

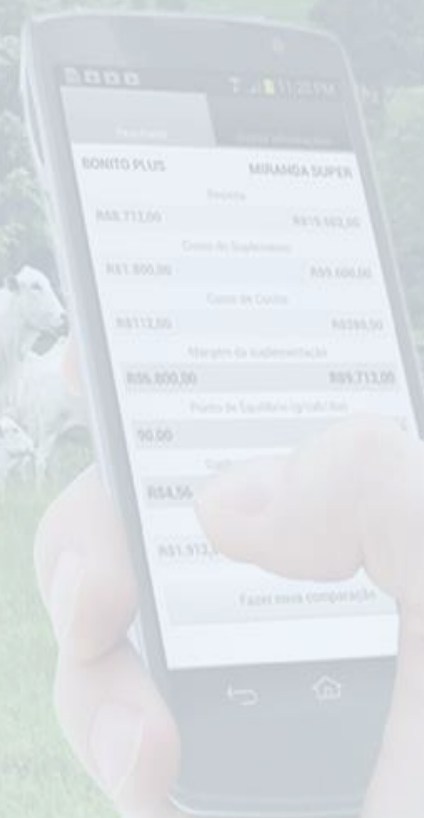
# ***Business Intelligence* Aplicado em Sistemas de Apoio à Tomada de Decisão de Produtores Rurais**

**Engenheiro MSc. Luciano Moraes Da Luz Brum**



# Roteiro

- **Introdução / Problemática**
- **Metodologia**
- **Execução e Implementação**
- **Resultados**
- **Trabalhos adicionais**





# Introdução



## Adoção de tecnologias em 2010:

- 80,8% das empresas brasileiras já utilizavam computadores;
- 76,9% já utilizavam a internet;



## Acesso à informação no campo:

- 1,46% dos estabelecimentos rurais possuíam acesso à internet em 2006;
- Em 2017, este percentual subiu para 26,10% (1,4 Milhões de propriedades);



## Volume de dados:

- Crescimento exponencial até 2020;
- Necessidade emergente de profissionais (*data scientists, data engineers, data architects, etc*);
- Políticas claras de privacidade, segurança da informação e propriedade intelectual;

# Introdução



## Tomada de decisão:

- Qual o melhor caminho?
- *Felling*, experiência, ...?
- Uso de dados, modelos, simuladores, sistemas?



FGC

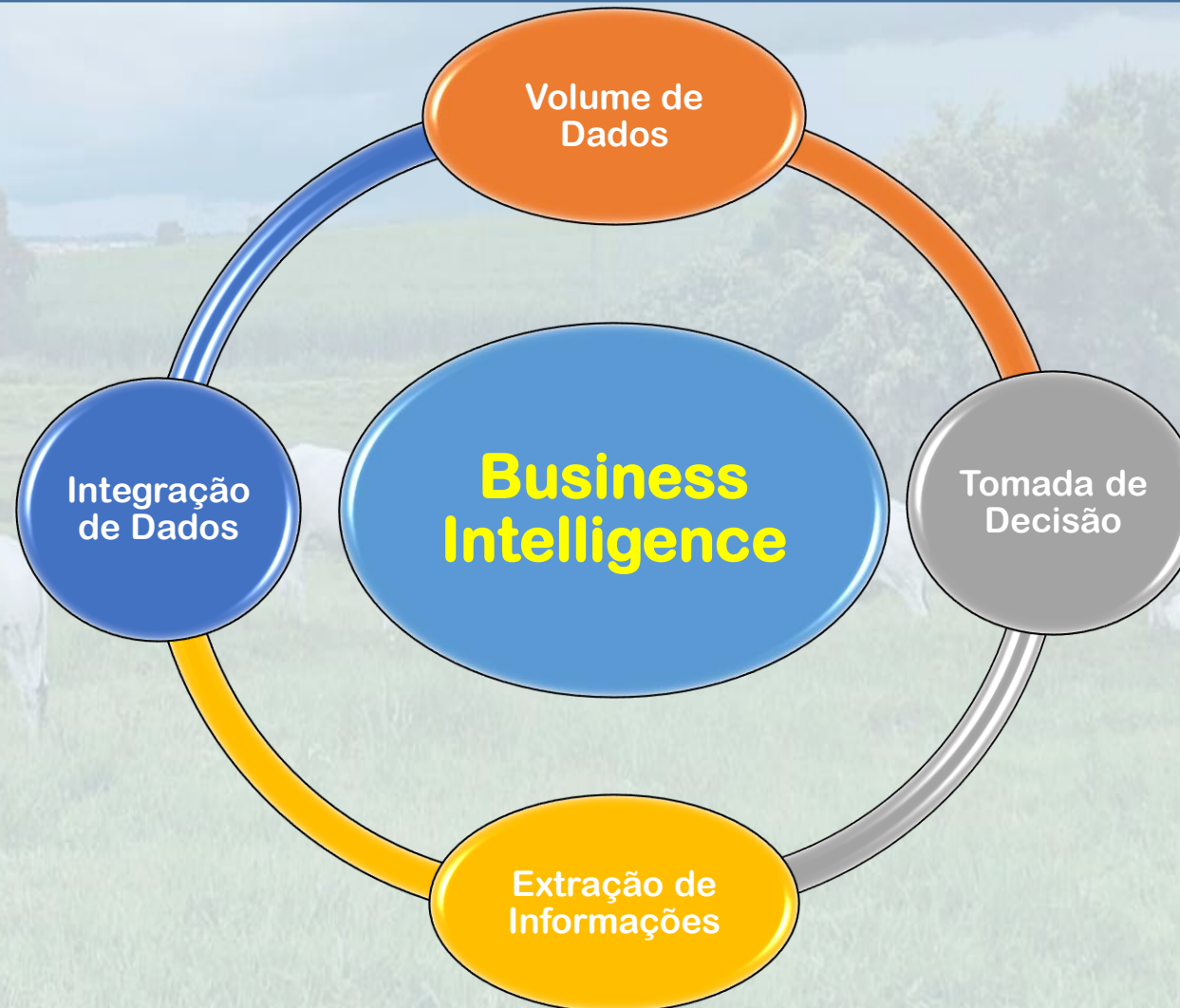
LS

## Sistemas de Apoio à Decisão (SAD):

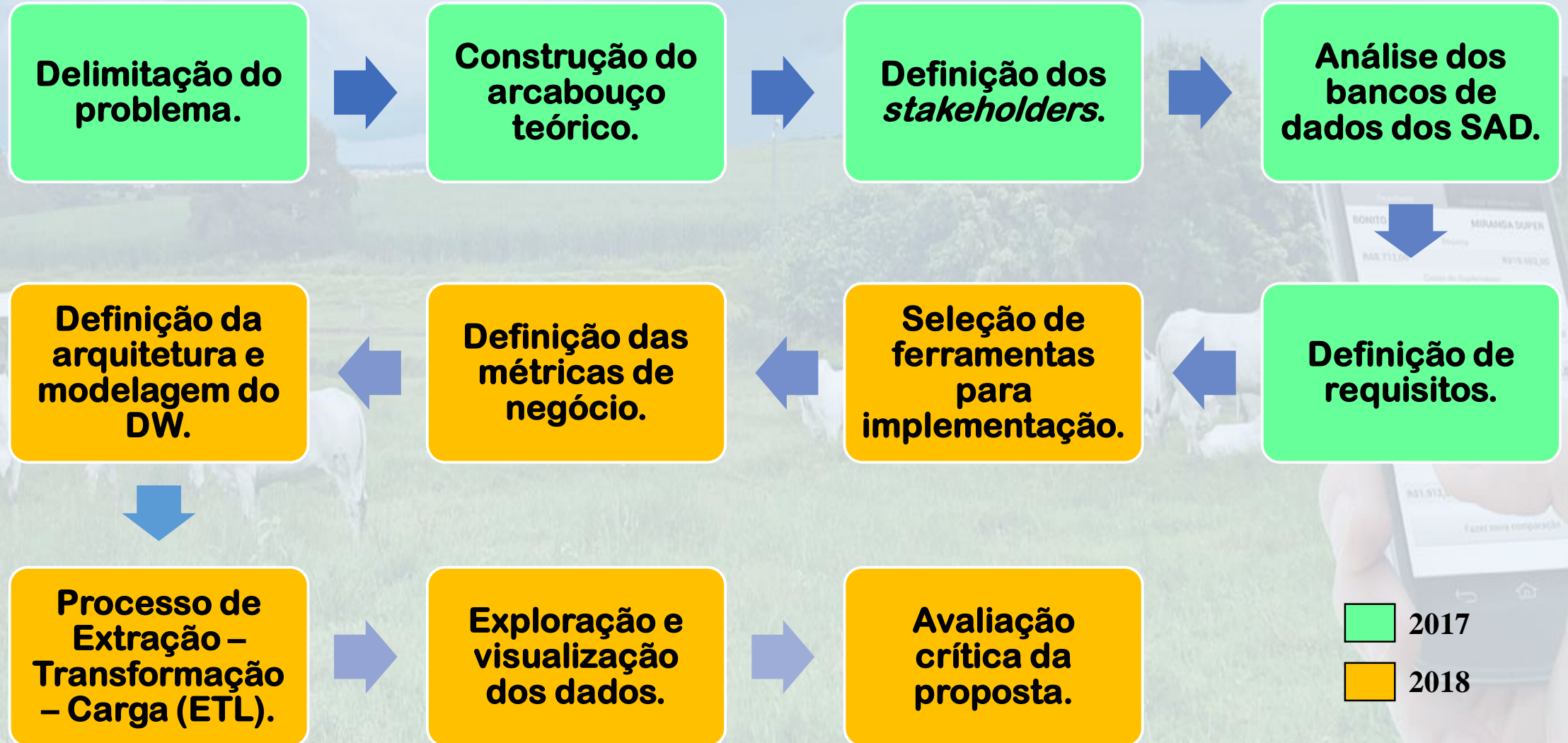
- Ferramenta de Gestão de Custos (FGC);
- *Livestock Sustainable* (LS);
- Necessidade de integração, organização e disponibilização das informações para *stakeholders*;



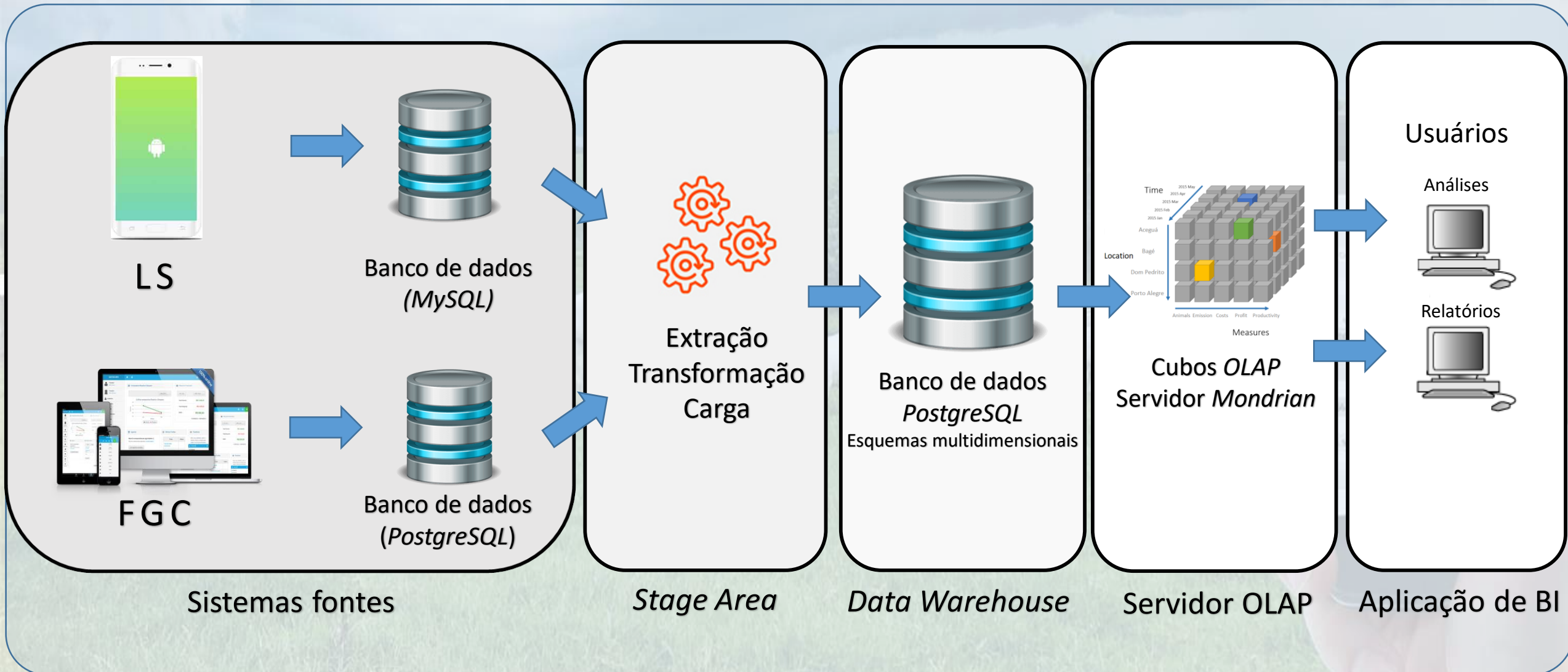
# Introdução



# Metodologia



# Metodologia





# Execução e Implementação





# Documento de Requisitos



## Documento de Requisitos do Sistema De *Business Intelligence* para Produtores Rurais v. 1.2

Bagé, 22 de novembro de 2018.

### Prefácio

Este documento foi elaborado para os clientes (especialistas da Empresa Brasileira de Pesquisa Agropecuária – Embrapa), desenvolvedores de software, profissionais de Tecnologia da Informação, Engenheiros e Cientistas de Computação e demais profissionais da tecnologia. Com este documento, espera-se dar subsídios para o desenvolvimento da proposta de integração de dados de diferentes Sistemas de Informação voltados à pecuária de corte, com o objetivo de subsidiar as decisões de pecuaristas brasileiros através das tecnologias de *Data Warehouse* e *Business Intelligence*.

### Histórico de Alterações

Data	Versão	Descrição	Autor
13/09/2018	1.0	Criação do documento de requisitos, descrição das fontes de dados ( <i>myBeef</i> , <i>Gestão de Custos</i> e <i>Livestock Sustainability</i> ), inclusão dos modelos conceituais ( <i>myBeef</i> e <i>Gestão de Custos</i> ), identificação dos requisitos, descrição dos requisitos funcionais e não-funcionais.	Luciano Moraes da Luz Brun
26/09/2018	1.1	Alteração do nome do sistema de Emissão para <i>Livestock Sustainability</i> . Mudança e ajustes nos modelos de banco de dados dos sistemas e acréscimo do modelo do <i>Livestock Sustainability</i> .	Luciano Moraes da Luz Brun
08/11/2018	1.2	Remoção do sistema <i>MyBeef</i> da integração. Inserção do <i>ETL</i> nos requisitos não-funcionais. Análise dos modelos conceituais dos sistemas fontes. Inserção de uma visão geral de arquitetura da solução.	Luciano Moraes da Luz Brun

### Sumário

1. INTRODUÇÃO	4
1.1 VISÃO GERAL DO DOCUMENTO	4
1.2 CONVENÇÕES, TERMOS E ABBREVIATURAS	4
1.2.1 Identificação dos requisitos	4
1.2.2 Prioridades dos requisitos	4
2. DESCRIÇÃO GERAL DO SISTEMA	5
2.1 PROBLEMA EXISTENTE	5
2.2 DESCRIÇÃO DOS SISTEMAS FONTES E DO SISTEMA DE BI	5
2.3 FUNÇÕES DO SISTEMA	7
2.4 VISÃO GERAL DA ARQUITETURA DA SOLUÇÃO PROPOSTA	8
3. REQUISITOS FUNCIONAIS	9
[RF001] Acesso ao sistema	9
[RF002] Acesso às informações do Data Warehouse	9
[RF003] Visualização das informações do Data Warehouse	9
[RF004] Geração de gráficos	10
[RF005] Geração de relatórios	10
4. REQUISITOS NÃO-FUNCIONAIS	11
[NF001] Usabilidade	11
[NF002] Desempenho	11
[NF003] Disponibilidade	11
[NF004] Segurança	12
[NF005] Hardware e Software	12
[NF006] Sistema ETL	12
5. GERENCIAMENTO DE REQUISITOS	12
6. REFERÊNCIAS	13

# Ferramentas para Implementação



## Suíte Pentaho CE v. 8.1.0.0\*

- **Pentaho Data Integration** (PDI) – módulo spoon;
- **Pentaho Schema Workbench** (PSW);
- **Pentaho Business Analytics Server** (PBAS);

\*CE = Community Edition.



CUTTING EDGE OPEN SOURCE ANALYTICS

## Saiku Analytics CE\*

- Análises OLAP;
- Visualização em tabelas e gráficos;
- Gráficos de linha, barras, pizza, *tree maps*, *sunburst*, entre outros;
- Exportação dos dados em .pdf, .csv e .xls;
- Exportação de gráficos em .pdf, .png e .jpeg;

\*CE = Community Edition.



## PostgreSQL v. 10.5

- Sistema Gerenciador de Banco de Dados (SGBD);
- Licença BSD\*;
- *Pgadmin* v. 3.6 (4.2);
- Já utilizado na FGC;

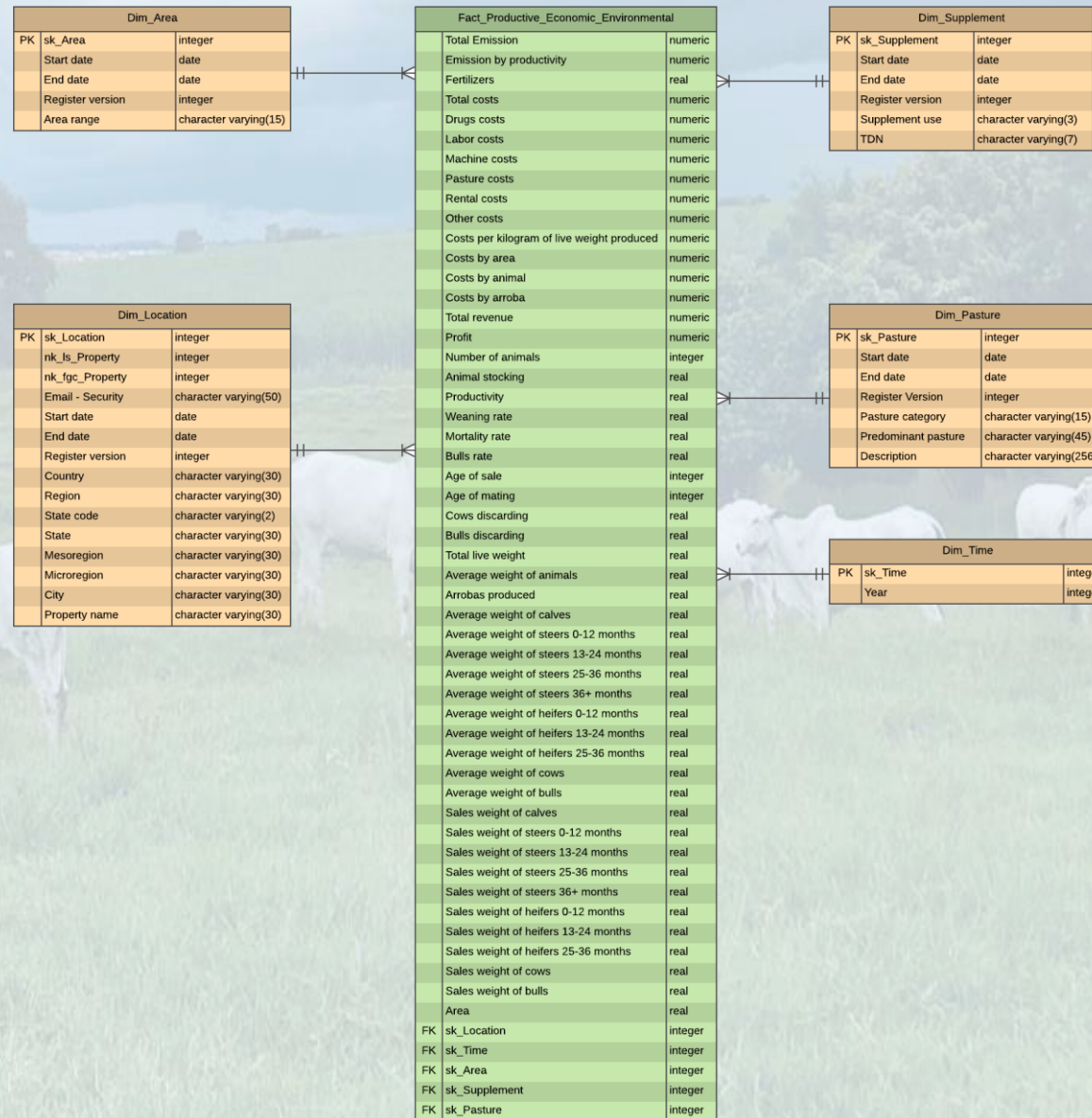
\*BSD = Berkeley Software Distribution.

**O que as ferramentas selecionadas têm em comum?**

**Open-source, gratuitas, documentação disponível e ampla comunidade de usuários;**



# Projeto do Data Warehouse



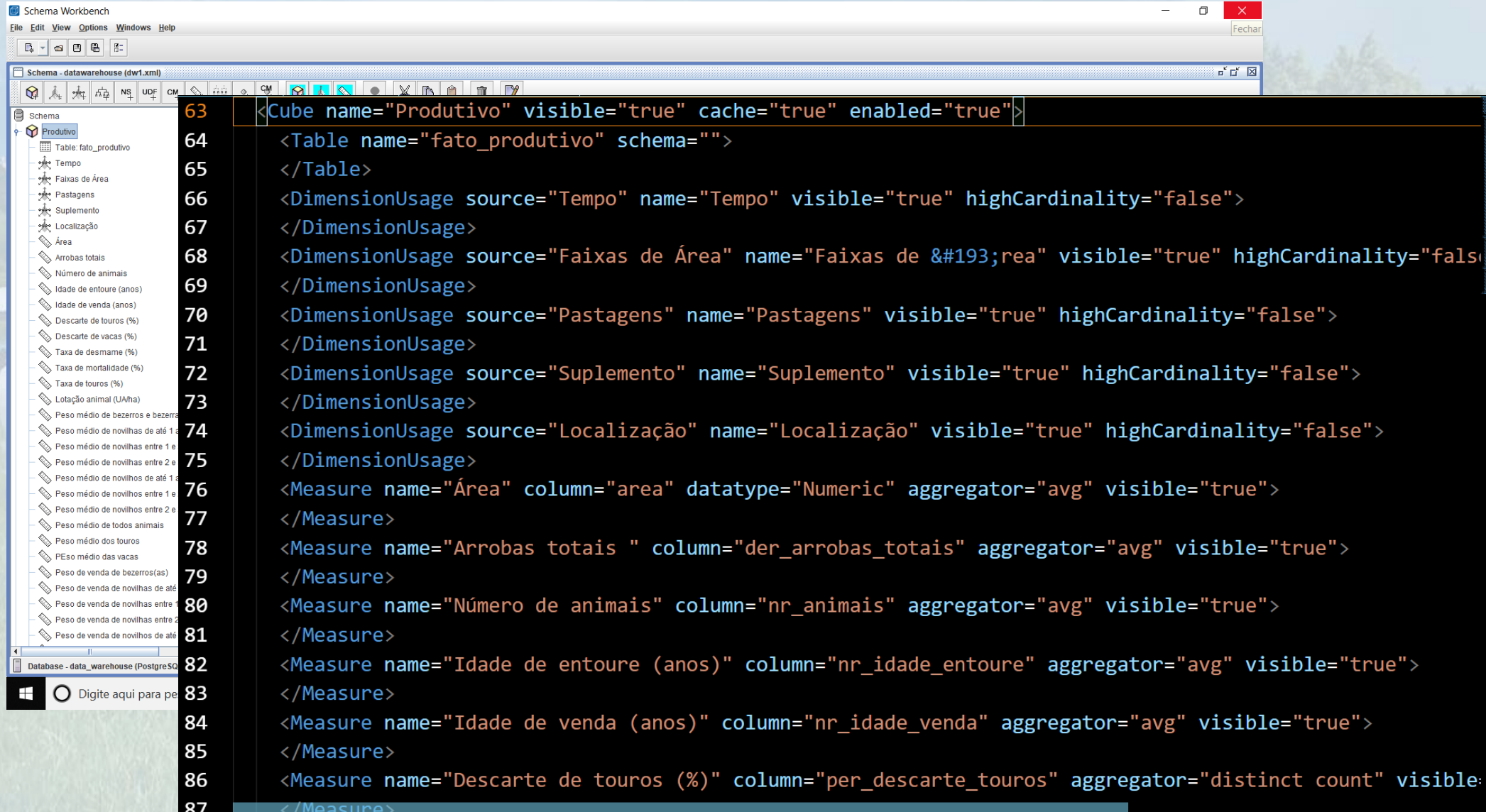


# Projeto do ETL

1. Execução das instâncias dos bancos de dados do FGC e LS;
2. Organização e importação dos dados das planilhas para as bases de dados;
3. Implementação das etapas do ETL separadas por diretórios;
4. Padrões de Nomenclatura (*Stage Area e transformations/Jobs*);
5. Criação de tabelas auxiliares;
6. Análise do tempo de execução e verificação dos resultados.



# Configuração de Cubos



The screenshot displays the Schema Workbench interface. On the left, a tree view shows the 'Schema' structure with a 'Produtivo' cube selected. The main editor area shows the XML configuration for this cube. The configuration includes a table named 'fato\_produtivo', several dimension usages for 'Tempo', 'Faixas de Área', 'Pastagens', 'Suplemento', and 'Localização', and several measures for 'Área', 'Arrobas totais', 'Número de animais', 'Idade de entoure (anos)', 'Idade de venda (anos)', and 'Descarte de touros (%)'. The measures are configured with various aggregators like 'avg' and 'distinct count'.

```
63 <Cube name="Produtivo" visible="true" cache="true" enabled="true">
64   <Table name="fato_produtivo" schema="">
65   </Table>
66   <DimensionUsage source="Tempo" name="Tempo" visible="true" highCardinality="false">
67   </DimensionUsage>
68   <DimensionUsage source="Faixas de Área" name="Faixas de Área" visible="true" highCardinality="false">
69   </DimensionUsage>
70   <DimensionUsage source="Pastagens" name="Pastagens" visible="true" highCardinality="false">
71   </DimensionUsage>
72   <DimensionUsage source="Suplemento" name="Suplemento" visible="true" highCardinality="false">
73   </DimensionUsage>
74   <DimensionUsage source="Localização" name="Localização" visible="true" highCardinality="false">
75   </DimensionUsage>
76   <Measure name="Área" column="area" datatype="Numeric" aggregator="avg" visible="true">
77   </Measure>
78   <Measure name="Arrobas totais" column="der_arrobas_totais" aggregator="avg" visible="true">
79   </Measure>
80   <Measure name="Número de animais" column="nr_animais" aggregator="avg" visible="true">
81   </Measure>
82   <Measure name="Idade de entoure (anos)" column="nr_idade_entoure" aggregator="avg" visible="true">
83   </Measure>
84   <Measure name="Idade de venda (anos)" column="nr_idade_venda" aggregator="avg" visible="true">
85   </Measure>
86   <Measure name="Descarte de touros (%)" column="per_descarte_touros" aggregator="distinct count" visible="true">
87   </Measure>
```

[illegible]



# Resultados

❑ Como a solução proposta pode auxiliar *stakeholders* em processos decisórios?

- ★ Propriedades maiores possuem um custo por área maior que propriedades menores?
- ★ Quais são as propriedades top 10 em produtividade e respectivos indicadores?
- ★ As propriedades top em produtividade possuem custos por área menores? E a taxa de desmame destas é maior?
- ★ É possível produzir mais emitindo menos? (taxa de desmame, produtividade e emissão por produtividade).

# Discussões

- ☛ Solução proposta atendeu os objetivos de **integrar, organizar e disponibilizar** informações de auxílio em processos decisórios na **pecuária de corte**;
- ☛ Três dimensões contempladas: **produtiva, econômica e ambiental**;
- ☛ Além da solução verificada na aplicação, foi implementada uma **arquitetura de BI completa** (sistemas fontes, ETL, DW e aplicação de BI, com respectiva **documentação**);
- ☛ **Fatores positivos:** reuniões constantes com *stakeholder* interno e uso de tecnologias **gratuitas**.
- ☛ **Dificuldades:** complexidade do funcionamento do LS, domínio do problema, ausência completa de documentação, sistemas fontes em desenvolvimento, problemas nas bases de dados dos fontes, diferenças de granularidade dos dados, gráficos com variáveis de diferentes magnitudes, alterações no *Saiku Analytics*.

# Publicações

## ➤ Publicações:

- BRUM, L. M. L.; DE PINHO, L. B.; CAMARGO, S. S. **Avaliação do Uso de Realidade Aumentada no Ensino de Arquitetura e Organização de Computadores.** *International Journal of Computer Architecture Education*, v. 6, p. 10-17, 2017.
- BRUM, L. M. L.; HEINEN, M. R.; CAMARGO, S. S. **Analysis of Student Performance Data from a Computer Architecture and Organization Course.** *Revista Novas Tecnologias na Educação*, v. 16, p. 1, 2018.
- BRUM, L. M. L.; LAMPERT, V. N.; CAMARGO, S. S. **Business Intelligence and Data Warehouse in Agrarian Sector: A Bibliometric Study.** *Journal of Agricultural Science*, v. 11, p. 353-368, 2019.
- BRUM, L. M. L.; LAMPERT, V. N.; CAMARGO, S. S.; EICKHOFF, F. A. **Aplicação de Técnicas de Mineração em Dados de Propriedades Leiteiras do Município de Derrubadas.** In: V Simpósio da Ciência do Agronegócio, 2017, Porto Alegre. *Anais do V Simpósio da Ciência do Agronegócio*, 2017. v. 1. p. 8-15.
- BRUM, L. M. L.; LAMPERT, V. N.; SILVA, G. M. **Análise de Aspectos Produtivos e Sociais da Atividade Leiteira Gaúcha e Brasileira.** In: X Salão Internacional de Ensino, Pesquisa e Extensão, 2018, Santana do Livramento. *Anais do X Salão Internacional de Ensino, Pesquisa e Extensão*, 2018.

## ➤ Produzidos, porém ainda não finalizados:

- **Análise da evasão de produtores da atividade leiteira do município de Derrubadas - RS** (Brum, L. M. L.; Ruelas, J. L. J.; Lampert, V. do N.; Eickhoff, F. A.; Camargo; S. da S.).
- **Data Integration of Livestock Information Systems through Data Warehousing** (Alonso, C. de M.; Brum, L. M. L.; Lampert, V. do N.; Camargo; S. da S.; Schafer, A. G.).



# Trabalhos Adicionais

- Desenvolvimento de um Aplicativo para coleta de dados de produtores de leite para a EMATER (Empresa de Assistência Técnica e Extensão Rural) e EMBRAPA (Empresa Brasileira de Pesquisa Agropecuária):
  - Tecnologias: Framework Ionic 3 (HTML, CSS e JavaScript).
- Desenvolvimento de um sistema online para apoio à decisão de produtores rurais da pecuária de corte:
  - Tecnologias: HTML, CSS, JavaScript, MySQL, Framework Yii 1.1.16.
- Análise de indicadores da produção leiteira do município de Derrubadas – RS:
  - Tecnologias: R e Árvores de Decisão (*rpart*).
- Desenvolvimento de trabalhos para estudantes de disciplinas de Arquiteturas de computadores:
  - Produção de vídeos didáticos para o *Moodle* e *Youtube*, análise do *feedback* dos vídeos, produção de TCC sobre o ensino de Arquitetura de Computadores com Realidade Aumentada e avaliação estatística e mineração dos dados de frequência, desempenho e atividades semipresenciais de três turmas de estudantes desta disciplina.

# Trabalhos Adicionais

- Aulas de Algoritmos, Laboratórios de Programação e Arquiteturas de Computadores ministradas na Universidade Federal do Pampa – campus Bagé, entre agosto de 2016 e março de 2018.
  - Um ano e meio de experiência com a linguagem C em ensino e dois anos de experiência adquiridos durante a graduação.
- Criação de contas no *github* e *youtube*, e criação de um site pessoal no estilo de blog (<http://lubrum.github.io>) para compartilhar tutoriais, códigos de projetos, artigos sobre tecnologias e Ciência de Dados, de maneira geral.

# **Aplicação de Técnicas de *Business Intelligence* em Sistemas de Apoio à Tomada de Decisão de Produtores Rurais**

**Engenheiro MSc. Luciano Moraes Da Luz Brum**

**Link para a apresentação: <>.**