontes possuem uma periodicidade que é medida apenas em ano;
  - iv) **fonte** - identifica a origem do dado, ou seja, de onde ele foi extraído;
- g) **município** - município do estado do Rio Grande do Sul
  - i) **mesorregião IBGE** - agrupamento de municípios classificados de acordo com critérios geográficos e de identidade determinados pelo IBGE. Correspondem ao maior nível na hierarquia de agregação territorial das regiões do IBGE contempladas pelo OBSSAN-RS;
  - ii) **microrregião IBGE** - agrupamento de municípios classificados de acordo com critérios geográficos e de identidade determinados pelo IBGE. Correspondem ao menor nível na hierarquia de agregação territorial das regiões do IBGE contempladas pelo OBSSAN-RS;
  - iii) **território rural** - agrupamento de municípios classificados de acordo com critérios geográficos e de identidade determinados pelo Ministério

do Desenvolvimento Agrário (MDA) em 2014. Correspondem ao maior nível na hierarquia de agregação territorial das regiões do MDA contempladas pelo OBSSAN-RS;

- iv) **micro território rural** - agrupamento de municípios classificados de acordo com critérios geográficos e de identidade determinados por pesquisa realizada pelo Núcleo de Estudos e Pesquisas em Segurança Alimentar e Nutricional em 2016. Correspondem ao menor nível na hierarquia de agregação territorial das regiões do MDA contempladas pelo OBSSAN-RS.

Dada a descrição das características dos dados e suas relações apresentadas pelo OBSSAN, desenvolveu-se um organograma, representado pela Figura 4.2. Sua função é auxiliar os próximos passos para o desenvolvimento do modelo, como por exemplo o modelo ER.

Figura 4.2 - Organograma dos dados do OBSSAN



Fonte: Autor, 2018.

A partir da avaliação das descrições e características do conjunto de dados, foram então desenvolvidas tabelas para definir os atributos necessários à modelagem dos dados a serem inseridos no sistema. Para tal, cada um dos elementos centrais apresentados no organograma da Figura 4.2 originou uma tabela no SGBD com os atributos para armazenar as

informações necessárias, atribuindo também a descrição do tipo de dado armazenado, limite de caracteres, bem como as relações com as demais tabelas, resultando em quatro tipos de tabelas, quais sejam:

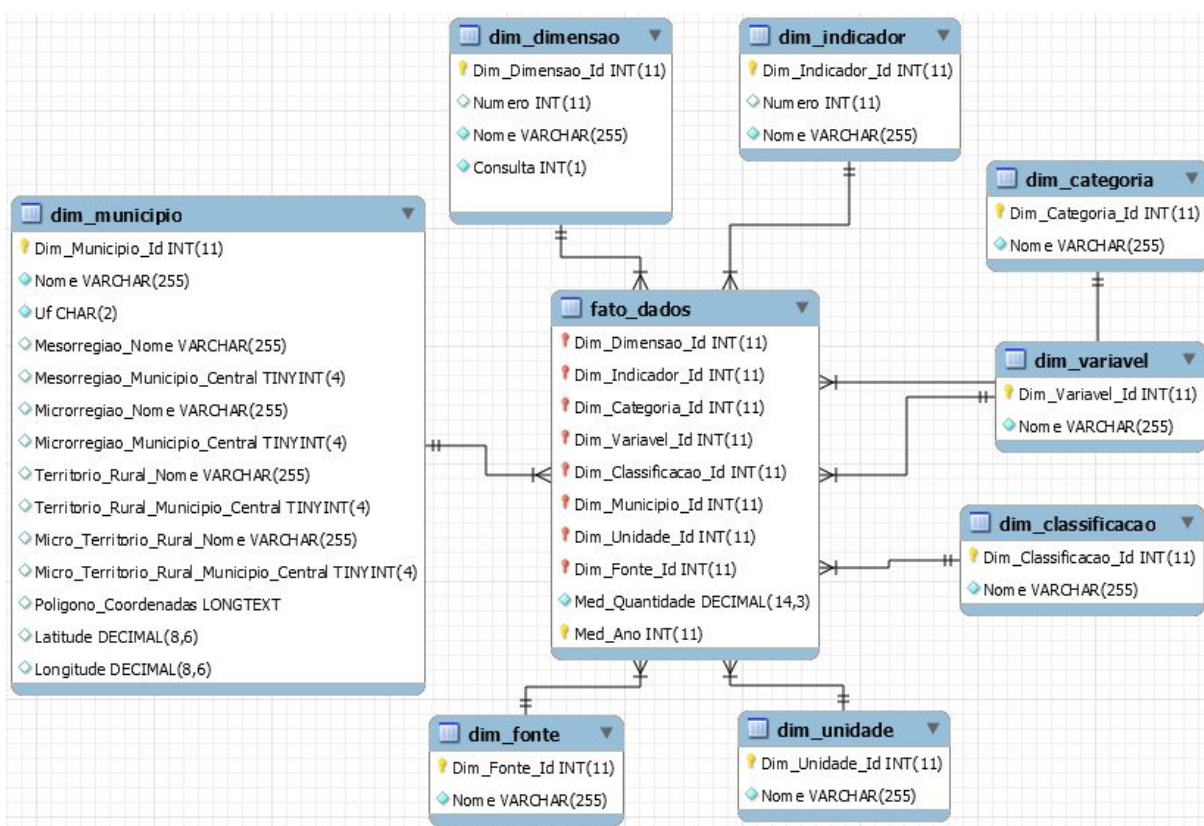
- a) Tipo A: armazena somente os “nomes” pertencentes a cada item. Este modelo foi aplicado nos itens “Categoria”, “Variável”, “Classificação”, “Subclassificação”, “Unidade” e “Fonte”;
- b) Tipo B: armazena os “nomes” e também o “número”, devido a necessidade de ordenamento dos dados armazenados nesta tabela. Este modelo foi aplicado no item “Indicador”;
- c) Tipo C: armazena os “nomes”, “número” e “consulta”, devido a necessidade de ocultar dados específicos das consultas padrão, visto que são utilizados apenas para geração de gráficos e mapas. Este modelo foi aplicado no item “Dimensão”;
- d) Tipo D: armazena os campos padrão de “nomes” e “UF”; bem como campos de localização contendo um conjunto de coordenadas geográficas para gerar o “Polígono” de cada município no mapa, e informações de “Latitude” e “Longitude” necessárias para exibir um ponto central em cada localização; e, por fim, são identificados também campos das agregações territoriais a que cada município pertence “Mesorregião”, “Microrregião”, “Território Rural”, “Microterritório Rural”, e “Região Funcional”, contendo os “nomes” de cada agregação, e o ponto do “município central” dos territórios para navegação nos mapas. Este modelo foi aplicado no item “Município”.

Dadas essas características, identificou-se que a melhor forma para armazenamento dos dados seria através de um *Data Warehouse*, o qual visando a performance foi modelado em formato estrela, representando os dados em formato multidimensional, permitindo apresentar as informações de uma forma visual, capaz de transformar seus dados brutos em informações perceptíveis, ou seja, transformar os dados em um conjunto visual que permita extrair informações de forma clara e objetiva para os seus usuários. O detalhamento das relações entre as tabelas é apresentado no ER (Figura 4.3) e também pelo dicionário de dados apresentado na Tabela 4.1 e Apêndice A.

No centro do modelo fica disposta a tabela fato, denominada aqui de “fato\_dados”. Ela é responsável por armazenar as medidas e também as chaves para as tabelas dispostas em sua volta.



Figura 4.3 - DW Obsan - Modelo Estrela



Fonte: Autor, 2018.

Tabela 4.1 - Dicionário de Dados: Entidades

Entidade	Descrição
fato_dados	Tabela fato do DW. Contêm os atributos: <ol style="list-style-type: none"> <li>“med_quantidade” - Armazena um dado bruto em formato decimal, suas características serão identificadas através da combinação das outras dimensões do DW;</li> <li>“ano” - Armazena o ano em que o dado foi coletado;</li> <li>“Dim_Dimensão_Id”, “Dim_Dim_Indicador_Id”, “Dim_Categoria_Id”, “Dim_Variavel_Id”, “Dim_Classificacao_Id”, “Dim_Municipio_Id”, “Dim_Unidade_Id” e “Dim_Fonte_Id” - Armazena as chaves para cada uma das dimensões apresentadas abaixo nesta tabela.</li> </ol>
dim_dimensao	Tabela dimensão. Contêm os atributos para armazenar as Dimensões do OBSSAN (agrupamento de indicadores sociais). Essa tabela armazena os campos: <ol style="list-style-type: none"> <li>“Dim_Dimensão_Id” - Chave primária;</li> <li>“Nome” - Armazena o nome da dimensão do OBSSAN;</li> </ol>

	<p>c) “Numero” - Armazena o número da dimensão do OBSSAN;</p> <p>d) “Consulta” - campo booleano, sua função é ocultar dados específicos das consultas padrão, visto que alguns dados são utilizados apenas para geração de gráficos e mapas. Valor padrão: mostrar nas consultas (true).</p>
dim_indicador	<p>Tabela dimensão. Armazena os Indicadores do OBSSAN. Contêm os atributos:</p> <p>a) “Dim_Indicador_Id” - Chave primária;</p> <p>b) “Nome” - Armazena o nome do indicador do OBSSAN;</p> <p>c) “Numero” - Armazena o número do indicador do OBSSAN.</p>
dim_categoria	<p>Tabela dimensão para armazenar as Categorias do OBSSAN. Contêm os atributos:</p> <p>a) “Dim_Categoria_Id” - Chave primária;</p> <p>b) “Nome” - Armazena o nome da categoria do OBSSAN;</p>
dim_variavel	<p>Tabela dimensão para armazenar as Variáveis do OBSSAN. Contêm os atributos:</p> <p>a) “Dim_Variavel_Id” - Chave primária;</p> <p>b) “Nome” - Armazena o nome da variável do OBSSAN;</p>
dim_classificacao	<p>Tabela dimensão para armazenar as classificações do OBSSAN. Contêm os atributos:</p> <p>a) “Dim_Classificação_Id” - Chave primária;</p> <p>b) “Nome” - Armazena o nome da classificação do OBSSAN.</p>
dim_unidade	<p>Tabela dimensão para armazenar as unidades de medida utilizadas pelo OBSSAN. Contêm os atributos:</p> <p>a) “Dim_Unidade_Id” - Chave primária;</p> <p>b) “Nome” - Armazena o nome da unidade de medida.</p>
dim_fonte	<p>Tabela dimensão para armazenar as fontes de dados. Contêm os atributos:</p> <p>a) “Dim_Fonte_Id” - Chave primária;</p> <p>b) “Nome” - Armazena o nome da fonte de dados.</p>
dim_municipio	<p>Tabela dimensão para armazenar as características necessárias para o OBSSAN dos municípios do RS. Contêm os campos:</p> <p>a) “Dim_Municipio_Id” - Chave primária;</p> <p>b) “Nome” - Armazena o nome do município;</p> <p>c) “Uf” - Armazena a unidade federativa do município;</p> <p>d) “Mesorregiao_Nome”, “Microrregiao_Nome”, “Territorio_Rural_Nome” e “Micro_Territorio_Rural_Nome” - campos com os nomes das agregações territoriais a que o município pertence;</p> <p>e) “Mesorregiao_Municipio_Central”, “Microrregiao_Municipio_Central”, “Territorio_Rural_Municipio_Central” e “Micro_Territorio_Rural_Municipio_Central” - campo booleano, para cada agregação é definido um município central, sua função é exibir um ponto central no polígono de cada uma das agregações no mapa;</p> <p>f) “Poligono_Coordenadas” - conjunto de coordenadas geográficas para gerar o “Polígono” do município nos mapas;</p> <p>g) “Latitude” e “Longitude” - coordenadas geográficas para exibir um ponto central dentro de cada município nos mapas.</p>

Fonte: Autor, 2018.

O modelo foi desenvolvido com a estrutura descrita visando que o mesmo seja extensível, ou seja, possa agregar novas Dimensões, Indicadores, Categorias e demais atributos determinados pelo OBSSAN. Essa tarefa poderá ser realizada livremente pelos usuários, a partir de uma planilha, desde que ela respeite e siga o modelo padrão estabelecido pelo autor (detalhamento na seção 5.1 do presente trabalho).

A partir do modelo de DW proposto é possível determinar como será realizada a inserção dos dados nas respectivas tabelas, bem como qual o Sistema de Gerenciamento de Banco de Dados (SGBD) para o armazenamento dos dados. Para tal, optou-se pelo uso do MySQL - um banco de dados relacional que utiliza a linguagem SQL (*Structured Query Language*) -, visto que é este o SGBD disponibilizado pelo CPD da UFRGS, onde o ambiente ficará disponível. Além disso, vale ressaltar que é *open source* e não possui nenhuma restrição de uso.

Para leitura da planilha de inserção de dados, foi desenvolvida uma *Stored Procedure* em MySQL, viabilizando a inclusão e atualização de dados, e garantindo a consistência das informações inseridas na tabela “fato\_dados” e suas tabelas dimensões. O script desenvolvido está detalhado no Apêndice B.

No presente capítulo foi apresentado o método de desenvolvimento do modelo conceitual para um sistema baseado em dados multidimensionais voltados para o monitoramento de Políticas Públicas. Observa-se que o modelo desenvolvido proporciona consistência ao banco de dados analisado neste trabalho, garantindo a confiabilidade das informações que serão apresentadas na ferramenta de visualização. Além disso, conforme descrito, o modelo conceitual da proposta é extensível, viabilizando a inserção de novos dados de maneira compatível com as características e relações dos dados apresentados pelo OBSSAN, conforme representado pela Figura 4.1 (página 29).

## **5. VALIDAÇÃO E AVALIAÇÃO DO MODELO**

Para a validação do modelo conceitual além da concepção do DW, foi desenvolvido um protótipo de um sistema Web para o OBSSAN, optou-se por não utilizar um sistema existente, pois foram estipulados diversos critérios visuais, funcionalidades e particularidades pelos membros do OBSSAN. Sendo assim, desenvolveu-se um ambiente administrativo (com acesso restrito) e um ambiente de visualização (com acesso público), os quais serão detalhados a seguir.

### **5.1. Gerenciamento dos dados no DW**

Devido aos usuários responsáveis pela coleta e gerenciamento dos dados estarem mais familiarizados com o Excel, foi desenvolvido um modelo de planilha para realizar a integração destes dados com o ambiente Web (Tabela 5.1).

Tabela 5.1 - Planilha de inserção e atualização de dados no DW

A	B	C	D	E	F	G	H	I	J
Dimensão	Indicador	Categoria	Variável	Unidade	Fonte	Classificação	Município	QTD.	Ano
Produção de Alimentos	Produção de alimentos (série anual)	Criação animal	Suíno	Cabeças	Pesquisa Pecuniária Municipal/ IBGE	Nenhuma	Taquari	16.984	2016
Renda e condições de vida	IDHM	Renda	IDHM Renda	Índice	Atlas do Desenvolvimento Humano/ PNUD/IP EA	Nenhuma	Taquari	0,74	2010
Acesso à alimentação adequada e saudável	Consumo alimentar	Adequação do peso ao nascer	Nenhuma	Crianças	Sistema de Informação de Atenção Básica/MS	Crianças nascidas com mais de 2500g	Taquari	27	2014
Saúde, nutrição e serviços relacionados	Mortalidade infantil	Menores de 1 ano de idade	Número de óbitos infantis	Óbitos	Sistema de Informações sobre Mortalidade/MS	Nenhuma	Taquari	3	2015
Educação	Taxa de analfabetismo	Todas	Pessoas não alfabetizadas	Pessoas	Censo Demográfico/IBGE	Todas	Taquari	1.328	2010
Políticas, programas e ações de SAN	Programa Bolsa Família	Benefícios concedidos	Famílias Beneficiárias do PBF	Famílias	Matriz de Informação Social/SAGI/MS	Nenhuma	Taquari	1.312	2017

Fonte: Autor, 2019.

No modelo de planilha criado, apresentado pela Tabela 5.1, são evidenciadas as colunas de “A” até “J”. Nas colunas de “A” até “H” constam as informações do dado, correspondendo a cadastros previamente inseridos manualmente através da interface Web do sistema (apresentada na Tabela 5.1). Atualmente, considerando todas as Dimensões do OBSSAN, enumeradas de 0 à 7, o volume de dados armazenados no sistema do ano 2000 até o ano de 2017 é de aproximadamente 2.496.554 registros, sendo que a Dimensão “Produção de Alimentos” representa 52% do total. Cabe ressaltar também, que caso o usuário altere qualquer termo textual, os dados são atualizados automaticamente.

Figura 5.1 - Área restrita - Cadastro de dimensões

Busca rápida <input type="text"/>				Cadastrar nova dimensão	Salvar selecionados	Excluir Selecionados
		Número * ▲	Nome *	Exibir na consulta *		
1	<input type="checkbox"/>	0	Dados demográficos	<input checked="" type="checkbox"/> Exibir na consulta		
2	<input type="checkbox"/>	1	Produção de Alimentos	<input checked="" type="checkbox"/> Exibir na consulta		
3	<input type="checkbox"/>	3	Renda e condições de vida	<input checked="" type="checkbox"/> Exibir na consulta		
4	<input type="checkbox"/>	4	Acesso à alimentação adequada e saudável	<input checked="" type="checkbox"/> Exibir na consulta		
5	<input type="checkbox"/>	5	Saúde, nutrição e serviços relacionados	<input checked="" type="checkbox"/> Exibir na consulta		
6	<input type="checkbox"/>	6	Educação	<input checked="" type="checkbox"/> Exibir na consulta		
7	<input type="checkbox"/>	7	Políticas, programas e ações de SAN	<input checked="" type="checkbox"/> Exibir na consulta		
8	<input type="checkbox"/>	901	Mapa inicial Vulnerabilidade	<input type="checkbox"/> Exibir na consulta		
9	<input type="checkbox"/>	902	Grafico inicial Vulnerabilidade	<input type="checkbox"/> Exibir na consulta		
10	<input type="checkbox"/>	904	Mapa inicial Políticas, programas e ações de SAN	<input type="checkbox"/> Exibir na consulta		
11	<input type="checkbox"/>	905	Mapa inicial popup Políticas, programas e ações de SAN	<input type="checkbox"/> Exibir na consulta		

\* Campo de preenchimento obrigatório

[1 a 11 de 11]

Fonte: Autor, 2019.

Deste modo, caso alguma das informações não conste no cadastro do sistema, será gerado um erro orientando o usuário administrador - que possui acesso restrito - a revisar a planilha e/ou cadastrar as respectivas informações no sistema previamente. Esta validação garante que não sejam registrados dados incorretos no Banco de Dados do sistema - como por exemplo aqueles relacionados a erros de digitação - que poderiam gerar inconsistências nas

informações apresentadas. Já nas colunas “I” e “J” são informados um valor para a quantidade e outro para o ano do dado. A combinação de todas as colunas gera um registro único, sendo que este registro pode ser alterado apenas através da importação da mesma combinação, somente com alteração do valor correspondente a coluna “I” (quantidade). Por outro lado, a alteração do conteúdo de qualquer uma das demais colunas irá gerar um novo registro.

De posse da planilha devidamente preenchida, e do cadastro das informações do dado realizado, os dados já podem ser importados para o sistema. Para isso os usuários devem entrar na área restrita do ambiente Web, e navegar até “Importação de dados” dentro do menu “Gerenciamento”. Na tela de importação (Figura 5.2) o usuário pode realizar o *upload* do arquivo, que deverá estar em formato de “Pasta de Trabalho do Microsoft Excel”, com as extensões em “XLS” ou “XLSX”, além disso, caso o usuário não possua um modelo de planilha em branco para preencher, nessa mesma tela o sistema disponibiliza um modelo em branco para *download*.

Figura 5.2 - Área restrita - Importação de dados

Click aqui para baixar um modelo em branco da planilha.'" data-bbox="214 490 808 688"/>

ENVIE A PLANILHA PARA REALIZAR A IMPORTAÇÃO

Arquivo

Selecionar Arquivo...

Arraste um arquivo e solte-o aqui

Enviar Cancelar

- O arquivo deverá estar em formato de "Pasta de Trabalho do Microsoft Excel", com as extensões em "XLS" ou "XLSX";  
- Os dados de todas as colunas, exceto "quantidade" e "ano" deverão estar previamente cadastrados;  
- [Click aqui](#) para baixar um modelo em branco da planilha.

Fonte: Autor, 2019.

Caso não exista incompatibilidade entre os dados da planilha com as informações dos dados previamente cadastrados no ambiente Web, o sistema insere ou atualiza os dados da tabela “dados”. Por questões de segurança, a exclusão de um ou mais dados inseridos é realizada através do sistema Web do OBSSAN, em uma tela independente criada especificamente para esta tarefa, conforme ilustrado na Figura 5.3.

Figura 5.3 - Área restrita - Cadastro de dimensões

SELECIONE OS DADOS QUE DESEJA EXCLUIR

Dimensão	1 - Produção de Alimentos
Indicador	8 - Produção de alimentos (série anual)
Categoria	
Variável	
Fonte	
Ano	2014

Excluir

A exclusão será realizada de acordo com a combinação dos campos preenchidos acima.

Fonte: Autor, 2019.

Nesta tela é possível realizar uma combinação de dados para exclusão, viabilizando a exclusão de todos os dados de um determinado ano; todos os dados de uma Dimensão; somente os dados de um determinado Indicador de uma Dimensão; ou também somente os dados de um determinado Indicador de uma Dimensão somente de um ano específico.

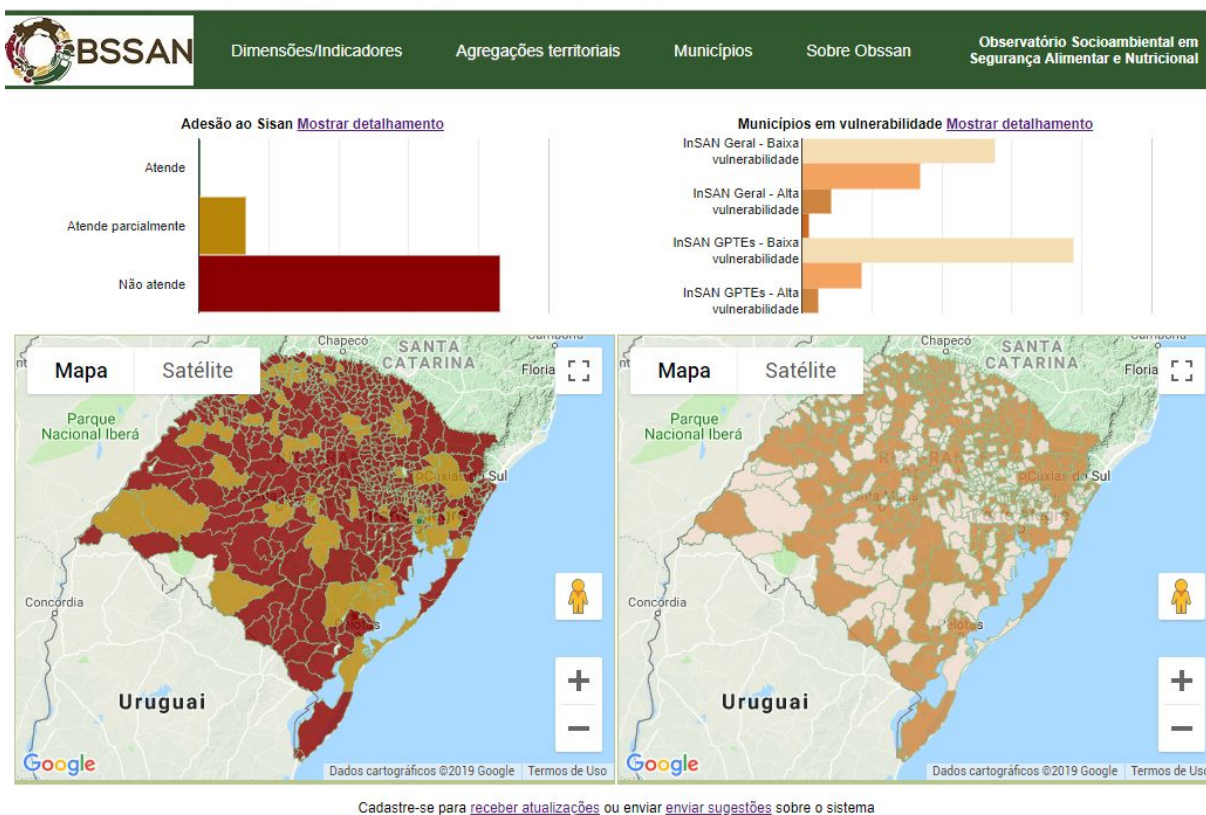
## 5.2. Ferramenta de visualização de dados

Nesta seção será apresentado o ambiente Web de visualização dos dados, desenvolvido como um protótipo neste trabalho. Para viabilizar a utilização do sistema pelo OBSSAN, uma aplicação desenvolvida a partir do modelo conceitual e o protótipo do ambiente Web aqui apresentados, será implementada futuramente através de uma parceria entre uma empresa responsável pelo desenvolvimento do ambiente, o OBSSAN e também a FEENG (Fundação Empresa Escola de Engenharia da Universidade Federal do Rio Grande do Sul).



Ao acessar o portal, nos deparamos com a tela inicial (Figura 5.4), onde são apresentados dois mapas e dois gráficos, demonstrando a situação do estado quanto a adesão ao SISAN e a vulnerabilidade em nível municipal; veremos essas representações em detalhes a seguir.

Figura 5.4 - Elementos da tela inicial do sistema



Fonte: Autor, 2018.

O primeiro mapa e gráfico (à esquerda) classificam os municípios segundo sua adesão ao SISAN (Sistema Nacional de Segurança Alimentar e Nutricional). Os municípios são classificados em: atendimento total, atendimento parcial e não atendimento aos critérios de adesão ao referido Sistema, conforme classificação informada pelo OBSSAN e sistematizada em formato numérico no Banco de Dados. Já o segundo bloco (à direita) refere-se à vulnerabilidade a insegurança alimentar e nutricional dos municípios do Rio Grande do Sul. Os municípios são classificados em quatro categorias para a população em geral (InSAN) e grupos tradicionais (GPTes): baixa vulnerabilidade, média vulnerabilidade, alta vulnerabilidade e muito alta vulnerabilidade.

O próximo recurso apresentado é a tela denominada “Dimensão/Indicadores” que pode ser acessada através do menu principal do ambiente. Neste ambiente é possível pesquisar pelos dados do OBSSAN segundo a dimensão de análise, visualizando as informações em forma de Planilha ou Cartograma. Conforme Figura 5.5, o usuário deverá parametrizar a visualização, selecionando uma dimensão, um indicador, uma categoria, uma unidade de medida, uma fonte, o ano e por fim o modo como deseja visualizar as informações.

Figura 5.5 - Filtros da Tela “Dimensões/Indicadores”

Dimensões/Indicadores x

SELECIONE OS DADOS QUE DESEJA VISUALIZAR

Acesse os manuais de [orientações gerais](#) e [criação de cartogramas](#) antes de realizar sua pesquisa

Dimensão \*

Indicador \*

Categoria \*

Unidade \*

Fonte \*

Ano \*

Visualização \*  Cartograma  Planilha

\* Campo de preenchimento obrigatório

Exibir

Fonte: Autor, 2019.

Por padrão o sistema sugere a forma de visualização em formato Planilha, e o resultado da parametrização abaixo é demonstrado pela Figura 5.6:

- a) Dimensão: Produção de Alimentos;
- b) Indicador: Produção de alimentos (série anual);
- c) Categoria: Criação animal;
- d) Unidade: Cabeças;
- e) Fonte: Pesquisa Pecuária Municipal/IBGE;
- f) Ano: 2016.

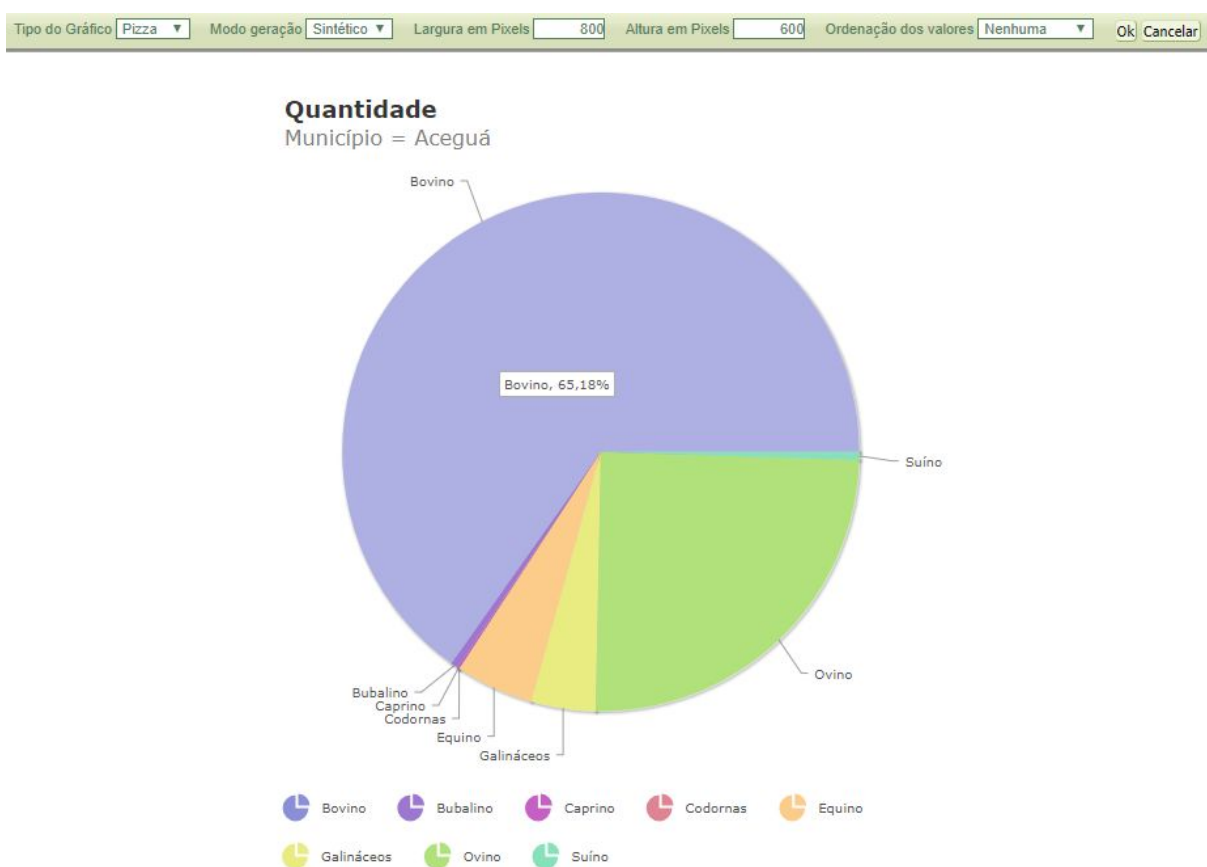
Figura 5.6 - Resultados em “modo planilha” da Tela “Dimensões/Indicadores”

Dimensões/Indicadores		
Dimensão: 1 - Produção de Alimentos		Categoria: Criação animal
Indicador: 8 - Produção de alimentos (série anual)		Fonte: Pesquisa Pecuniária Municipal/IBGE
		Unidade: Cabeças
		Ano: 2016
<span>Exportação</span> <span>Conf. Gráficos</span> <span>Configuração</span> <span>Quebras</span> <span>Voltar</span>		
Município	Variável	Quantidade
Aceguá	Bovino	101.267,00
	Bubalino	714,00
	Caprino	202,00
	Codornas	0,00
	Equino	7.581,00
	Galináceos	6.253,00
	Ovino	38.551,00
	Suíno	786,00
	<b>Total</b>	<b>155.354,00</b>
Agudo	Bovino	15.357,00
	Bubalino	3,00
	Caprino	256,00
	Codornas	100,00
	Equino	83,00
	Galináceos	50.684,00
	Ovino	658,00
	Suíno	8.466,00
	<b>Total</b>	<b>75.607,00</b>

Fonte: Autor, 2019.

A partir da Planilha exibida pelo sistema, o usuário ainda conta com a funcionalidade de geração de gráficos. Clicando nos ícones verdes ao lado das linhas de “Total” ou colunas da Planilha, o sistema gera gráficos que podem ser personalizados pelo usuário. Na Figura 5.7 é apresentado um exemplo de gráfico, na qual podem ser observadas as opções de personalização, contando com o tipo de gráfico, medidas e ordenação dos valores.

Figura 5.7 - Resultados em “modo gráfico” da Tela “Dimensões/Indicadores”



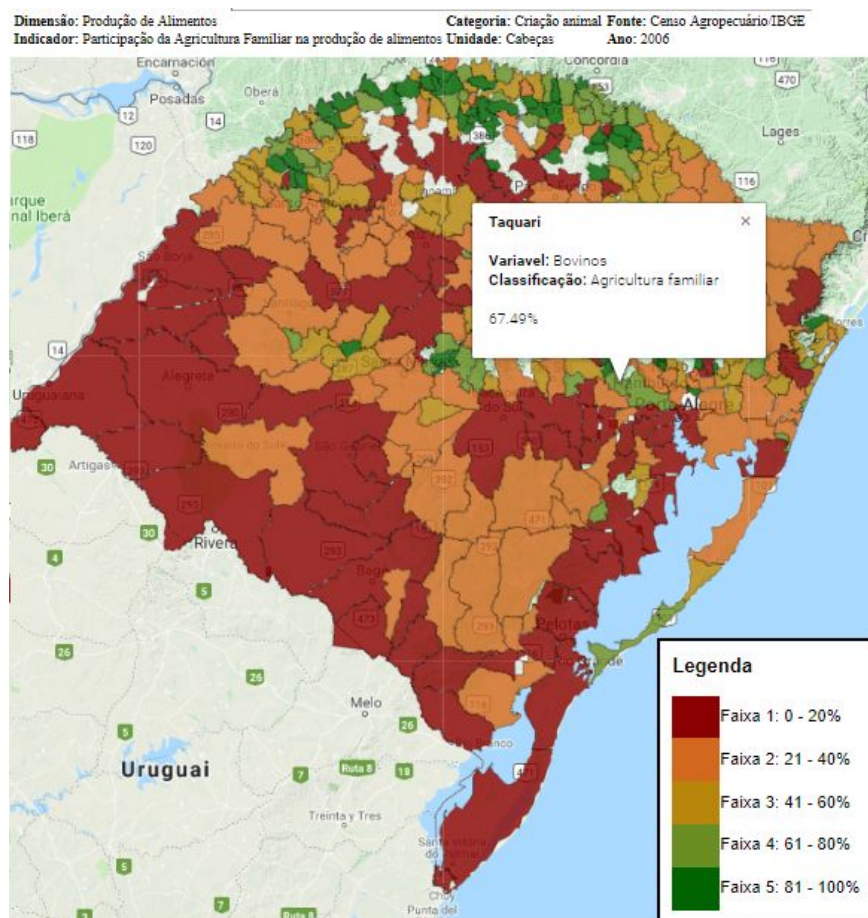
Fonte: Autor, 2019.

Caso o usuário optar pelo modo de visualização Cartograma, são abertos dois parâmetros adicionais que definem a totalização selecionada (simples ou agrupada) e a medida a ser representada (Percentual ou Quantidade). A totalização pode ser definida por itens isolados (apenas Variável, e apenas Classificação) ou pela combinação dos mesmos, sendo que o OBSSAN disponibiliza um Manual para o usuário orientando sobre as combinações indicadas para geração de cartogramas relevantes. O resultado da parametrização abaixo é ilustrado pela Figura 5.8:

- a) Dimensão: Produção de Alimentos;
- b) Indicador: Participação da Agricultura Familiar na Produção de Alimentos;

- c) Categoria: Criação animal;
- d) Unidade: Cabeças;
- e) Fonte: Censo Agropecuário/IBGE;
- f) Ano: 2006;
- g) Variável+Classificação: Bovinos (variável) e Agricultura Familiar (Classificação).
- h) Medida: Percentual.

Figura 5.8 - Resultado em formato de cartograma da Tela “Dimensões/Indicadores”



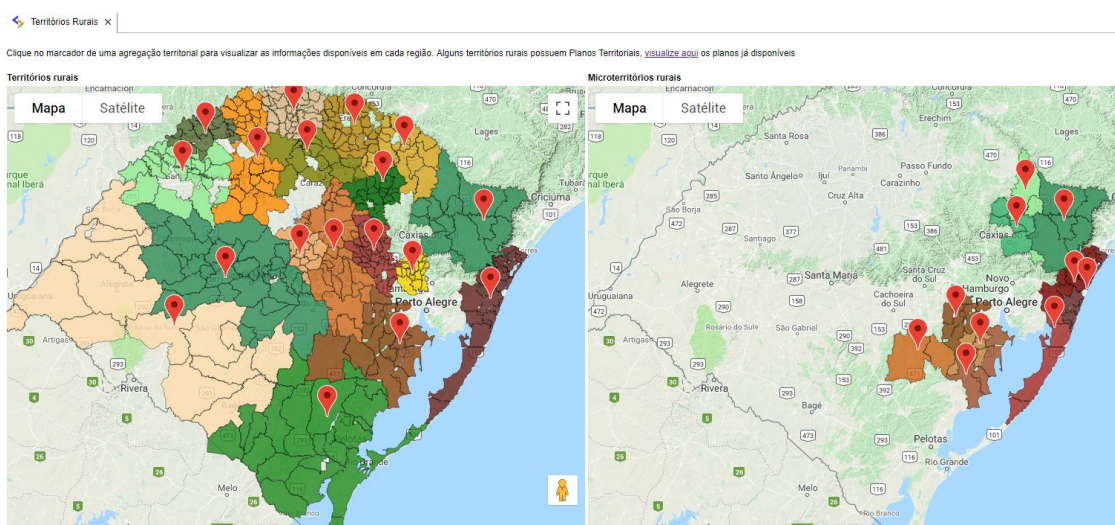
Fonte: Autor, 2019.

O cartograma gerado classifica cada município através de uma escala de cores descrita na legenda. Por padrão, o sistema exibe uma escala de cinco (05) faixas de percentuais, com cores que remetem a um semáforo, representando visualmente a situação de cada município em relação ao parâmetro pesquisado. No exemplo apresentado na Figura 5.8 observa-se que os municípios com baixos percentuais de participação da Agricultura Familiar na criação de

bovinos aparecem em tons de vermelho/laranja, enquanto aqueles com percentuais maiores são visualizados em tons amarelos/verdes. Nesta funcionalidade o usuário pode ainda clicar sobre um município e visualizar o detalhamento do dado pesquisado.

A segunda opção do menu principal, disposto na parte superior do ambiente, o usuário pode acessar a tela “Agregações territoriais”, disponibilizando duas opções de agregações, quais sejam: “Regionais IBGE” e “Territórios rurais”. Após a opção por uma das agregações, conforme apresenta a Figura 5.9, são exibidos dois cartogramas representando as subdivisões da agregação selecionada. A opção Regionais IBGE exibe as subdivisões “Mesorregiões” e “Microrregiões”; já a outra opção apresenta subdivisões em “Territórios rurais” e “Microterritórios rurais”.

Figura 5.9 - Cartogramas da tela “Agregações territoriais”



Fonte: Autor, 2019.

Para cada agregação exibida nos cartogramas, o ambiente permite que o usuário clique em um marcador, exibindo mais detalhes sobre o conjunto dos municípios em uma nova tela. Após selecionar alguns filtros semelhantes aos apresentados na tela “Dimensão/Indicadores” (detalhados anteriormente na Figura 5.5), o sistema exibe uma planilha com os resultados conforme demonstra a Figura 5.10.

Figura 5.10 - Detalhamento das “Agregações territoriais”

Territórios Rurais -> Microterritórios rurais -> Campos  
 Dimensão: 7 - Políticas, programas e ações de SAN Categoria: Benefícios concedidos Fonte: Matriz de Informação Social/SAGI/MDS  
 Indicador: 3 - Programa Bolsa Família Unidade: Famílias Ano: 2017

Exportação Conf. Gráficos Configuração Quebras Voltar

Município	Famílias Beneficiárias do Programa Bolsa Família	
	Quantidade	Total
Bom Jesus	1.070	1.070
Cambará do Sul	257	257
Jaquirana	295	295
São Francisco de Paula	1.035	1.035
São José dos Ausentes	224	224
<b>Total Geral</b>	<b>2.881</b>	<b>2.881</b>

Caso observar algum erro nos dados ou desejar sugerir outra opção de visualização, [preencha este formulário](#) para contatar a equipe do OBSSAN.

Fonte: Autor, 2019.

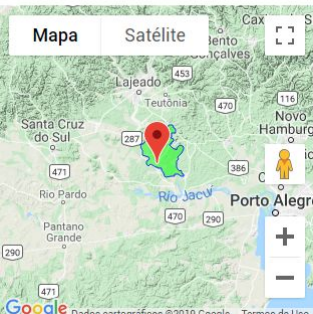
A última tela disponível no ambiente é denominada “Municípios”. A primeira visualização que o usuário observa ao clicar sobre o menu principal apresenta uma listagem simples para seleção de um município. Após esta opção, o sistema exibe uma tela (ilustrada pela Figura 5.11) com informações sobre o município selecionado, um mapa para navegação e ainda alguns links para aprofundar sua pesquisa e visualizar dados em diferentes formatos.

Figura 5.11 - Tela “Municípios”

Taquari - RS

- Relatório municipal (Indicadores OBSSAN/RS)
- Acessar dados atualizados de Políticas Públicas
- Plano de SAN municipal
- Tabela de dados do município

NAVIGUE ABAIXO ATRAVÉS DO STREET VIEW



Faça uma busca rápida utilizando uma PALAVRA-CHAVE, ou explore as DIMENSÕES abaixo para visualizar os dados do município

Busca rápida

- Dimensão : 1 - Produção de Alimentos
- Dimensão : 3 - Renda e condições de vida
- Dimensão : 4 - Acesso à alimentação adequada e saudável
- Dimensão : 5 - Saúde, nutrição e serviços relacionados
- Dimensão : 6 - Educação
- Dimensão : 7 - Políticas, programas e ações de SAN
- Dimensão : 0 - Dados demográficos

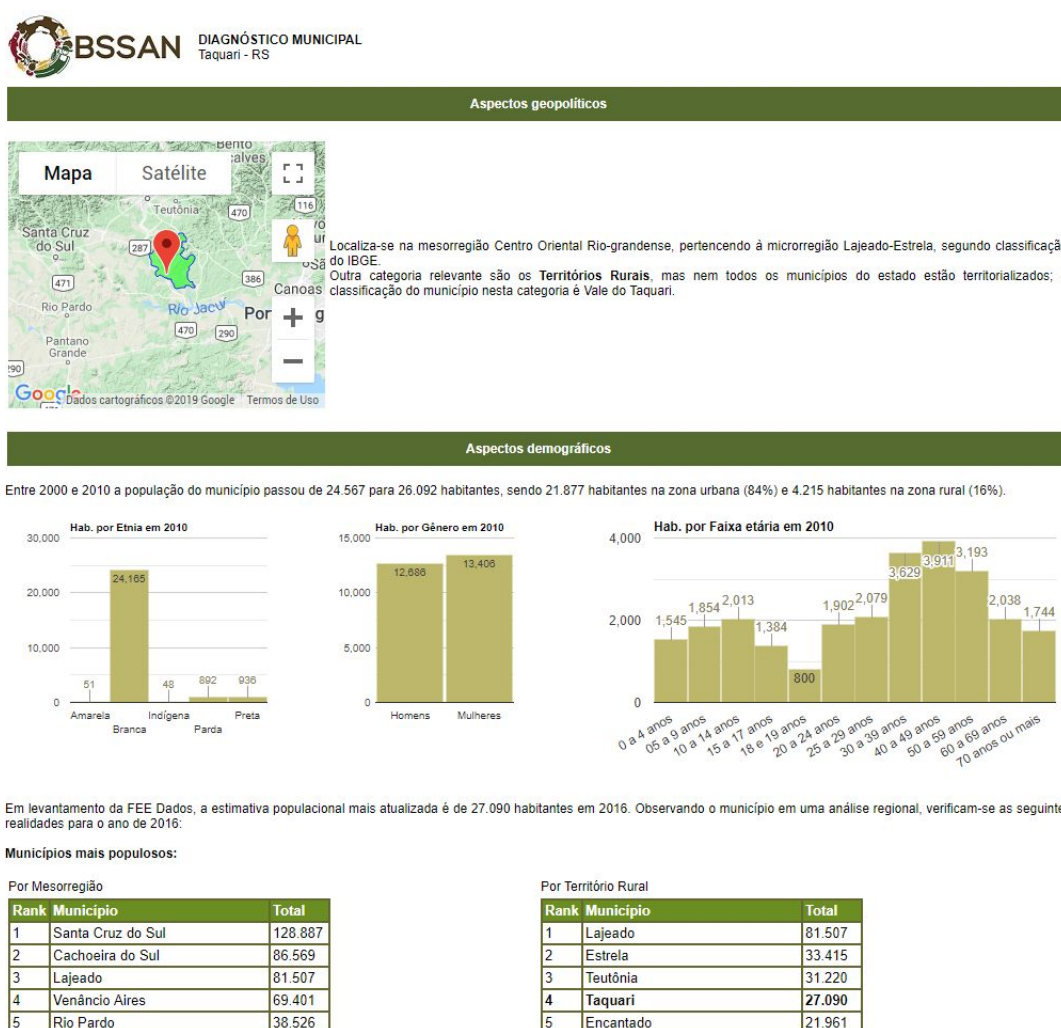
[1 a 860 de 860]

Fonte: Autor, 2019.

Entre as opções de navegação, estão a visualização de dados municipais em formato de Relatório (disponível para download em PDF) e também em Tabela (disponível para

download em XLS ou CSV). No bloco superior esquerdo da tela ainda são disponibilizados links que direcionam para ambientes externos, nos quais o usuário pode acessar dados atualizados de Políticas Públicas do município (gerados por um sistema independente, de autoria do Governo Federal), e também o Plano de SAN municipal, quando disponível (organizados em uma nuvem gerenciada pelo próprio OBSSAN). Dentre essas opções, cabe ressaltar a função “Relatório municipal”, a qual apresenta uma síntese que contempla as diferentes opções de visualização de dados do sistema. O resultado é apresentado parcialmente na Figura 5.12, e contempla um relatório detalhado com dados de todos os Indicadores do OBSSAN, visualizados em formato de textos, gráficos, tabelas e cartogramas. Esta visualização conta com ferramentas gratuitas para a geração dos resultados, quais sejam: HTML; PHP; MySQL; JavaScript; Google Maps; e Google Charts.

Figura 5.12 - Tela “Municípios” - Relatório municipal



Fonte: Autor, 2019.



Ainda no bloco esquerdo da tela “Municípios” (vide Figura 5.2), o sistema exibe na parte inferior uma visualização em formato de mapa, o qual é proveniente das API do Google Maps e possibilita a navegação pelo município através de seu recurso “Street View”. Por fim, no bloco direito da tela observa-se um menu similar ao disponibilizado nas demais telas de pesquisa de Dimensões e Indicadores, oportunizando que o usuário busque por dados específicos sobre o município selecionado.

### **5.3. Avaliação da ferramenta**

Dada a relevância de desenvolvimento de uma ferramenta intuitiva e prática para uso pelos diferentes públicos atendidos pelo OBSSAN, bem como uma análise apurada sobre o trabalho desenvolvido, aplicou-se um método de averiguação do nível de usabilidade do sistema. A mensuração da efetividade e usabilidade do produto foi realizada a partir da escala SUS (System Usability Scale), testando o sistema com membros da equipe do OBSSAN.

Segundo Brooke ([s.d.]), usabilidade não existe em nenhum sentido absoluto; só pode ser definido com referência a contextos específicos, de modo que surge a necessidade de medidas gerais que possam ser usadas para comparar a usabilidade em vários contextos. A Escala SUS apresenta-se como uma escala de usabilidade confiável e de baixo custo que pode ser usada para avaliações globais de usabilidade de sistemas (BROOKE, [s.d]).

A escolha pelo SUS se deve ao fato dele apresentar um balanço interessante entre ser cientificamente apurado e ao mesmo tempo não ser extremamente longo para o usuário nem para o pesquisador. O método foi criado por John Brooke em 1986, e pode ser usado para avaliar produtos, serviços, hardware, software, websites, aplicações , e qualquer outro tipo de interface, avaliando critérios de efetividade (os usuários conseguem completar seus objetivos?), eficiência (quanto esforço e recursos são necessários para isso?) e satisfação (a experiência foi satisfatória?).

Para a avaliação do modelo desenvolvido, foi disponibilizado um formulário, onde, inicialmente os participante responderam dados referente a sua identificação, tais quais:

- Idade;
- Sexo: Feminino; Masculino; ou Prefiro não informar;

- Perfil: Membro da equipe do OBSSAN e/ou vinculado ao PGDR; Profissional, acadêmico ou professor de TI (analista, desenvolvedor de software ou banco de dados); ou Público geral;
- Nível de conhecimento com computadores: escala de 1 a 10;
- Conhecimento com os dados do sistema: escala de 1 a 10;
- Conhecimento com o sistema: escala de 1 a 10.

Em seguida, para testar a usabilidade qualitativa pelo usuário, foram determinadas quatro (04) tarefas, sendo cada uma para uma tela do sistema. São elas:

- Tarefa 1: Na tela inicial do sistema identifique a quantidade de municípios no RS que não atendem aos critérios de adesão ao SISAN.
- Tarefa 2: Na tela Dimensões/Indicadores faça uma pesquisa buscando descobrir qual foi a participação da Agricultura Familiar na criação de aves de algum município do Rio Grande do Sul, em 2006.
- Tarefa 3: Através da navegação na tela de Agregações Territoriais, pesquise o valor de recursos repassados para a Política Nacional de Alimentação Escolar em 2016 no Território Rural Litoral.
- Tarefa 4: Na tela Municípios realize o download da Tabela de dados para acessar todas as informações disponíveis de um município em formato editável (XLS).

Por fim, após a realização das tarefas o participante é então convidado a responder um questionário simples de 10 itens, com opções de resposta organizadas em uma escala de 1 a 5 - onde 1 significa Discordo Completamente e 5 significa Concordo Completamente.  
Questionário:

1. Eu acho que gostaria de usar esse sistema com frequência.
2. Eu acho o sistema desnecessariamente complexo.
3. Eu achei o sistema fácil de usar.
4. Eu acho que precisaria de ajuda de uma pessoa com conhecimentos técnicos para usar o sistema.
5. Eu acho que as várias funções do sistema estão muito bem integradas.
6. Eu acho que o sistema apresenta muita inconsistência.
7. Eu imagino que as pessoas aprenderão como usar esse sistema rapidamente.
8. Eu achei o sistema atrapalhado de usar.

9. Eu me senti confiante ao usar o sistema.
  10. Eu precisei aprender várias coisas novas antes de conseguir usar o sistema.
- Texto livre para críticas e/ou sugestões.

Devido a quantidade limitada de usuários ativos na equipe do OBSSAN, a avaliação foi aplicada com 7 (sete) usuários do sistema. Após todos terem realizado as tarefas e respondido ao questionário as respostas foram convertidas em uma pontuação, visando obter a representação quantitativa da percepção do usuário sobre a usabilidade do sistema.

O método pressupõe uma média de 68 pontos, sendo que resultados abaixo deste valor representam problemas sérios de usabilidade no produto. Para essa avaliação, o perfil dos participantes estão representados na Tabela 5.2, enquanto os resultados das questões e cálculo da média de SUS Score podem ser encontrados na Tabela 5.3.

O cálculo é realizado a partir das respostas dos usuários no questionário. Para as respostas ímpares (1, 3, 5, 7, 9), é subtraído “1” da pontuação que o usuário atribuiu à resposta, já para as respostas pares (2, 4, 6, 8 10), é diminuído a pontuação que o usuário atribuiu de 5(5-x). Por fim, para obter o SUS Score, é somado o valor de cada uma das perguntas e multiplicado por 2,5.

Tabela 5.2 - Perfil dos participantes da avaliação de usabilidade

<b>Participante</b>	<b>Idade</b>	<b>Sexo</b>	<b>Nível de conhecimento com computadores</b>	<b>Conhecimento com os dados do sistema</b>	<b>Conhecimento com o sistema</b>
p1	42	Masculino	8	9	9
p2	29	Feminino	8	10	9
p3	35	Feminino	7	9	8
p4	30	Feminino	5	3	5
p5	41	Feminino	6	5	5
p6	23	Masculino	7	8	8
p7	46	Masculino	7	8	7

Fonte: Autor, 2019.

Tabela 5.3 - Resultados da avaliação de usabilidade

<b>Participante</b>	<b>q1</b>	<b>q2</b>	<b>q3</b>	<b>q4</b>	<b>q5</b>	<b>q6</b>	<b>q7</b>	<b>q8</b>	<b>q9</b>	<b>q10</b>	<b>SUS Score</b>
p1	5	1	4	1	5	1	4	2	4	2	87,5
p2	5	2	4	1	4	1	5	2	5	1	90,0
p3	4	2	4	1	4	2	4	2	4	1	80,0
p4	4	2	3	2	3	2	3	2	3	1	67,5
p5	2	3	2	4	3	1	2	3	2	1	47,5
p6	4	2	3	2	4	1	5	2	5	1	82,5
p7	5	2	3	2	4	1	2	2	2	2	67,5
<b>Média do SUS Score</b>											<b>74,6</b>

Fonte: Autor, 2019.

Observa-se que a média do *SUS Score* obtida a partir da avaliação de usabilidade aplicada foi de 74,6 pontos, portanto, 5,6 pontos acima da pontuação mínima esperada. Cabe notar que os usuários que atingiram menores escores informaram também níveis mais baixos de conhecimento com computadores e com os dados do sistema.

## 6. CONCLUSÃO

O presente trabalho se propôs a desenvolver e validar um modelo conceitual para o Observatório Socioambiental em Segurança Alimentar e Nutricional (OBSSAN) do estado Rio Grande do Sul (RS), concretizando um *Data Warehouse* (DW) multidimensional e extensível que pudesse viabilizar, por sua vez, uma ferramenta de análise de dados em um ambiente Web. Através da revisão da literatura, bem como de trabalhos relacionados a temática abordada, observou-se que a ausência de um Modelo Conceitual pode resultar em um sistema com limitações significativas para sua ampliação, e, também, para a garantia de funcionalidade e atendimento dos objetivos e requisitos a que se propõe no decorrer de todo o processo.

A partir do reconhecimento das informações trabalhadas pela instituição atendida, identificando as características e hierarquias entre os dados, foram definidos os atributos necessários à modelagem dos dados a serem inseridos no sistema. No decorrer do trabalho, identificou-se que a melhor forma para armazenamento dos dados seria através de um *Data Warehouse* (DW) modelado em formato estrela. Representando os dados em formato multidimensional, este modelo proporciona que sistemas apresentem as informações de uma forma visual, capaz de transformar seus dados brutos em informações perceptíveis; ou seja, transformar os dados em um conjunto visual que permita extrair informações de forma clara e objetiva para os seus usuários. Para a escolha do Sistema de Gerenciamento de Banco de Dados (SGBD) foram consideradas opções open source e sem restrições de uso, definindo-se pelo desenvolvimento de uma Stored Procedure em MySQL, o qual é disponibilizado pelo CPD da UFRGS, onde o ambiente ficará disponível.

Observou-se que o modelo desenvolvido proporciona consistência ao banco de dados analisado neste trabalho, garantindo a confiabilidade das informações que serão apresentadas na ferramenta de visualização. Além disso, o modelo desenvolvido proporcionou o desenvolvimento de um DW multidimensional e extensível, viabilizando a inserção de novos dados de maneira compatível com as características e relações dos dados apresentados pelo OBSSAN, e, portanto, resolvendo uma fragilidade importante observada em trabalhos anteriores. Cabe notar, ainda, que os resultados observados reiteram as conclusões de autores da temática de Modelagem Conceitual, indicando a relevância de garantir a etapa de elaboração do Modelo Conceitual previamente ao desenvolvimento de um DW e de uma

ferramenta de visualização, oportunizando um resultado mais eficiente, confiável e adequado às necessidades do usuário final da aplicação.

A concretização do DW em um ambiente Web, por sua vez, viabilizou a utilização do sistema pelo OBSSAN, assim como por seus respectivos usuários finais. Nesse sentido, ressalta-se que quando tratamos de portais que disponibilizam conjuntos de dados com um número expressivo de registros, a simples apresentação das informações em uma tabela torna-se insuficiente para subsidiar as análises que os usuários demandam. Desta forma, o desenvolvimento de ferramentas eficientes de visualização de informações pode trazer ganhos relevantes para os usuários de sistemas de dados abertos relacionados à Políticas Públicas no país, contribuindo para aprimorar a gestão pública e fortalecer o controle social.

Os dados disponibilizados pelo OBSSAN oferecem um grande potencial para extrair informações que podem ser utilizadas pelos mais diversos órgãos públicos e também de toda a população em geral. Através do Modelo Conceitual apresentado, o presente trabalho proporcionou a implementação de um *Data Warehouse* (DW), bem como um protótipo de ambiente Web sólido para o desenvolvimento de uma aplicação eficiente voltada ao monitoramento de políticas públicas alimentada pelos dados secundários do OBSSAN. A pesquisa de usabilidade demonstrou a adequação do protótipo desenvolvido para os usuários finais aos quais este trabalho foi direcionado. Todavia, observou-se que para a utilização do sistema por públicos com pouca habilidade em informática e conhecimento sobre a temática, são necessárias adaptações da ferramenta e/ou disponibilização de capacitações prévias.

Para trabalhos futuros, percebe-se a importância de realizar mais estudos visando aprimorar a ferramenta de visualização, proporcionando métodos de pesquisa mais dinâmicos, permitindo que o usuário selecione outras categorias, variáveis, classificações, indicadores e/ou dimensões para ampliar as possibilidades de uso nas diferentes questões que a Política Pública abrange. Também demonstrou-se relevante o estudo sobre as formas de visualização dos dados no ambiente Web voltado para o usuário final, proporcionando o desenvolvimento de aplicações que facilitem a interpretação dos dados pelos usuários e, por consequência, sua aplicação na rotina de gestores públicos e atores do controle social em Políticas Públicas do país.

## REFERÊNCIAS

BARR, C.E.; KOMOROWSKI, H.J.; PATTISON-GORDON, E. P.; GREENES, R.A.

*Conceptual Modeling for The Unified Medical Language System*. SCAMC: Inc. (1988).

BERNASCONI A; CERI S; CAMPI A.; MASSEROLI M. *Conceptual Modeling for*

*Genomics: Building an Integrated Repository of Open Data*. 2017. Springer International Publishing AG 2017.

BRASIL 2006. Congresso Nacional. **Lei 11.346, de 15 de setembro de 2006**. Lei Orgânica de Segurança Alimentar e Nutricional (LOSAN). Brasília, DF, 2006.

BRASIL 2011. **Lei Nº 12.527, de 18 de Novembro de 2011**. Regula o acesso a informações e dá outras providências. Disponível em

<[http://www.planalto.gov.br/ccivil\\_03/\\_ato2011-2014/2011/lei/112527.htm](http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/112527.htm)>. Acesso em setembro de 2018.

BROOKE, J. **SUS: A quick and dirty usability scale**. In.: Jordan, P.W., Thomas, B.,

Weerdmeester, B.A., McClelland, I.L. (eds.). *Usability Evaluation in Industry*. pp. 189-194.

Taylor & Francis, PA (1996).

CASTRO, L.P.M. **Sistema de Apoio ao Projeto OBSSAN**. Instituto de Informática UFRGS, Porto Alegre, 2014.

COUGO P. **Modelagem Conceitual e Projeto de Bancos de Dados**. Rio de Janeiro, Editora Campus. 1997.

EMBLEY, D.W.; MOK, W.Y. *Mapping Conceptual Models to Database Schemas*. In:

Embley, D.W.; Thalheim, B. (eds.) *Handbook of Conceptual Modeling: Theory, Practice, and Research Challenges*. pp. 123-163. Springer, Heidelberg (2011).

KIMBALL R.; ROSS M. *The Data Warehouse ETL Toolkit. Third Edition*, 2013.

MACHADO, F. N. R. **Projeto de Data Warehouse: Uma Visão multidimensional**. São Paulo: Érica, 2000, 248 p.

MIZOGUCHI, Riichiro; IKEDA, Mitsuru. *Towards Ontology Engineering. Proceedings of the Joint Pacific Asian Conference on Experts Systems/Singapore International Conference on Intelligent Systems*. pg. 259-266. Singapura, 1997.

MIZUSAKI, L.E.P. **Grafos de Avaliação: um modelo conceitual para avaliação escolar apoiada por computador**. Dissertação (Mestrado) – Universidade Federal do Rio Grande do Sul. Programa de Pós-Graduação em Computação. Porto Alegre, BR – RS, 2016.

OLIVÉ, A. *Conceptual Modeling of Information Systems*. Springer Berlin Heidelberg New York, 2007.

OPEN KNOWLEDGE BRASIL. **Projetos com dados aberto**. Disponível em <<https://br.okfn.org/projetos/>>. Acesso em setembro de 2018.

OPEN KNOWLEDGE FOUNDATION. *Open Data Handbook*. Disponível em: <<http://opendatahandbook.org/guide/sr/>>. Acesso em maio de 2018.

PORTAL BRASILEIRO DE DADOS ABERTOS. **O que são dados abertos?**. Disponível em <<http://dados.gov.br/pagina/dados-abertos>>. Acesso em abril de 2018.

RIGO, L.M. **Sistema de Apoio à Decisão com Apresentação Georreferenciada de Dados no âmbito do Projeto Agroflorestas**. Instituto de Informática UFRGS, Porto Alegre, 2012.

RIGON, V. **Sistema Web de Acompanhamento do Projeto Agroflorestas**. Instituto de Informática UFRGS, Porto Alegre, 2013.



SIQUEIRA, A.C. **Indicadores de SAN e sua avaliação nos municípios**. Apresentação de oficina. Curso de Fortalecimento do SISAN e da Educação Alimentar e Nutricional no RS. Porto Alegre, dez/2017. (Curso ministrado/Seminário).

SIQUEIRA, A.C.; WIVES, D.G.; MEDEIROS, T. M.; SCHREIBER, N.M.; SILVA, L.X.; COELHO-DE-SOUZA, G.P. **Observatório Socioambiental em Segurança Alimentar e Nutricional: análise dos indicadores de produção de alimentos em nível municipal no Rio Grande do Sul**. REDES (SANTA CRUZ DO SUL. ONLINE), v. 21, p. 49-62, 2016.

SIQUEIRA, A.C.; WIVES, D.G.; ROCHA, B; Silva, L.X.; COELHO-DE-SOUZA, G.P. **Observatório Socioambiental em Segurança Alimentar e Nutricional do Rio Grande do Sul: Histórico, metodologias e resultados parciais**. 2015. (Apresentação de Trabalho/Seminário).

THE PROJECT CARTOON. ***How Projects Really Work (Brazilian Portuguese Version)***. Disponível em: <<http://www.projectcartoon.com/cartoon/611>>. Acesso em janeiro de 2019.

TRYFONA, N.; JENSEN, C.S. ***Conceptual Data Modeling for Spatiotemporal Applications***. GeoInformatica 3:3, 245-268 (1999).

VASSILIADIS, P.; SIMITSIS, A.; SKIADOPOULOS, S. ***Conceptual Modeling for ETL Processes***. DOLAP'02: Virginia, USA (2002).

## APÊNDICE A - DICIONÁRIO DE DADOS: ATRIBUTOS

Tabela 4.2 - Dicionário de Dados: Atributos

<b>Entidade: fato_dados</b>						
<b>Atributo</b>	<b>Tipo de dado (tamanho)</b>	<b>P K</b>	<b>N N</b>	<b>U Q</b>	<b>AI</b>	<b>Descrição</b>
Dim_Dimensao_Id	Inteiro (11)	x	x			Código da Dimensão
Dim_Indicador_Id	Inteiro (11)	x	x			Código do Indicador
Dim_Categoria_Id	Inteiro (11)	x	x			Código da Categoria
Dim_Variavel_Id	Inteiro (11)	x	x			Código da Variável
Dim_Classificacao_Id	Inteiro (11)	x	x			Código da Classificação
Dim_Municipio_Id	Inteiro (11)	x	x			Código do Município
Dim_Unidade_Id	Inteiro (11)	x	x			Código da Unidade
Dim_Fonte_Id	Inteiro (11)	x	x			Código da Fonte
Med_Ano	Inteiro (4)	x	x			Ano do dado
Med_Quantidade	Decimal (14,3)		x			Quantidade do dado
<b>Entidade: dim_dimensao</b>						
<b>Atributo</b>	<b>Tipo de dado (tamanho)</b>	<b>P K</b>	<b>N N</b>	<b>U Q</b>	<b>AI</b>	<b>Descrição</b>
Dim_Dimensao_Id	Inteiro (11)	x	x		x	Código auto incrementável único de identificação
Numero	Inteiro (11)		x	x		Número OBSSAN da dimensão
Nome	Caractere (255)			x		Nome da dimensão
Consulta	Inteiro (1)					Parâmetro que define se a dimensão irá aparecer nas consultas, "1" = sim.
<b>Entidade: dim_indicador</b>						
<b>Atributo</b>	<b>Tipo de dado (tamanho)</b>	<b>P K</b>	<b>N N</b>	<b>U Q</b>	<b>AI</b>	<b>Descrição</b>
Dim_Indicador_Id	Inteiro (11)	x	x		x	Código auto incrementável único de identificação
Numero	Inteiro (11)		x	x		Número OBSSAN do indicador
Nome	Caractere (255)			x		Nome do indicador
<b>Entidade: dim_categoria</b>						
<b>Atributo</b>	<b>Tipo de dado (tamanho)</b>	<b>P K</b>	<b>N N</b>	<b>U Q</b>	<b>AI</b>	<b>Descrição</b>
Dim_Categoria_Id	Inteiro (11)	x	x		x	Código auto incrementável único de identificação

Nome	Caractere (255)			x		Nome da categoria
<b>Entidade: dim_variavel</b>						
Atributo	Tipo de dado (tamanho)	P K	N N	U Q	AI	Descrição
Dim_Variavel_Id	Inteiro (11)	x	x		x	Código auto incrementável único de identificação
Nome	Caractere (255)			x		Nome da variável
<b>Entidade: dim_classificacao</b>						
Atributo	Tipo de dado (tamanho)	P K	N N	U Q	AI	Descrição
Dim_Classificacao_Id	Inteiro (11)	x	x		x	Código auto incrementável único de identificação
Nome	Caractere (255)			x		Nome da classificação
<b>Entidade: dim_unidade</b>						
Atributo	Tipo de dado (tamanho)	P K	N N	U Q	AI	Descrição
Dim_Unidade_Id	Inteiro (11)	x	x		x	Código auto incrementável único de identificação
Nome	Caractere (255)			x		Nome da unidade de medida
<b>Entidade: dim_fonte</b>						
Atributo	Tipo de dado (tamanho)	P K	N N	U Q	AI	Descrição
Dim_Fonte_Id	Inteiro (11)	x	x		x	Código auto incrementável único de identificação
Nome	Caractere (255)			x		Nome da fonte (origem do dado)
<b>Entidade: dim_municipio</b>						
Atributo	Tipo de dado (tamanho)	P K	N N	U Q	AI	Descrição
Dim_Municipio_Id	Inteiro (11)	x	x			Código IBGE único do município
Nome	Caractere (255)		x	x		Nome do município
Uf	Caractere (2)		x			Unidade federativa do município
Mesorregiao_Nome	Caractere (255)			x		Nome da mesorregião que o município pertence

Mesorregiao_Municipio_Central	Inteiro (1)				Se dado = "1" então é o município central da mesorregião
Microrregiao_Nome	Caractere (255)			x	Nome da microrregião que o município pertence
Microrregiao_Municipio_Central	Inteiro (1)				Se dado = "1" então é o município central da microrregião
Territorio_Rural_Nome	Caractere (255)			x	Nome da território rural que o município pertence
Territorio_Rural_Municipio_Central	Inteiro (1)				Se dado = "1" então é o município central do território rural
Micro_Territorio_Rural_Nome	Caractere (255)			x	Nome do micro território rural que o município pertence
Micro_Territorio_Rural_Municipio_Central	Inteiro (1)				Se dado = "1" então é o município central do micro território rural
Poligono_Coordenadas	Texto		x		Armazena o conjunto de coordenadas para formar um polígono do município
Latitude	Decimal (8,6)		x		Latitude central do município
Longitude	Decimal (8,6)		x		Longitude central do município

## APÊNDICE B - *STORED PROCEDURE* PARA INSERÇÃO DE DADOS

```
DROP PROCEDURE IF EXISTS sp_insertupdate_dados;
DELIMITER $$
CREATE PROCEDURE sp_insertupdate_dados(
    IN p_dimensao_nome      VARCHAR(255),
    IN p_indicador_nome    VARCHAR(255),
    IN p_categoria_nome    VARCHAR(255),
    IN p_variavel_nome     VARCHAR(255),
    IN p_classificacao_nome VARCHAR(255),
    IN p_subclassificacao_nome VARCHAR(255),
    IN p_municipio_nome    VARCHAR(255),
    IN p_municipio_uf     VARCHAR(255),
    IN p_unidade_nome     VARCHAR(255),
    IN p_fonte_nome       VARCHAR(255),
    IN p_med_quantidade   DECIMAL(14,3),
    IN p_med_ano          INT(11)
)
BEGIN
    DECLARE vDim_Dimensao_Id INT DEFAULT (SELECT Dim_Dimensao_Id
    FROM dim_dimensao WHERE UPPER(nome) = UPPER(p_dimensao_nome));

    DECLARE vDim_Indicador_Id INT DEFAULT (SELECT Dim_Indicador_Id
    FROM dim_indicador WHERE UPPER(nome) = UPPER(p_indicador_nome));

    DECLARE vDim_Categoria_Id INT DEFAULT (SELECT Dim_Categoria_Id
FROM dim_categoria WHERE UPPER(nome) = UPPER(p_categoria_nome));

    DECLARE vDim_Variavel_Id INT DEFAULT (SELECT Dim_Variavel_Id
    FROM dim_variavel WHERE UPPER(nome) = UPPER(p_variavel_nome));

    DECLARE vDim_Classificacao_Id INT DEFAULT (SELECT
    Dim_Classificacao_Id FROM dim_classificacao WHERE UPPER(nome) =
    UPPER(p_classificacao_nome));

    DECLARE vDim_Subclassificacao_Id INT DEFAULT (SELECT
    Dim_Subclassificacao_Id FROM dim_subclassificacao WHERE UPPER(nome) =
    UPPER(p_subclassificacao_nome));

    DECLARE vDim_Municipio_Id INT DEFAULT (SELECT Dim_Municipio_Id
    FROM dim_municipio WHERE UPPER(nome) = UPPER(p_municipio_nome)
    AND uf = p_municipio_uf);

    DECLARE vDim_Unidade_Id INT DEFAULT (SELECT Dim_Unidade_Id
    FROM dim_unidade WHERE UPPER(nome) = UPPER(p_unidade_nome));

    DECLARE vDim_Fonte_Id INT DEFAULT (SELECT Dim_Fonte_Id
    FROM dim_fonte WHERE UPPER(nome) = UPPER(p_fonte_nome));

    DECLARE vInsertUpdate TEXT DEFAULT "";
```

**DECLARE EXIT HANDLER FOR 1048**

**BEGIN**

**SELECT**

'ERRO' **AS** sp\_insertupdate\_fato\_status,  
'REGISTRO NÃO ENCONTRADO' **AS** erro\_mensagem,

**IF**(p\_dimensao\_nome = " **or** p\_dimensao\_nome **IS NULL**,  
'NULO', **IF**((**SELECT** Nome **FROM** dim\_dimensao **WHERE** nome =  
p\_dimensao\_nome) **IS NULL**, **UPPER**(p\_dimensao\_nome),  
**null**)) **AS** Dimensao,

**IF**(p\_indicador\_nome = " **or** p\_indicador\_nome **IS NULL**,  
'NULO', **IF**((**SELECT** Nome **FROM** dim\_indicador **WHERE** nome =  
p\_indicador\_nome) **IS NULL**, **UPPER**(p\_indicador\_nome), **null**)) **AS**  
Indicador,

**IF**(p\_categoria\_nome = " **or** p\_categoria\_nome **IS NULL**, 'NULO',  
**IF**((**SELECT** Nome **FROM** dim\_categoria **WHERE** nome =  
p\_categoria\_nome) **IS NULL**, **UPPER**(p\_categoria\_nome), **null**)) **AS**  
Categoria,

**IF**(p\_variavel\_nome = " **or** p\_variavel\_nome **IS NULL**, 'NULO',  
**IF**((**SELECT** Nome **FROM** dim\_variavel **WHERE** nome =  
p\_variavel\_nome) **IS NULL**, **UPPER**(p\_variavel\_nome), **null**)) **AS**  
Variavel,

**IF**(p\_classificacao\_nome = " **or** p\_classificacao\_nome **IS NULL**, 'NULO',  
**IF**((**SELECT** Nome **FROM** dim\_classificacao **WHERE** nome =  
p\_classificacao\_nome) **IS NULL**, **UPPER**(p\_classificacao\_nome),  
**null**)) **AS** Classificacao,

**IF**(p\_municipio\_nome = " **or** p\_municipio\_nome **IS NULL** **or**  
p\_municipio\_uf = " **or** p\_municipio\_uf **IS NULL**, 'NULO',  
**IF**((**SELECT** Nome **FROM** dim\_municipio **WHERE** nome = p\_municipio\_nome  
**AND**  
Uf = p\_municipio\_uf) **IS NULL**, **CONCAT**(**UPPER**(p\_municipio\_nome), '/',  
**UPPER**(p\_municipio\_uf)), **null**)) **AS** Municipio,

**IF**(p\_unidade\_nome = " **or** p\_unidade\_nome **IS NULL**, 'NULO',  
**IF**((**SELECT** Nome **FROM** dim\_unidade **WHERE** nome =  
p\_unidade\_nome) **IS NULL**, **UPPER**(p\_unidade\_nome), **null**)) **AS** Unidade,

**IF**(p\_fonte\_nome = " **or** p\_fonte\_nome **IS NULL**, 'NULO', **IF**((**SELECT**  
Nome **FROM** dim\_fonte **WHERE** nome = p\_fonte\_nome) **IS NULL**,  
**UPPER**(p\_fonte\_nome), **null**)) **AS** Fonte;

**END;**

**IF (**

**(SELECT COUNT(\*) FROM fato\_dados WHERE**  
Dim\_Dimensao\_Id = vDim\_Dimensao\_Id **AND**

```

Dim_Indicador_ID = vDim_Indicador_Id AND
Dim_Categoria_Id = vDim_Categoria_Id AND
Dim_Variavel_Id = vDim_Variavel_Id AND
Dim_Classificacao_Id = vDim_Classificacao_Id AND
Dim_Subclassificacao_Id = vDim_Subclassificacao_Id AND
Dim_Municipio_Id = vDim_Municipio_Id AND
Dim_Unidade_Id = vDim_Unidade_Id AND
Dim_Fonte_Id = vDim_Fonte_Id AND
Med_Ano = p_med_ano
    ) > 0
)
THEN SET vInsertUpdate = "UPDATE";
ELSE SET vInsertUpdate = "INSERT";
END IF;
REPLACE fato_dados (
(
    Dim_Dimensao_Id,
    Dim_Indicador_Id,
    Dim_Categoria_Id,
    Dim_Variavel_Id,
    Dim_Classificacao_Id,
    Dim_Subclassificacao_Id,
    Dim_Municipio_Id,
    Dim_Unidade_Id,
    Dim_Fonte_Id,
    Med_Quantidade,
    Med_Ano
) VALUES (
    vDim_Dimensao_Id,
    vDim_Indicador_Id,
    vDim_Categoria_Id,
    vDim_Variavel_Id,
    vDim_Classificacao_Id,
    vDim_Subclassificacao_Id,
    vDim_Municipio_Id,
    vDim_Unidade_Id,
    vDim_Fonte_Id,
    p_med_quantidade,
    p_med_ano
);
SELECT 'OK' AS sp_insertupdate_fato_status, vInsertUpdate AS InsertUpdate;
END $$
DELIMITER ;

```