



UNIVERSIDADE ESTADUAL DO NORTE DO PARANÁ
CAMPUS LUIZ MENEGHEL

EDUARDO CRISTOVO DE FREITAS AGUIAR

**IMPLANTANDO A METODOLOGIA AGILE ROLAP -
UMA METODOLOGIA ÁGIL DE DESENVOLVIMENTO
PARA SISTEMAS DE INTELIGÊNCIA DO NEGÓCIO.**

Bandeirantes
2016

Eduardo Cristovo de Freitas Aguiar

**IMPLANTANDO A METODOLOGIA AGILE ROLAP -
UMA METODOLOGIA ÁGIL DE DESENVOLVIMENTO
PARA SISTEMAS DE INTELIGÊNCIA DO NEGÓCIO.**

Trabalho de Conclusão de Curso submetido à
Universidade Estadual do Norte do Paraná,
como requisito parcial para a obtenção do
grau de Bacharel em Sistemas de Informação.

Orientador: Prof. Dr. André Luís Andrade
Menolli

Bandeirantes

2016

Eduardo Cristovo de Freitas Aguiar

Implantando a metodologia Agile ROLAP - Uma metodologia ágil de desenvolvimento para sistemas de inteligência do negócio.

Trabalho de Conclusão de Curso submetido à Universidade Estadual do Norte do Paraná, como requisito parcial para a obtenção do grau de Bacharel em Sistemas de Informação.

COMISSÃO EXAMINADORA

Prof. Dr. André Luís Andrade Menolli
UENP – *Campus* Luiz Meneghel

Prof. Me. Glauco Carlos Silva
UENP – *Campus* Luiz Meneghel

Prof. Me. Ricardo Gonçalves Coelho
UENP – *Campus* Luiz Meneghel

Bandeirantes, ___ de _____ de 2016

Dedico este trabalho a minha família que sempre esteve presente e incentivou meus estudos.

AGRADECIMENTOS

Agradeço ao meu Orientador pela ajuda e boa orientação no desenvolvimento deste trabalho.

Agradeço a fundação Araucária e Carlos Chagas, pois, este trabalho foi iniciado durante meu período de bolsa PIBIC.

Também devo muito aos professores que sanaram minhas dúvidas, sem eles o desenvolvimento deste trabalho provavelmente não seria possível, agradeço principalmente ao professor Glauco Carlos Silva pelas dúvidas sanadas.

Apesar de alguns professores terem participado mais ativamente do desenvolvimento deste trabalho, também é necessário agradecer a todos os professores que participaram do meu aprendizado durante este período de graduação, portanto, obrigado a todos.

“Estamos afogados por informação, mas morrendo de fome por conhecimento.” **(John Nalsbett)**

RESUMO

Tomar uma decisão importante a respeito de seu negócio não costuma ser algo fácil para um gestor. As ferramentas de BI podem auxiliar o gestor nesse processo, providenciando informações a respeito do estado da organização como um todo, de modo a facilitar a tomada da decisão. Porém, a implantação de um ambiente de BI pelas metodologias tradicionais é custosa e trabalhosa. Para solucionar este problema foram desenvolvidas metodologias ágeis de desenvolvimento de BI. Dentre elas uma nova metodologia denominada Agile ROLAP visa implantar ambientes de BI de forma mais rápida e fácil em pequenas e médias empresas. No entanto, para que isso seja possível é necessário implementar ferramentas que dêem suporte a essa implantação e testá-las. Atualmente a metodologia Agile ROLAP conta com ferramentas que foram desenvolvidas para auxiliar sua implantação, todavia, estas ferramentas ainda carecem de mais testes. Por isso este trabalho pretende implantar a metodologia Agile ROLAP utilizando das ferramentas desenvolvidas para a mesma, de modo a testar a implantação da metodologia e suas ferramentas.

Palavras-chave: Data Warehouse; Metodologia Ágil Rolap; ETL.

ABSTRACT

Make an important decision about your business is usually not an easy thing for a manager. BI tools can assist the manager in this process, providing information about the status of the organization as a whole, in order to facilitate the decision making process. However, the implementation of a BI environment by traditional methods is costly and labor intensive. To solve this problem agile methodologies of BI development have been developed. Among them a new methodology called Agile ROLAP aims to deploy BI environments more quickly and easily in small and medium enterprises. However, for this to be possible it is necessary to implement tools that support this deploy and test them. Currently the methodology Agile ROLAP has tools that are designed to assist its implementation, however, these tools still need further testing. Therefore this work intends to deploy Agile ROLAP methodology using the tools developed for it, in order to test the implementation of the methodology and tools.

Key-words: Data Warehouse; Methodology Agile Rolap; ETL.

LISTA DE ILUSTRAÇÕES

Figura 1: Exemplo de modelo dimensional (SOUZA, 2014).	21
Figura 2: Processo de Implantação de um Data Warehouse Adaptado de BIReady (2011).	23
Figura 3: Estrutura Agile ROLAP (SOUZA, 2014).	30
Figura 4: Passos para a implantação do Agile ROLAP adaptados de Souza; Menolli; Coelho, (2014).	31
Figura 5: Exemplo de transformação no Kettle (PENTAHO, 2016).	35
Figura 6: Organização dos <i>plugins</i> Agile ROLAP.	36
Figura 7: Fato avaliacao aluno.	43
Figura 8: Fato avaliacao aluno_professor.	46
Figura 9: Fato avaliacao docente.	46
Figura 10: Fato avaliacao estagiario.	46
Figura 11: Fato avaliacao tecnico.	47
Figura 12: Configuração do <i>plugin</i> base de dados estrangeira.	48
Figura 13: Configuração da dimensão tempo no <i>plugin</i> tempo.	49
Figura 14: Configuração da dimensão professor no <i>plugin</i> dimensão.	50
Figura 15: Fato avaliacao aluno na aba de configuração SQL.	51
Figura 16: Fato avaliacao aluno na aba de configuração Dimensão.	51
Figura 17: Fato avaliacao aluno na aba de configuração Dimensão de Tempo.	52
Figura 18: Fato avaliacao aluno na aba de configuração Medidas.	52
Figura 19: Dimensão de tempo ano configurada no <i>plugin</i> dimensão lógica de tempo.	54
Figura 20: Configuração dimensão professor no <i>plugin</i> dimensão lógica.	55
Figura 21: Cubo Avaliacao aluno na aba de configuração Cubo.	56
Figura 22: Cubo Avaliacao aluno na aba de configuração Medidas.	57
Figura 23: Configuração do <i>plugin</i> SchemaOut.	58
Figura 24: Exemplo de consulta no cubo Avaliacao aluno_professor.	59
Figura 25: Exemplo de consulta no cubo Avaliacao aluno.	60
Figura 26: Exemplo de consulta no cubo Avaliacao docente.	60
Figura 27: Campo coluna do <i>plugin</i> Dimensão lógica.	62
Figura 28: Visualizando as colunas disponíveis no <i>plugin</i> Dimensão lógica.	62
Figura 29: Campo coluna do <i>plugin</i> Dimensão lógica após visualizar as colunas disponíveis.	62
Figura 30: Erro ao salvar <i>plugins</i> com configurações em branco.	63
Figura 31: Erro ao deletar tabelas da dimensão e tentar inserir uma nova tabela. ..	64
Figura 32: Nome incorreto para a janela do <i>plugin</i> Dimensão lógica de tempo.	64
Figura 33: <i>Plugin</i> cubo com nome de janela incorreto.	65
Figura 34: Os <i>plugins</i> Dimensão lógica de tempo e Dimensão lógica não distinguem corretamente dimensões tempo de não tempo.	65
Figura 35: Mensagem de warning ao ligar dois passos.	67
Figura 36: Janela não ajustável da ferramenta Fato.	68
Figura 37: Janela não ajustável da ferramenta Fato.	69
Figura 38: Fases Agile ROLAP adaptados de Souza(2014).	70
Figura 39: Configuração da dimensão disciplina no <i>plugin</i> dimensão.	83
Figura 40: Configuração da dimensão curso no <i>plugin</i> dimensão.	84
Figura 41: Configuração da dimensão campus no <i>plugin</i> dimensão.	84

Figura 42: Fato avaliacao aluno_professor na aba de configuração SQL.	85
Figura 43: Fato avaliacao aluno_professor na aba de configuração Dimensão.	85
Figura 44: Fato avaliacao aluno_professor na aba de configuração Dimensão de Tempo.	86
Figura 45: Fato avaliacao aluno_professor na aba de configuração Medidas.	86
Figura 46: Fato avaliacao docente na aba de configuração SQL.	86
Figura 47: Fato avaliacao docente na aba de configuração Dimensão.	87
Figura 48: Fato avaliacao docente na aba de configuração Dimensão de Tempo. ...	87
Figura 49: Fato avaliacao docente na aba de configuração Medidas.	87
Figura 50: Fato avaliacao estagiario na aba de configuração SQL.	88
Figura 51: Fato avaliacao estagiario na aba de configuração Dimensão.	88
Figura 52: Fato avaliacao estagiario na aba de configuração Dimensão de Tempo.	88
Figura 53: Fato avaliacao estagiario na aba de configuração Medidas.	89
Figura 54: Fato avaliacao tecnico na aba de configuração SQL.	89
Figura 55: Fato avaliacao tecnico na aba de configuração Dimensão.	89
Figura 56: Fato avaliacao tecnico na aba de configuração Dimensão de Tempo. ...	90
Figura 57: Fato avaliacao tecnico na aba de configuração Medidas.	90
Figura 58: Configuração dimensão disciplina no <i>plugin</i> dimensão lógica.	91
Figura 59: Configuração dimensão curso no <i>plugin</i> dimensão lógica.	92
Figura 60: Configuração dimensão campus no <i>plugin</i> dimensão lógica.	93
Figura 61: Cubo Avaliacao aluno_professor na aba de configuração Cubo.	94
Figura 62: Cubo Avaliacao aluno_professor na aba de configuração Medidas.	95
Figura 63: Cubo Avaliacao docente na aba de configuração Cubo.	95
Figura 64: Cubo Avaliacao docente na aba de configuração Medidas.	96
Figura 65: Cubo Avaliacao estagiario na aba de configuração Cubo.	96
Figura 66: Cubo Avaliacao estagiario na aba de configuração Medidas.	97
Figura 67: Cubo Avaliacao tecnico na aba de configuração Cubo.	97
Figura 68: Cubo Avaliacao tecnico na aba de configuração Medidas.	98

LISTA DE QUADROS

Quadro 1: Medidas do fato avaliacao aluno.	43
Quadro 2: Medidas do fato avaliacao aluno_professor.	74
Quadro 3: Medidas do fato avaliacao docente.	75
Quadro 4: Medidas do fato avaliacao estagiario	79
Quadro 5: Medidas do fato avaliacao tecnico.....	80

LISTA DE ABREVIATURAS E SIGLAS

BI	Business Intelligence
DW	Data Warehouse
ETL	Estração, Transformação e Carga (Extract, Transform and Load)
FDW	Foreign Data Wrapper
MDX	Multi Dimensional Expression
OLAP	On-Line Analytical Processing
ROLAP	Relational On-Line Analytical Processing
SQL	Structured Query Language
UENP	Universidade Estadual do Norte do Paraná
XML	Extensible Markup Language

SUMÁRIO

LISTA DE ILUSTRAÇÕES.....	8
LISTA DE QUADROS	10
LISTA DE ABREVIATURAS E SIGLAS	11
1. INTRODUÇÃO	15
1.1 Problema	15
1.2 Justificativa	16
1.3 Objetivos	17
1.3.1 Objetivo Geral	17
1.3.2 Objetivos Específicos	17
1.4 Organização do trabalho	17
2. Revisão de literatura	19
2.1 BI	19
2.2 Data Warehouse	20
2.2.1 Processo de ETL.....	23
2.3 Ferramentas OLAP	28
2.4 Agile ROLAP	29
3. Ferramentas utilizadas.....	34
3.1 PostgreSQL	34
3.2 Kettle - Data integration	34
3.3 Plugins Agile ROLAP	35
3.4 Mondrian.....	38
3.5 Pivot4j.....	39
4. Metodologia.....	40
5. Desenvolvimento	42
5.1 Acesso a base de dados	42
5.1.1 Levantamento dos dados e requisitos – 1° Passo.....	42
5.1.2 Acesso a fonte de dados – 2° Passo.....	42
5.1.3 Criação da base de dados intermediária – 3° Passo	47
5.2 Modelo físico	48
5.2.1 Criação das dimensões de tempo – 4° Passo.....	49

5.2.2 Criação das dimensões convencionais – 5° Passo	49
5.2.3 Criação dos fatos – 6° Passo.....	50
5.3 Modelo lógico.....	52
5.3.1 Criação das hierarquias e níveis das dimensões de tempo – 7° Passo	53
5.3.2 Criação das hierarquias e níveis das dimensões convencionais – 8° Passo.....	54
5.3.3 Criação dos cubos – 9° Passo.....	55
5.3.4 Exportação do Schema XML – 10° Passo.....	57
6. Resultados	59
6.1 Erros encontrados.....	61
7. Discussão	66
7.1 Sugestões de melhoria	66
8. Conclusão e trabalhos futuros.....	71
Referências	72
Apêndice A – Medidas dos fatos.....	74
Apêndice B – Configuração das dimensões.....	83
Apêndice C – Configuração dos fatos.....	85
Apêndice D – Configuração das dimensões lógicas	91
Apêndice E – Configuração dos cubos	94

1. INTRODUÇÃO

Muitas organizações geram empregos e modificam a economia e sociedade ao seu redor, e quando bem administradas podem gerar altos lucros para seus proprietários e investidores.

No entanto, para uma empresa gerar lucro é importante que estas tomem decisões estratégicas adequadas. Porém, o ato de tomar decisões não costuma ser fácil, afinal os erros e acertos do administrador podem afetar não só a ele, mas a organização como um todo.

Uma decisão geralmente é tomada com base em informações, mas o problema é que geralmente as informações que poderiam auxiliar na tomada de decisões se encontram dispersas e fragmentadas, o que faz com que os executivos tomem decisões baseadas em informações escassas. Tal medida é inaceitável para empresas que querem ser competitivas no mercado em que atuam, pois a sucessão de decisões erradas podem levar à falência da mesma (ANGELONI, 2003).

Para minimizar estes problemas, uma técnica que pode ser empregada é a Inteligência do Negócio. Inteligência de Negócio ou *Business Intelligence* (BI) é o conjunto de tecnologias orientadas a disponibilizar informação e conhecimento para as empresas, permitindo uma melhor visão do que está ocorrendo e contribuindo para uma melhor tomada de decisão (Machado, 2010).

As informações que podem auxiliar na tomada de decisão são geradas pelas próprias empresas, em suas transações diárias. Um sistema de BI auxilia os administradores com ferramentas capazes de agrupar estes dados dispersos pelos vários setores da organização e permitir uma análise sobre eles. As principais ferramentas e tecnologias utilizadas pelos sistemas de BI são os *Data Warehouses* (DW) e as ferramentas *On-Line Analytical Processing* (OLAP), que surgiram no começo dos anos 90, como novas ferramentas que foram a base dos sistemas de BI (Shim et al, 2002).

A criação de um sistema de BI, resumidamente, consiste em criar um DW a partir dos dados transacionais da organização, e utilizar de uma ferramenta OLAP para visualizá-los e analisá-los.

1.1 Problema

Os dados que são armazenados no DW seguem a modelagem dimensional, porém, as bases organizacionais são modeladas de maneira relacional e para que esta conversão de modelos aconteça é necessário passar pelo processo de ETL, processo este que pode levar cerca de 80% do tempo necessário para o desenvolvimento de um DW (Menolli, 2006).

O mercado evolui rapidamente e se fazem necessárias metodologias de desenvolvimento de DWs mais ágeis e adaptáveis as mudanças, com base nisto foi desenvolvida a metodologia Agile ROLAP, que reduz o processo de ETL, diminuindo o tempo necessário para a criação ou alteração de um DW (SOUZA; MENOLLI; COELHO, 2014).

Além deste problema da metodologia tradicional de DW, o *Agile ROLAP* também tem o objetivo de solucionar problemas presentes em outras metodologias ágeis, como por exemplo, o Agile BI, que apresenta como principal problema a impossibilidade de se usar boa parte das ferramentas OLAP existentes, concebidas para a metodologia tradicional de DW (SOUZA, 2014).

A metodologia *Agile ROLAP* já soluciona o problema com as ferramentas OLAP, e foram desenvolvidas diversas ferramentas auxiliares para a implantação da metodologia *Agile ROLAP*, porém somente algumas destas foram de fato testadas e utilizadas durante a implantação da metodologia no trabalho de Souza (2014). Dessa forma, este trabalho pretende implantar a metodologia Agile ROLAP utilizando das ferramentas Base de Dados Estrangeira, Fato, Dimensão, Tempo, Feriado, XMLOut, Cubo, Dimensão Lógica, Dimensão Lógica de Tempo e Schema Out, que foram desenvolvidas para auxiliar na implantação da metodologia Agile ROLAP.

1.2 Justificativa

Os *Data Warehouses* são tradicionalmente desenvolvidos utilizando-se do processo de *Extraction Transformation and Load* (ETL) para transformar uma base de dados Relacional em uma base dimensional, processo que leva um longo tempo, porém, este tempo é amenizado devido a ferramentas de ETL que auxiliam neste processo.

Metodologias ágeis para desenvolvimento de Data Warehouse estão disponíveis no mercado hoje, porém, estas não contam com tantas ferramentas auxiliares para acelerar seu processo, a partir deste problema foi elaborado o projeto “Agile ROLAP, uma metodologia ágil de desenvolvimento de DW”, que tem por objetivo desenvolver ferramentas para auxiliar a implantação da metodologia.

O projeto desenvolveu e continua desenvolvendo mais e mais ferramentas que possam auxiliar a implantação da metodologia Agile ROLAP. Porém dentre estas ferramentas desenvolvidas somente algumas foram utilizadas em uma implantação real da metodologia, necessitando de mais testes e análises para certificar que as ferramentas desenvolvidas cumprem seus requisitos, ou expor quais funções poderiam ser melhoradas.

1.3 Objetivos

1.3.1 Objetivo Geral

Este trabalho tem como objetivo principal implantar a metodologia Agile ROLAP utilizando dos *plugins* Base de Dados Estrangeira, Fato, Dimensão, Tempo, Feriado, XMLOut, Cubo, Dimensão Lógica, Dimensão Lógica de Tempo e Schema Out, desenvolvidos para auxiliar a implantação da metodologia Agile Rolap.

1.3.2 Objetivos Específicos

- Definir uma base de dados para implantar a metodologia Agile Rolap;
- Implantar a metodologia Agile Rolap à base selecionada, utilizando os *plugins* desenvolvidos;
- Fazer uma análise sobre a implantação realizada para identificar erros e possíveis melhorias na metodologia e nos *plugins* utilizados.

1.4 Organização do trabalho

Este trabalho segue a seguinte organização, no capítulo 2 o leitor irá encontrar os conceitos necessários para o entendimento do trabalho. No capítulo 3

são descritas as ferramentas utilizadas no trabalho e porque elas foram utilizadas e não suas concorrentes. No capítulo 4 é descrita a metodologia de desenvolvimento seguida pelo presente trabalho. O capítulo 5 é descreve como foi desenvolvido este trabalho passo a passo. Já no capítulo 6 são apresentados os resultados obtidos através da realização do trabalho. O capítulo 7 procura colocar em discussão tópicos que não são necessariamente resultados do trabalho, mas poderiam se tornar. E por último a conclusão do autor sobre o trabalho no capítulo 8.

2. Revisão de literatura

A revisão da literatura encontra-se dividida em cinco seções centradas nos temas: BI, DW, ETL, OLAP e Agile ROLAP. Na primeira seção o leitor irá encontrar uma breve Introdução e contextualização ao tema BI. Na segunda é explicado o processo de criação de um DW. Na terceira são abordadas as ferramentas OLAP de visualização de dados e seus tipos. Na quarta e última seção é apresentada a metodologia Agile ROLAP, que é metodologia ágil de desenvolvimento de BI que este trabalho tem por objetivo implantar.

2.1 BI

Segundo Calazans (2012) o uso da informação é considerado um dos recursos mais importantes para garantir a competitividade nos dias atuais. A informação é criada, tratada e transferida às organizações, sendo a eficiência desse processo a garantia do bom desempenho e do deslocamento da organização para patamares mais elevados de competitividade.

Por outro lado, quando estas informações não são aproveitadas de maneira a proporcionar decisões mais rápidas e assertivas a empresa pode sofrer sérias consequências, como: Não se sobressair sobre seus concorrentes, perder oportunidades de mercado, perder seus clientes atuais, não identificar indivíduos que seriam clientes em potencial e etc...

Tomar decisões baseadas somente em opiniões pessoais é inaceitável atualmente, pois a sucessão de decisões erradas pode causar à falência da empresa (ANGELONI, 2003). Uma decisão baseada em informações concretas tende a ser mais assertiva, porém, conseguir informações úteis a tomada de decisão não é algo simples. É nesse momento que um sistema de *Business Intelligence* (BI) se mostra um recurso indispensável para ajudar os administradores da empresa a melhorar sua tomada de decisão.

A implantação de ambientes de BI permite que as organizações utilizem seus próprios dados para gerar conhecimento que auxiliará no processo administrativo de seu negócio, pois tal tecnologia oferece as organizações uma maior quantidade e qualidade de informações relevantes ao processo decisório (ANTONELLI, 2009).

O termo BI foi usado pela primeira vez em 1989 por Howard Dressner, em seguida seu colega de pesquisa Gartner Group, utilizou um termo genérico para descrever conceitos e métodos que permitiam melhorar a tomada de decisão de negócios usando o apoio baseado em fatos. Devido à repercussão sofrida, o termo acabou sendo amplamente adotado por profissionais de apoio a tomada de decisão (NEGASH e GRAY, 2008).

Uma outra definição do termo BI é dada por Machado (2010) Inteligência de Negócio ou Business Intelligence (BI) é o conjunto de tecnologias orientadas a disponibilizar informação e conhecimento para as empresas, permitindo uma melhor visão do que está ocorrendo, contribuindo para uma melhor tomada de decisão.

Segundo Machado (2010) BI se trata de um conjunto de tecnologias orientadas a disponibilizar informação, porém tecnologias é um termo demasiadamente genérico. Por isso, as principais dentre estas ferramentas serão abordadas neste trabalho nas próximas seções.

Segundo Shim et al (2012) as principais ferramentas e tecnologias utilizadas pelos sistemas de BI são os Data Warehouses (DW) e as ferramentas On-Line Analytical Processing (OLAP), que surgiram no começo dos anos 90, como novas ferramentas que foram a base dos sistemas de BI. Sendo assim, as próximas seções serão respectivamente DW e OLAP.

2.2 Data Warehouse

Para gerar informações úteis aos administradores, os sistemas de BI precisam de dados históricos sobre as atividades de sua organização. Para manter o histórico de dados, muitas empresas utilizam o DW, que é o armazém de dados que os sistemas de BI utilizam para consultas.

O DW é modelado para auxiliar na tomada de decisão e, portanto possui certas características que o diferenciam de outras formas de armazenamento. Na visão de Immon (2000) elas são:

- **Orientação por Assunto:** as informações ficam agrupadas por assuntos que auxiliam na tomada de decisões pelas organizações. Geralmente representam as atividades realizadas por elas;

- **Varição no Tempo:** os dados armazenados no DW são um retrato dos dados da organização no momento em que foram capturados. Caso os dados sejam alterados na base organizacional, são mantidos os dados de ambos os tempos no DW, antes da alteração e depois da mesma.
- **Integração:** as informações podem ser provenientes de diversas fontes de dados, necessitando a padronização de tais dados antes da carga no DW garantindo unicidade de informações;
- **Dados Não Voláteis:** no DW são executadas somente as operações de carga e consulta dos dados, não havendo a atualização ou exclusão dos dados já carregados.

Além destas características, também há diferenças nos objetivos dos bancos de dados transacionais com o DW. Um banco de dados transacional é construído com o objetivo de manter a consistência dos dados de cada transação. O DW possui como principais objetivos a simplicidade e a velocidade de execução de consultas, e sempre consulta um grupo grande de transações, e não uma transação isolada, visto que o objetivo é controlar as transações realizadas por um grupo de indivíduos.

Devido aos diferentes objetivos, o DW também segue uma modelagem de dados específica. Os bancos transacionais seguem a modelagem entidade-relacionamento, enquanto o DW segue a modelagem dimensional. Esta modelagem considera três aspectos principais, fatos, dimensões e medidas. Um exemplo desta modelagem é demonstrado pela figura 1:

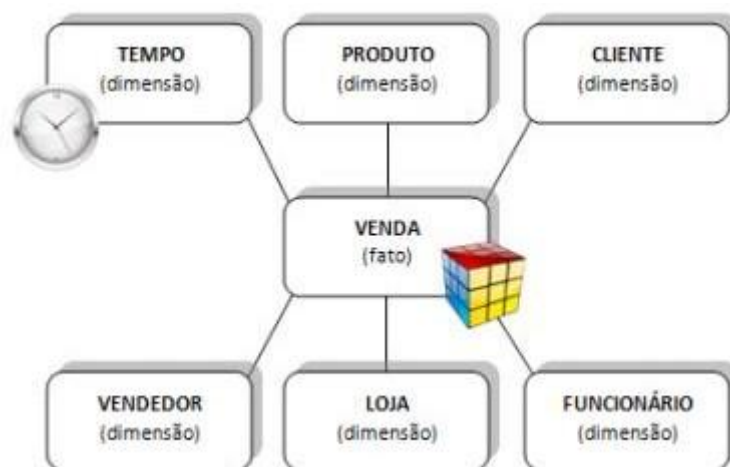


Figura 1: Exemplo de modelo dimensional (SOUZA, 2014).

- Fato: A transação a ser observada. Geralmente são vendas, empréstimos, estoque e etc. Segundo Machado (2010, p.79) “cada fato representa um item, uma transação ou um evento de negócio e é utilizado para analisar o processo de negócio de uma empresa”.
- Dimensão: Participantes do fato. Geralmente uma dimensão é criada para responder alguma dessas perguntas: Quem vendeu? Para quem foi vendido? O que foi vendido? Quando foi vendido? As dimensões são tabelas que permitem que os dados sejam analisados de diferentes formas, são elas que darão o contexto para que seja possível analisar os assuntos de interesse da uma organização. Essas tabelas relatam dados descritivos sobre um determinado elemento do DW, normalmente não havendo atributos numéricos (INMON, 2000).
- Medida: A medida é a quantificação daquilo que está sendo analisado. Medidas comuns são: preço, lucro, quantidade, dias de atraso... As medidas de um fato são atributos que procuram revelar a performance do negócio estudado, são tidas como medidas atributos que armazenam o total vendido em uma compra, o número de itens comprados, o estoque disponível, o custo de uma venda, etc(Souza; Coelho, 2014).

O DW possui dados das bases transacionais da organização, porém seguindo uma nova perspectiva e uma nova modelagem. Deste modo, o papel do DW é integrar dados de diferentes fontes, formatos e modelos (KIMBALL et al., 1998). Para que esta integração seja possível, é necessária a extração, transformação e carga destes dados, processo denominado ETL (*Extraction, Transformation and Load*).

A Figura 2 ilustra o processo de implantação de um *Data Warehouse*, onde os diferentes sistemas da organização passam pelo processo de ETL para serem integrados no DW.

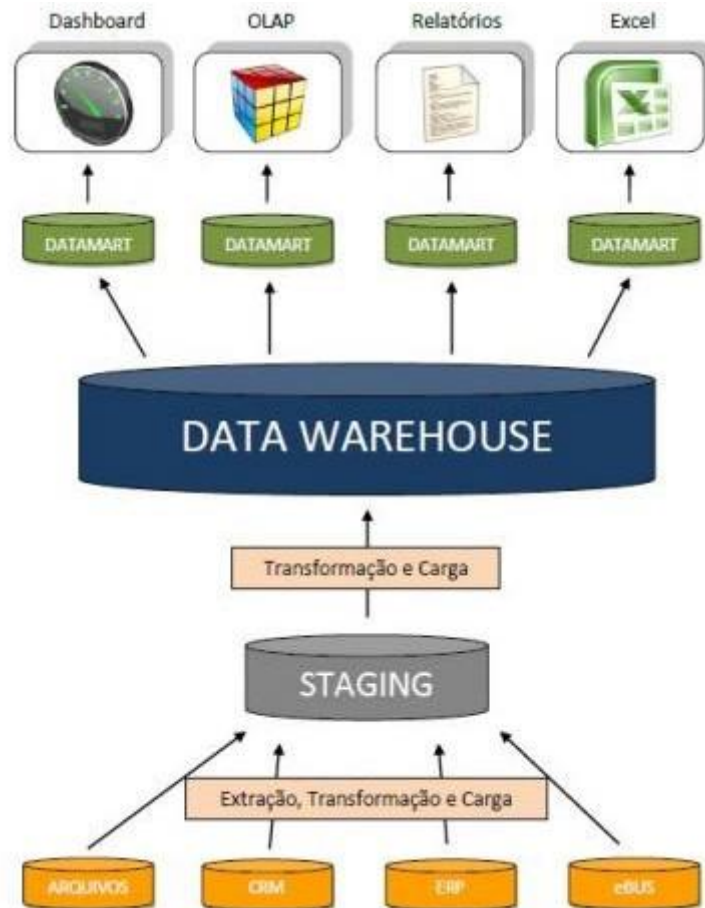


Figura 2: Processo de Implantação de um Data Warehouse Adaptado de BIReady (2011).

Segundo Kimball e Caserta (2004) O ETL é a base do *data warehouse*. Um sistema ETL devidamente projetado extrai os dados dos sistemas de origem, reforça a qualidade dos dados e padrões de consistência, conforma os dados de forma que fontes distintas podem ser usadas juntas. Devido a sua importância, o processo de ETL será o tema da próxima subseção, sendo melhor detalhado nela.

2.2.1 Processo de ETL

O processo de ETL é tido como a base do DW, ele é o processo que demanda mais tempo na criação do DW, e ao mesmo tempo o processo que mais desenvolve o DW. Extração, Transformação e Carga são os três passos que dão nome ao processo, porém durante o processo de ETL há vários sub-processos, como por exemplo:

- Integração de diversas fontes de dados;

- Padronização de dados;
- Criação das tabelas de tempo;
- Discretização de dados;
- Definição do nível de granularidade;
- Criação do modelo dimensional;
- Configurações e regras de carga;
- Carga dos dados no modelo dimensional.

Integração de diversas fontes de dados: Uma das características do DW é a possibilidade de integrar fontes de dados distintas. A ideia da junção das bases é simples, porém fazer isso não é. Os próximos sub-processos é que tornam esta integração possível.

Padronização dos dados: O desenvolvimento de software é um trabalho que depende absolutamente das pessoas que o projetam, nenhuma pessoa é igual a outra, assim como os softwares projetados por elas dificilmente seriam idênticos, por isso a padronização é necessária.

É muito comum encontrar bancos de dados em que os nomes das tabelas são resumidos, por exemplo "categ" ou "prod" que apesar de não estarem completos são o suficiente para entender que se referem as tabelas "categoria" e "produto". Agora imagine duas bases, uma com as tabelas "categ" e "prod" e outra com as tabelas sem o resumo "categoria" e "produto", ao juntar estas bases qual nome deveria ser utilizado?

Este foi um exemplo simples, porém ele se estende a nomes de colunas ("sexo" ou "gênero", por exemplo), ou até mesmo aos dados contidos nelas ("masculino e feminino" ou "m e f"). Agora imagine nomenclaturas e abreviações específicas do negócio da empresa contratante e desconhecidas pelos programadores.

Criação das tabelas de tempo: Um dado sempre presente em uma transação é o tempo, seja pela necessidade de saber quando a venda ocorreu caso ocorram problemas com a venda ou para a regulamentação de notas fiscais e prazos de garantia. Este é um dado essencial no DW, pois o interessante é observar o desenvolvimento da empresa ao longo do tempo, bem como as mudanças nos clientes.

Este dado geralmente está presente como uma das colunas na tabela de vendas ou serviços do banco da organização. No DW é interessante criar uma tabela específica para este dado, pois, caso contrário não seria possível observar todas as mudanças ocorridas em um determinado período de tempo. Isto será melhor detalhado no próximo sub-processo, discretização de dados.

Discretização de dados: O grande objetivo do DW muitas das vezes é descobrir o perfil de clientes e no que estes perfis podem auxiliar a empresa a conseguir mais clientes, vender mais, manter os clientes que possui e etc...

O problema dos bancos de dados transacionais para a análise da informação nesse aspecto é que eles só permitem a visualização de cada transação isoladamente, e não de um grupo de transações pré definido. Pesquisas do tipo "O que os homens entre 25-35 compram?" não pode ser respondidas pelos bancos transacionais, os dados estão com critérios muito restritos para a pesquisa. É possível ver o que homens de 25,26,27...35 anos compram, porém isoladamente nos bancos transacionais, o grupo entre 25-35 não existe. Pesquisas com critérios tão restritivos podem não ser interessantes para se traçar um perfil, afinal, no prazo de um ano o perfil de compra dificilmente muda tão drasticamente.

Esta criação de grupos de pesquisa e critérios que unem estes grupos é papel da discretização de dados. O mesmo vale para datas, por exemplo, é fácil consultar qual mês daquele ano foi mais lucrativo, ou qual dia do mês costuma ser mais lucrativo. Agora pense em um grupo não definido no banco de dados, como por exemplo qual foi o bimestre que mais vendeu? Qual o semestre que mais vendeu? A primeira, segunda, terceira ou quarta semana do mês é que vendem mais? Os períodos mais lucrativos são os que obtiveram promoção ou nossas promoções não surtiram o efeito desejado?

É pela necessidade da discretização de datas que o passo anterior faz-se necessário, pois sem uma nova tabela não é possível este agrupamento.

Definição do nível de granularidade: A granularidade representa o nível de detalhe dos dados armazenados no DW, quanto maior for o nível de detalhes menor será a granularidade e quanto maior for a granularidade menor será o nível de detalhes (INMON, 2005).

Para entender melhor a granularidade, pense na granularidade como o tamanho do grão. Se queremos mais detalhes, devemos na construção do DW salvar os dados na menor migalha possível, por exemplo, detalhando:

- Quem comprou, de que país, estado e cidade é o comprador, quantos anos ele possui, sexo, etc;
- Data da compra em ano, mês, dia, hora, minuto e segundo:
- Quem vendeu, idade, sexo, etc;
- Aonde vendeu (A empresa pode ter várias filiais):
- O que foi vendido, categoria do item vendido, etc.

O maior grão (a junção de todas as migalhas) deve ser considerado o total de vendas, por isso quanto maior a granularidade menor o detalhamento.

Este sub-processo limita o quão detalhadas poderão ser as pesquisas, e ao mesmo tempo diminui o armazenamento necessário para manter os dados no banco, excluindo detalhes desnecessários.

Criação do modelo dimensional: Para a criação do modelo dimensional é que se passa por todos estes sub-processos, é no modelo dimensional que os dados integrados, padronizados e discretizados são armazenados.

Como visto na seção DW, a modelagem dimensional se baseia em fatos, dimensões e medidas. Um modelo dimensional possui o formato parecido com o de uma estrela, com o fato (transação ou evento a ser analisado) sendo o centro e as dimensões se ligando a ele para responder as perguntas: Quem comprou ou vendeu? Quando foi feita a transação? O quê foi vendido ou comprado? Onde, em qual loja ou cidade?

Já as medidas são os atributos do fato que permitem que ele seja medido, respondendo as perguntas: Quanto foi vendido? Quanto lucraramos? Porque aquele cliente comprou? Quanto esperamos vender desta vez? Se lançarmos um produto desta categoria, quantos clientes provavelmente comprariam? O que os clientes que vão a loja na segunda feira compram de diferente dos que vão na sexta feira?

Para responder tantas perguntas como as citadas nos parágrafos anteriores é necessária uma enorme quantia de dados e conseqüentemente um processamento enorme para juntar todos estes dados. O modelo dimensional não é utilizado por acaso, para Kimball, et al (2008) a modelagem dimensional é definida como um método de modelagem de dados para aumentar o desempenho de consultas analíticas e para facilitar seu uso, por meio de um desenho lógico para a estruturação dos dados, sendo intuitivo para os usuários de negócio.

Deste modo, este sub-processo é que cria o banco de dados em que serão armazenados os dados integrados, padronizados e discretizados para serem analisados, utilizando um modelo de análise mais simples e ágil.

Configurações e regras de carga: Antes de armazenar os dados no modelo dimensional é importante definir como estes dados serão coletados, neste sub-processo são definidas rotinas de carga, e criação de chaves substitutas.

Sempre na criação do DW há uma carga de dados enorme, pois o banco do DW está vazio e os dados de diversos bancos devem ser inseridos nele. Porém, isto não é o fim do trabalho, cada nova transação deverá ser utilizada como base de pesquisa no futuro, sendo necessária a constante atualização dos dados do DW.

No entanto, não é interessante atualizar estes dados a todo momento, pois caso esta atualização ocorra em um momento de pico da empresa, tanto o sistema transacional quanto o DW podem sofrer lentidão devido ao volume de dados sendo atualizado. É por esse motivo que são necessárias rotinas de carga e regras de carga.

Já quanto a criação de chaves substitutas, elas são necessárias para que o DW registre tudo que de fato ocorreu ao longo do tempo com precisão. Supondo que o cliente João se case, nos bancos transacionais basta atualizar o status do cliente João de solteiro para casado. O problema disto é que as compras do João quando ainda solteiro não poderiam ser distinguidas de suas compras depois de casado, pois no banco transacional só o existe o João casado, que foi atualizado.

As chaves substitutas são usadas para distinguir o João casado do João solteiro, criando um novo cliente no DW, caso algum campo que determine o perfil seja alterado, como por exemplo, o ato de se casar.

Por ultimo, mas não menos importante, o termo genérico "configurações" se refere a praticamente todos os passos anteriores, pois eles são configurações pré-carga de dados.

Carga dos dados no modelo dimensional: Este é o ultimo dos sub-processos do processo de ETL, sendo ele a inserção e atualização dos dados no DW a partir dos dados transacionais. Após a conclusão deste ultimo sub-processo pode-se dizer que o DW está pronto, agora só restando a consulta aos dados nele contidos.

Assim que os dados estão disponíveis para consulta eles são acessados por ferramentas OLAP (*Online Analytical Processing*) que são as ferramentas

responsáveis por apresentar os dados ao usuário e permitir que o usuário realize diversas formas de consulta.

2.3 Ferramentas OLAP

De acordo com Machado (2010, p.85) “ferramentas OLAP são as aplicações às quais os usuários finais têm acesso para extrair os dados de suas bases e construir os relatórios capazes de responder às suas questões gerenciais”. Em outras palavras, são as ferramentas que se conectam ao DW e permitem a visualização e análise dos dados nele contidos.

As ferramentas OLAP foram projetadas para trabalhar com a modelagem dimensional dos DWs, e por isso geralmente possuem as seguintes funcionalidades:

- **Slide and Dice:** Com o Slice and Dice o usuário pode facilmente trocar as dimensões que estão sendo analisadas, ou a ordem em que estão sendo analisadas.
- **Drill Down and Roll Up:** Estas opções permitem que o usuário aumente ou diminua o nível de detalhe das dimensões, por exemplo, buscar por ano, mês ou dia na dimensão de tempo.
- **Geração de gráficos e relatórios:** O sistema de BI como um todo também tem como função expor o desempenho do negócio analisado, sendo assim é natural a necessidade da geração de gráficos e relatórios de acordo com os dados selecionados.

Estas são as três funcionalidades mais comuns, porém é claro que alguns OLAPs mais completos podem possuir mais opções. Apesar das ferramentas OLAP disponíveis no mercado geralmente possuírem funcionalidades muito parecidas, há uma quarta funcionalidade que divide estas ferramentas em tipos de ferramentas OLAP, esta funcionalidade é a conexão com o banco de dados.

Os tipos de OLAP existentes são vários e a tendência é que com o tempo mais possibilidades surjam, porém para os fins deste trabalho serão apresentados somente três tipos, ROLAP, MOLAP e HOLAP:

- **Multidimensional Online Analytical Processing (MOLAP):** OLAPs para bancos de dados multidimensionais, que utilizam a linguagem MDX (*Multi-Dimensional Expression*) para consultas.

- **Relational Online Analytical Processing (ROLAP):** OLAPs para bancos de dados relacionais. Estes OLAPs acessam a base de dados relacional e convertem a linguagem SQL (*Structured Query Language*) dos bancos relacionais na linguagem MDX.
- **Hybrid Online Analytical Processing (HOLAP):** OLAPs híbridos, que suportam tanto conexões a bancos relacionais quanto bancos multidimensionais.

2.4 Agile ROLAP

A metodologia Agile ROLAP é uma metodologia ágil de desenvolvimento de BI. Ela foi baseada na metodologia Agile BI, porém, procura amenizar os problemas da mesma (SOUZA; MENOLLI; COELHO, 2014),

O grande problema da metodologia Agile BI é a impossibilidade de utilizá-la em conjunto com as ferramentas OLAP projetadas para os DWs tradicionais. O Agile ROLAP aproveita a ágil implantação da metodologia Agile BI e resolve o problema com as ferramentas OLAP.

A metodologia Agile ROLAP acessa o banco de dados ainda em formato relacional, e através de um servidor ROLAP as consultas SQL são convertidas em MDX. Isto possibilita que a metodologia Agile ROLAP seja utilizada com qualquer OLAP baseado em MDX.

A Figura 3 mostra como é organizada a estrutura Agile ROLAP.

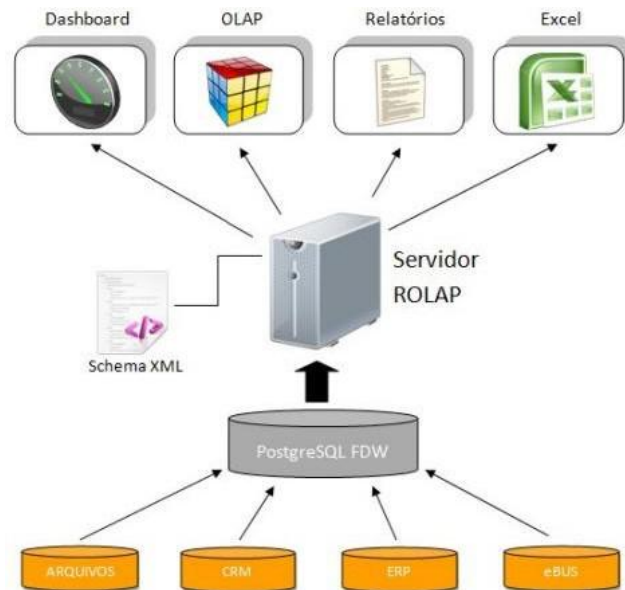


Figura 3: Estrutura Agile ROLAP (SOUZA, 2014).

Já quanto a redução no tempo e investimento necessários para a criação de um sistema de BI, o ato de acessar as bases ainda em formato relacional reduz o processo de ETL e consequentemente o tempo e investimento necessários para se obter retorno sobre o investimento em um sistema de BI (SOUZA; MENOLLI; COELHO, 2014).

A metodologia Agile ROLAP passa por 10 passos para ser implantada completamente. Estes passos são:

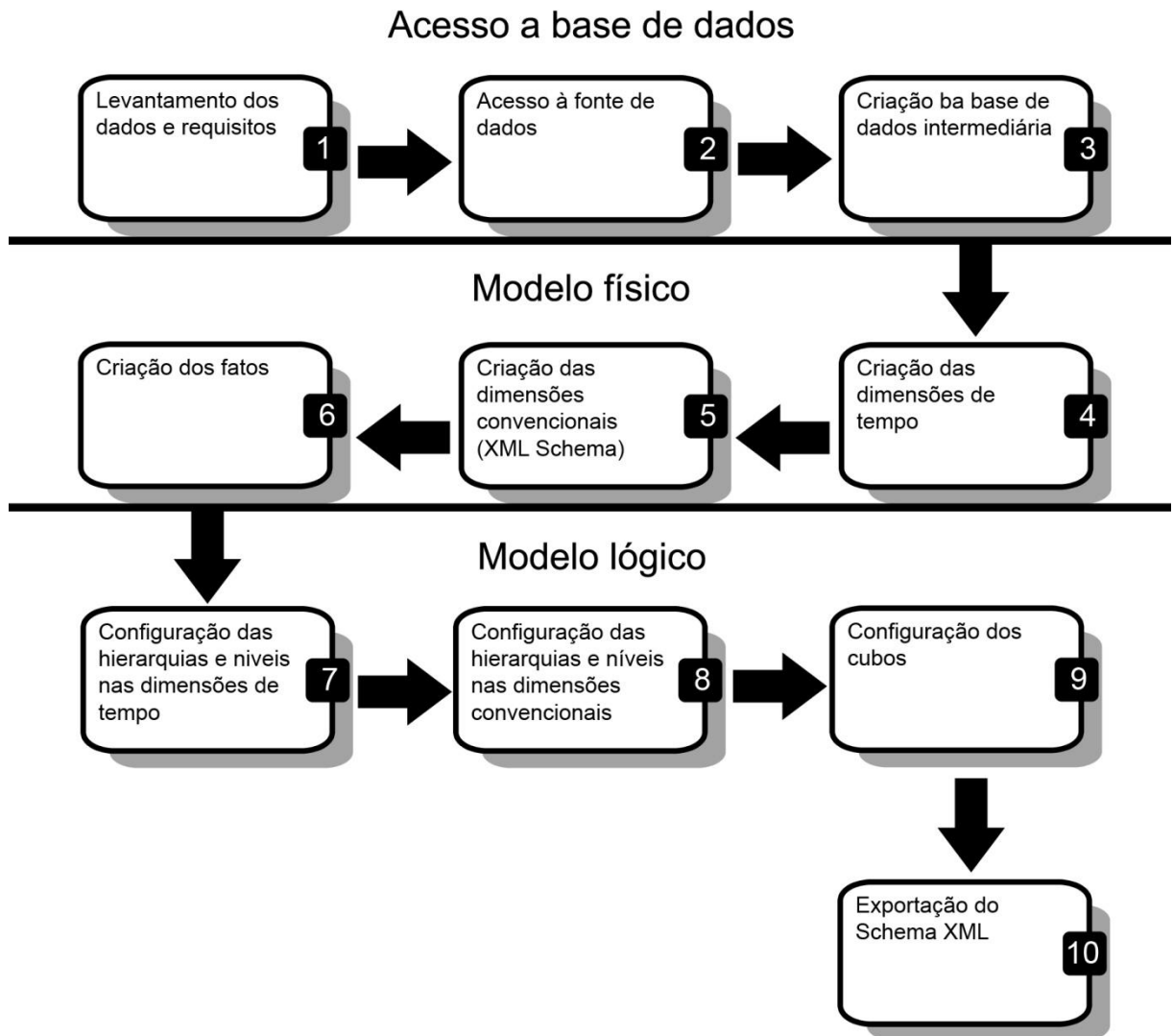


Figura 4: Passos para a implantação do Agile ROLAP adaptados de Souza; Menolli; Coelho, (2014).

1º Passo – **Levantamento dos dados e requisitos:** Todo sistema a ser construído passa por esta etapa, caso contrário não se sabe o que deve ser construído.

2º Passo – **Acesso a fonte de dados:** É nesta etapa que a equipe de desenvolvimento tem acesso aos dados da organização. A partir deste momento a equipe de desenvolvimento poderá dizer se possui dados suficientes para atender aos requisitos demandados pela organização ou não.

3º Passo – **Criação da base de dados intermediária:** O Agile ROLAP utiliza dos dados da organização ainda em formato relacional e sem transformações. Porém, para não sobrecarregar o banco de dados da organização e evitar alterações

no banco transacional, é recomendada a criação de uma base de dados intermediária através de um *Foreign Data Wrapper* (FDW).

4° Passo – **Criação das dimensões de tempo:** Assim como em um DW tradicional, no Agile ROLAP a criação de tabelas tempo é necessária, pois, assim é possível se fazer uma análise temporal eficiente e menos restritiva.

Também no quarto passo se inicia a construção do *Schema XML*, arquivo de configuração que faz a ligação entre o banco de dados e o servidor ROLAP. Mais especificamente são configuradas as tabelas de tempo no *Schema XML*.

5° Passo – **Criação das dimensões convencionais:** Este passo se trata do mapeamento das dimensões convencionais para o *Schema XML*, visto que não há a necessidade da criação de tabelas físicas para as dimensões convencionais na metodologia Agile ROLAP.

6° Passo – **Criação dos fatos:** Os fatos na metodologia Agile ROLAP são gerados como uma *View Materializada*. Portanto, este passo se refere a criação de *Views* de fato.

7° Passo – **Criação das hierarquias e níveis das dimensões de tempo:** No quarto passo a configuração XML das dimensões tempo foi iniciada, e no sétimo ela é concluída, configurando as hierarquias e níveis das dimensões de tempo.

8° Passo – **Criação das hierarquias e níveis das dimensões convencionais:** No quinto passo se deu início a configuração do *Schema XML* para as dimensões convencionais, neste passo essa configuração é finalizada com a adição das hierarquias e níveis.

9° Passo – **Criação dos cubos:** O nono passo transforma os fatos em cubos no *Schema XML*, adicionando medidas para analisar o negócio e informando quais dimensões farão parte daquele fato.

10° Passo - **Exportação do *Schema XML*:** Neste passo as configurações geradas pelos outros passos são exportadas para o arquivo *Schema XML*. A partir deste ponto é possível ligar a base de dados ao sistema OLAP e fazer consultas, dando como implantado o sistema de BI segundo a metodologia Agile ROLAP.

Principais vantagens e desvantagens: Esta metodologia apresenta como principais vantagens a redução no tempo e investimento necessários para a construção de um DW. A redução no tempo se deve ao fato da minimização do processo de ETL, enquanto que a redução no investimento se deve a própria

redução no tempo. Muitas empresas de desenvolvimento de software calculam o custo do sistema de acordo com o tempo de trabalho necessário para seu desenvolvimento.

Como desvantagem principal a metodologia Agile ROLAP apresenta a velocidade de execução das consultas. Segundo Kimball e Ross (2002) a modelagem dimensional tem como foco a simplicidade, para que possa ser entendida e processada de maneira simples. Assim um banco de dados modelado de maneira dimensional, na maioria dos casos, irá executar consultas mais rapidamente do que um banco modelado de maneira relacional, no entanto, essa desvantagem não é suficiente para inviabilizar a metodologia Agile ROLAP.

A metodologia Agile ROLAP não veio para substituir o DW tradicional, porém, segundo Souza, Menolli e Coelho (2014) ela tem como objetivo atender a empresas de menor porte e possibilitar que essas empresas consigam usufruir das tecnologias de BI com menor investimento e mais rapidamente. Papel que ela tende a ser mais eficaz em desempenhar do que a metodologia tradicional.

3. Ferramentas utilizadas

Esta seção faz uma breve descrição das ferramentas utilizadas no trabalho e o motivo pelo qual foram escolhidas dentre as outras disponíveis.

3.1 PostgreSQL

O PostgreSQL é um dos bancos de dados relacionais mais utilizados no mundo atualmente e segundo a DB Engines (2016) ele é o quarto banco de dados relacional mais popular. Nas palavras de seus criadores “o PostgreSQL é um poderoso banco de dados relacional. Ele possui mais de 15 anos de ativo desenvolvimento e uma bem testada arquitetura que conquistou uma forte reputação por confiabilidade, integridade e exatidão” (POSTGRESQL, 2016).

O PostgreSQL foi escolhido como banco de dados a ser utilizado neste trabalho por diversos fatores, como por exemplo:

- É um banco de dados gratuito;
- É um banco relacional. Já que metodologia Agile ROLAP foi projetada para bancos relacionais este era um requisito para a implantação da metodologia;
- Ele possui a tecnologia *Foreign Data Wrapper*, recomendada para a criação da base de dados intermediária da metodologia Agile ROLAP.
- A ferramenta Base de dados estrangeira foi desenvolvida para funcionar com o *Foreign Data Wrapper*, e a alteração do banco de dados poderia impossibilitar os testes da ferramenta.

3.2 Kettle - Data integration

O Data integration, também chamado de Kettle, é uma ferramenta java de código aberto da Pentaho e ela tem como objetivo auxiliar o desenvolvedor de DW no processo de ETL. Sua utilização baseia-se na transformação, onde o usuário deve criar uma transformação e dentro da transformação definir quais passos ou etapas serão necessárias para que aquela transformação ocorra.

Os passos ou etapas necessárias para que uma transformação ocorra foram transformados em *plugins* na ferramenta Kettle, onde a cada *plugin* configurado o usuário está um passo mais próximo de realizar a transformação desejada. Além de configurar é necessário ligar os plugins como se fossem passos, de maneira que o programa entenda qual a ordem dos passos. Um exemplo de transformação pode ser observado na Figura 5, onde inicia-se a transformação com a entrada de dados (*plugins de input*) ocorrem as transformações, e finalmente os dados são exportados através dos *plugins de saída (outputs)*.

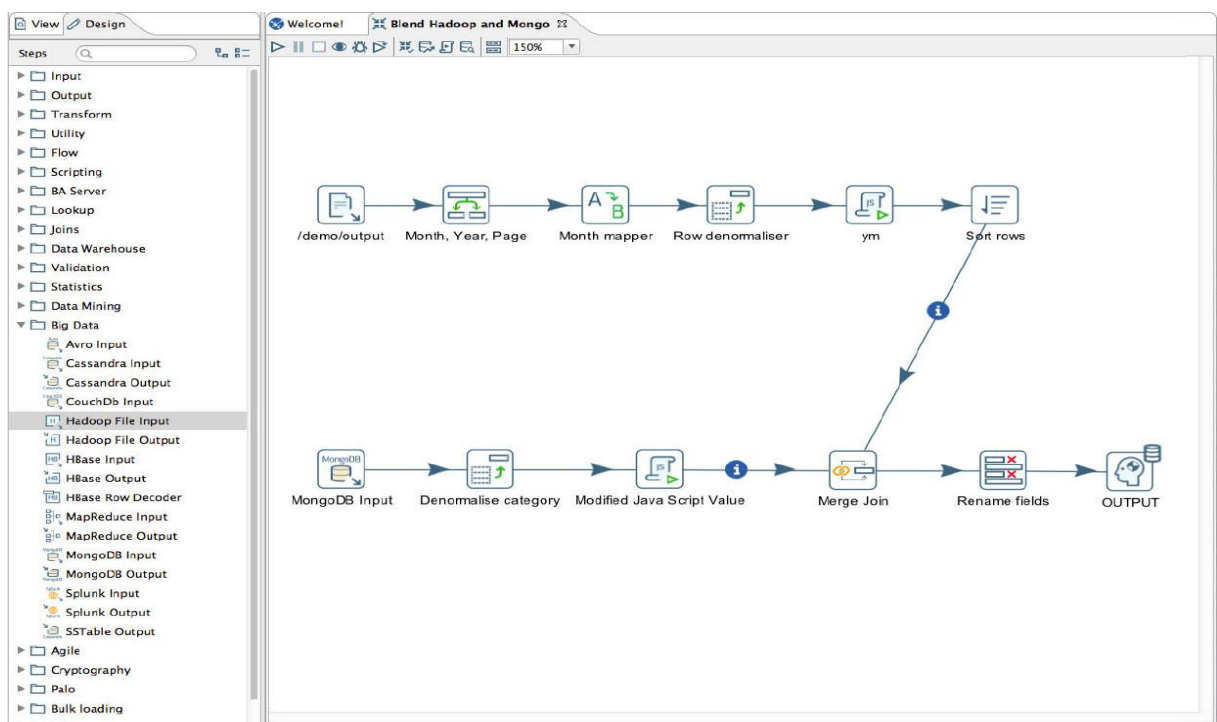


Figura 5: Exemplo de transformação no Kettle (PENTAHO, 2016).

Já quanto ao motivo deste trabalho escolher o *Data integration* para ser utilizado, ele não foi escolhido, ele foi necessário para o trabalho. Isto se deve ao fato das ferramentas desenvolvidas para auxiliar a implantação da metodologia Agile ROLAP se tratarem de *plugins* da ferramenta *Data integration*. Estes *plugins* serão detalhados na próxima subseção.

3.3 Plugins Agile ROLAP

Os plugins desenvolvidos para auxiliar a metodologia Agile ROLAP foram desenvolvidos para a ferramenta *Data integration* da Pentaho, e foram desenvolvidos em um projeto de pesquisa do centro de ciências tecnológicas da UENP¹. Esta escolha foi uma tentativa de aproximar as metodologias tradicional e Agile ROLAP, bem como aproveitar a estrutura da ferramenta *Data integration*.

Os *plugins* desenvolvidos foram Base de dados estrangeira, Dimensão, Tempo, Feriado, Fato, XMLOut, Dimensão lógica, Dimensão lógica de tempo, Cubo e SchemaOut. Com exceção dos *plugins* Feriado e XMLOut, cada *plugin* desenvolvido deveria auxiliar um passo da metodologia Agile ROLAP.

Para trocar informações e criar artefatos os *plugins* deveriam ser ligados dentro da ferramenta *Data integration* como na Figura 6:

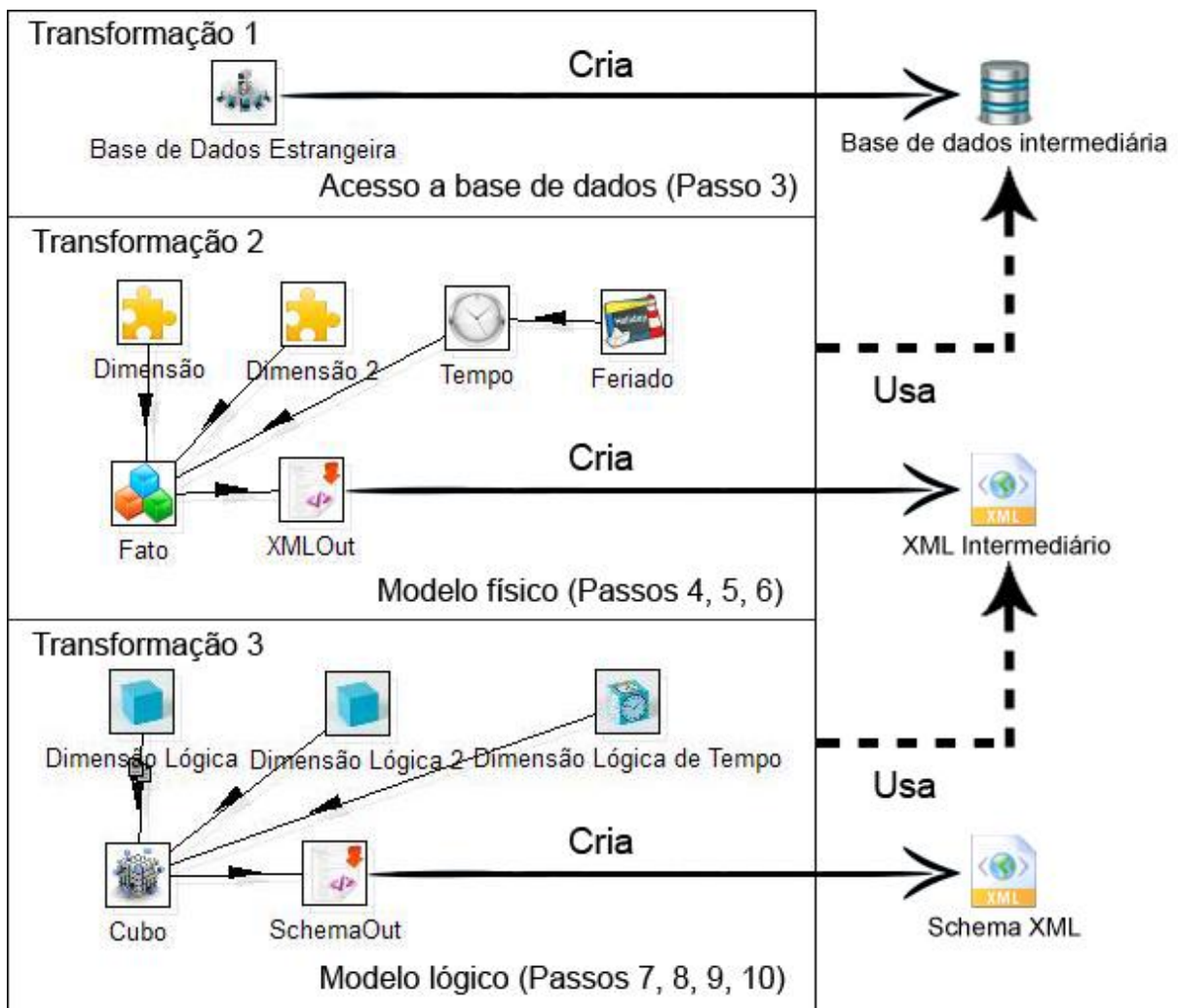


Figura 6: Organização dos *plugins* Agile ROLAP.

Agora uma explicação de cada *plugin* individualmente:

¹ Agile ROLAP – <http://projetos.uenp.edu.br/agilerolap/>

Base de dados estrangeira: Este plugin é responsável por auxiliar na execução do passo três da metodologia Agile ROLAP, criação da base de dados intermediária. Para utilizar este *plugin* o usuário deve previamente adicionar ao PostgreSQL a extensão *postgresFDW* e criar um *server foreign data wrapper*.

Já quanto à configuração da ferramenta em si, basta configurar a conexão da base de dados de origem e destino e depois selecionar as tabelas que irão para o FDW.

Feriado: Esta ferramenta não possui um passo diretamente associado a ela na metodologia Agile ROLAP, porém, pode-se dizer que ela faz parte do passo quatro, pois ela é uma configuração das dimensões de tempo. Nesta ferramenta é possível adicionar quais dias serão feriados e uma descrição para cada feriado.

Tempo: Ferramenta referente ao passo quatro da metodologia Agile ROLAP, criação das dimensões de tempo. A ferramenta tempo cria uma tabela de tempo com 24 formatos de data, onde a menor granularidade é referente aos dias.

Dimensão: A ferramenta Dimensão permite que o usuário crie um mapeamento de tabelas que gerariam uma dimensão no modelo dimensional. Esta ferramenta destina-se a auxiliar o usuário na execução do passo cinco da metodologia Agile ROLAP, criação das dimensões convencionais.

Fato: Diretamente relacionada com o passo seis da metodologia. Esta ferramenta liga as configurações geradas nos *plugins* anteriores e tenta facilitar a geração de uma tabela de fato, ou mais tecnicamente, uma *view* materializada em forma de fato. O motivo pela qual é criada uma *view* materializada e não uma tabela pode ser encontrado no trabalho de Souza (2014), mas, em termos de configuração da ferramenta o usuário dificilmente sentirá diferença neste detalhe técnico.

A utilização da ferramenta consiste em selecionar as chaves das dimensões que se ligariam com o fato e apontar quais chaves pertencem a quais dimensões.

XMLOut: Este *plugin* tem por objetivo gerar um arquivo XML intermediário que liga a parte física com a lógica. Ele não é um passo Agile ROLAP, pois ele foi gerado visando separar o desenvolvimento de *plugins* em duas partes, de forma que os desenvolvedores pudessem desenvolver e testar seus *plugins* sem a necessidade que o *plugin* do outro programador estivesse pronto.

O XML gerado por este *plugin* é utilizado pelos *plugins* Dimensão, Dimensão lógica de tempo, Cubo e SchemaOut. Deste modo todas as configurações físicas

são reconhecidas pelo mapeamento contido nesse XML, não necessitando de conexão com o banco de dados para os próximos *plugins*.

Dimensão lógica de tempo: A dimensão lógica de tempo permite a configuração de hierarquias e níveis para as dimensões de tempo, tarefa descrita pelo passo sete da metodologia.

Dimensão lógica: Permite a configuração de hierarquias e níveis das dimensões convencionais, como descrito no passo oito da metodologia Agile ROLAP.

Cubo: A ferramenta cubo é responsável pela configuração de medidas e agregadores, bem como informar quais dimensões serão usadas no cubo. Estas tarefas são descritas pelo passo nove da metodologia Agile ROLAP.

SchemaOut: Este é o último *plugin* e o último passo da metodologia Agile ROLAP. O *plugin* SchemaOut exporta as configurações de todos os *plugins* anteriores (exceto Base de dados estrangeira) no formato *Mondrian Schema*. O *plugin* Base de dados estrangeira é o único a não ser integrado no SchemaOut, isto se deve a seu papel se tratar da criação da base de dados intermediária, não tendo relação com o Mondrian Schema ou o modelo dimensional.

Ao passar por este último passo/*plugin* pode-se dizer que a metodologia Agile ROLAP está implantada.

3.4 Mondrian

Como descrito na metodologia Agile ROLAP, para utilizar uma aplicação OLAP em um banco de dados relacional é necessário um servidor ROLAP. O servidor ROLAP é responsável por converter consultas SQL em consultas MDX, onde o SQL é linguagem entendida pelos bancos de dados relacionais e MDX é a linguagem entendida pelos bancos multidimensionais e ferramentas OLAP.

O Mondrian será o servidor ROLAP deste trabalho. Assim como a ferramenta *Data integration* o Mondrian pertence à Pentaho e não foi escolhido para o trabalho, ele foi necessário. A própria metodologia Agile ROLAP cita a exportação de um arquivo de configuração do Mondrian como sendo um dos passos da metodologia. Conseqüentemente os *plugins* geram como artefato final este arquivo de configuração, denominado Mondrian Schema ou Schema XML. Dito isto fica

evidente que este trabalho não poderia ser realizado sem a utilização do Mondrian como servidor ROLAP.

3.5 Pivot4j

O Pivot4j é uma ferramenta OLAP que utiliza como servidor ROLAP o Mondrian. O Pivot4j poderia ser considerado uma ferramenta OLAP incompleta por alguns usuários. Isto por seu baixo tratamento de erros, visto que a ferramenta possui visíveis bugs em sua versão atual (versão 0.9). Outras opções gratuitas seriam o Jpivot e o openL, porém, ambas foram descontinuadas, restando o Pivot4j como melhor opção. Por isso o Pivot4j foi a ferramenta OLAP utilizada neste trabalho, porém, nada impede que ela seja substituída no futuro, pois, desde que a ferramenta OLAP aceite o Mondrian como servidor ela poderá ser utilizada.

4. Metodologia

Tartuce (2006) definiu a metodologia como sendo um corpo de regras e procedimentos, estabelecidos para realizar uma pesquisa científica. Dito isto, esta seção tem o objetivo de mostrar quais as regras e procedimentos estão sendo aplicados a esta pesquisa.

A pesquisa científica costuma ser dividida em quatro pontos de vista:

1. Quanto à sua natureza;
2. Quanto à abordagem do problema;
3. Quanto aos seus objetivos ou fins;
4. Quanto aos seus procedimentos técnicos;

Cada um destes quatro pontos de vista serão abordados em um sub-tópico.

Além dos quatro sub-tópicos já descritos, também há um quinto sub-tópico denominado “Etapas”, ele descreve passo a passo como será desenvolvido o trabalho.

Natureza: Esta pesquisa procura resolver o problema da escassez de testes das ferramentas auxiliares da metodologia Agile ROLAP

Do ponto de vista da natureza da pesquisa, esta é uma pesquisa aplicada, pois, a atividade descrita no parágrafo anterior compreende a definição de pesquisa aplicada segundo Barros e Lehfeld (2000, p.78), a motivação da pesquisa aplicada é a produção do conhecimento obtido para aplicação dos seus resultados com o objetivo de “contribuir para fins práticos, visando a solução mais ou menos imediata do problema encontrado com a realidade”.

Abordagem do problema: Do ponto de vista da abordagem do problema, a pesquisa é qualitativa, devido às variáveis a serem observadas:

- A efetividade da implantação da metodologia Agile ROLAP;
- O que pode ser melhorado na implantação da metodologia e suas ferramentas.

Ambas variáveis são difíceis de serem mensuradas e consideradas subjetivas, e para Minayo (2001) a pesquisa qualitativa preocupa-se com aspectos da realidade que não podem ser quantificados.

Objetivos: De acordo com Gil, (2007)

Uma pesquisa exploratória tem como objetivo proporcionar maior familiaridade com o problema, com vistas a torná-lo mais explícito ou a

construir hipóteses. A grande maioria dessas pesquisas envolve: (a) levantamento bibliográfico; (b) entrevistas com pessoas que tiveram experiências práticas com o problema pesquisado; e (c) análise de exemplos que estimulem a compreensão.

Do ponto de vista dos objetivos desta pesquisa, ela pode ser classificada como exploratória, dado o fato que, para esta pesquisa foi necessário uma pesquisa bibliográfica em artigos e monografias.

Procedimentos técnicos: Quanto aos procedimentos técnicos, esta pesquisa pode ser considerada bibliográfica, pois, “A pesquisa bibliográfica é feita a partir do levantamento de referências teóricas já analisadas, e publicadas por meios escritos e eletrônicos, como livros, artigos científicos, páginas de web sites” (FONSECA, 2002, p. 32).

Etapas: Por fim, para que este trabalho seja metodologicamente científico, é necessário definir como ele será desenvolvido em seus mínimos detalhes, explicitando quais passos serão dados e quais caminhos serão seguidos, desde sua iniciação até a sua conclusão.

As etapas que definem os passos metodológicos deste trabalho são:

- Definir qual base de dados será utilizada no trabalho;
- Criar o modelo dimensional da base de dados;
- Utilizar as ferramentas auxiliares;
- Implantar a metodologia;
- Explicitar os problemas encontrados ao utilizar as ferramentas;
- Explicitar os problemas encontrados na implantação da metodologia;
- Expor o que se concluiu a partir do trabalho.

5. Desenvolvimento

O presente capítulo descreve como foi realizada a implantação da metodologia Agile ROLAP neste trabalho. Para a implantação foi utilizada a base de avaliação institucional da UENP. Esta base de dados contém os as respostas das avaliações institucionais realizadas anualmente pelos alunos, professores, estagiários e técnicos, onde são avaliados diversos aspectos referentes à instituição.

O capítulo encontra-se dividido em três seções centradas nos temas: Acesso a base de dados, Modelo físico e Modelo lógico, cada uma representando uma parte da metodologia Agile ROLAP, como representado pela figura 4.

5.1 Acesso a base de dados

O acesso a base de dados se trata do início da implantação da metodologia, onde são levantados os dados e requisitos, ocorre o acesso a fonte de dados e é criada uma base de dados intermediária onde será implantada de fato a metodologia. Os próximos tópicos falam de como cada um dos passos desta etapa foram executados neste trabalho.

5.1.1 Levantamento dos dados e requisitos – 1º Passo

Com o objetivo de ajudar na melhoria da Universidade Estadual do Norte do Paraná e seguindo a sugestão do professor Glauco Carlos Silva, a base de dados escolhida para este trabalho foi a base de avaliação institucional da UENP. Deste modo a realização deste trabalho não só auxiliaria na melhoria da metodologia Agile ROLAP, como construiria um ambiente de BI que poderia auxiliar os administradores da UENP a tomar decisões mais assertivas.

5.1.2 Acesso a fonte de dados – 2º Passo

A base foi disponibilizada em forma de um arquivo Microsoft Office Access, sendo posteriormente transferida para uma base de dados postgresSQL versão 9.3.

Esta migração não alterou ou transformou em nada a estrutura ou dados contidos na base original. Esta migração ocorreu, pois o postgres versão 9.3 possuía a tecnologia FDW, que é uma tecnologia não obrigatória, mas recomendada para a implantação da metodologia Agile ROLAP.

Após a migração da base ainda em formato relacional e sem alterações, foi necessário desenhar o modelo dimensional da base, pois, mesmo não possuindo um modelo dimensional fisicamente, a metodologia Agile ROLAP ainda precisa simulá-lo para que a base seja interpretada corretamente pela ferramenta ROLAP a ser utilizada para visualizar os dados posteriormente.

Durante este *design* do modelo dimensional foram definidos cinco fatos. Eles serão apresentados a seguir em forma de figuras. Também é importante dizer que os caracteres especiais (acentos e cedilhas) presentes nas medidas foram descartados, pois a ferramenta ROLAP utilizada neste trabalho não os interpreta corretamente.

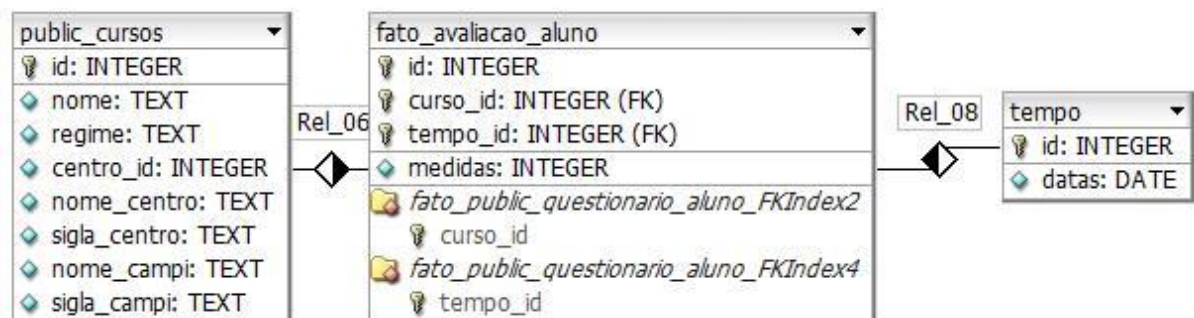


Figura 7: Fato avaliacao aluno.

A Figura 7 mostra como foi modelado o fato avaliacao aluno. Suas medidas como se pode observar não são mostradas. Isto se deve ao fato de elas serem muitas e sua inserção no modelo poderia atrapalhar a legibilidade da figura. Por este motivo elas serão apresentadas em forma de quadros como no quadro 1, que representa as medidas do fato avaliacao aluno.

Quadro 1: Medidas do fato avaliacao aluno.

Coluna	Nome	Formula	Agregador
su_q1	Media estagios e projetos de extensao atendem e beneficiam a comunidade externa.	coluna!=0	avg

su_q2	Media meios de comunicacao da UENP atendem as necessidades de divulgacao.	coluna!=0	avg
su_q3	Media UENP incentiva a iniciacao cientifica.	coluna!=0	avg
su_q4	Media imagem publica da UENP na sociedade.	coluna!=0	avg
su_q5	Media contribuicao UENP para a sociedade.	coluna!=0	avg
si_q1	Media acessa portal da UENP frequentemente.	coluna!=0	avg
si_q2a	Media salas sao arejadas.	coluna!=0	avg
si_q2b	Media salas sao iluminadas.	coluna!=0	avg
si_q2c	Media salas tem boa acustica.	coluna!=0	avg
si_q2d	Media sao amplas para comportar o numero de alunos.	coluna!=0	avg
si_q2e	Media salas apresentam mobiliario adequado.	coluna!=0	avg
si_q2f	Media salas dispoem de recursos audiovisuais suficientes.	coluna!=0	avg
si_q3a	Media biblioteca possui acervo com qualidade.	coluna!=0	avg
si_q3b	Media biblioteca possui numero suficiente de exemplares.	coluna!=0	avg
si_q3c	Media biblioteca possui instalacoes satisfatorias.	coluna!=0	avg
si_q3d	Media eficiencia sistema de busca de livros da biblioteca.	coluna!=0	avg
si_q4	Media numero de equipamentos atendem aos de alunos.	coluna!=0	avg
si_q5	Media condicoes dos equipamentos atendem aos de alunos.	coluna!=0	avg
si_q6	Media laboratorios de informatica atendem as necessidades.	coluna!=0	avg

si_q7a	Media limpeza salas de aula.	coluna!=0	avg
si_q7b	Media conservacao salas de aula.	coluna!=0	avg
si_q7c	Media limpeza do campus.	coluna!=0	avg
si_q7d	Media conservacao do campus.	coluna!=0	avg
si_q7e	Media limpeza dos banheiros do campus.	coluna!=0	avg
si_q7f	Media conservacao dos banheiros do campus.	coluna!=0	avg
si_q7g	Media iluminacao do campus.	coluna!=0	avg
si_q7h	Media seguranca do campus.	coluna!=0	avg
sa_q1a	Media atuacao do coordenador de colegiado de curso.	coluna!=0	avg
sa_q1b	Media atuacao do diretor de centro de estudos.	coluna!=0	avg
sa_q1c	Media atuacao da diretoria do campus / unidade.	coluna!=0	avg
sa_q1d	Media atuacao do Vice-reitor em exercicio da reitoria.	coluna!=0	avg
sa_q2	Media atendimento das secretarias academicas.	coluna!=0	avg
sa_q3	Media atendimento da telefonista.	coluna!=0	avg
sa_q4	Media atendimento do protocolo.	coluna!=0	avg
sa_q5	Media atendimento do servicos de fotocopia.	coluna!=0	avg
sa_q6	Media atendimento da biblioteca.	coluna!=0	avg
cpa_q1	Media avaliacao da avaliacao.	coluna!=0	avg
chave_ acesso	Quantidade alunos.		distinct-count
chave_ acesso	Quantidade alunos finalizaram algum questionario.	fim is not null	distinct-count
chave_ acesso	Quantidade alunos nao finalizaram algum questionario.	fim is null	diistinct-count

Para visualizar todas as medidas dos outros fatos o leitor deve consultar o apêndice A. As Figuras 8, 9, 10 e 11 representam os outros quatro fatos restantes:

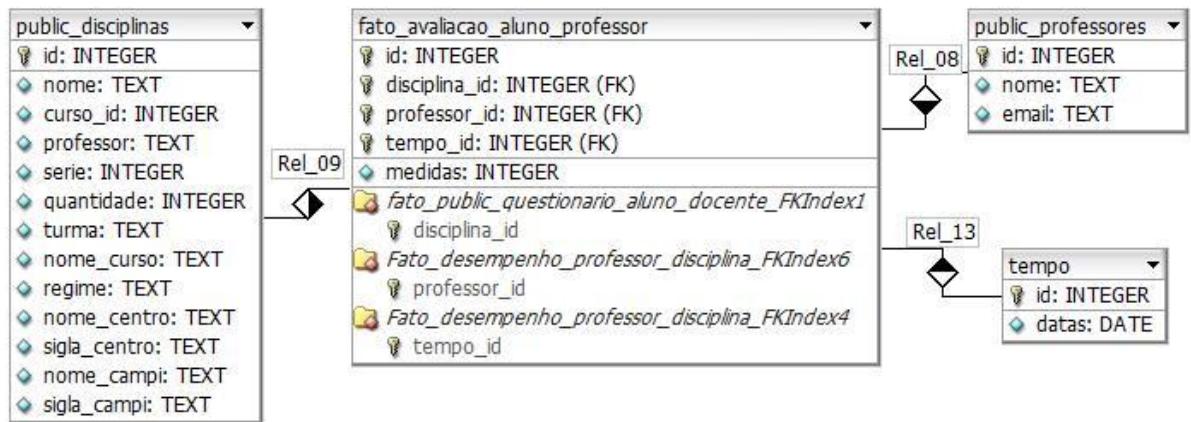


Figura 8: Fato avaliacao aluno_professor.



Figura 9: Fato avaliacao docente.

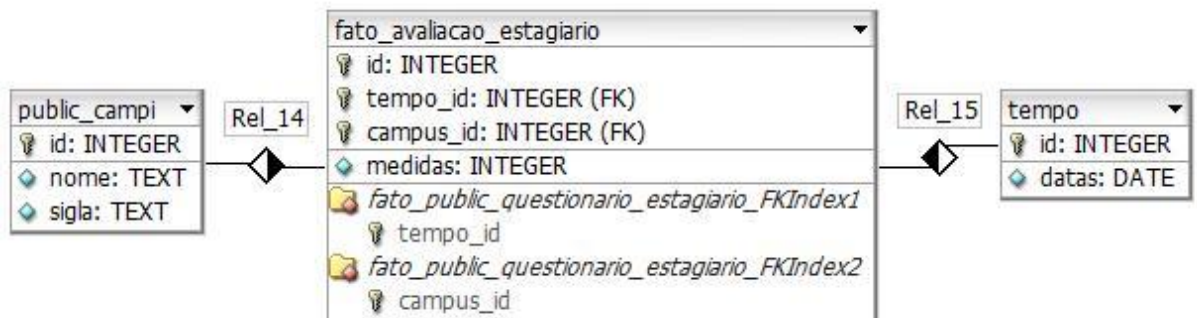


Figura 10: Fato avaliacao estagiario.

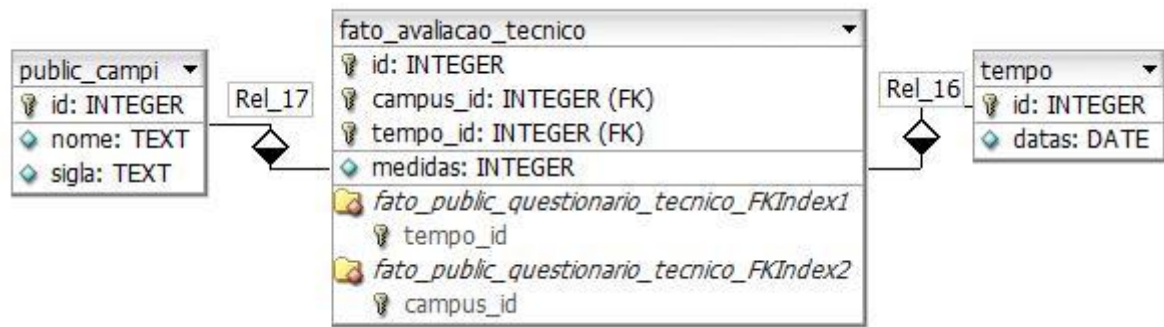


Figura 11: Fato avaliacao tecnico.

5.1.3 Criação da base de dados intermediária – 3º Passo

Toda a base de avaliação institucional da UENP foi transferida para um postgresFDW, ou seja, foram criadas tabelas estrangeiras no postgresFDW que referenciam as tabelas originais. Deste modo a base original não seria alterada e o processamento das análises do ambiente de BI poderia ser realizado em outro computador.

Para auxiliar este passo foi criado o *plugin* base de dados estrangeira. A utilização deste *plugin* é mostrada pela Figura 12:

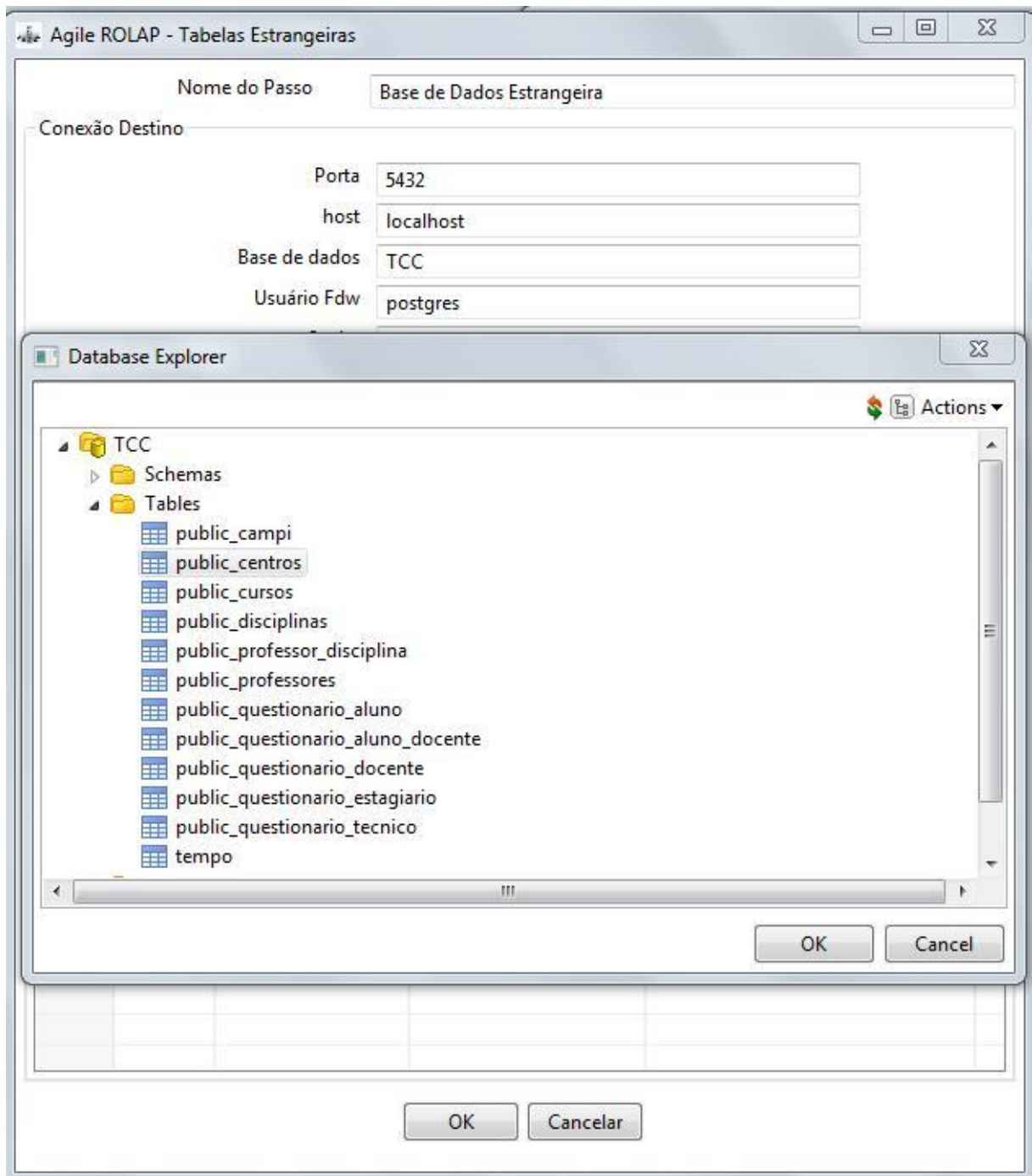


Figura 12: Configuração do *plugin* base de dados estrangeira.

5.2 Modelo físico

A etapa modelo físico pode ser interpretada como a simulação da existência do modelo dimensional. Nesta etapa ocorre o mapeamento das tabelas e colunas que gerariam as dimensões e fatos em um modelo físico real. Deste modo, quando o usuário dá o nome a uma dimensão, por exemplo, ele está dando nome ao

mapeamento que geraria aquela dimensão no modelo dimensional real, caso ele existisse.

Esta etapa é dividida nos passos criação das dimensões de tempo, criação das dimensões convencionais e criação dos fatos. Mantendo a fidelidade a metodologia Agile ROLAP estes passos serão os próximos tópicos desta seção, onde é descrito como estes passos foram executados neste trabalho.

5.2.1 Criação das dimensões de tempo – 4º Passo

Foi criada uma única dimensão de tempo, pois na base de avaliação institucional o único momento em que o tempo é um dado é o ano em que a avaliação foi realizada.

O *plugin* dedicado a auxiliar na execução deste passo se chama tempo e a configuração do mesmo será apresentada a seguir pela Figura 13.

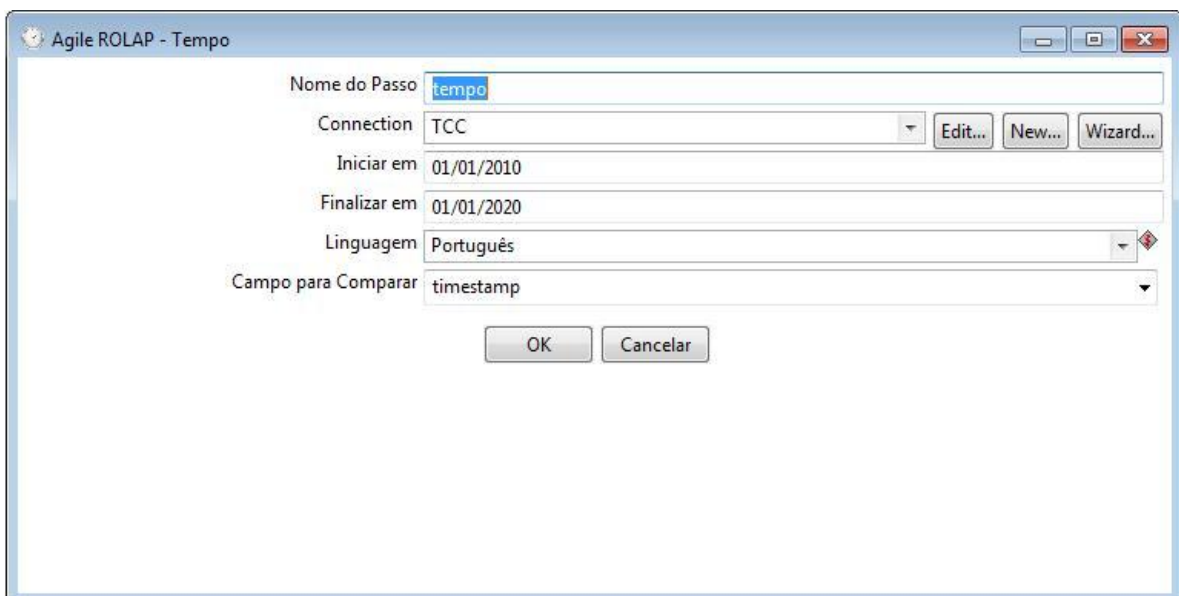


Figura 13: Configuração da dimensão tempo no *plugin* tempo.

5.2.2 Criação das dimensões convencionais – 5º Passo

Ao todo foram definidas quatro dimensões convencionais, elas são professor, disciplina, curso e campus. O *plugin* que auxilia na criação destas dimensões se

chama Dimensão e as configurações realizadas nele para criar a dimensão professor serão mostradas pela Figura 14:

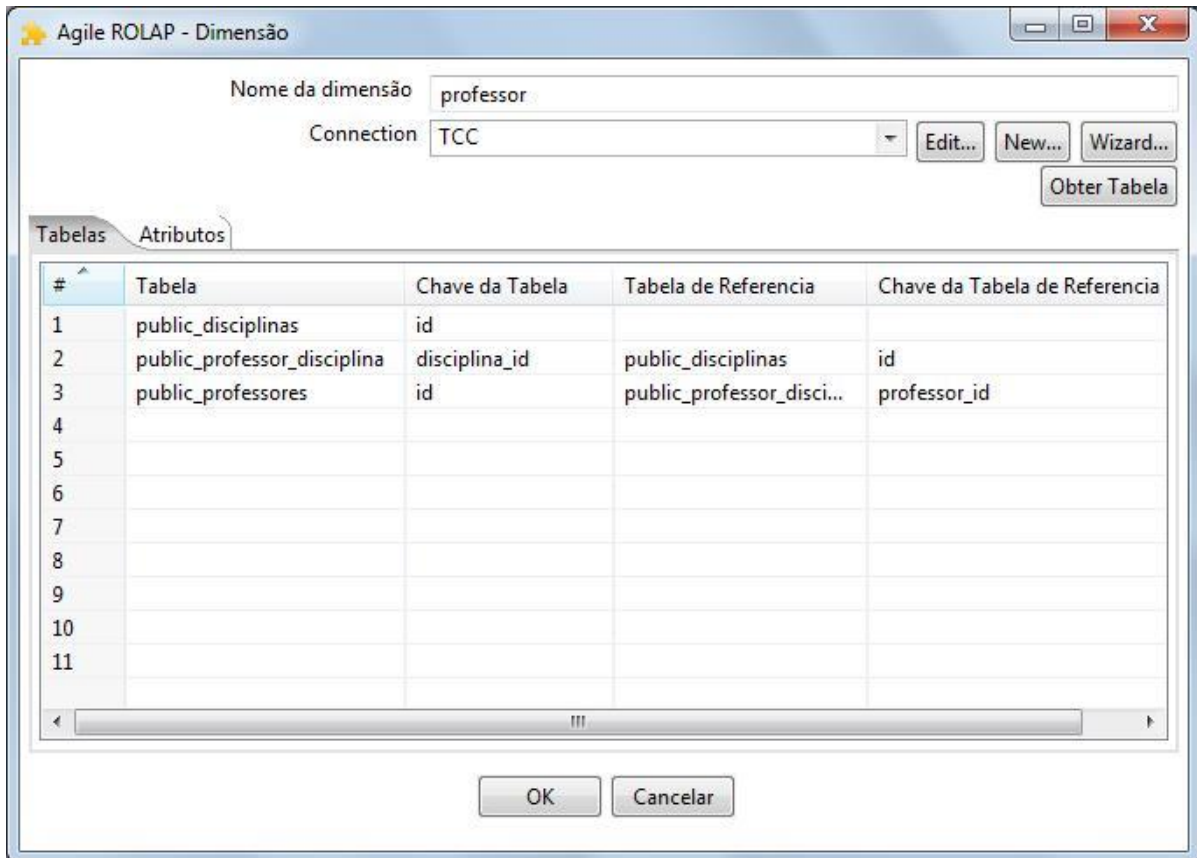


Figura 14: Configuração da dimensão professor no *plugin* dimensão.

As configurações das dimensões disciplina, curso e campus foram movidas para o apêndice B.

5.2.3 Criação dos fatos – 6º Passo

No modelo dimensional foram definidos cinco fatos, avaliacao aluno, avaliacao docente, avaliacao aluno_docente, avaliacao técnico e avaliacao estagiário. Para a implantação da metodologia Agile ROLAP estes cinco fatos construídos utilizando-se do *plugin* fato. A forma como ele foi configurado será mostrado a seguir:

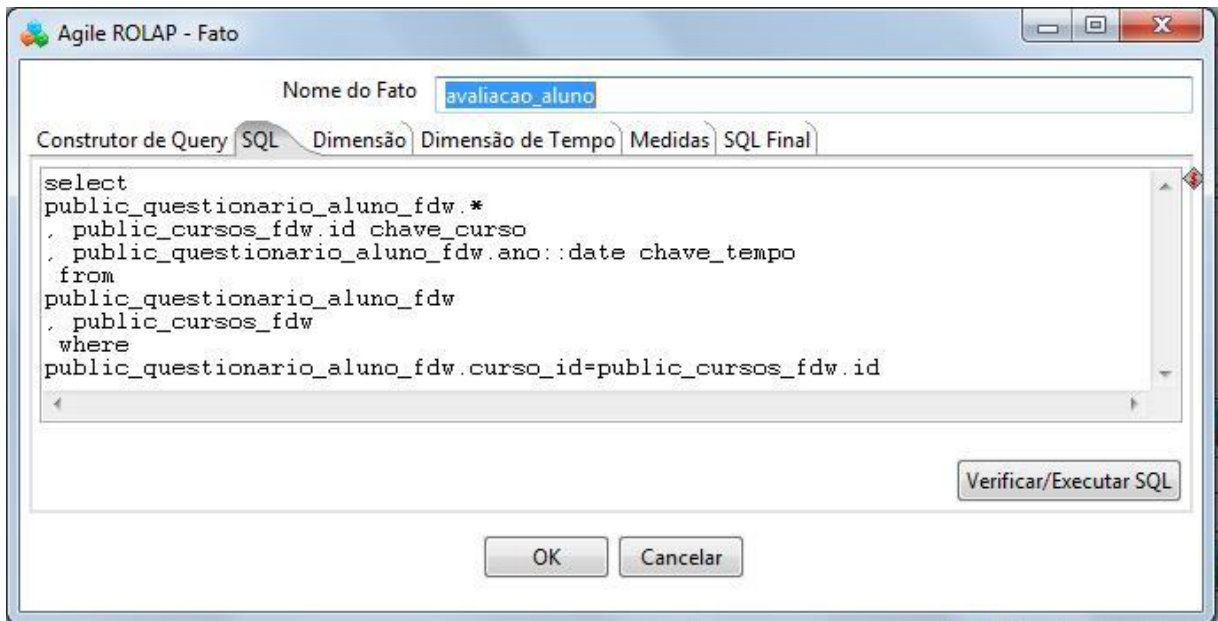


Figura 15: Fato avaliacao aluno na aba de configuração SQL.

A Figura 15 mostra como foi configurada a aba de configuração SQL do *plugin* fato. Este SQL é responsável por selecionar as chaves das dimensões e medidas que farão parte do fato.

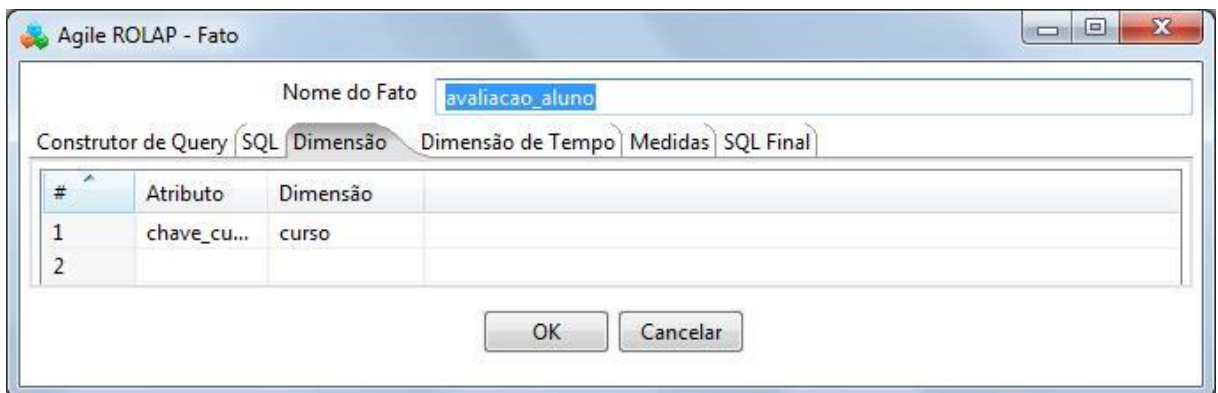


Figura 16: Fato avaliacao aluno na aba de configuração Dimensão.

Na Figura 16 é possível visualizar como é feita a ligação entre qual chave estrangeira do fato irá representar qual dimensão.

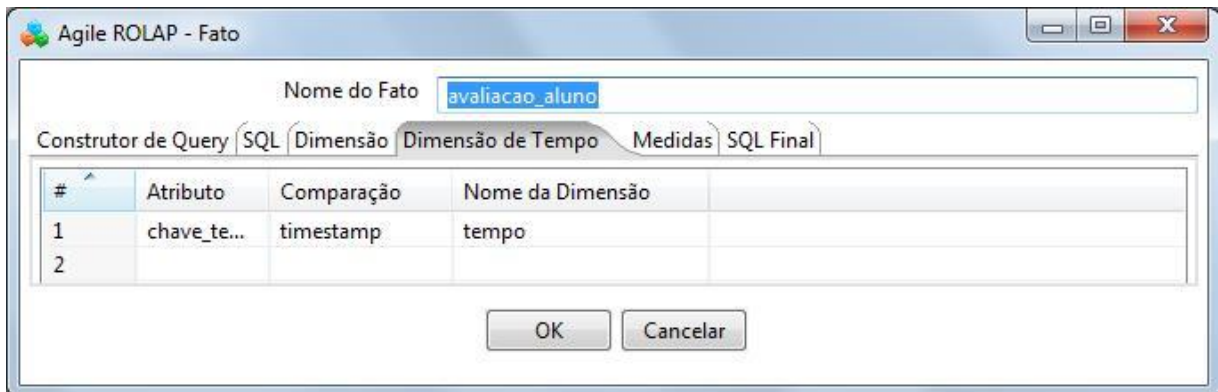


Figura 17: Fato avaliacao aluno na aba de configuração Dimensão de Tempo.

Na Figura 17 é configurada a ligação das chaves estrangeiras com as dimensões, porém diferente da Figura 16 desta vez a configuração é para as dimensões de tempo.

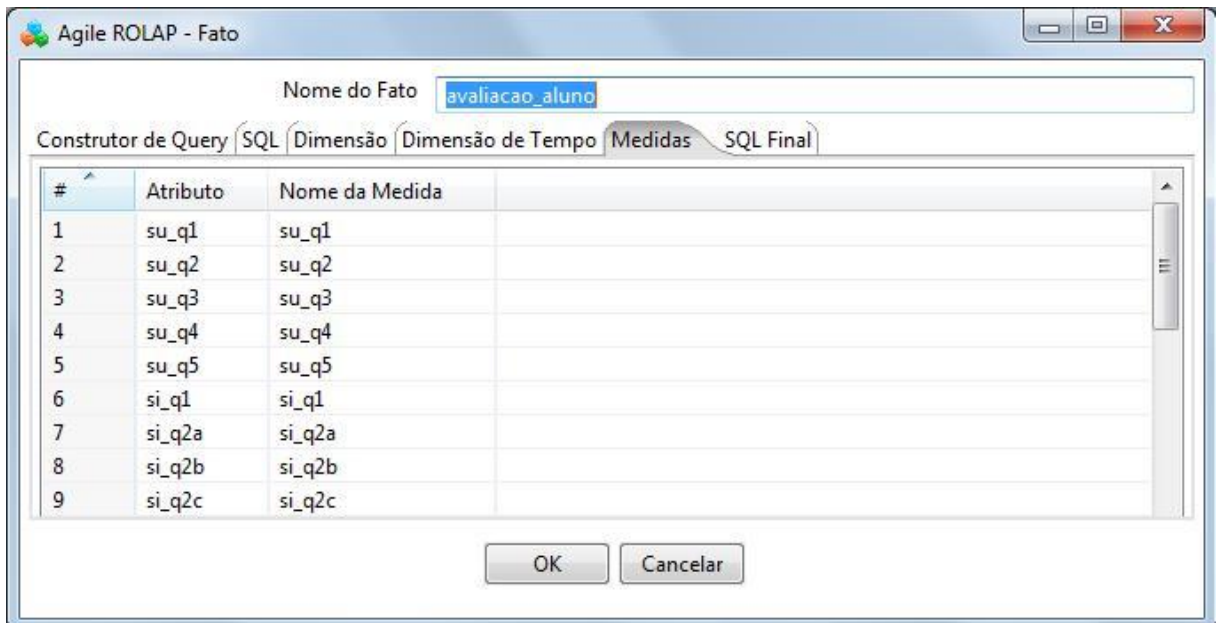


Figura 18: Fato avaliacao aluno na aba de configuração Medidas.

A aba de configuração medidas mostrada pela Figura 18 foi a ultima a ser configurada, sua configuração permite dizer quais dos campos selecionados para o fato serão medidas.

A configuração dos outros quatro fatos definidos no modelo dimensional está presente no apêndice C.

5.3 Modelo lógico

O modelo lógico é a parte em que se configura como tudo aparecerá na ferramenta OLAP a ser utilizada. O mapeamento gerado no modelo físico serve de estrutura para esta configuração. Nenhuma tabela ou coluna não selecionada no modelo físico pode ser utilizada, pois esta etapa não tem conexão com o banco de dados.

Esta é a última etapa e ao fim dela a metodologia Agile ROLAP estará completamente implantada. Seus passos são criação das hierarquias e níveis das dimensões de tempo, criação das hierarquias e níveis das dimensões convencionais, criação dos cubos e exportação do Schema XML. Esses serão os tópicos seguintes a este texto, explicando como o modelo lógico foi criado neste trabalho.

5.3.1 Criação das hierarquias e níveis das dimensões de tempo – 7º Passo

A dimensão tempo configurada anteriormente possuía o nome de tempo, mas ao analisá-la mais atentamente optou-se por mudar seu nome neste passo para Ano, devido ao ano ser sua menor granularidade e dar mais sentido a análise deste modo. A Figura 19 retrata como ela foi configurada em seu referente *plugin*.

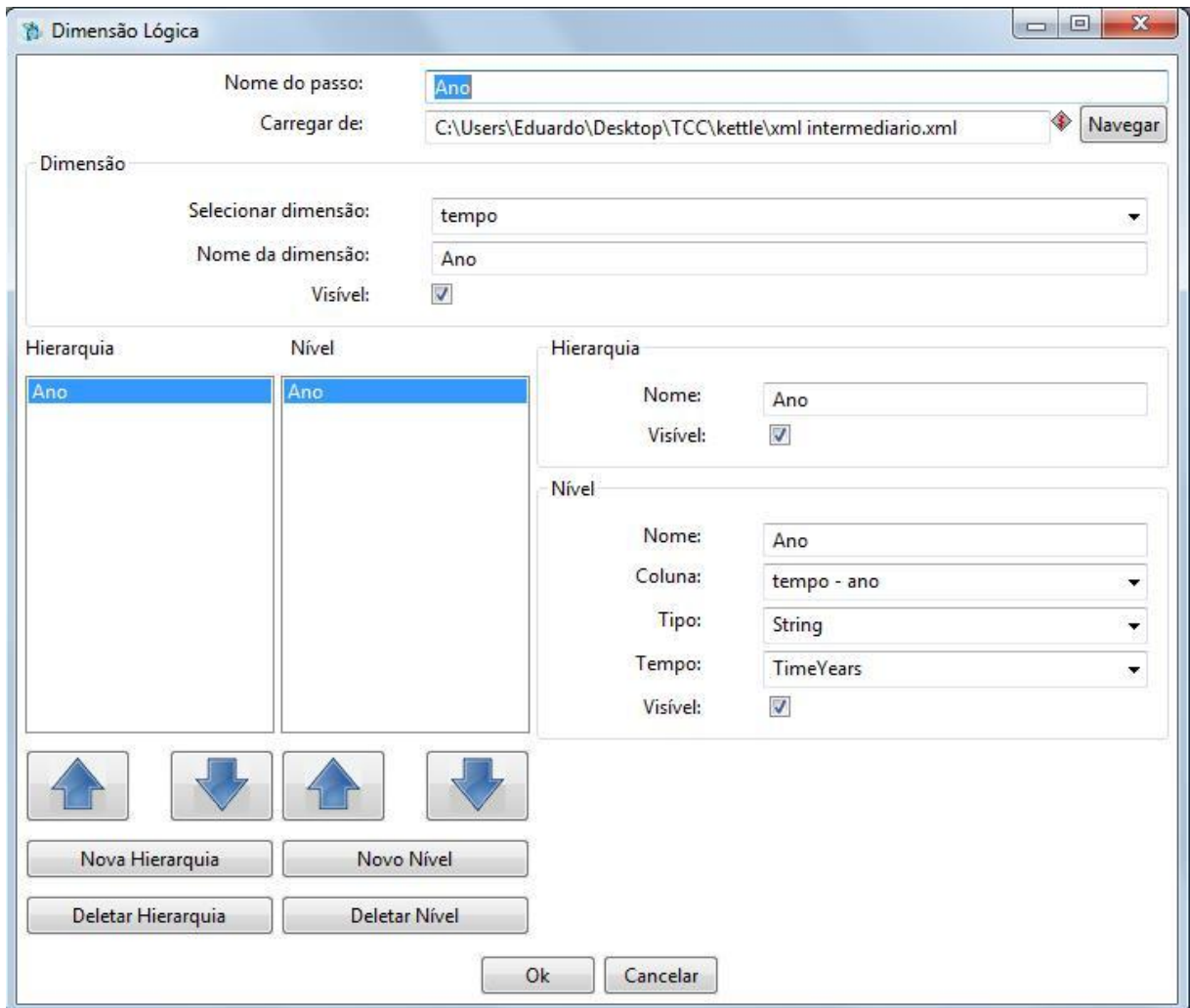


Figura 19: Dimensão de tempo ano configurada no *plugin* dimensão lógica de tempo.

5.3.2 Criação das hierarquias e níveis das dimensões convencionais – 8º Passo

Neste passo foram configuradas as hierarquias e níveis das dimensões Disciplina, Professor, Curso e Campus. Cada dimensão teve somente uma hierarquia, com exceção da dimensão Disciplina, que teve uma hierarquia denominada Disciplina e outra denominada Disciplina_simples. A hierarquia Disciplina_simples se fez necessária para realizar análises mais resumidas a respeito do fato avaliação aluno_professor.

A reprodução destas configurações foi feita no *plugin* Dimensão Lógica. A Figura 20 mostra como o *plugin* foi configurado para a dimensão professor:

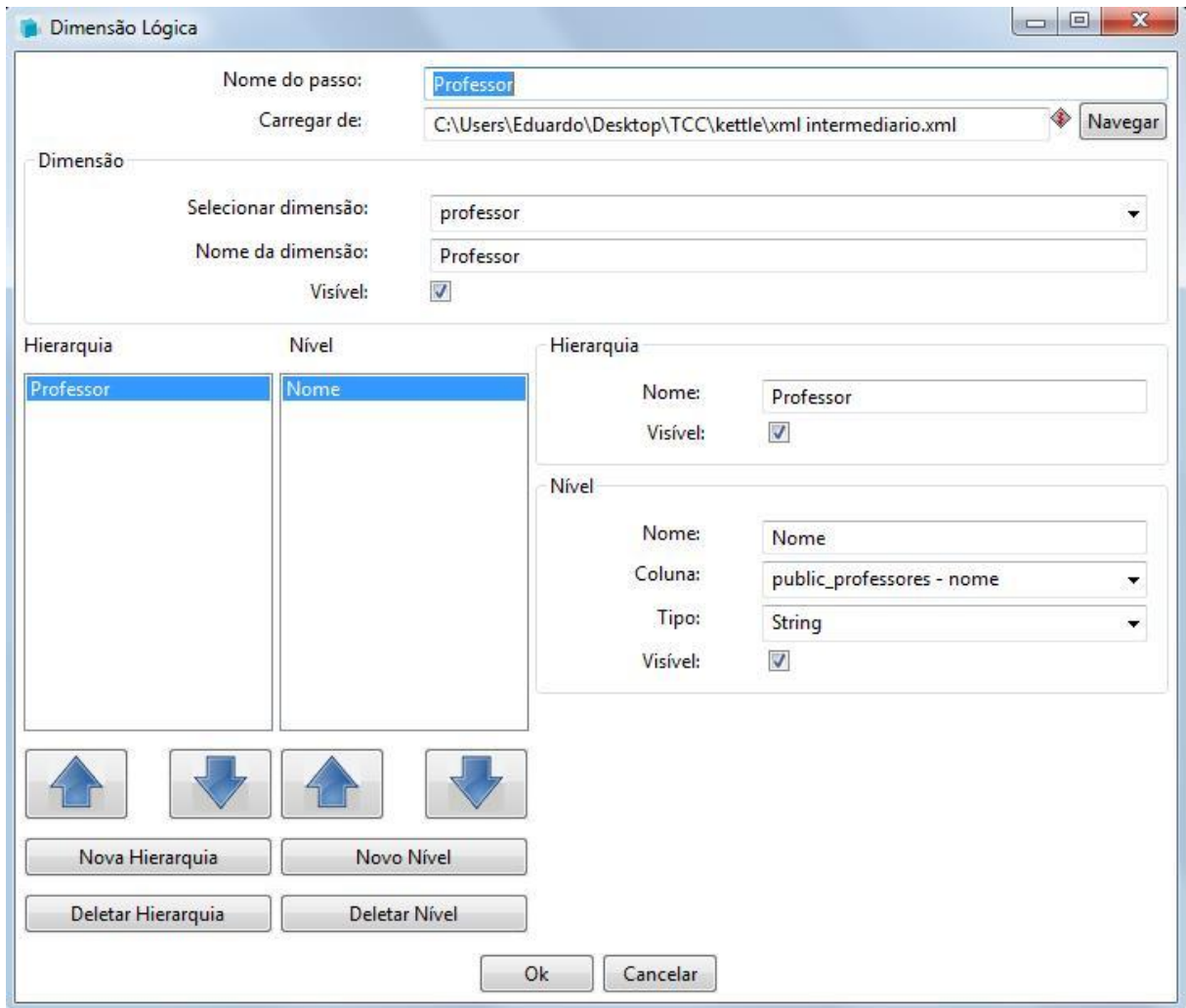


Figura 20: Configuração dimensão professor no *plugin* dimensão lógica.

A configuração das demais dimensões lógicas encontra-se no apêndice D.

5.3.3 Criação dos cubos – 9º Passo

Neste trabalho cada fato está diretamente ligado a um Cubo, deste modo, os cubos criados foram Avaliação aluno, Avaliação aluno_docente, Avaliação docente, Avaliação tecnico e Avaliação estagiario. O *plugin* destinado a auxiliar o usuário neste passo chama-se Cubo e ele foi configurado da seguinte maneira:

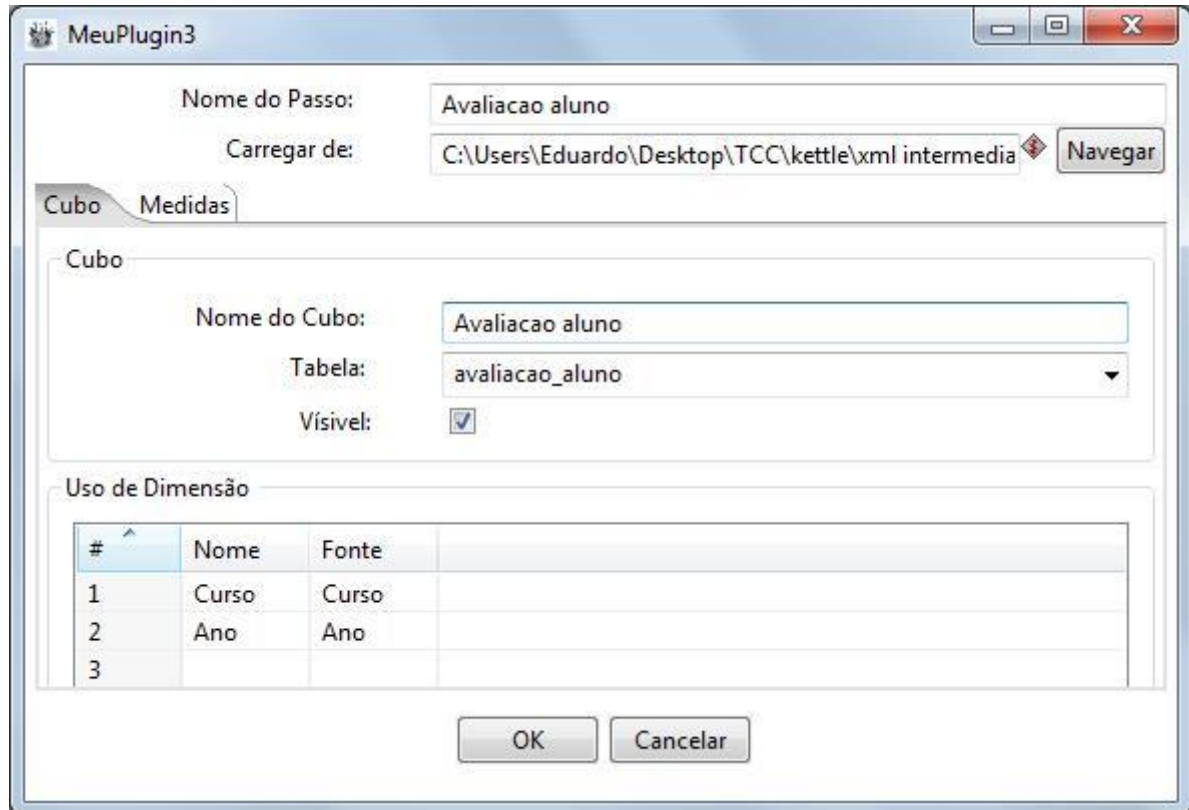


Figura 21: Cubo Avaliacao aluno na aba de configuração Cubo.

A Figura 21 demonstra como foi a configuração do cubo avaliacao aluno em relação ao nome do cubo e quais dimensões farão parte do cubo.

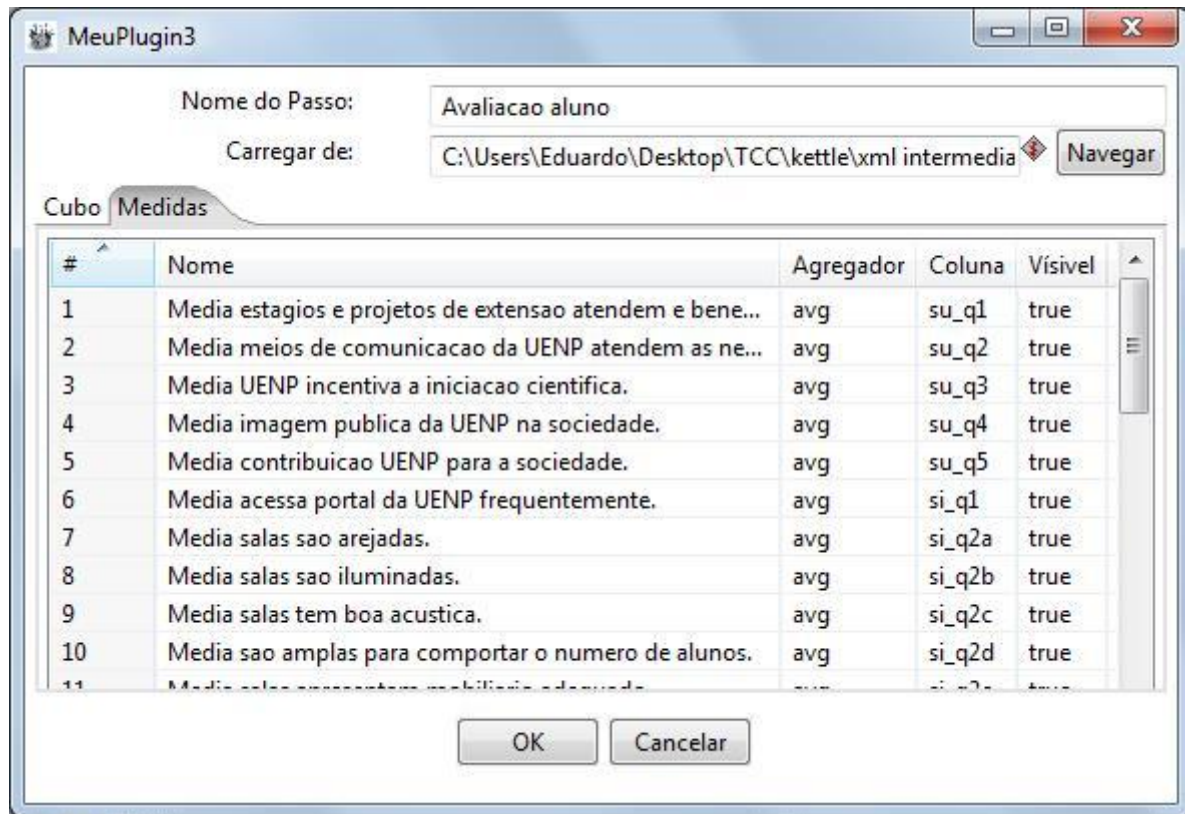


Figura 22: Cubo Avaliacao aluno na aba de configuração Medidas.

Já a Figura 22 mostra como foi a configuração das medidas do cubo avaliacao aluno. Aqui é possível configurar qual agregador será utilizado, o nome da medida, a coluna correspondente e se ela será visível no cubo ou não.

Se o leitor analisar novamente as medidas mostradas no quadro 1 irá identificar que existem formulas para cada medida. Estas formulas não foram configuradas no *plugin* Cubo, elas precisaram ser criadas manualmente no Schema XML, pois a ferramenta não possui tais opções de configuração.

A configuração dos cubos Avaliacao aluno_docente, Avaliacao docente, Avaliacao tecnico e Avaliacao estagiario estará disponível no apêndice E.

5.3.4 Exportação do Schema XML – 10º Passo

A exportação do *Schema XML* é feita através do *plugin* SchemaOut, que exporta as configurações feitas anteriormente no modelo XML legível para o servidor ROLAP Mondrian. Sua interface é bastante simples como demonstrada na figura a seguir, porém internamente é um *plugin* complexo, pois une as configurações de todos os *plugins* anteriores.

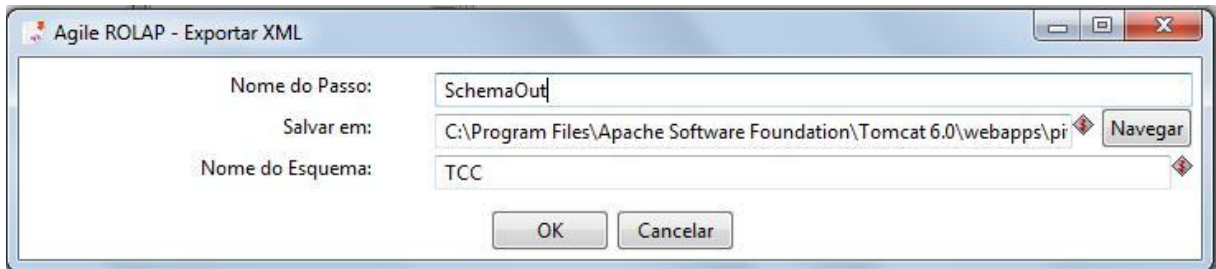


Figura 23: Configuração do *plugin* SchemaOut.

Ao executar este décimo passo a metodologia Agile ROLAP está completamente implantada, faltando agora somente utilizar de uma ferramenta de análise para visualizar os dados.

Algumas análises realizadas no sistema BI implantado de acordo com a metodologia Agile ROLAP serão mostradas na seção resultados. Porém, é importante lembrar que o objetivo deste trabalho é a implantação da metodologia, de modo a testar o funcionamento da mesma e de suas ferramentas, mas não faz parte dos objetivos deste trabalho analisar a base de avaliação institucional.

6. Resultados

O principal resultado deste trabalho sem dúvida é a metodologia Agile ROLAP implantada e funcionando, por isso neste capítulo serão apresentados alguns exemplos de análises que poderiam ser realizadas no sistema implantado.

The screenshot shows an OLAP tool interface with the following components:

- Top Bar:** Contains buttons for 'Mostrar Pais', 'Ocultar Spans', 'Não Vazio', 'Trocar Eixos', 'Drill Posição', 'Membro', 'Substituir', 'Através', 'Properties', and 'Agg.'.
- Navegador OLAP:** Shows the selected cube 'Avaliacao aluno_professor'.
- Estrutura do Cubo:** A tree view showing the hierarchy: 'Disciplina_simples' > 'Professor' > '(All)'. Other dimensions include 'Nome' and 'Ano'.
- Estrutura Pivot:** Shows the pivot table structure with columns for 'Measures' (Media geral aluno, Media geral professor, Media relacao professor alunos) and rows for 'Professor' > '(All)'.
- Resultado da Consulta:** A table displaying the query results:

Professor	Media geral aluno	Media geral professor	Media relacao professor alunos.
All Professors	4,147	4,198	4,224
- Consulta MDX:** A query editor showing the MDX query:


```
1 SELECT NON EMPTY {[Measures].[Media geral aluno], [Measures].[Media geral professor], [Measures].[Media relacao professor alunos.]} ON COLUMNS, NON EMPTY Hierarchize({[Professor].[All Professors]}) ON ROWS FROM [Avaliacao aluno_professor]
```
- Footer:** Shows 'Rodar' and 'Restabelecer' buttons, and a status bar indicating 'Tempo de execução da Consulta : 18 msec'.

Figura 24: Exemplo de consulta no cubo Avaliacao aluno_professor.

A Figura 24 ilustra um exemplo de consulta no cubo Avaliacao aluno_professor, onde se tenta estabelecer uma conexão entre as notas dadas aos professores com a média de notas dadas a relação professor-aluno ou a auto-avaliação do aluno na disciplina. Consultas como essa poderia revelar o grau de imparcialidade dos alunos ao avaliar a disciplina e o professor que ministra a disciplina.

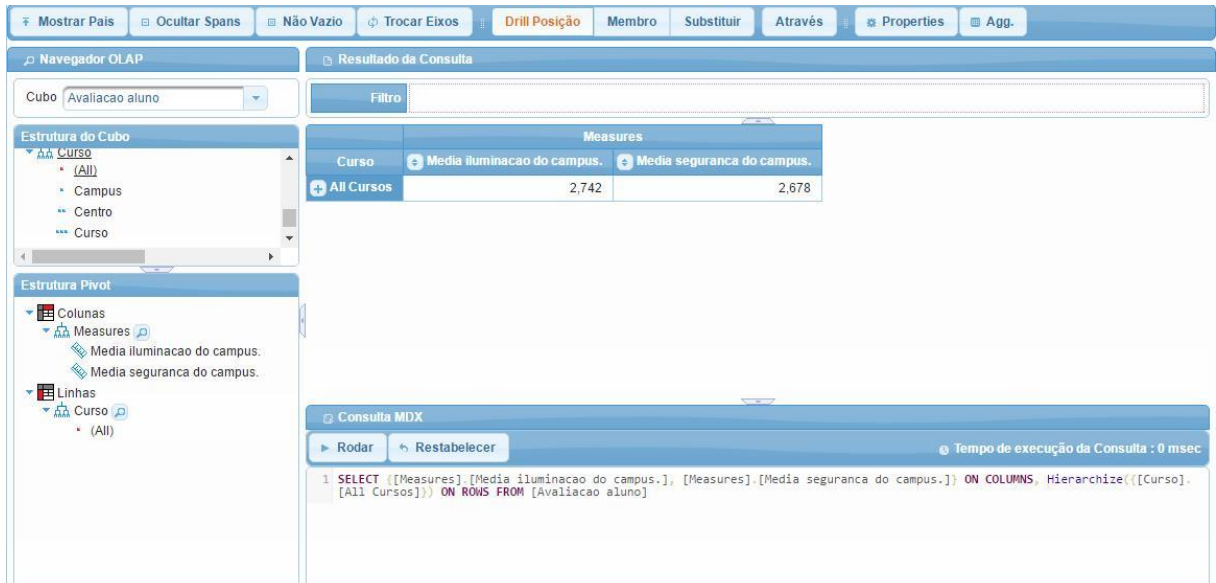


Figura 25: Exemplo de consulta no cubo Avaliacao aluno.

A Figura 25 é um exemplo de consulta no cubo Avaliacao aluno, onde os gestores poderiam procurar por ligações entre as notas dadas a segurança do campus de acordo com a iluminação do mesmo, visto que vários dos cursos são noturnos.

Pesquisas como essa poderiam direcionar os investimentos da universidade para a melhoria da iluminação e não a contratação de mais seguranças ou construção de muros e cercas.

A mesma consulta também poderia ser realizada de acordo com a opinião dos professores ao invés dos alunos, para confirmar as suspeitas dos gestores. Esta consulta sobre os professores é ilustrada pela Figura 26:

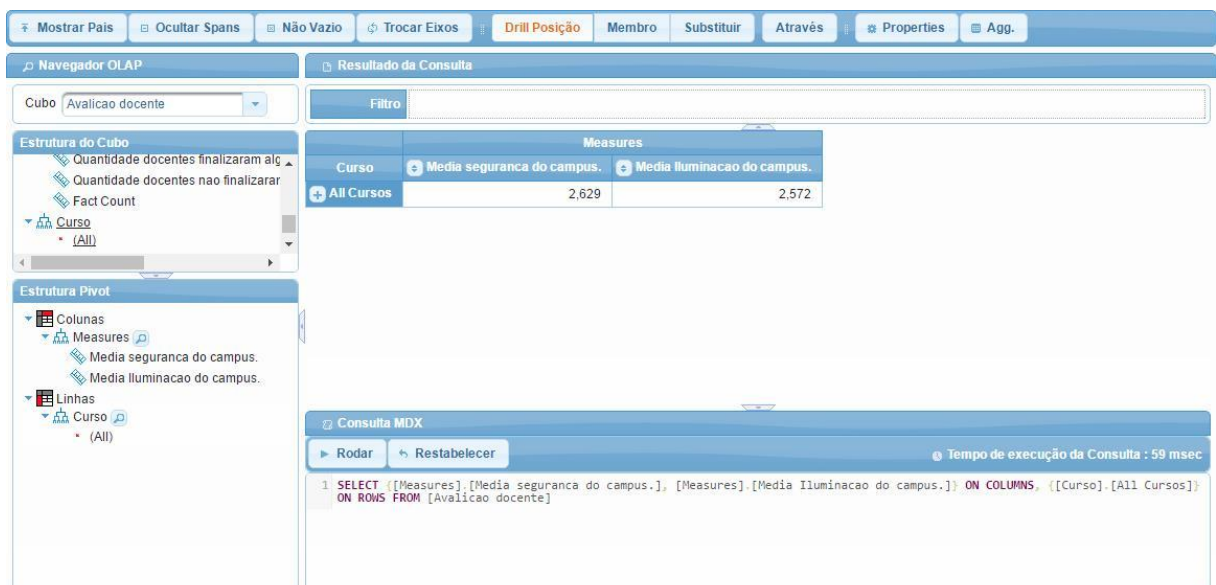


Figura 26: Exemplo de consulta no cubo Avaliacao docente.

As consultas apresentadas e o que elas poderiam indicar são exemplos de consultas que poderiam interessar aos gestores da UENP, e não afirmações sobre o que elas de fato indicam. É importante lembrar que o objetivo deste trabalho é implantar um ambiente de BI seguindo a metodologia Agile ROLAP, e não minerar dados na base de avaliação institucional.

Além da metodologia implantada e um sistema de BI da base de avaliação institucional, este trabalho também permitiu que fossem identificados diversos erros nos *plugins* desenvolvidos para auxiliar a metodologia Agile ROLAP.

Dentre os erros identificados alguns destes já eram de conhecimento da equipe de desenvolvimento Agile ROLAP e já estavam sendo corrigidos em uma nova versão dos *plugins* que ainda não se encontra finalizada. Por isso, este trabalho considerará como resultados somente os erros que não eram de conhecimento da equipe.

Desta forma, este trabalho apresenta como resultados:

- A identificação de treze erros nos *plugins* Agile ROLAP;
- Um sistema de BI implantado seguindo a metodologia Agile ROLAP;
- Um sistema de BI que permite a análise das avaliações institucionais da UENP.

Todos os erros encontrados no trabalho (de prévio conhecimento da equipe Agile ROLAP ou não) são listados na seção 6.1 Erros encontrados.

6.1 Erros encontrados

Durante a implantação da metodologia Agile ROLAP em alguns momentos os *plugins* utilizados para auxiliar a implantação demonstraram funcionamento diferente do esperado, algumas vezes não cumprindo seu objetivo. Esta seção destina-se a relatar quais as situações em que houve comportamento incorreto ou inesperado por parte dos *plugins*. Estes erros são:

- A aba de configuração medidas do *plugin* Fato não salva a coluna selecionada, mas sim o nome dado a medida, causando erro no *SchemaXML* por referenciar uma coluna que não existe.
- Ao alterar o nome do passo em um dos *plugins* o *plugin* do passo seguinte não atualiza automaticamente o nome do passo alterado,

fazendo com que o nome desatualizado seja uma referência a algo inexistente e consequentemente gerando erros nos próximos passos.

- Antes de executar a *query* que cria a *view* materializada o *plugin* Fato exibe uma janela de aviso com a *query* que irá ser executada, no entanto esta janela fica minimizada ao invés de ser exibida, fazendo com que o usuário pense que o sistema ainda está carregando.
- Ao tentar visualizar as opções de um *combo-box* nos *plugins* Dimensão lógica e Dimensão lógica de tempo a opção que antes estava selecionada desaparece como demonstrado abaixo pelas figuras 27, 28 e 29:

Coluna:

Figura 27: Campo coluna do *plugin* Dimensão lógica.

Coluna:

Tipo:

Visível:

- public_disciplinas - id
- public_disciplinas - nome
- public_disciplinas - curso_id
- public_disciplinas - professor
- public_disciplinas - serie
- public_disciplinas - quantidade_alunos
- public_disciplinas - turma
- public_professor_disciplina - professor_id
- public_professor_disciplina - disciplina_id
- public_professores - id
- public_professores - nome
- public_professores - email

Figura 28: Visualizando as colunas disponíveis no *plugin* Dimensão lógica.

Coluna:

Figura 29: Campo coluna do *plugin* Dimensão lógica após visualizar as colunas disponíveis.

- A aba SQL final do *plugin* Fato deveria auxiliar o usuário a re-executar somente a *query* que gera a *view* materializada do fato, porém ao clicar no botão gerar sql final o *plugin* gera uma sql diferente da mostrada ao clicar no botão executar do programa *Kettle*,
- O *plugin* Base de dados estrangeira não executa ao executar o *Kettle*, mas sim ao clicar no botão ok da seleção de tabelas. Além disso, as tabelas já selecionadas ficam em uma *jTable*. O problema é que este *jTable* possui colunas que não fazem sentido para o usuário e provavelmente não deveria ser mostradas.

- Os *plugins* Dimensão, Base de dados estrangeira, Fato e Tempo retornam erro caso o usuário tente salvá-los em branco. O erro apresentado é mostrado na Figura 30:

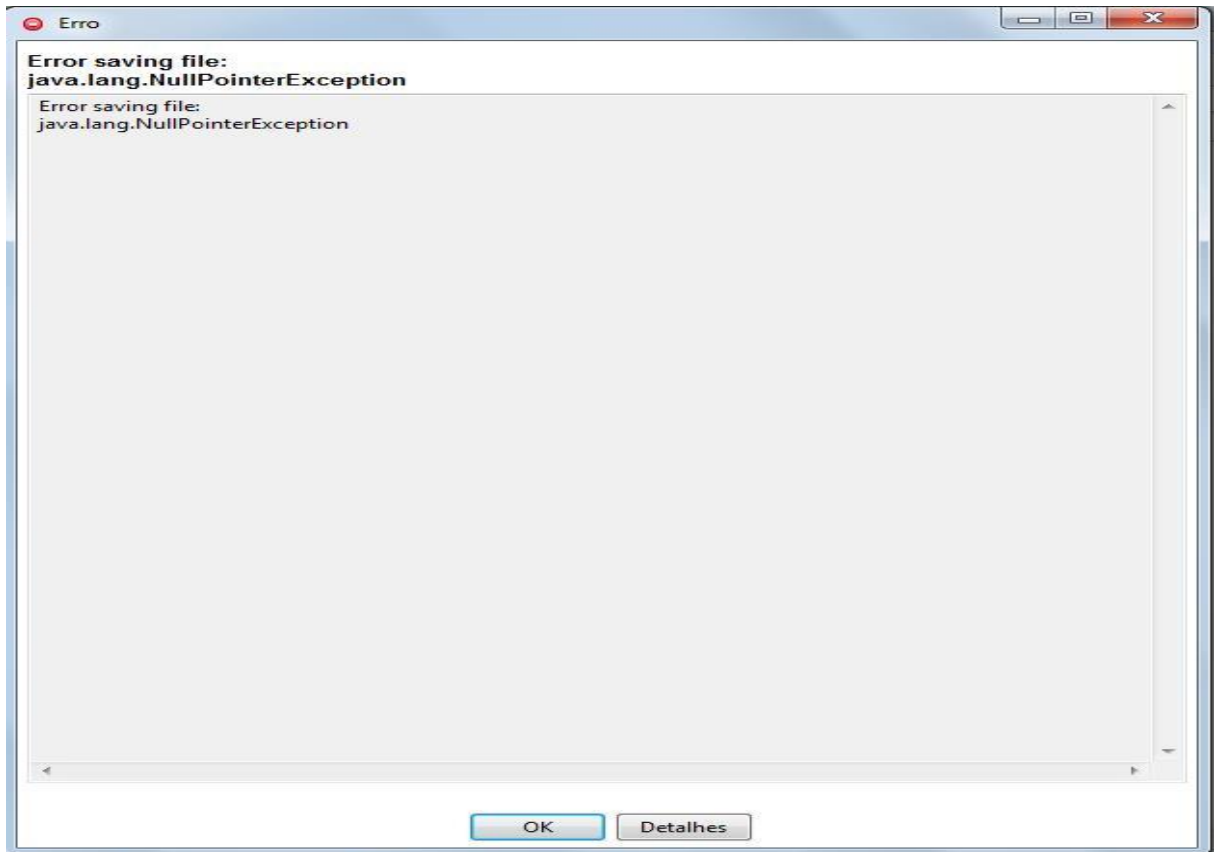


Figura 30: Erro ao salvar *plugins* com configurações em branco.

- No *plugin* cubo na aba de medidas a coluna visível aparece com as opções *true* e *false*, o que deveria estar em português.
- Os *plugins* não adicionam o nome do *schema* antes do nome das tabelas, isso resulta em erro de referência quando utiliza-se qualquer *schema* que não seja *public*.
- A ideia de se utilizar o *postgresFDW* é utilizar de uma base de dados externa, porém o *plugin* Base de dados estrangeira não importa os dados corretamente caso o usuário tente criar um *postgresFDW* de uma base externa.
- Por questões de segurança também considerei neste trabalho o fato do campo senha do *plugin* Base de dados estrangeira ser visível um erro.
- O *plugin* base de dados estrangeira não importa a tabela corretamente algumas vezes porque ele converte o tipo *text* para *varchar(45)*,

resultando em erro na importação caso a tabela possua alguma ocorrência daquele texto com mais de 45 caracteres.

- Ao gerar o *schemaXML* a chave estrangeira das dimensões de tempo é preenchida incorretamente pelos *plugins*, resultando no não funcionamento das dimensões de tempo na ferramenta OLAP.
- Ao excluir uma tabela no *plugin* Dimensão e tentar inserir uma nova tabela logo em seguida o *plugin* apresenta erro fatal e a ferramenta Kettle precisa ser finalizada. A imagem do erro é representada pela Figura 31.

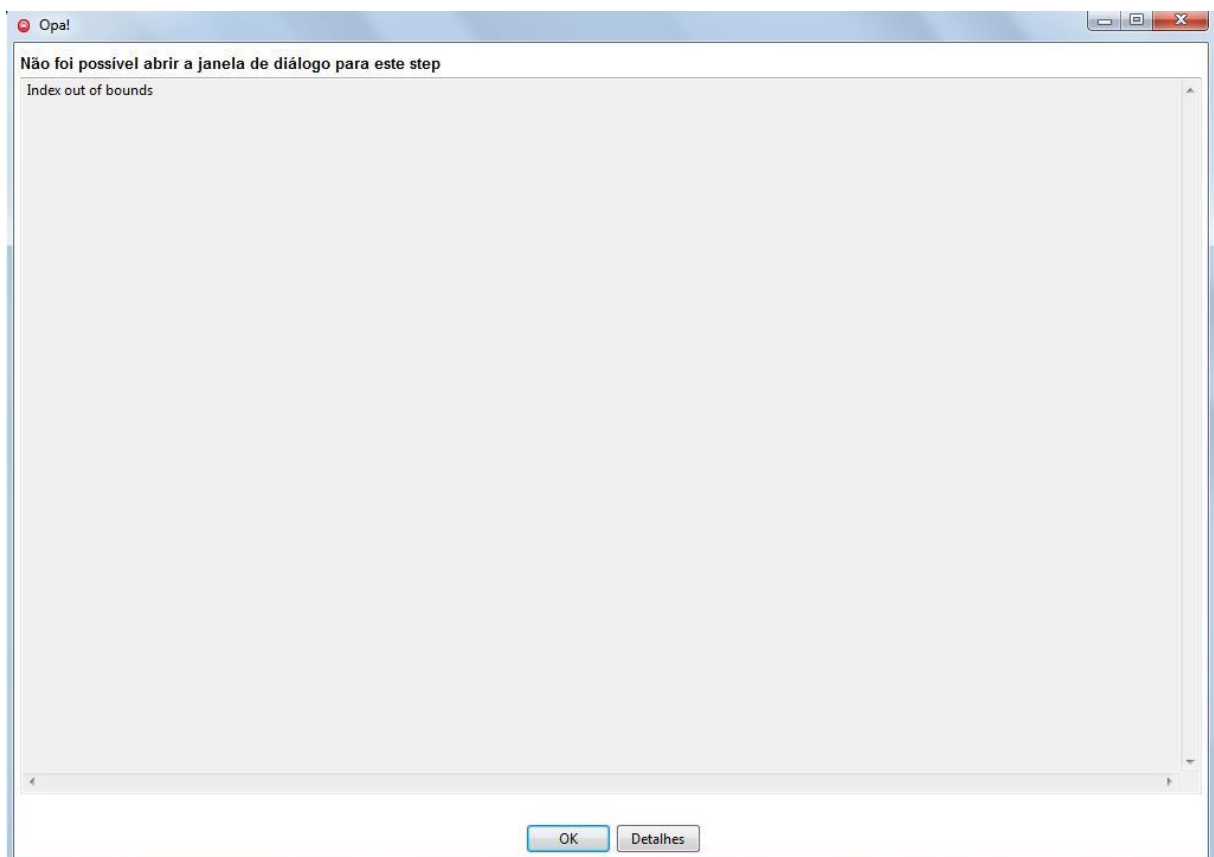


Figura 31: Erro ao deletar tabelas da dimensão e tentar inserir uma nova tabela.

- O *plugin* Dimensão lógica de tempo está com o nome de sua janela incorreto. O nome apresentado é Dimensão lógica e não Dimensão lógica de tempo, como apresentado na Figura 32:



Figura 32: Nome incorreto para a janela do *plugin* Dimensão lógica de tempo.

- O mesmo problema de nome de janela incorreto ocorre com o *plugin* Cubo, como mostrado na Figura 33:

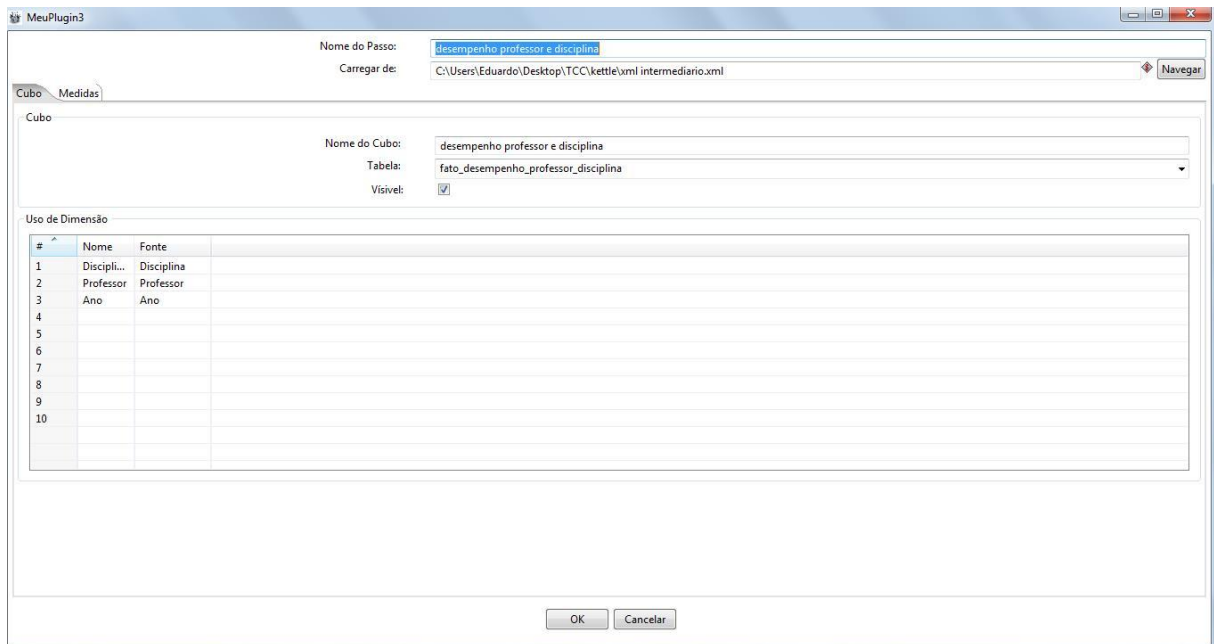


Figura 33: *Plugin* cubo com nome de janela incorreto.

- No *plugin* Dimensão lógica de tempo é possível selecionar dimensões que não são de tempo. O inverso também ocorre com o *plugin* Dimensão lógica, onde é possível selecionar e configurar dimensões de tempo. Este erro é ilustrado pela Figura 34:

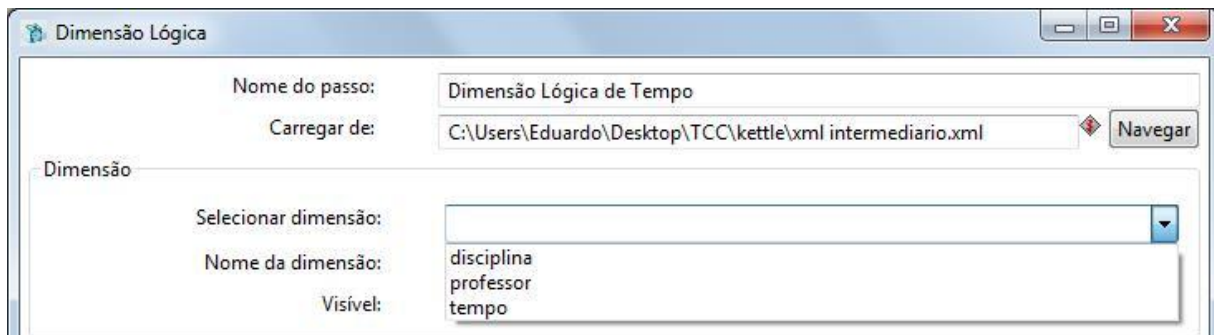


Figura 34: Os *plugins* Dimensão lógica de tempo e Dimensão lógica não distinguem corretamente dimensões tempo de não tempo.

- Baixo tratamento de erros. Isto é válido para todos os *plugins* utilizados, pois nenhum deles avisa o usuário quando o mesmo está cometendo algum erro.

7. Discussão

Dois dos *plugins* auxiliares da metodologia Agile ROLAP não foram citados durante a implantação deste trabalho, eles são XMLOut e Feriado. Entretanto, ambos foram utilizados. O *plugin* XMLOut não foi citado pois não está relacionada a nenhum passo da implantação Agile ROLAP, mas sim a integração dos passos 4,5,6 – 7,8,9,10. Já o *plugin* Feriado não foi utilizado na implantação da base de dados Avaliação institucional, pois a granularidade mínima de tempo da base utilizada era de anos e não de dias. Todavia, este *plugin* foi testado com uma base de dados secundária denominada “pagila”. A base de dados pagila² é uma base gratuita disponibilizada para download pela pgFoundry.

Além dos erros apresentados na seção 6.1, também foi possível observar situações em que os *plugins* Agile ROLAP funcionaram bem, mas poderiam funcionar melhor com alguns ajustes. Nestes casos seria incorreto afirmar que o *plugin* possui um erro. Por isso a próxima seção irá relatar situações em que o sistema funcionou, porém seria agradável se ele fosse um pouco diferente. Ou seja, serão sugestões de melhorias para os *plugins* Agile ROLAP e para a metodologia Agile ROLAP.

7.1 Sugestões de melhoria

Começando com as sugestões de melhoria para os *plugins* Agile ROLAP temos:

- O sistema é bastante redundante. Algumas vezes os campos são até destinados a funções distintas, porém na prática eles geralmente recebem o mesmo preenchimento.
- Uma opção para diminuir a perda de produtividade gerada pela redundância seria o auto-preenchimento e o pré-preenchimento. O auto-preenchimento e pré-preenchimento também poderiam ser aplicados para evitar erros, como por exemplo auto-preenchimento do campo “Tipo” em um nível, coisa que aumentaria a produtividade pois seria uma preocupação a menos para o usuário. Exemplos de campos

² Pagila - <http://pgfoundry.org/projects/dbsamples/>

que poderiam se encaixar no auto-preenchimento, pré-preenchimento ou remoção de redundância seriam “campo para comparar”, “nome”, “selecionar” e “nome do passo”.

- Quando dois passos são ligados é comum que apareça uma mensagem de *warning* dizendo que o padrão deles é diferente. Por ser uma mensagem que não auxilia o usuário de forma alguma, e ainda pode confundi-lo, ela não deveria existir. A forma como a mensagem aparece é mostrada na Figura 35:

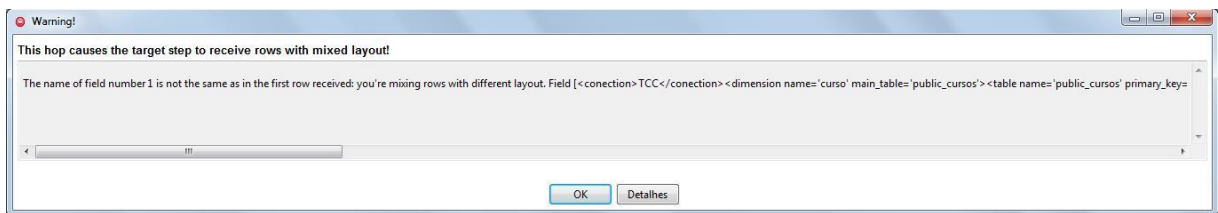
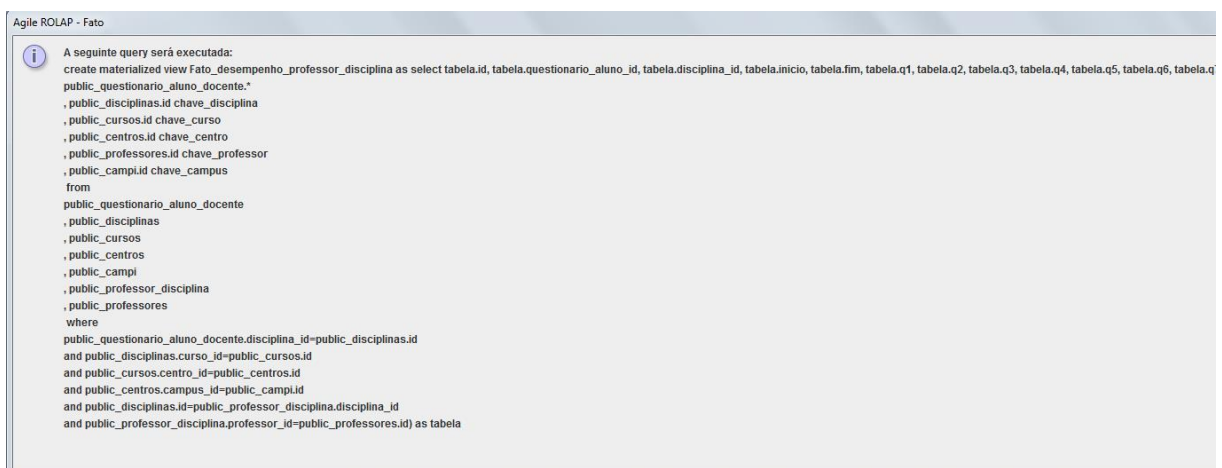


Figura 35: Mensagem de warning ao ligar dois passos.

- O *plugin* Fato não avisa quando a *view* materializada foi criada com sucesso. Como relatado na seção de erros encontrados, os *plugins* não possuem tratamento de erros, e também não há aviso de sucessos.
- O *plugin* base de dados estrangeira não avisa o usuário se a importação da tabela FDW foi bem sucedida ou não. Novamente problema com as mensagens do sistema.
- Como usuário das ferramentas implantando um sistema de BI, posso dizer que é normal e frequente esquecer qual o próximo nível, dimensão ou fato que o usuário havia pensado em implantar. Inconscientemente sempre que isto acontecia utilizava dos campos de seleção dos *plugins* para consultar quais as opções possíveis e me lembrar de quais eu já havia utilizado. Isto em termos de utilização significava que eu sempre pulava o primeiro campo “nome” preenchia o campo com “seleção” para me lembrar do nome e só depois voltava para o campo “nome” para preenche-lo. Dito isto a próxima sugestão é a inversão destes campos, sempre colocando os campos selecionáveis antes dos campos de nome.
- Na base de avaliação institucional a menor granularidade de tempo era ano, porém o *plugin* Tempo tem como menor granularidade dia e não

possui nenhuma opção de granularidade. A sugestão agora é inserir opções de granularidade no *plugin* tempo, visto que no momento ele não se enquadra as possíveis necessidades dos usuários.

- O *plugin* Fato não permite a escrita de sql direta, há a obrigatoriedade de se utilizar o *query builder*. Isto é algo muito restritivo e deveria ser removido, visto que o mais comum é que o usuário teste a *query* em seu banco de dados para depois copiá-la para o sistema.
- Uma das etapas que mais tomou tempo durante a implantação da metodologia Agile ROLAP com a base de avaliação institucional foi a criação de medidas. Apesar disso não havia nenhum *plugin* dedicado especificamente às medidas, e os que possuíam alguma configuração de medidas eram muito limitados, não permitindo a criação de medidas calculadas e nem configuração de formatos de saída. Um *plugin* dedicada somente a criação de medidas deveria ser construída.
- No *plugin* Fato antes de se executar a *query* que cria a *view* materializada que servirá como Fato, o *plugin* exibe uma janela com a *query* que será executada. O problema desta janela é que ela não se ajusta a tela, fazendo com que o usuário tenha que arrastá-la para executá-la caso a *query* seja muito grande. Um exemplo deste problema será mostrado nas Figuras 36 e 37:



```

A seguinte query será executada:
create materialized view Fato_desempenho_professor_disciplina as select tabela.id, tabela.questionario_aluno_id, tabela.disciplina_id, tabela.inicio, tabela.fim, tabela.q1, tabela.q2, tabela.q3, tabela.q4, tabela.q5, tabela.q6, tabela.q
public_questionario_aluno_docente.*
, public_disciplinas.id chave_disciplina
, public_cursos.id chave_curso
, public centros.id chave_centro
, public_professores.id chave_professor
, public_campi.id chave_campus
from
public_questionario_aluno_docente
, public_disciplinas
, public_cursos
, public_centros
, public_campi
, public_professor_disciplina
, public_professores
where
public_questionario_aluno_docente.disciplina_id=public_disciplinas.id
and public_disciplinas.curso_id=public_cursos.id
and public_cursos.centro_id=public_centros.id
and public_centros.campus_id=public_campi.id
and public_disciplinas.id=public_professor_disciplina.disciplina_id
and public_professor_disciplina.professor_id=public_professores.id) as tabela

```

Figura 36: Janela não ajustável da ferramenta Fato.

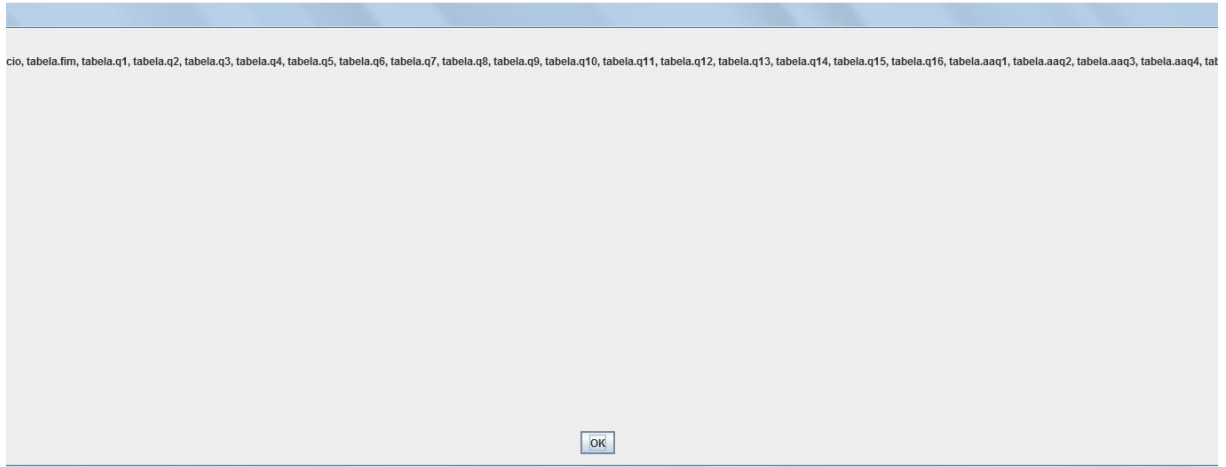


Figura 37: Janela não ajustável da ferramenta Fato.

Já quanto à metodologia Agile ROLAP sugere-se neste trabalho que dois passos sejam adicionados a metodologia, criação do modelo dimensional e criação de medidas. A Figura 4 que ilustra as fases Agile ROLAP após adaptada ficaria da seguinte maneira:

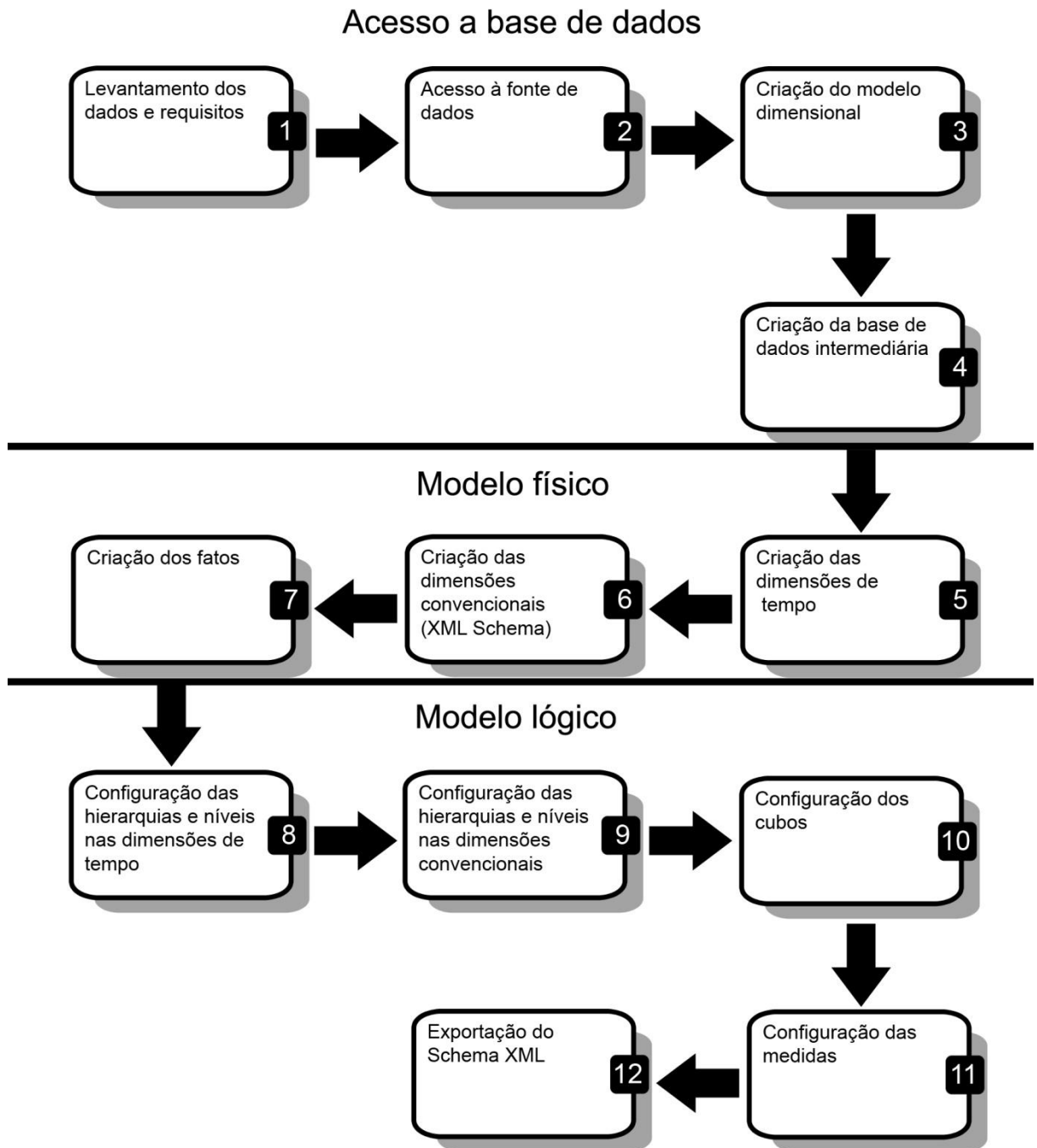


Figura 38: Fases Agile ROLAP adaptados de Souza(2014).

8. Conclusão e trabalhos futuros

A proposta deste trabalho era implantar um sistema de BI utilizando a metodologia Agile ROLAP e seus *plugins*. Isto foi realizado com sucesso, mostrando a validade da metodologia, e a utilidade dos *plugins* Agile ROLAP, que auxiliaram efetivamente na implantação da metodologia, criando um ambiente de BI para a análise das avaliações institucionais da UENP.

Apesar de úteis os *plugins* Agile ROLAP ainda precisam de correções, principalmente no que diz respeito ao *feedback* para o usuário. Espera-se que os problemas apresentados neste trabalho sejam resolvidos na próxima versão funcional dos *plugins*.

Já quanto à metodologia Agile ROLAP, diria que ela é uma boa opção para projetos que não envolvem a integração de diferentes bases de dados, ou pouca integração. Isto se deve a metodologia tradicional de desenvolvimento ter uma integração mais completa e melhor definida através do processo de ETL, processo que a metodologia Agile ROLAP procura minimizar ao máximo para proporcionar uma rápida implantação.

Apesar da desvantagem na integração e desvantagem de desempenho das consultas, a metodologia Agile ROLAP ainda é muito válida. Ambas as desvantagens são minimizadas por seu público alvo, visto que ela procura atender empresas de pequeno porte, e estas geralmente possuem poucos dados (Isto torna as consultas mais ágeis) e poucas bases para serem integradas.

E por fim, como trabalhos futuros propõe-se:

- Testar os *plugins* em sua nova versão;
- Integrar todos os anos que ocorreram avaliação institucional e preparar o sistema para as avaliações institucionais futuras.
- Minerar dados úteis a universidade utilizando do sistema implantado.
- Desenvolver um *plugin* para geração de medidas.

Referências

- ANGELONI, Maria Terezinha “Elementos **intervenientes na tomada de decisão**” Ci. Inf., Brasília, v. 32, n. 1, p. 17-22, 2003. Disponível em: <http://www.scielo.br/pdf/ci/v32n1/15969.pdf>. Acesso em: 20 maio 2014.
- ANTONELLI, Ricardo Adriano “**Conhecendo o Business Intelligence (BI): Uma Ferramenta de Auxílio à Tomada de Decisão**” Revista TECAP, v.3, n.3, p. 79-85, 2009. Disponível em: <http://revistas.utfpr.edu.br/pb/index.php/CAP/article/viewFile/933/544>. Acesso em: 20 maio 2014.
- BARROS, A. J. S. e LEHFELD, N. A. S. **Fundamentos de Metodologia: Um guia para a iniciação científica**. 2 ed. São Paulo: Makron Books, 2000.
- CALAZANS, Angélica Toffano Seidel “**Conceito e uso da informação organizacional e informação estratégica**” Transinformação, v.18, n. 1, 2012. Disponível em: <http://periodicos.puccampinas.edu.br/seer/index.php/transinfo/article/download/683/663>. Acesso em: 20 maio 2014
- DB Engines. “**DB-Engines Ranking**”. Disponível em: <<http://db-engines.com/en/ranking>>. Acesso em: 02/06/2016.
- FONSECA, J. J. S. **Metodologia da pesquisa científica**. Fortaleza UEC, 2002. Apostila.
- GIL, A. C. **Como elaborar projetos de pesquisa**. 4. ed. São Paulo: Atlas, 2007.
- INMON, W. H.. **Building the Data Warehouse**. 3. ed. Canada: John Wiley & Sons, 2002. 412 p.
- INMON, William H. “**Building the Data Warehouse**” Wiley Publishing 2005 4 ed.
- INMON, William H. “**Data Warehouse and Software Development**” 2000
- KIMBALL, Ralph; CASERTA, Joe. “**The Data Warehouse ETL Toolkit: Practical Techniques for Extracting, Cleaning, Conforming, and Delivering Data**”. John Wiley & Sons 2004
- KIMBALL, R.; REEVES L.; ROSS M.; THORNTHWAITE W. **The Data Warehouse Lifecycle Toolkit: Expert Methods for Designing, Developing, and Deploying Data Warehouses**. John Wiley & Sons Inc., 1998.
- KIMBALL, Ralph; ROSS, Margy. **The Data Warehouse Toolkit: Guia completo para modelagem dimensional**. 2. ed. Rio de Janeiro: Campus, 2002. 493 p. Tradução de: Ana Beatriz Tavares e Daniela Lacerda.
- KIMBALL, Ralph; ROSS, Margy; THORNTHWAITE, Warren; MUNDY, Joy; BECKER, Bob. **The Data Warehouse Lifecycle Toolkit**. Second Edition. Indianapolis: Wiley Publishing, 2008.
- MACHADO, F. N. R. (2010) **Tecnologia e Projeto de Data Warehouse: uma visão multidimensional**, 5 ed., Érica, São Paulo.
- MENOLLI, A. L. A. (2006) **A Data Warehouse Architecture in Layers for Science and Technology**, Proceedings of the Eighteenth International Conference on

- Software Engineering Knowledge Engineering (SEKE'2006), San Francisco, CA, USA.
- MINAYO, M, C, S. (Org.). **Pesquisa Social: teoria método e criatividade**. Petrópolis, Vozes, 2001.
- NEGASH, Salomão; GRAY, Paul. “**Business Intelligence**”. Springer Berlin Heidelberg, 2008.
- POSTGRESQL. “**About us**”. Disponível em: <<https://www.postgresql.org/about/>>. Acesso em: 07/07/2016.
- SHIM, J. P., Warkentin, M., Courtney, J., Power, D. J., Sharda, R. e Carlsson, C. (2002) **Past, Present, And Future Of Decision Support Technology**, Decision Support System, V. 33, N. 2, P. 111-126.
- SOUZA, Elielson B. de; MENOLLI, André L. Andrade; COELHO, Ricardo G.. **Uma Metodologia Agile ROLAP para Implantação de Ambientes de Inteligência de Negócios**. 2014. Disponível em: <<http://projetos.uenp.edu.br/agilerolap/images/DOCUMENTOS/ERBD.pdf>>. Acesso em: 09 ago. 2015.
- SOUZA, Elielson Barbosa de; COELHO, Ricardo Gonçalves. Um Método Agile ROLAP para implantação de Ambientes de Inteligência de Negócios. 2014. 102 f. TCC (Graduação) - Curso de Sistemas de Informação, Centro de Ciências Tecnológicas, Universidade Estadual do Norte do Paraná, Bandeirantes, 2014.
- TARTUCE, T. J. A. **Métodos de pesquisa**. Fortaleza: UNICE – Ensino Superior, 2006. Apostila

Apêndice A – Medidas dos fatos

Quadro 2: Medidas do fato avaliacao aluno_professor.

Coluna	Nome	Formula	Agregador
q1	Media professor apresenta plano de ensino.	coluna!=0	avg
q2	Media professor cumpre conteudo programatico.	coluna!=0	avg
q3	Media professor varia tecnicas de ensino.	coluna!=0	avg
q4	Media professor utiliza diversos recursos.	coluna!=0	avg
q5	Media aplicacao pratica do conteudo.	coluna!=0	avg
q6	Media dominio teorico professor.	coluna!=0	avg
q7	Media dominio pratico professor.	coluna!=0	avg
q8	Media dominio pratica clinica professor.	coluna!=0	avg
q9	Media professor domina a disciplina.	coluna!=0	avg
q10	Media relacao professor alunos.	coluna!=0	avg
q11	Media aulas compreensíveis.	coluna!=0	avg
q12	Media professor soluciona duvidas durante a aula.	coluna!=0	avg
q13	Media professor atende extra-classe.	coluna!=0	avg
q14	Media avaliacoes coerentes com conteudo.	coluna!=0	avg
q15	Media extensao da prova adequada.	coluna!=0	avg
q16	Media professor pontual.	coluna!=0	avg
q17	Media professor nao falta.	coluna!=0	avg
q18	Media professor incentiva os	coluna!=0	avg

	alunos a participar de eventos.		
aa_q1	Media conhecimento aluno para acompanhar a disciplina.	coluna!=0	avg
aa_q2	Media motivacao aluno para acompanhar a disciplina.	coluna!=0	avg
aa_q3	Media cumprimento horario aluno.	coluna!=0	avg
aa_q4	Media participacao aluno na classe e extra-classe.	coluna!=0	avg
questionario_aluno_id	Quantidade alunos.		distinct-count
questionario_aluno_id	Quantidade alunos finalizaram algum questionario.	fim is not null	distinct-count
questionario_aluno_id	Quantidade alunos nao finalizaram algum questionario.	fim is null	diistinct-count
q1,q2...q18	Media geral professor.	$(Q1+q2...+q18)/18$	avg
aa_q1...aa_q4	Media geral aluno.	$(aa_q1+...+aa_q4)/4$	avg

Quadro 3: Medidas do fato avaliacao docente.

Coluna	Nome	Formula	Agregador
au_q1	Media divulgacao das normas.	coluna!=0	avg
au_q2	Media mudancas a partir das decisoes do colegiado de curso.	coluna!=0	avg
au_q3	Media praticas institucionais que estimulam a melhoria do ensino.	coluna!=0	avg
au_q4	Media relacao grade, demandas sociais, necessidades alunos.	coluna!=0	avg
au_q5	Media integracao pesquisa, ensino e extensao.	coluna!=0	avg
au_q6a	Media condicoes e incentivo para pesquisa pelo professor.	coluna!=0	avg
au_q6b	Media apoio participacao de eventos	coluna!=0	avg

	cientificos de apresentacao.		
au_q6c	Media relacao entre pesquisa e desenvolvimento da regioao.	coluna!=0	avg
au_q6d	Media incentivo a realizacao de eventos cientificos.	coluna!=0	avg
au_q6e	Media tramite dos projetos de pesquisa na UENP.	coluna!=0	avg
au_q6f	Media execucao dos projetos de pesquisa na UENP.	coluna!=0	avg
au_q6g	Media integracao pesquisa com a atividade de ensino da UENP.	coluna!=0	avg
au_q6h	Media politicas criacao e expansao pos-graduacao lato sensu.	coluna!=0	avg
au_q6i	Media politicas criacao e expansao pos-graduacao stricto sensu.	coluna!=0	avg
au_q6j	Media politicas manutencao pos-graduacao stricto sensu.	coluna!=0	avg
au_q7a	Media condicoes e incentivo para desenvolvimento de projetos de extensao.	coluna!=0	avg
au_q7b	Media tramite projetos de extensao na UENP.	coluna!=0	avg
au_q7c	Media dos projetos de extensao na UENP.	coluna!=0	avg
au_q7d	Media relevancia dos projetos de extensao UENP para desenvolvimento local.	coluna!=0	avg
au_q7e	Media integracao extensao e atividades de ensino da UENP.	coluna!=0	avg
au_q7f	Media incentivo a participacao de alunos em projetos de extensao.	coluna!=0	avg
au_q7g	Media incentivo realizacao eventos extensao.	coluna!=0	avg

au_q7h	Media apoio UENP na participação em eventos de extensão.	coluna!=0	avg
au_q7i	Media incentivo a realização de eventos culturais.	coluna!=0	avg
au_q8	Media políticas de qualificação e titulação docente.	coluna!=0	avg
au_q9	Media políticas atendimento portadores necessidades especiais.	coluna!=0	avg
au_q10	Media políticas atendimento discentes.	coluna!=0	avg
au_q11	Media acesso ao telefone.	coluna!=0	avg
au_q12	Media acesso a microcomputadores.	coluna!=0	avg
au_q13	Media atendimento das secretarias.	coluna!=0	avg
au_q14	Media atendimento da telefonista.	coluna!=0	avg
au_q15	Media atendimento do protocolo.	coluna!=0	avg
au_q16	Media atendimento do serviço de fotocópia.	coluna!=0	avg
au_q17	Media atendimento da biblioteca.	coluna!=0	avg
au_q18a	Media atuação do coordenador de colegiado de curso.	coluna!=0	avg
au_q18b	Media atuação do diretor de centro de estudos.	coluna!=0	avg
au_q18c	Media atuação da diretoria do campus.	coluna!=0	avg
au_q18d	Media atuação do Vice-reitor em exercício da reitoria.	coluna!=0	avg
au_q19a	Media Pro-Reitoria (PROAF).	coluna!=0	avg
au_q19b	Media Pro-Reitoria (PROPG).	coluna!=0	avg
au_q19c	Media Pro-Reitoria (PROEC).	coluna!=0	avg
au_q19d	Media Pro-Reitoria (PROGRAD).	coluna!=0	avg
au_q19e	Media Pro-Reitoria (PRORH).	coluna!=0	avg
au_q19f	Media Pro-Reitoria (PROPAV).	coluna!=0	avg
au_q20	Media satisfação com as atividades que desenvolve.	coluna!=0	avg
au_q21	Media condições espaço físico onde	coluna!=0	avg

	desenvolve atividades de permanencia.		
au_q22	Media condicoes espaco fisico onde desenvolve atividades de ensino.	coluna!=0	avg
au_q23	Media condicoes espaco fisico onde desenvolve atividades de pesquisa e extensao.	coluna!=0	avg
au_q24a	Media salas sao arejadas.	coluna!=0	avg
au_q24b	Media salas sao iluminadas.	coluna!=0	avg
au_q24c	Media salas tem boa acustica.	coluna!=0	avg
au_q24d	Media salas sao amplas para comportar o numero de alunos.	coluna!=0	avg
au_q24e	Media salas apresentam mobiliario adequado.	coluna!=0	avg
au_q24f	Media salas dispoem de recursos audiovisuais.	coluna!=0	avg
au_q25	Media sala permanencia professores para atendimento alunos está adequada.	coluna!=0	avg
au_q26a	Media biblioteca possui acervo com qualidade.	coluna!=0	avg
au_q26b	Media biblioteca possui numero suficiente de exemplares.	coluna!=0	avg
au_q26c	Media possui instalacoes para a leitura e estudo satisfatorias.	coluna!=0	avg
au_q27	Media numero dos equipamentos disponiveis atendem aos de alunos.	coluna!=0	avg
au_q28	Media condicoes dos equipamentos disponiveis atendem aos de alunos.	coluna!=0	avg
au_q29	Media situacao dos laboratorios de informatica.	coluna!=0	avg
au_q30a	Media limpeza das salas de aula.	coluna!=0	avg
au_q30b	Media conservacao das salas de aula.	coluna!=0	avg
au_q30c	Media limpeza do campus.	coluna!=0	avg
au_q30d	Media conservacao do campus.	coluna!=0	avg

au_q30e	Media limpeza dos banheiros do campus.	coluna!=0	avg
au_q30f	Media conservacao dos banheiros do campus.	coluna!=0	avg
au_q30g	Media Iluminacao do campus.	coluna!=0	avg
au_q30h	Media seguranca do campus.	coluna!=0	avg
au_q31	Media participaco nas decisao da UENP.	coluna!=0	avg
au_q32	Media clareza funcionamento administrativo UENP.	coluna!=0	avg
au_q33	Media cumprimento (PDI) da UENP.	coluna!=0	avg
au_q34	Media imagem e contribuicao UENP na sociedade.	coluna!=0	avg
cpa_q1	Media avaliacao da avaliacao.	coluna!=0	avg
chave_acesso	Quantidade docentes.		distinct-count
chave_acesso	Quantidade docentes finalizaram algum questionario.	fim is not null	distinct-count
chave_acesso	Quantidade docentes nao finalizaram algum questionario.	fim is null	diistinct-count

Quadro 4: Medidas do fato avaliacao estagiario

Coluna	Nome	Formula	Agregador
ae_q1	Media oferta de cursos e palestras pela instituicao.	coluna!=0	avg
ae_q2	Media divulgacao das normas sobre estagio.	coluna!=0	avg
ae_q3	Media pagamento em dia da bolsa auxilio.	coluna!=0	avg
ae_q4	Media agamento em dia do vale transporte.	coluna!=0	avg
ae_q5	Media reducao da carga horaria em epoca de prova.	coluna!=0	avg
ae_q6	Media cumprimento do recesso remunerado.	coluna!=0	avg

ae_q7a	Media atuacao do Reitor em exercicio na reitoria.	coluna!=0	avg
ae_q7b	Media atuacao da Diretoria do Campus.	coluna!=0	avg
ae_q7c	Media atuacao do Supervisor/Chefe Imediato/Diretor de Centro.	coluna!=0	avg
ae_q7d	Media atuacao do Gestor de Estagio.	coluna!=0	avg
ae_q8	Media satisfacao com o estagio.	coluna!=0	avg
ae_q9	Media condicoes espaco onde desenvolve atividades.	coluna!=0	avg
ae_q10a	Media limpeza do setor onde trabalha.	coluna!=0	avg
ae_q10b	Media conservacao do setor onde trabalha.	coluna!=0	avg
ae_q11	Media clareza funcionamento administrativo da UENP.	coluna!=0	avg
ae_q12	Media equipamentos e materiais disponiveis atividades.	coluna!=0	avg
ae_q13	Media imagem publica e contribuicoes da UENP na sociedade.	coluna!=0	avg
cpa_q1	Media avaliacao da avaliacao.	coluna!=0	avg
chave_acesso	Quantidade estagiarios.		distinct-count
chave_acesso	Quantidade estagiarios finalizaram algum questionario.	fim is not null	distinct-count
chave_acesso	Quantidade estagiarios nao finalizaram algum questionario.	fim is null	diistinct-count

Quadro 5: Medidas do fato avaliacao tecnico

Coluna	Nome	Formula	Agregador
au_q1	Media programa qualificacao profissional.	coluna!=0	avg
au_q2	Media melhoria qualidade de vida na periodo de trabalho.	coluna!=0	avg
au_q3	Media valorizacao do servidor.	coluna!=0	avg

au_q4	Media programa de qualificacao profissional e melhoria da qualidade de vida.	coluna!=0	avg
au_q5	Media nota quantidade de pessoal que trabalha no seu setor.	coluna!=0	avg
au_q6	Media divulgacao das normas.	coluna!=0	avg
au_q7	Media circulacao de informacao na UENP.	coluna!=0	avg
au_q8	Media acesso a equipamentos de comunicacao.	coluna!=0	avg
au_q9	Media acesso a equipamentos de informacao.	coluna!=0	avg
au_q10	Media protocolo, fluxo e distribuicao de documentos.	coluna!=0	avg
au_q11a	Media atuacao da diretoria do campus.	coluna!=0	avg
au_q11b	Media atua do Vice-reitor em exercicio da reitoria.	coluna!=0	avg
au_q12a	Media pro-Reitoria (PROAF).	coluna!=0	avg
au_q12b	Media pro-Reitoria (PROPG).	coluna!=0	avg
au_q12c	Media pro-Reitoria (PROEC).	coluna!=0	avg
au_q12d	Media pro-Reitoria (PROGRAD).	coluna!=0	avg
au_q12e	Media pro-Reitoria (PRORH).	coluna!=0	avg
au_q12f	Media pro-Reitoria (PROPAV).	coluna!=0	avg
au_q13	Media satisfacao com as atividades que desenvolve.	coluna!=0	avg
au_q14	Media condicoes espaço fisico onde desenvolve atividades.	coluna!=0	avg
au_q15a	Media limpeza setor onde trabalha.	coluna!=0	avg
au_q15b	Media conservacao setor onde trabalha.	coluna!=0	avg
au_q15c	Media limpeza do campus.	coluna!=0	avg
au_q15d	Media conservacao do campus.	coluna!=0	avg
au_q15e	Media limpeza dos banheiros do campus.	coluna!=0	avg

au_q15f	Media conservacao dos banheiros do campus.	coluna!=0	avg
au_q15g	Media iluminacao do campus.	coluna!=0	avg
au_q15h	Media seguranca do campus.	coluna!=0	avg
au_q16	Media participacao nas decisos da UENP.	coluna!=0	avg
au_q17	Media clareza funcionamento administrativo da UENP.	coluna!=0	avg
au_q18	Media equipamentos e materiais disponiveis para atividades.	coluna!=0	avg
au_q19	Media imagem e contribuicao UENP na sociedade.	coluna!=0	avg
au_q20	Media cumprimento (PDI) da UENP.	coluna!=0	avg
cpa_q1	Media avaliacao da avaliacao.	coluna!=0	avg
chave_acesso	Quantidade tecnicos.		distinct-count
chave_acesso	Quantidade tecnicos finalizaram algum questionario.	fim is not null	distinct-count
chave_acesso	Quantidade tecnicos nao finalizaram algum questionario.	fim is null	diistinct-count

Apêndice B – Configuração das dimensões

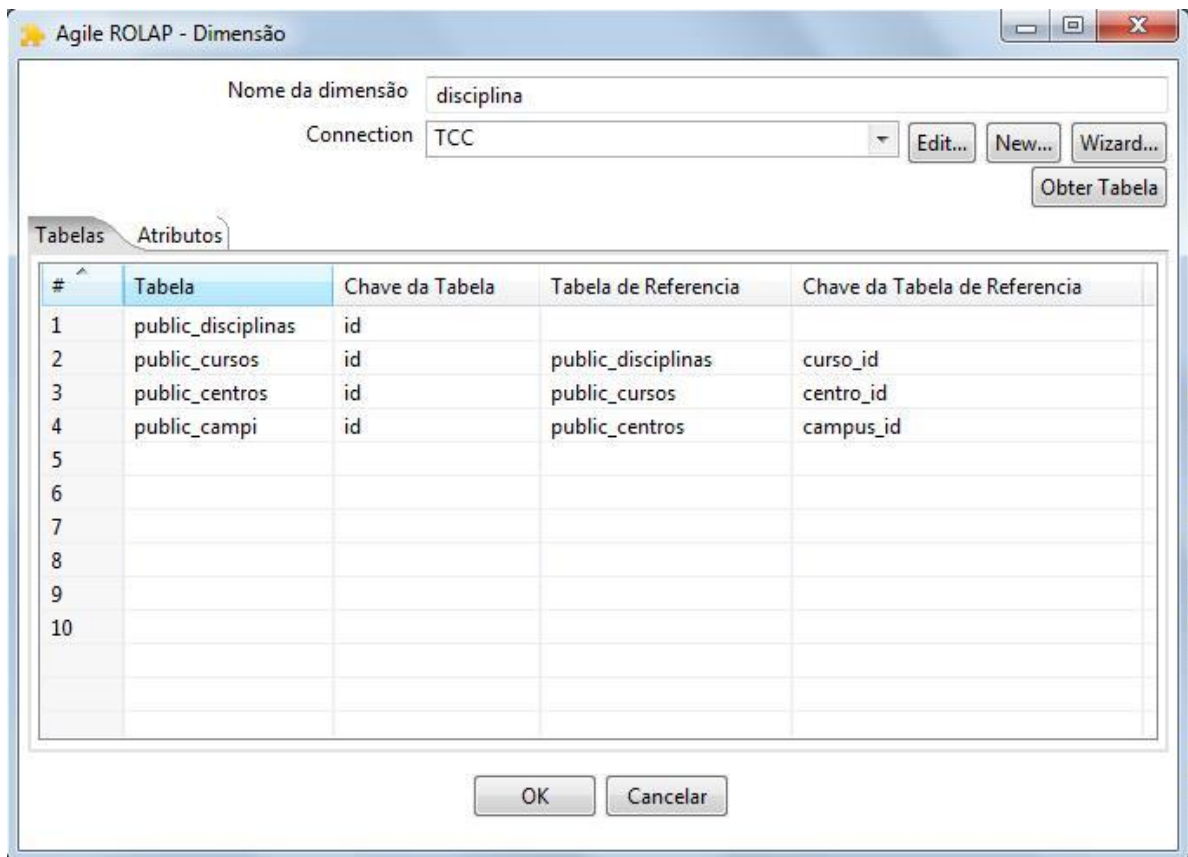


Figura 39: Configuração da dimensão disciplina no *plugin* dimensão.

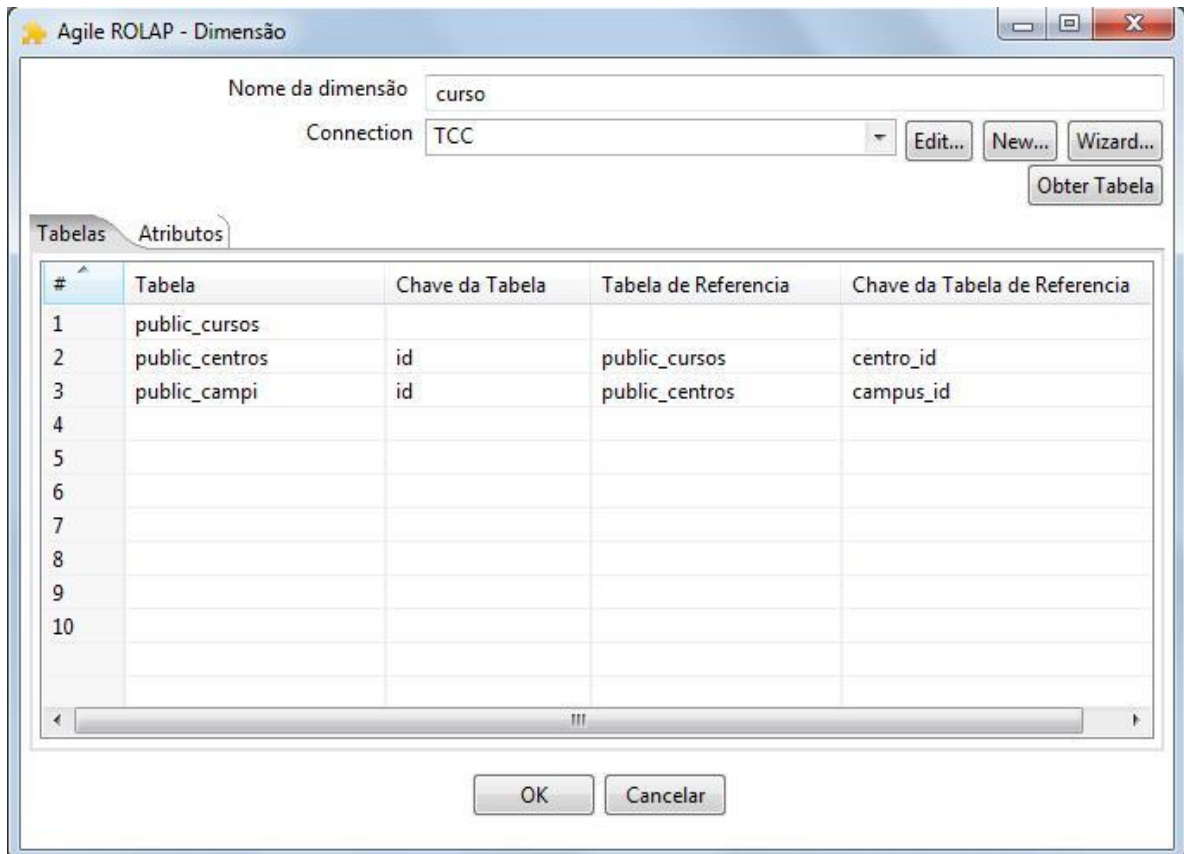


Figura 40: Configuração da dimensão curso no *plugin* dimensão.

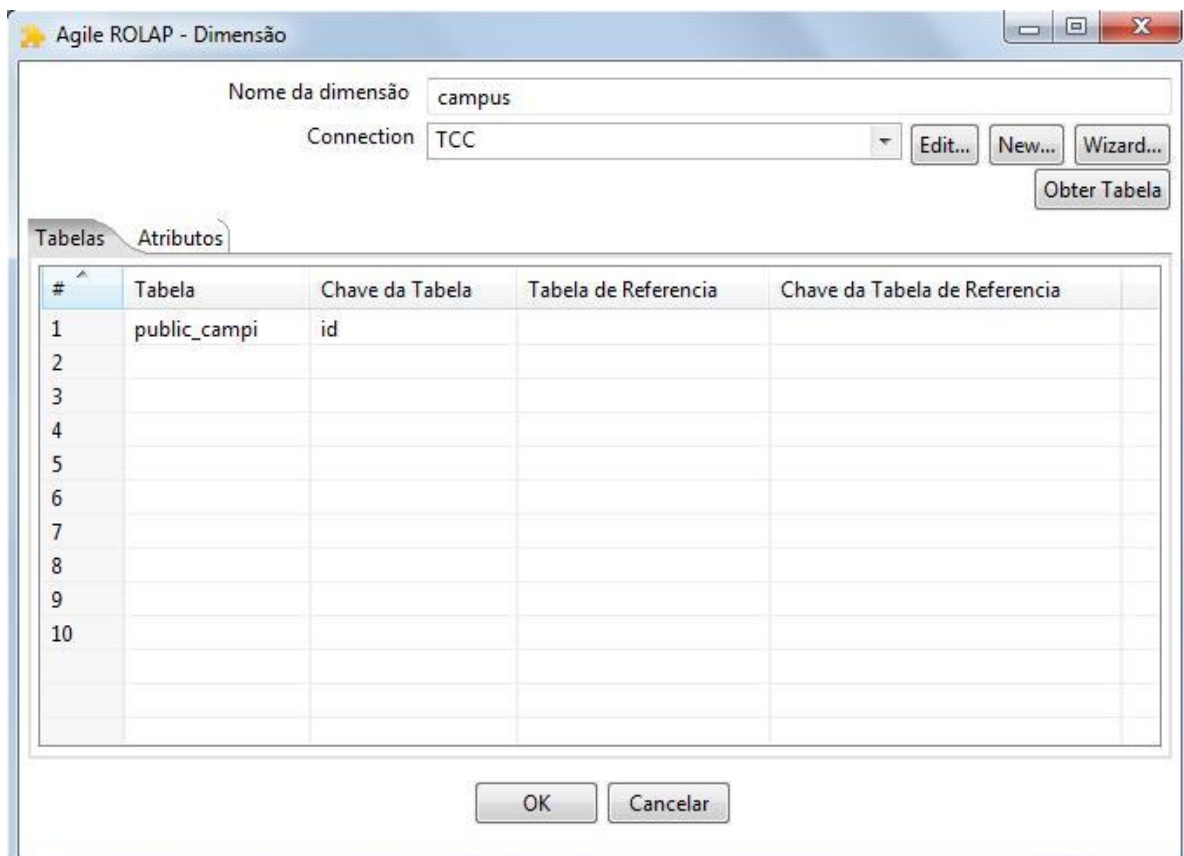


Figura 41: Configuração da dimensão campus no *plugin* dimensão.

Apêndice C – Configuração dos fatos

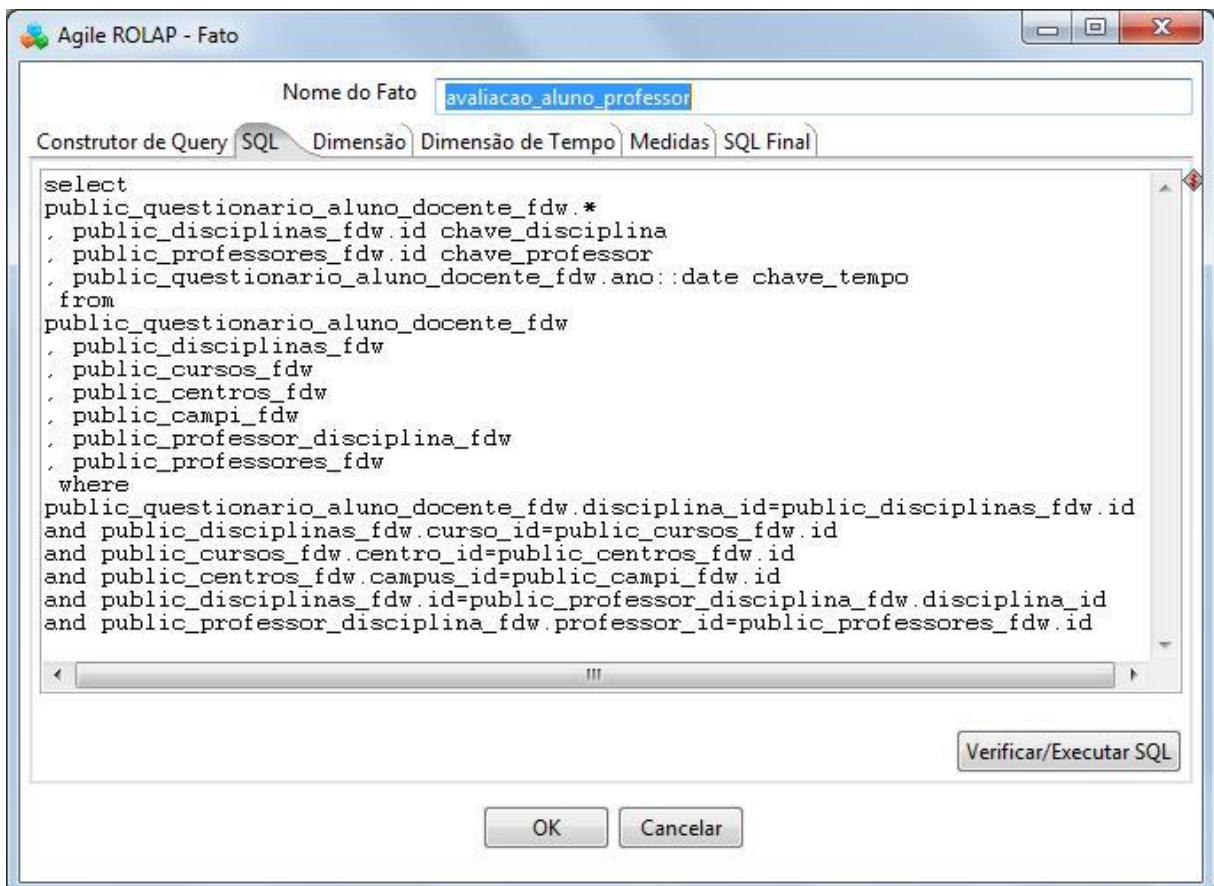


Figura 42: Fato avaliacao aluno_professor na aba de configuração SQL.

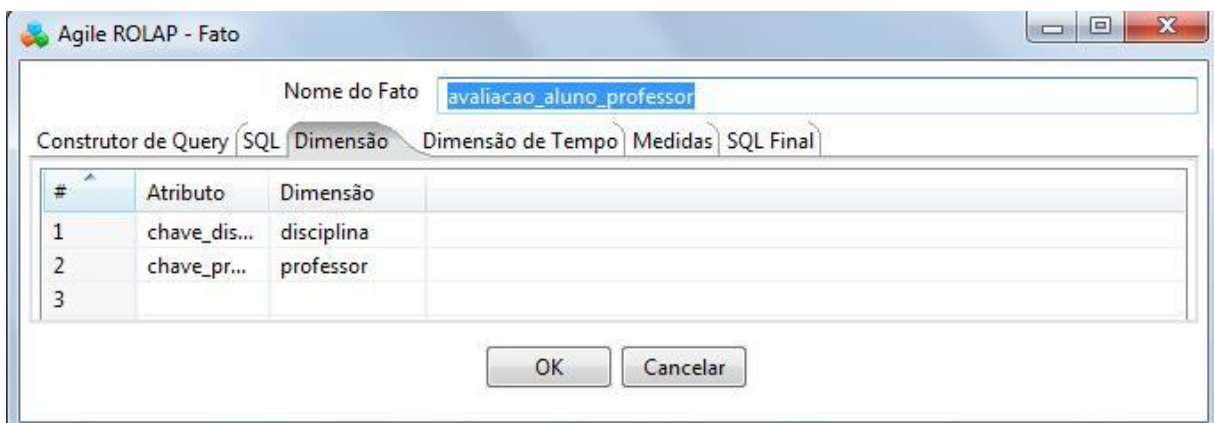


Figura 43: Fato avaliacao aluno_professor na aba de configuração Dimensão.

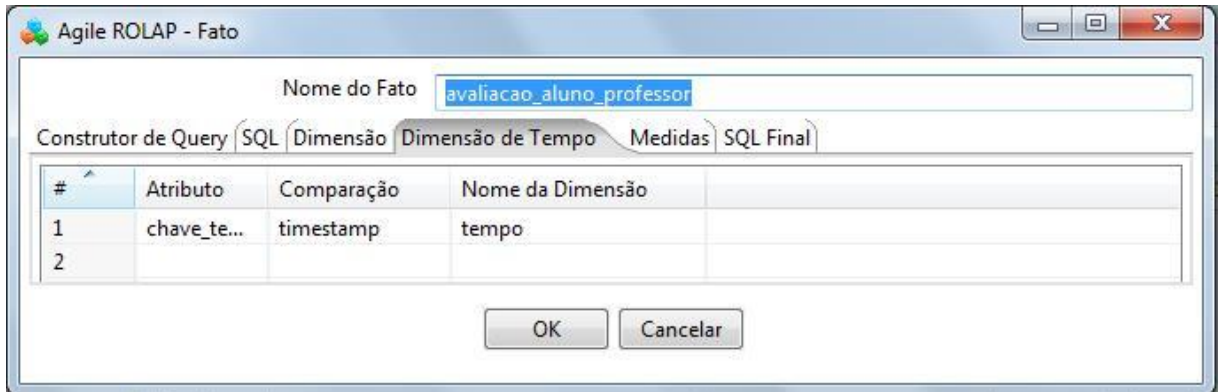


Figura 44: Fato avaliacao aluno_professor na aba de configuração Dimensão de Tempo.

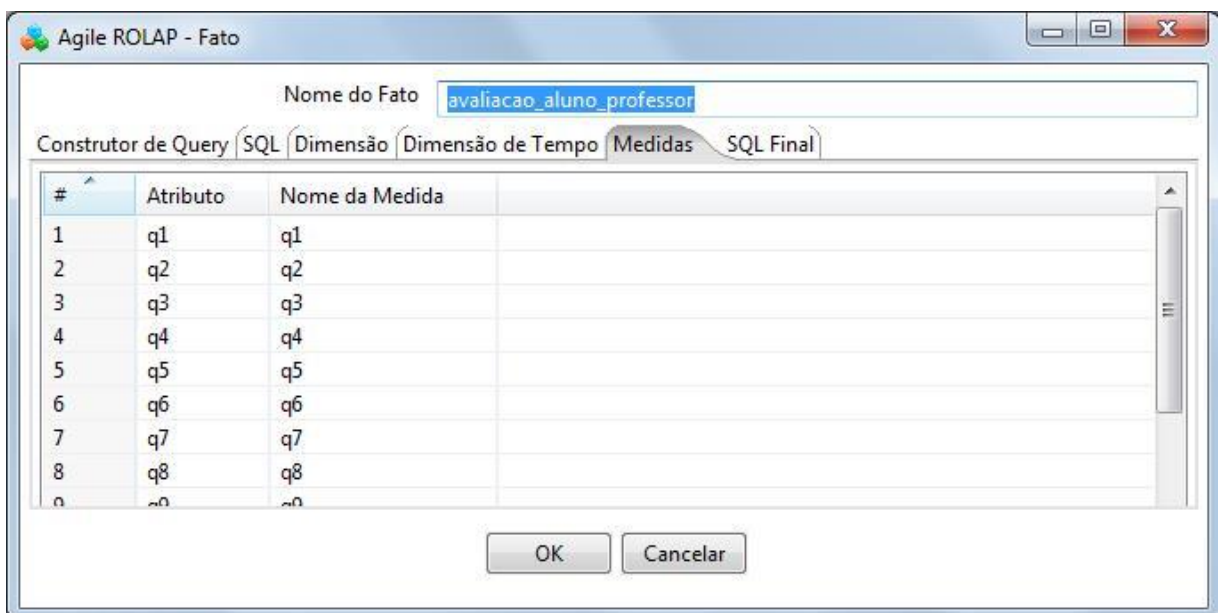


Figura 45: Fato avaliacao aluno_professor na aba de configuração Medidas.

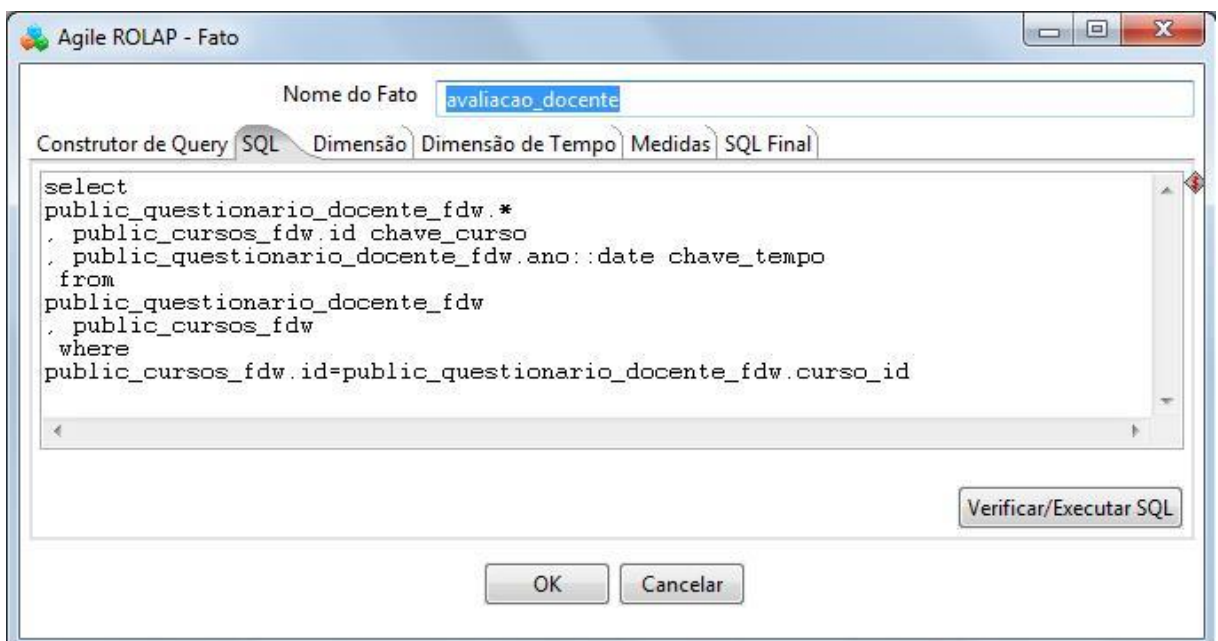


Figura 46: Fato avaliacao docente na aba de configuração SQL.

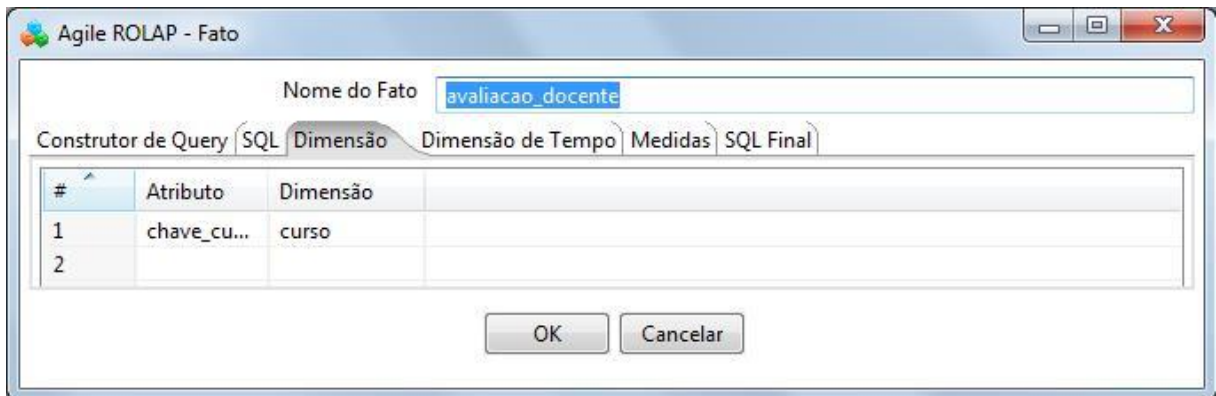


Figura 47: Fato avaliacao docente na aba de configuração Dimensão.

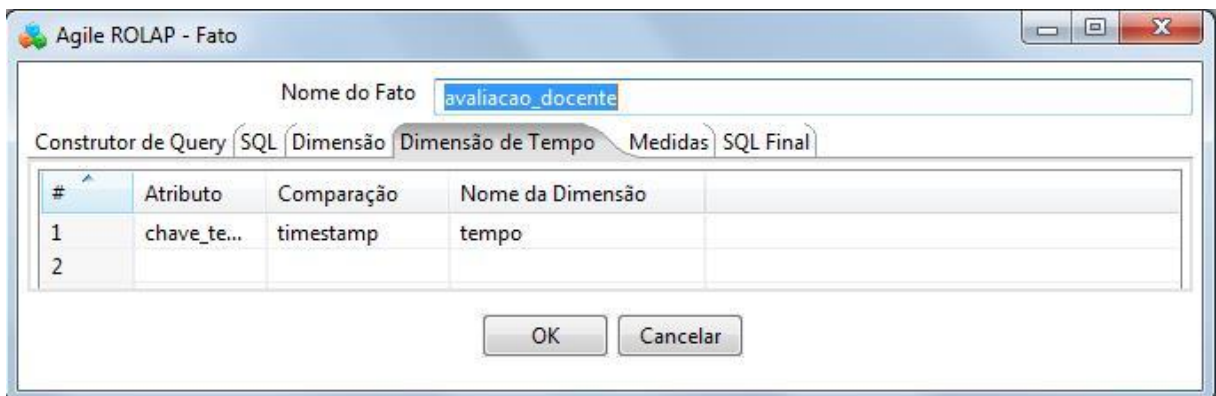


Figura 48: Fato avaliacao docente na aba de configuração Dimensão de Tempo.

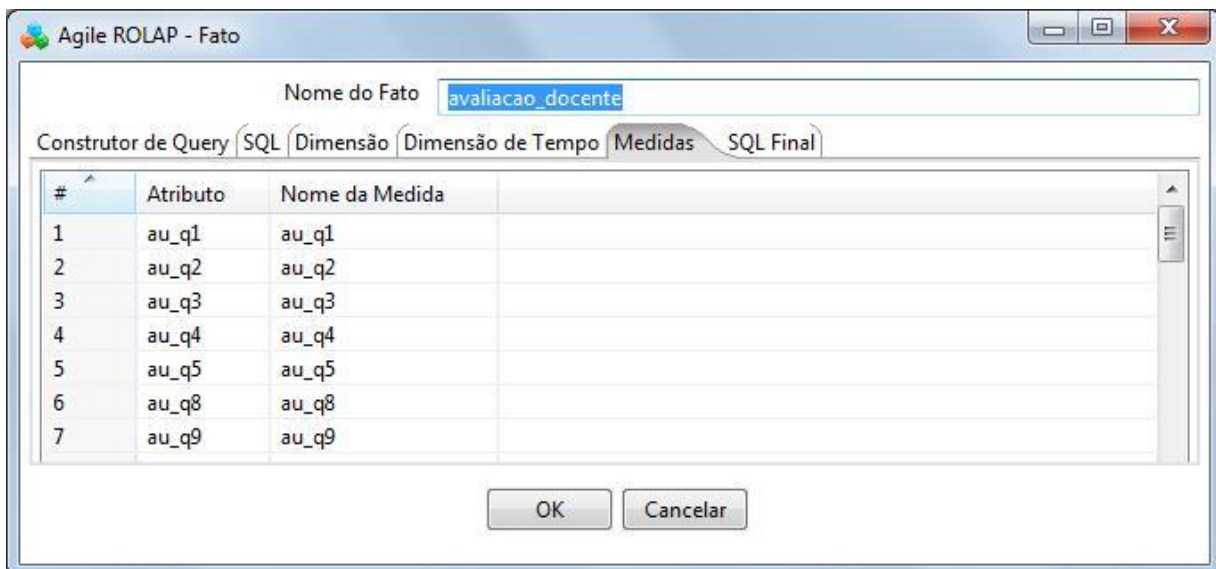


Figura 49: Fato avaliacao docente na aba de configuração Medidas.

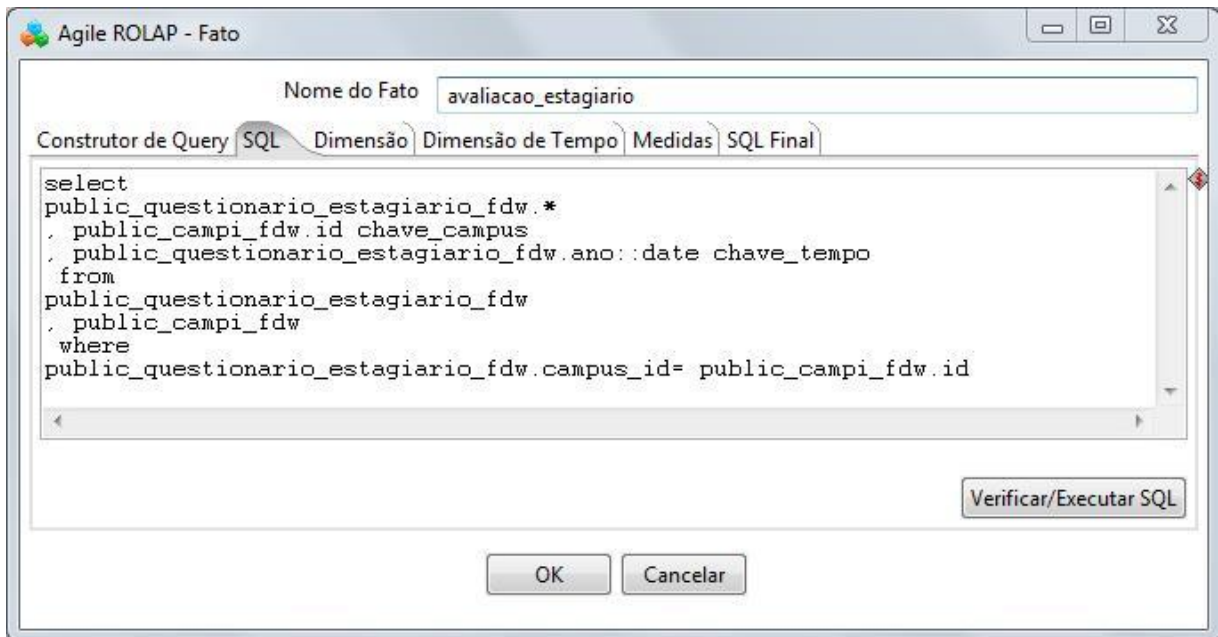


Figura 50: Fato avaliacao estagiario na aba de configuração SQL.

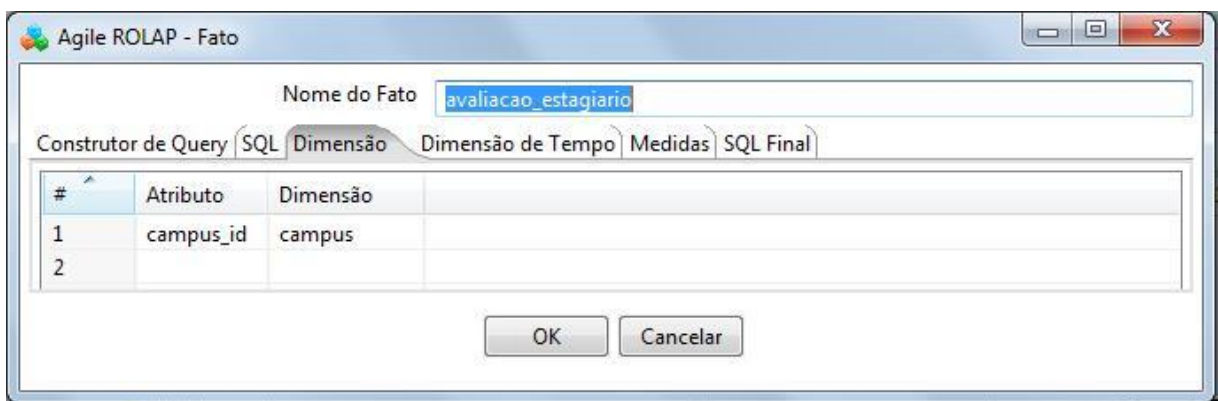


Figura 51: Fato avaliacao estagiario na aba de configuração Dimensão.

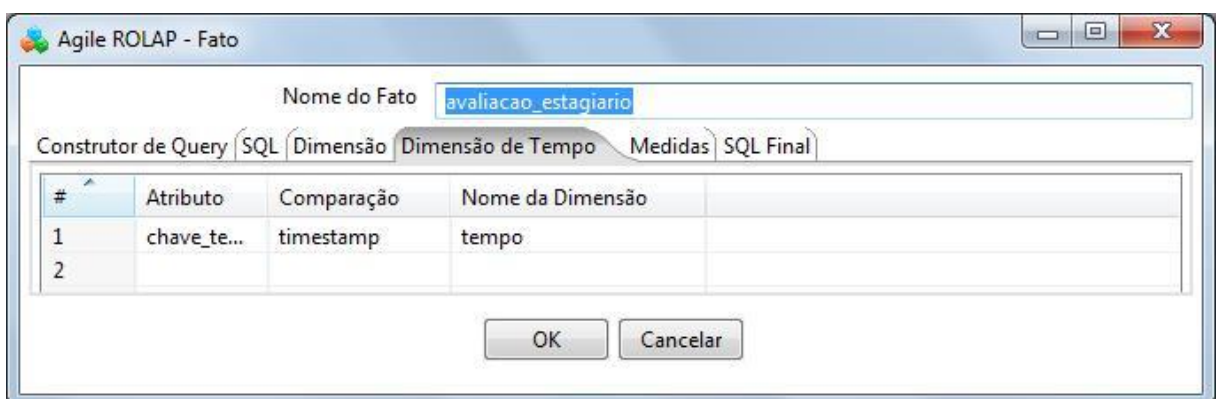


Figura 52: Fato avaliacao estagiario na aba de configuração Dimensão de Tempo.

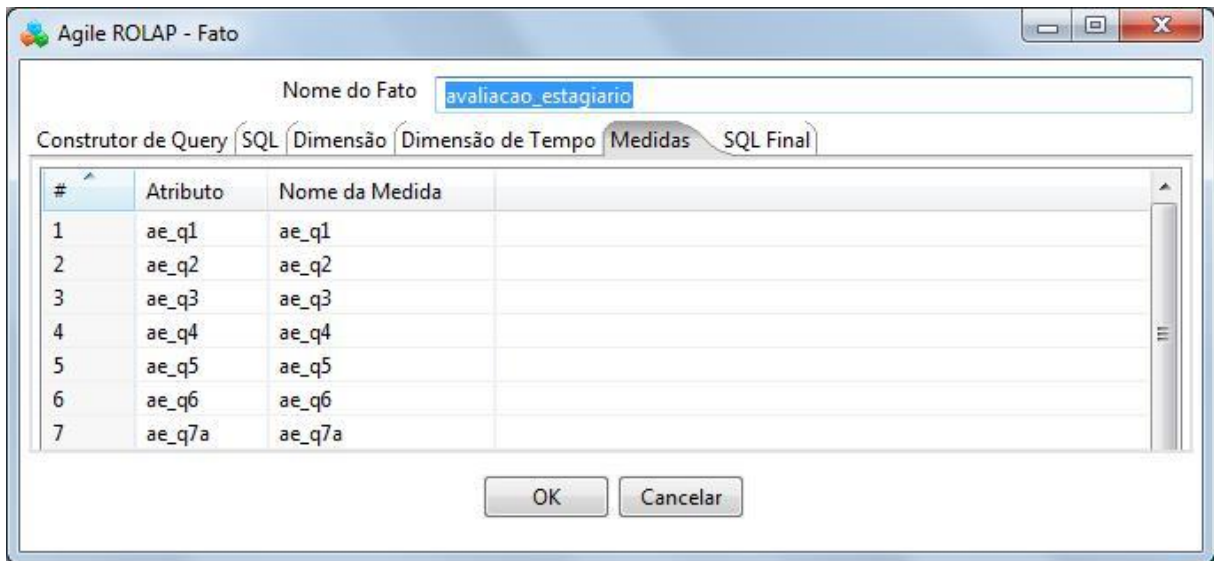


Figura 53: Fato avaliacao estagiario na aba de configuração Medidas.

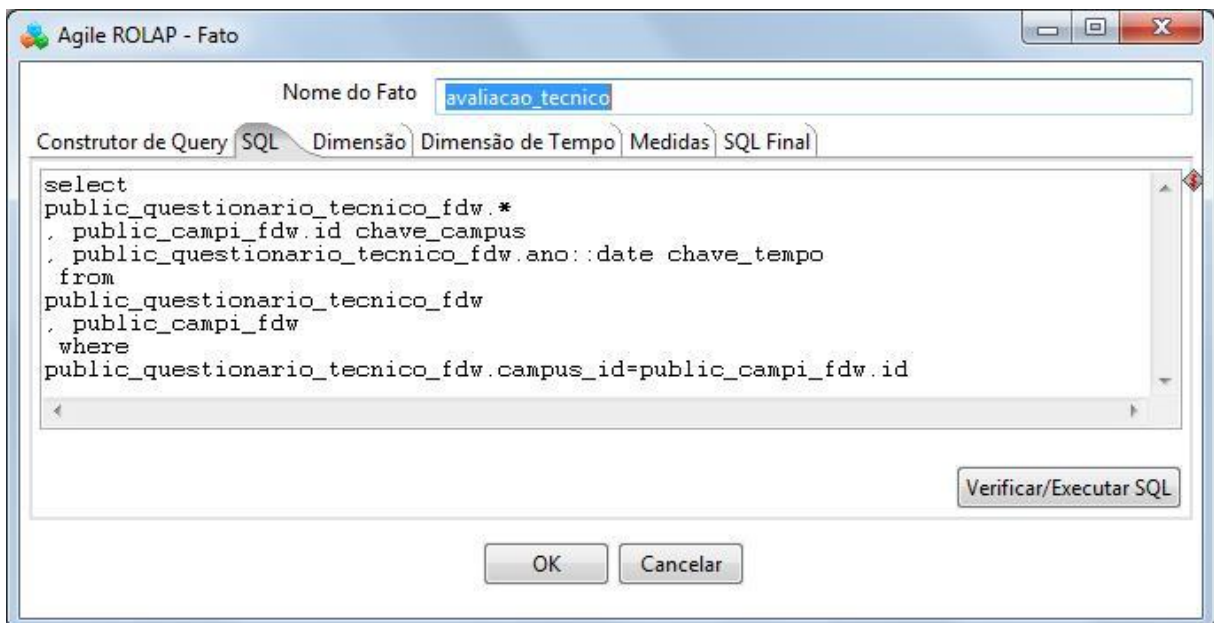


Figura 54: Fato avaliacao tecnico na aba de configuração SQL.

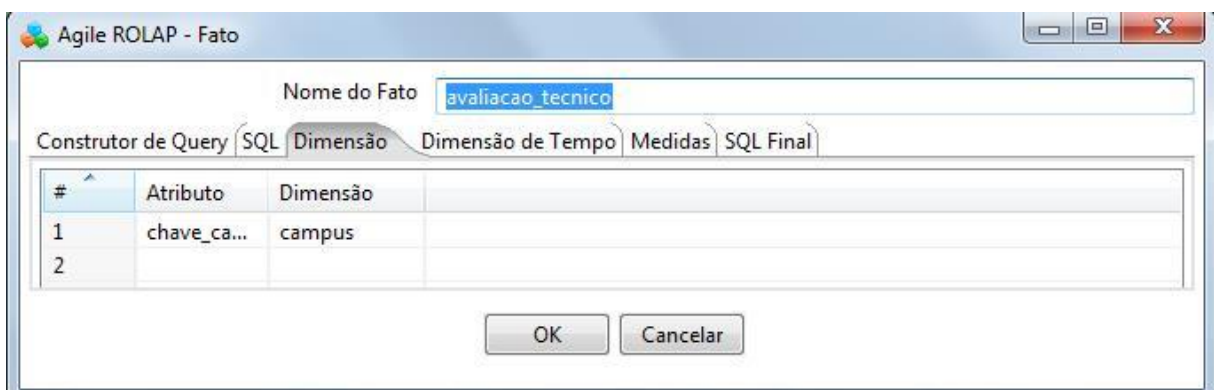


Figura 55: Fato avaliacao tecnico na aba de configuração Dimensão.

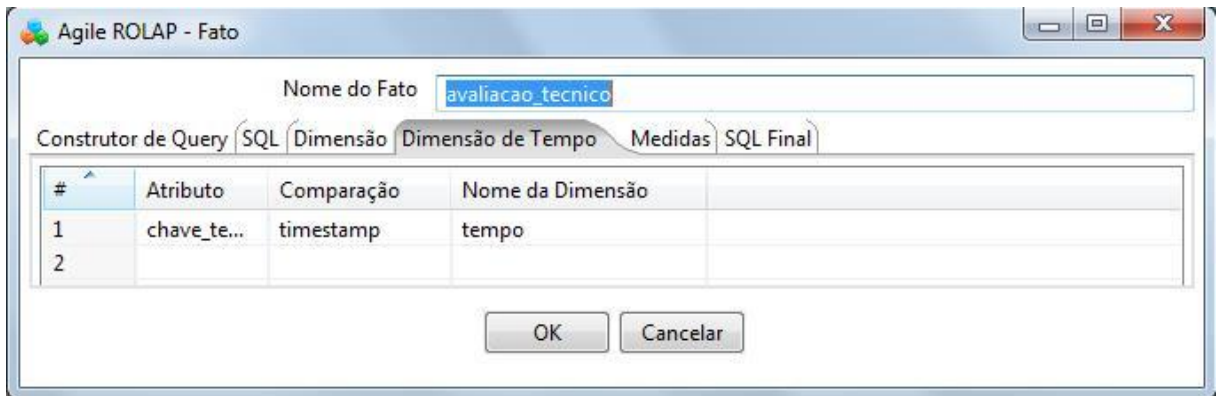


Figura 56: Fato avaliacao tecnico na aba de configuração Dimensão de Tempo.

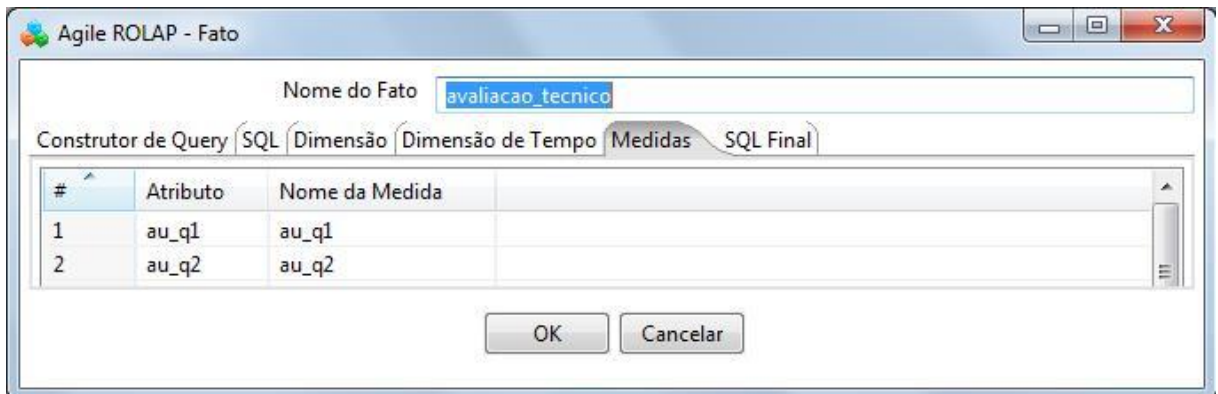


Figura 57: Fato avaliacao tecnico na aba de configuração Medidas.

Apêndice D – Configuração das dimensões lógicas

The screenshot shows the 'Dimensão Lógica' dialog box with the following configuration:

- Nome do passo:** Disciplina
- Carregar de:** C:\Users\Eduardo\Desktop\TCC\kettle\xml intermediario.xml
- Dimensão:**
 - Selecionar dimensão:** disciplina
 - Nome da dimensão:** Disciplina
 - Visível:**
- Hierarquia:**
 - Hierarquia (Left):** Disciplina, Disciplina_simples
 - Nível (Right):** Campus, Centro, Curso, Regime, Serie, Disciplina, Turma
- Hierarquia (Right):**
 - Nome:** Disciplina
 - Visível:**
- Nível (Right):**
 - Nome:** Campus
 - Coluna:** public_campi_fdw - nome
 - Tipo:** String
 - Visível:**
- Buttons:**
 - Up/Down arrows for Hierarquia and Nível
 - Nova Hierarquia, Novo Nível, Deletar Hierarquia, Deletar Nível
 - Ok, Cancelar

Figura 58: Configuração dimensão disciplina no *plugin* dimensão lógica.

Dimensão Lógica

Nome do passo:

Carregar de:

Dimensão

Selecionar dimensão:

Nome da dimensão:

Visível:

Hierarquia	Nível
Curso	Campus
	Centro
	Curso
	Regime

Hierarquia

Nome:

Visível:

Nível

Nome:

Coluna:

Tipo:

Visível:

Figura 59: Configuração dimensão curso no *plugin* dimensão lógica.

The image shows a software window titled "Dimensão Lógica" with the following configuration fields:

- Nome do passo:**
- Carregar de:**
- Dimensão**
 - Selecionar dimensão:**
 - Nome da dimensão:**
 - Visível:**
- Hierarquia**
 -
- Nível**
 -
- Hierarquia (Detailed)**
 - Nome:**
 - Visível:**
- Nível (Detailed)**
 - Nome:**
 - Coluna:**
 - Tipo:**
 - Visível:**

At the bottom, there are four arrow buttons (up, down, up, down), four buttons labeled "Nova Hierarquia", "Novo Nível", "Deletar Hierarquia", and "Deletar Nível", and "Ok" and "Cancelar" buttons.

Figura 60: Configuração dimensão campus no *plugin* dimensão lógica.

Apêndice E – Configuração dos cubos

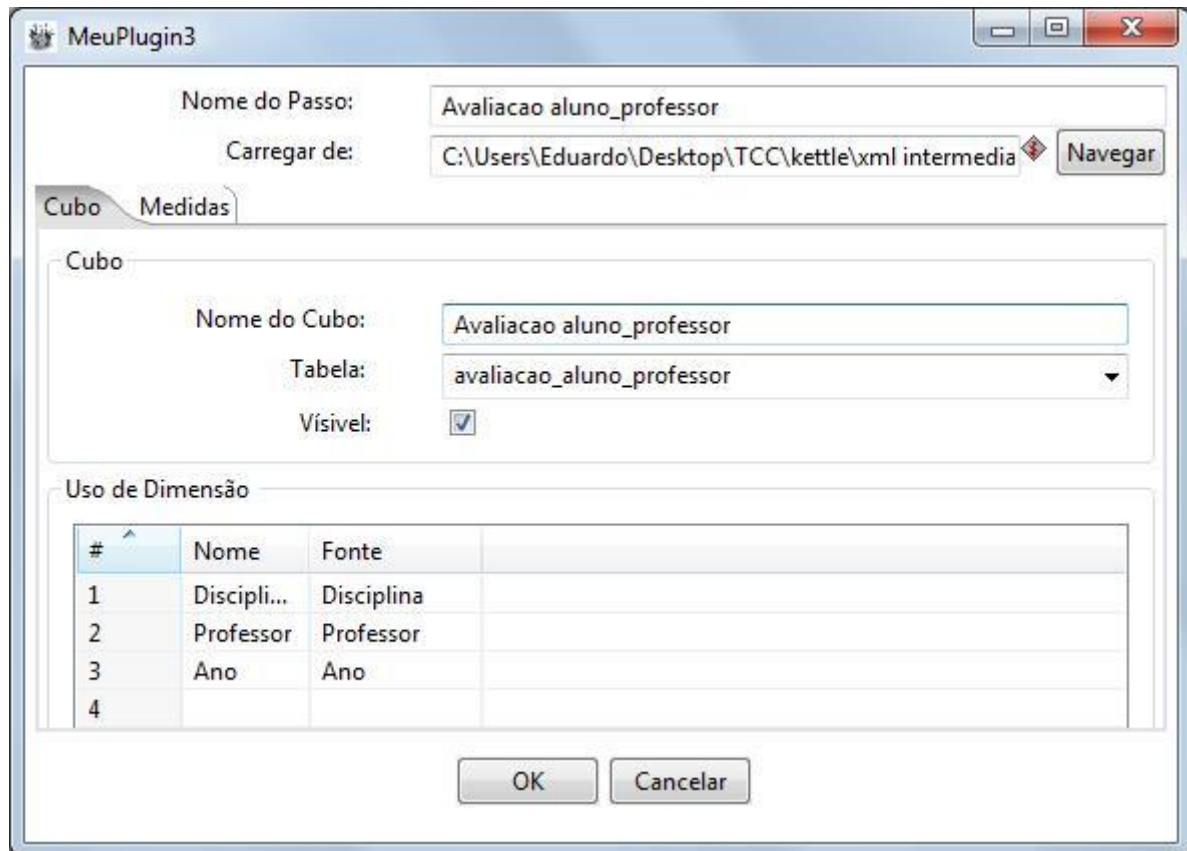


Figura 61: Cubo Avaliacao aluno_professor na aba de configuração Cubo.

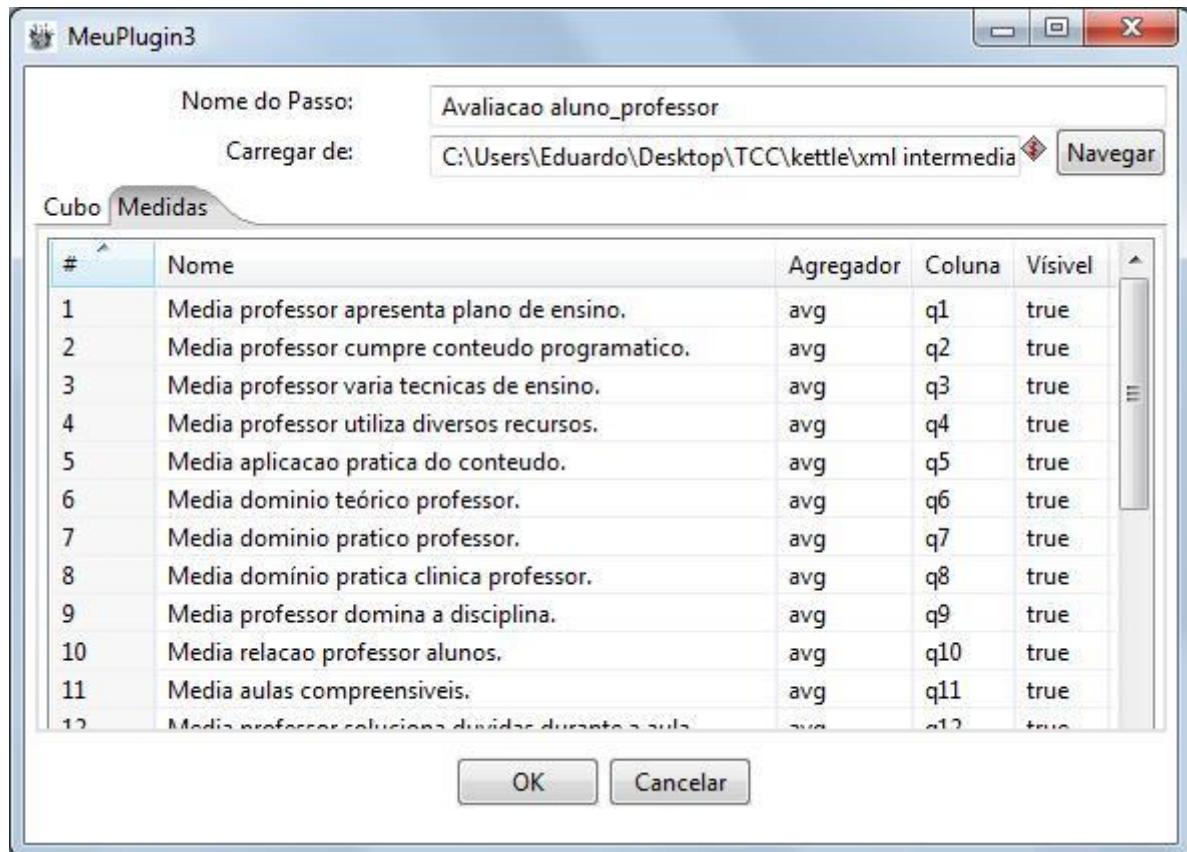


Figura 62: Cubo Avaliacao aluno_professor na aba de configuração Medidas.

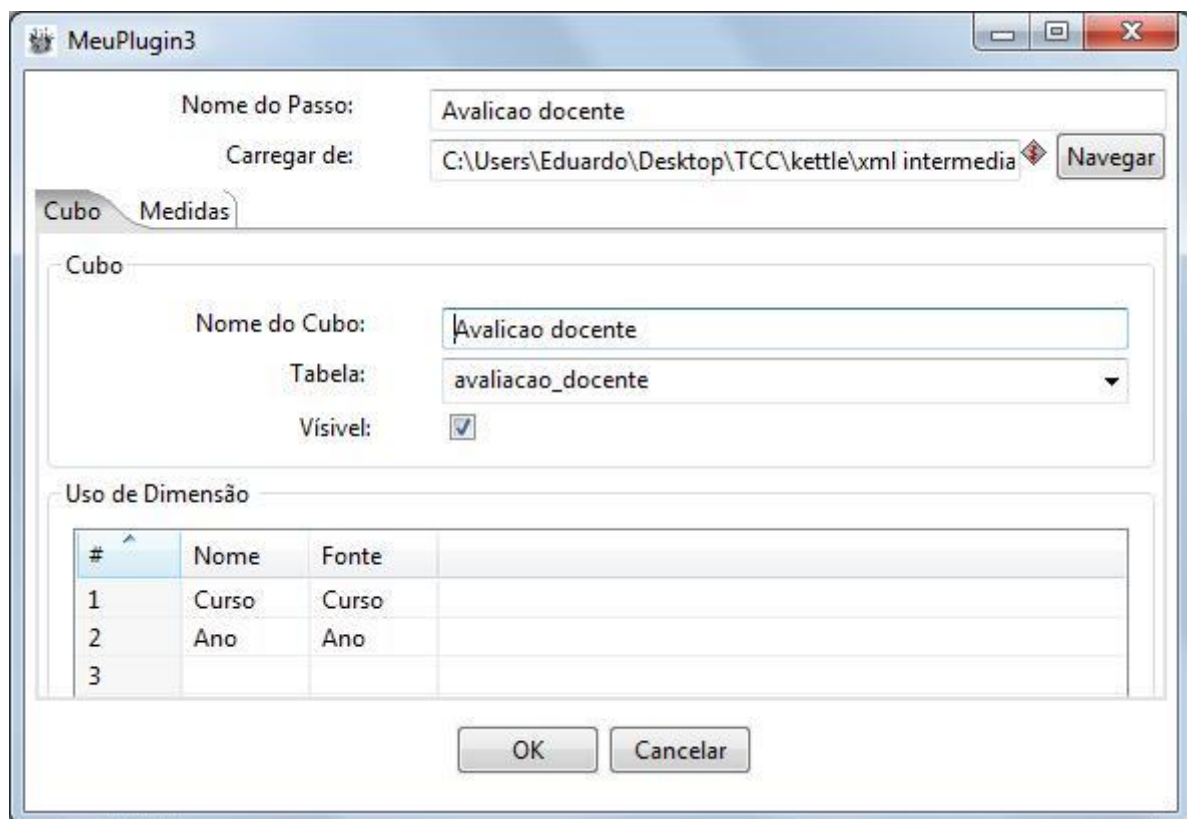


Figura 63: Cubo Avaliacao docente na aba de configuração Cubo.

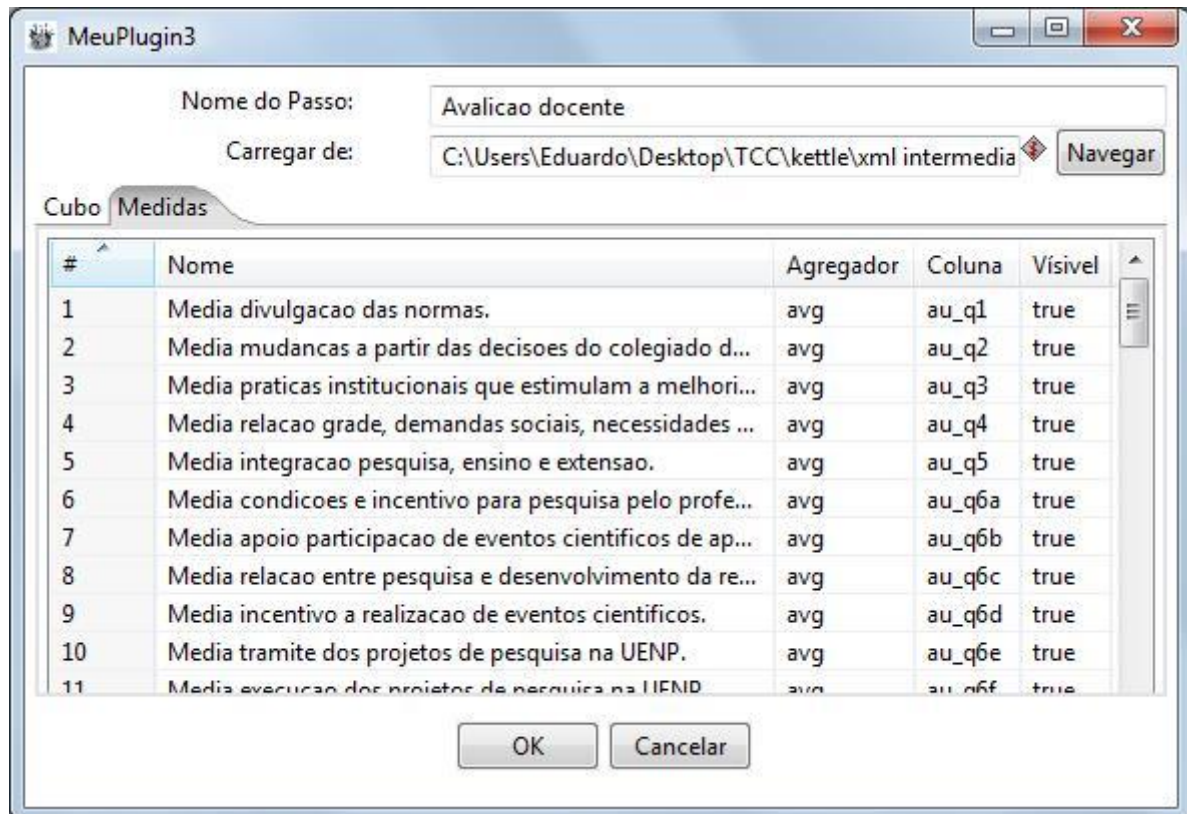


Figura 64: Cubo Avaliacao docente na aba de configuração Medidas.

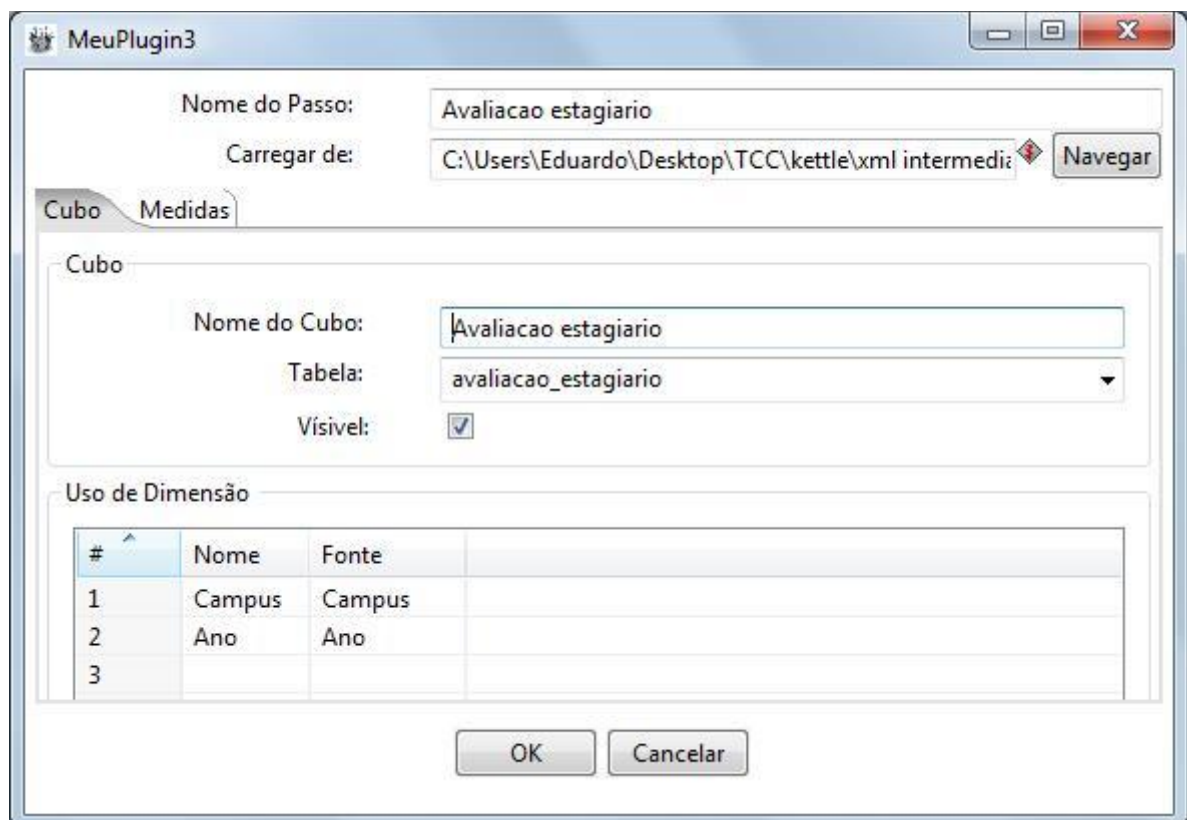


Figura 65: Cubo Avaliacao estagiario na aba de configuração Cubo.

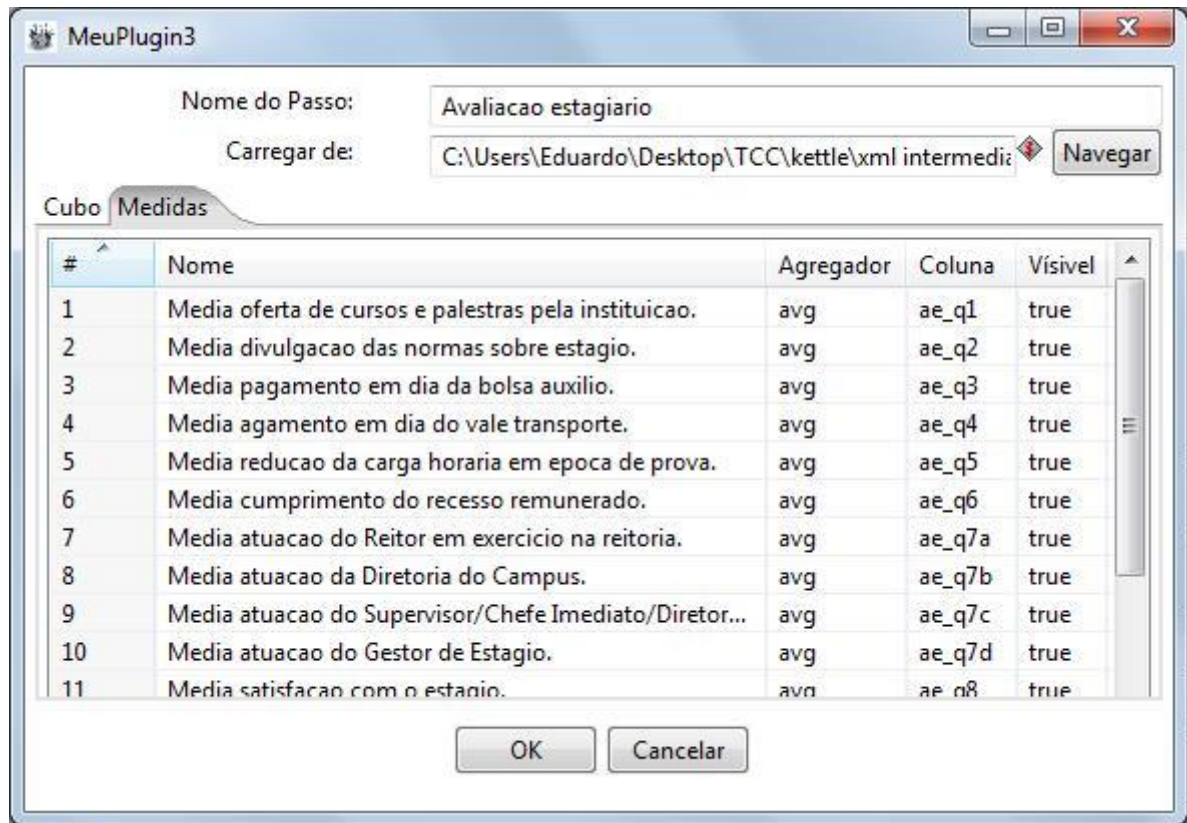


Figura 66: Cubo Avaliacao estagiario na aba de configuração Medidas.

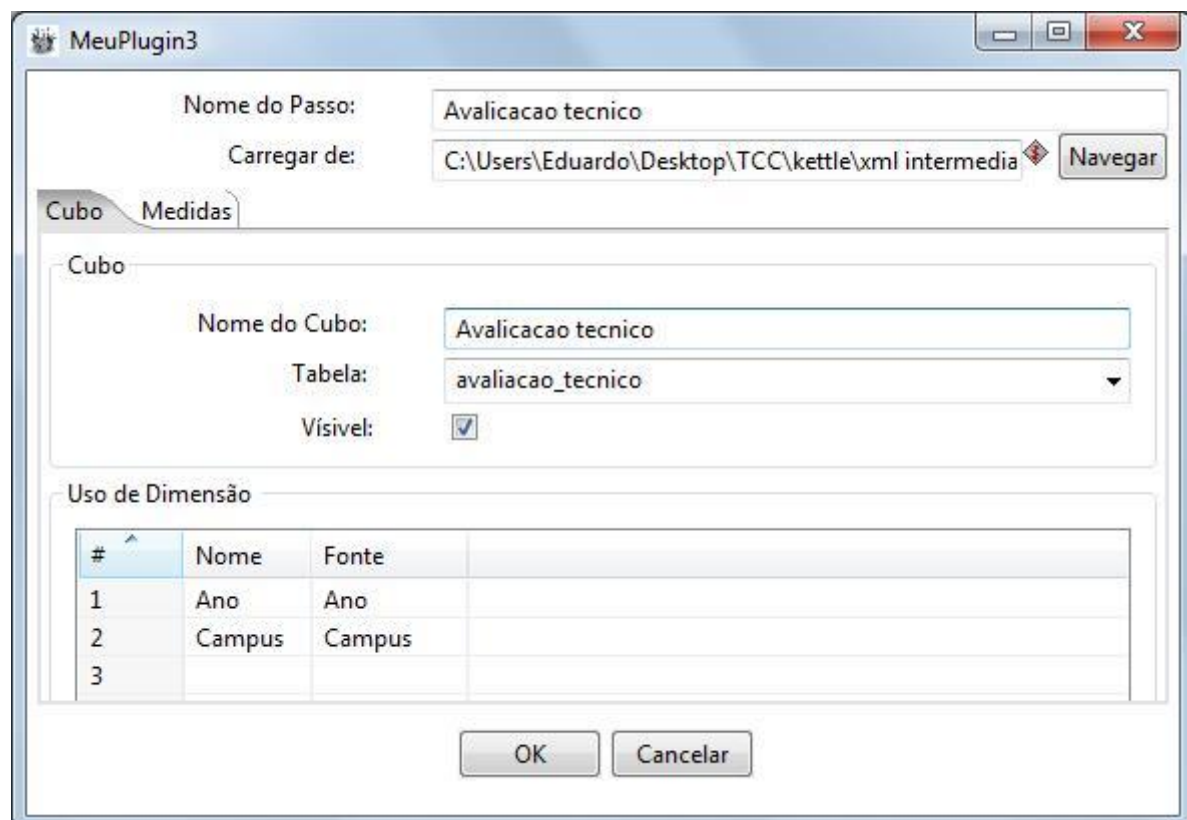


Figura 67: Cubo Avaliacao tecnico na aba de configuração Cubo.

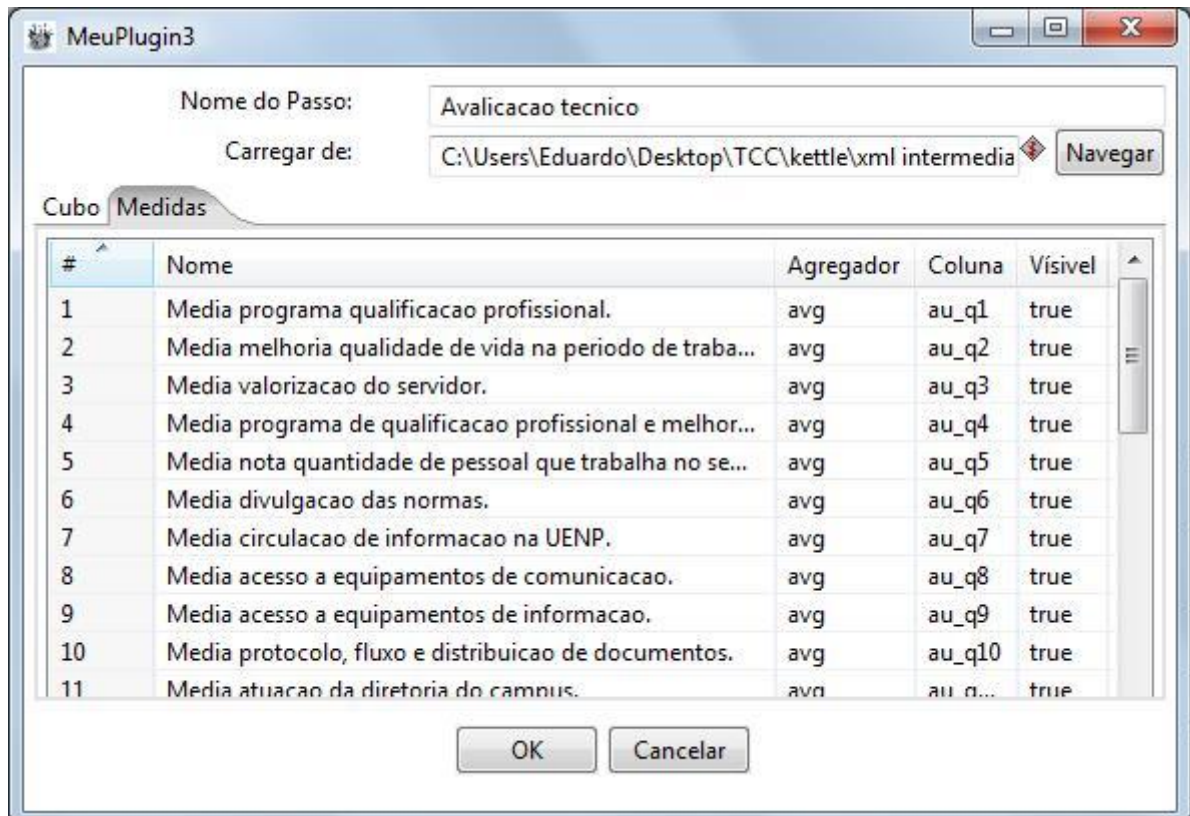


Figura 68: Cubo Avaliacao tecnico na aba de configuração Medidas.