



**FUNDAÇÃO
ARAUCÁRIA**

GESIS
Grupo de Estudos de
Sistemas de Informação

**INSTITUTO
FEDERAL**
Paraná
Campus
Palmas

DESENVOLVIMENTO DE BUSINESS INTELLIGENCE PARA O PERFIL DA PRODUÇÃO CIENTÍFICA DO IFPR CAMPUS PALMAS

Acadêmica: Talita Fatima Paula Monteiro

Orientadora: Prof^a. Me. Lauriana Paludo

Palmas/PR – 29, nov. 2017



Desenvolvimento de Business Intelligence para o Perfil da Produção Científica do IFPR campus Palmas

1. INTRODUÇÃO
2. PROBLEMA
3. TRABALHOS SIMILARES
4. OBJETIVO
5. FUNDAMENTAÇÃO TEÓRICA
6. METODOLOGIA
7. MATERIAIS E MÉTODOS
8. PROJETO
9. RESULTADOS
10. CONSIDERAÇÕES FINAIS



INTRODUÇÃO

- ❖ Os **Sistemas de Apoio a Decisão** viram fornecer apoio em decisões através de informações obtidas de diversas fontes de dados. A implementação mais utilizada da arquitetura é a de BI (TAVARES, 2008; CECI, 2014).
- ❖ **Business Intelligence** para Turban (2009) e Machado (2013) é um termo “guarda-chuva” que inclui arquiteturas, ferramentas, bancos de dados, aplicações e metodologias cujos principais objetivos são permitir o acesso e manipulação dos dados fornecendo aos gestores a capacidade de realizar análises mais adequadas.
- ❖ As técnicas aplicadas de BI neste trabalho foram as de *data warehouse*, processamento analítico *on-line*, extração, transformação e carga de dados.



INTRODUÇÃO

- ❖ Dentre os principais objetivos dos Institutos Federais está a **realização de pesquisas**, aplicadas a produção cultural, empreendedorismo, cooperativismo, desenvolvimento científico e tecnológico e desenvolvimento de técnicas e tecnologias (Lei nº 11.892/2008).
- ❖ Regime de trabalho dos docentes efetivos (Lei 5.596/87, Decreto Federal 94.664, Portaria 475 e Lei 11.784/08)
- ❖ Órgãos públicos como o CNPq, CAPES, FINEP, Ministério da Ciência e Tecnologia, Fundações de Amparo à Pesquisa e as próprias instituições através de suas pró-reitorias de pesquisa e extensão fazem o uso de indicadores e demais informações consistentes e confiáveis para auxiliar em seus processos decisórios.



PROBLEMA

- ❖ Comitê de Pesquisa e Extensão do IFPR campus Palmas (Portaria 3, de 9/6/2009), responsável por formular, subsidiar e acompanhar a política de pesquisa científica e extensão no âmbito da instituição;
- ❖ Ausência de um sistema de informação no IFPR campus Palmas que gerencie a produção ligada aos projetos dos docentes;
- ❖ A gestão da pesquisa e extensão do campus tem frequentemente recorrido a consultas por e-mail, a coordenadores dos cursos e aplicação de auditorias internas, a fim de levantar informações sobre a produção científica do campus e utilizar os dados em seus processos decisórios.



TRABALHOS SIMILARES

- ❖ PPGS da Universidade Federal do Pernambuco (SANTANA *et al*, 2011);
- ❖ Ensino a Distância (EaD) Amadeu (FILHO *et al*, 2015);
- ❖ UFSM de Francisco Westphalen (MUNARETTO *et al*, 2017);
- ❖ *ScriptLattes* (MENA-CHALCO e CESAR JUNIOR, 2013)
- ❖ *LattesDataXplorer* (DIAS, 2016)



OBJETIVO GERAL

Desenvolver um ambiente de apoio à decisão através da aplicação de técnicas de *Business Intelligence* para identificar o perfil da produção científica e atuação dos pesquisadores do IFPR campus Palmas

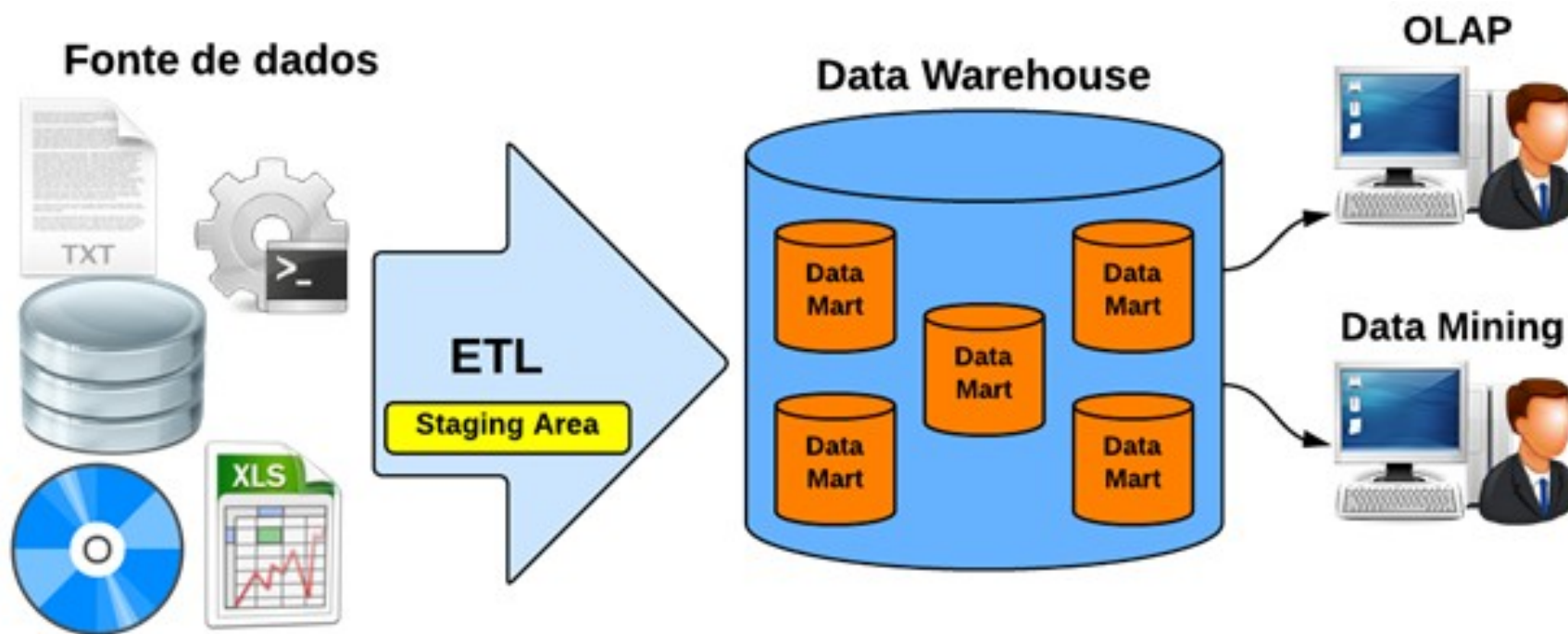


FUNDAMENTAÇÃO TEÓRICA

- ❖ **Sistemas de Apoio a Decisão** (TAVARES, 2008; SILBERSCHATZ, KORTH e SUDARSHAN, 2006; CECI, 2012 *apud* TURBAN, 1990; HEINZLE, 2010).
- ❖ **Business Intelligence** (SANZÕES, OLIVEIRA e BAPTISTA, 2006; LOH, 2014; TURBAN, 2009; MACHADO, 2013).
- ❖ **Data Warehouse** (INMON, 1997; MACHADO, 2013; KIMBALL, 1998)
- ❖ **OLAP** - Processamento Analítico *On-line* (COLAÇO JUNIOR, 2004; MACHADO, 2013, SANZÕES, OLIVEIRA e BAPTISTA, 2006)



ARQUITETURA DE BI



Fonte: ELIAS, 2014



METODOLOGIA DO TRABALHO

Estudo de caso
dos Institutos
Federais e do
COPE - IFPR
campus Palmas

Revisão
Bibliográfica sobre
BI, DW e
OLAP

Identificação das
Fontes de Dados

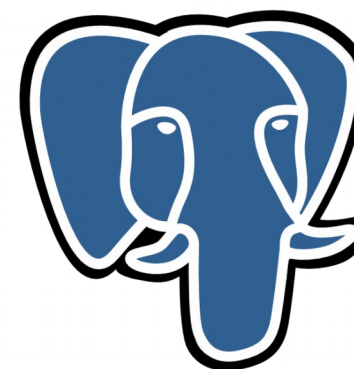
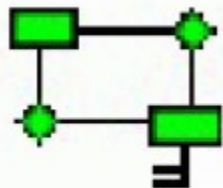
Desenvolvimento
do DW

OLAP



MATERIAIS & MÉTODOS

Stela **EXPERTA**



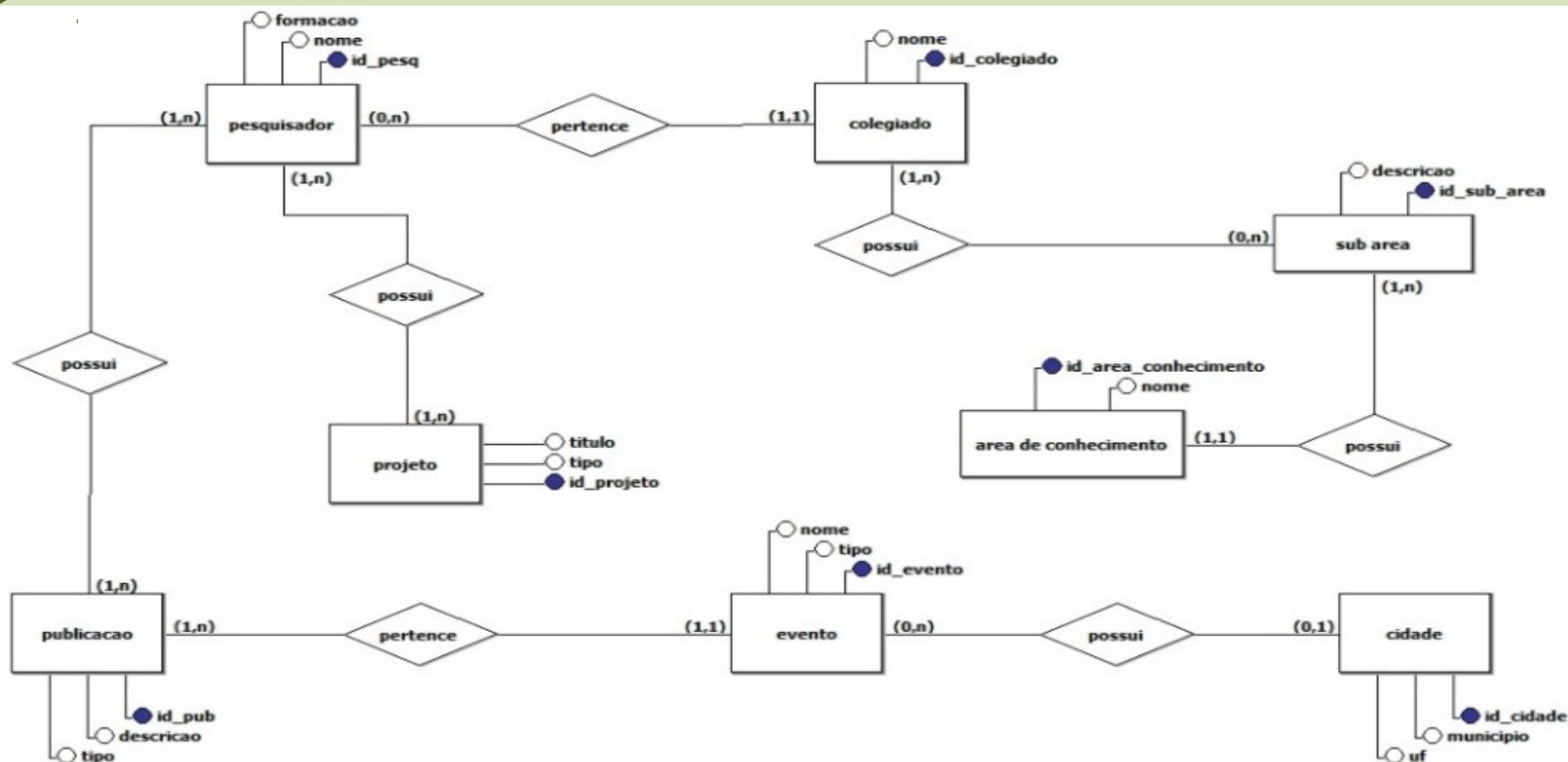
Postgre**SQL**



CUTTING EDGE OPEN SOURCE ANALYTICS

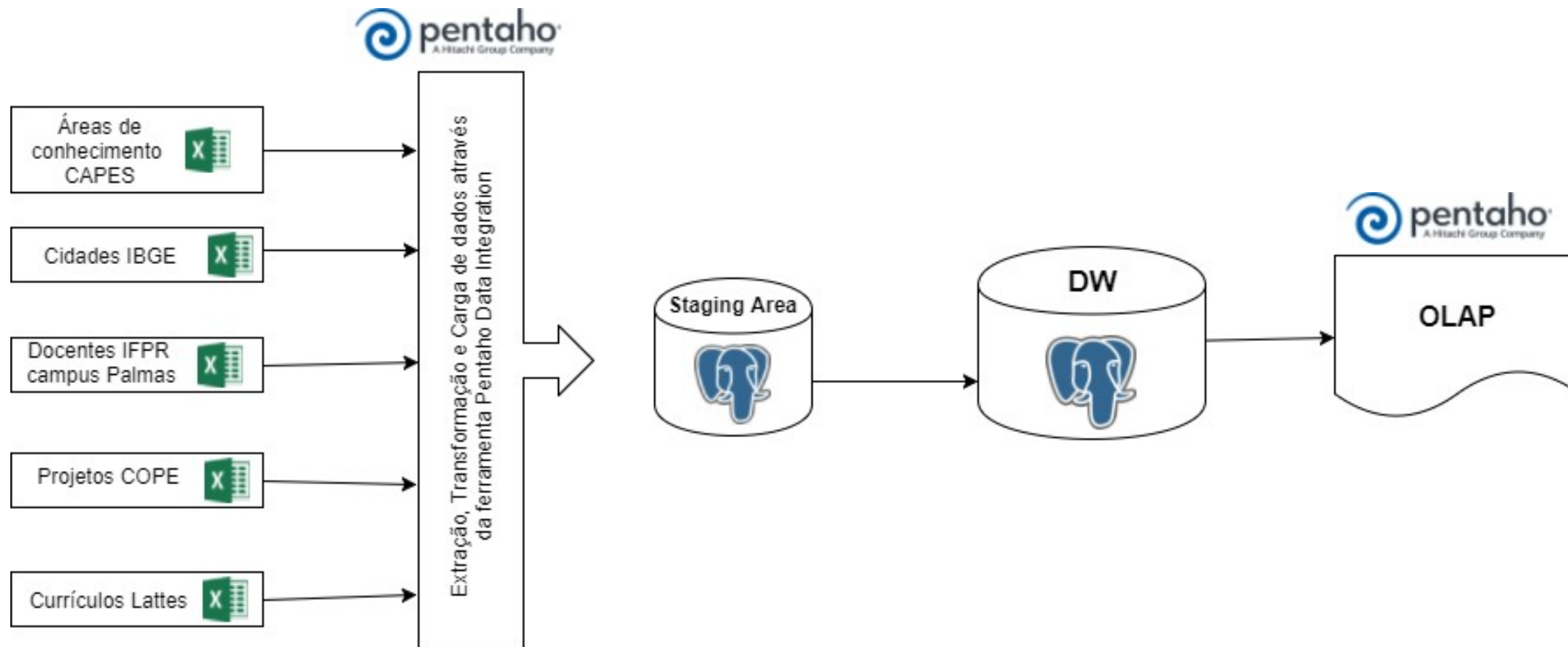


MODELAGEM ENTIDADE-RELACIONAMENTO (ER)



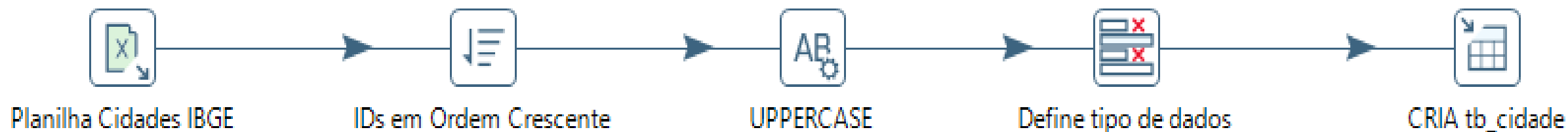


ARQUITETURA DO BI DE PRODUÇÃO CIENTÍFICA





ETL DA TABELA DE CIDADES





SQL DA TABELA DE COLEGIADOS

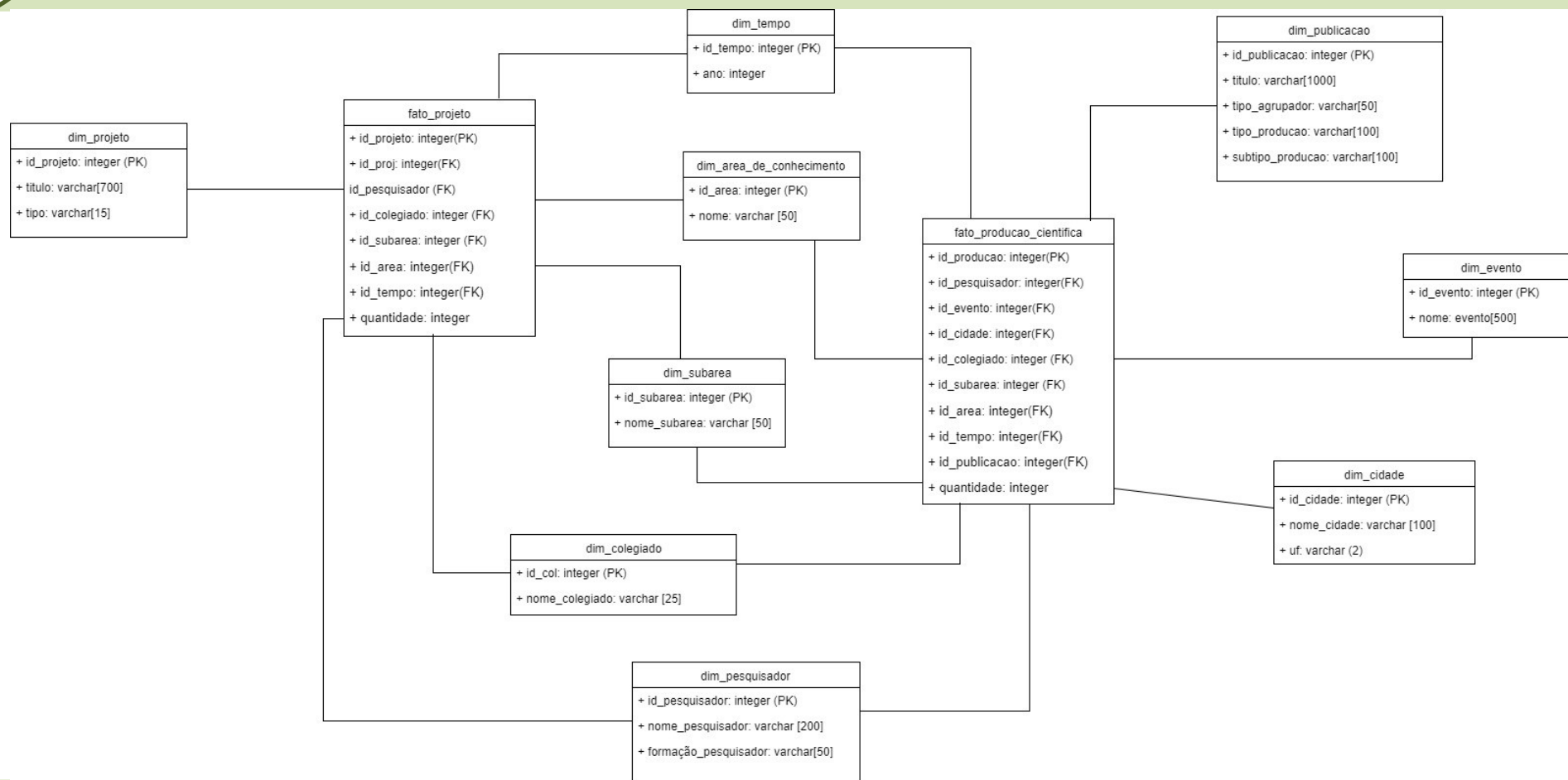
bi_cope on postgres@PostgreSQL 9.6

```
1 CREATE TABLE public.tb_colegiado
2 (
3     id_col bigint NOT NULL ,
4     descricao character varying(40) NOT NULL,
5     sub_area bigint NOT NULL,
6     CONSTRAINT pk_tb_colegiado PRIMARY KEY (id_col),
7     CONSTRAINT tb_colegiado_sub_area_fkey FOREIGN KEY (sub_area)
8 )
9
10 insert into tb_colegiado (descricao, sub_area) values ('ADMINISTRAÇÃO',51), ('ARTES VISUAIS', 74),
11 ('CIÊNCIAS BIOLÓGICAS', 8), ('CIÊNCIAS CONTÁBEIS', 51), ('DIREITO', 50), ('EDUCAÇÃO FÍSICA', 42),
12 ('ENFERMAGEM', 37), ('ENGENHARIA AGRÔNOMICA', 43), ('ENGENHARIA CIVIL', 21), ('FARMÁCIA', 36),
13 ('LETRAS PORTUGUÊS/INGLÊS', 73), ('PEDAGOGIA', 69), ('QUÍMICA',6), ('SISTEMAS DE INFORMAÇÃO', 3),
14 ('MULTIDISCIPLINAR', 75)
```




MODELAGEM DIMENSIONAL DW

MODELO ESTRELA (KIMBAL, 1997)





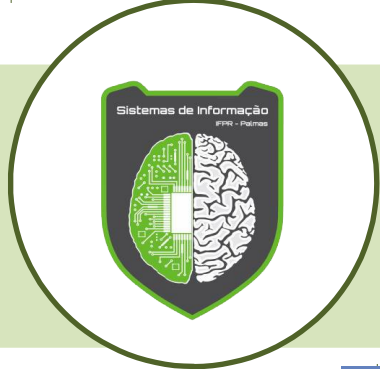
CRIAÇÃO DAS DIMENSÕES DO DW

```
1  create table dim_cidade(  
2      id_cidade INTEGER PRIMARY KEY,  
3      nome_cidade VARCHAR(100),  
4      uf VARCHAR(2)  
5  )  
6  
7  INSERT INTO dim_cidade(id_cidade, nome_cidade, uf)  
8      SELECT id_cidade, municipio, uf FROM  
9      dblink ('user=postgres password=root dbname=bi_cope' :: text,  
10         'SELECT id_cidade, municipio, uf FROM tb_cidade':: text, false)  
11      tb_cidade(id_cidade integer, municipio character varying, uf character varying)  
12
```



CRIAÇÃO DA TABELA FATO DO DW

```
108 create table fato_producao_cientifica(  
109     id_fato bigserial primary key,  
110     id_area_conhecimento bigint references dim_area_de_conhecimento (id_area),  
111     id_cidade bigint references dim_cidade (id_cidade),  
112     id_colegiado bigint references dim_colegiado (id_colegiado),  
113     id_evento bigint references dim_evento (id_evento),  
114     id_pesquisador bigint references dim_pesquisador (id_pesquisador),  
115     id_producao bigint references dim_producao (id_producao),  
116     id_projeto bigint references dim_projeto (id_projeto),  
117     id_tempo bigint references dim_tempo (id_tempo),  
118     quantidade int,  
119 )  
120
```



CRIAÇÃO CUBOS OLAP – SCHEMA WORKBENCH

Schema - Producao Cientifica (producao-cientifica-olap_v3.mondrian.xml)*

Schema

- Producao Cientifica
 - Table: fato_producao
 - Tempo
 - Tipo Publicacao
 - Pesquisador
 - Colegiado
 - Cidade
 - Area de Conhecimento
 - Sub Area de Conhecimento
 - Evento
 - Quantidade Publicacao
 - CM Percentual Publicacao
- Projetos

Measure for 'Producao Cientifica' Cube

Attribute	Value
name	Quantidade Publicacao
description	
aggregator	distinct-count
column	id_publicacao
formatString	
datatype	Integer
formatter	
caption	Quantidade Publicação
visible	<input checked="" type="checkbox"/>

Database - dw_cope (PostgreSQL)



PUC PAINEL DE CONTROLE DO USUÁRIO

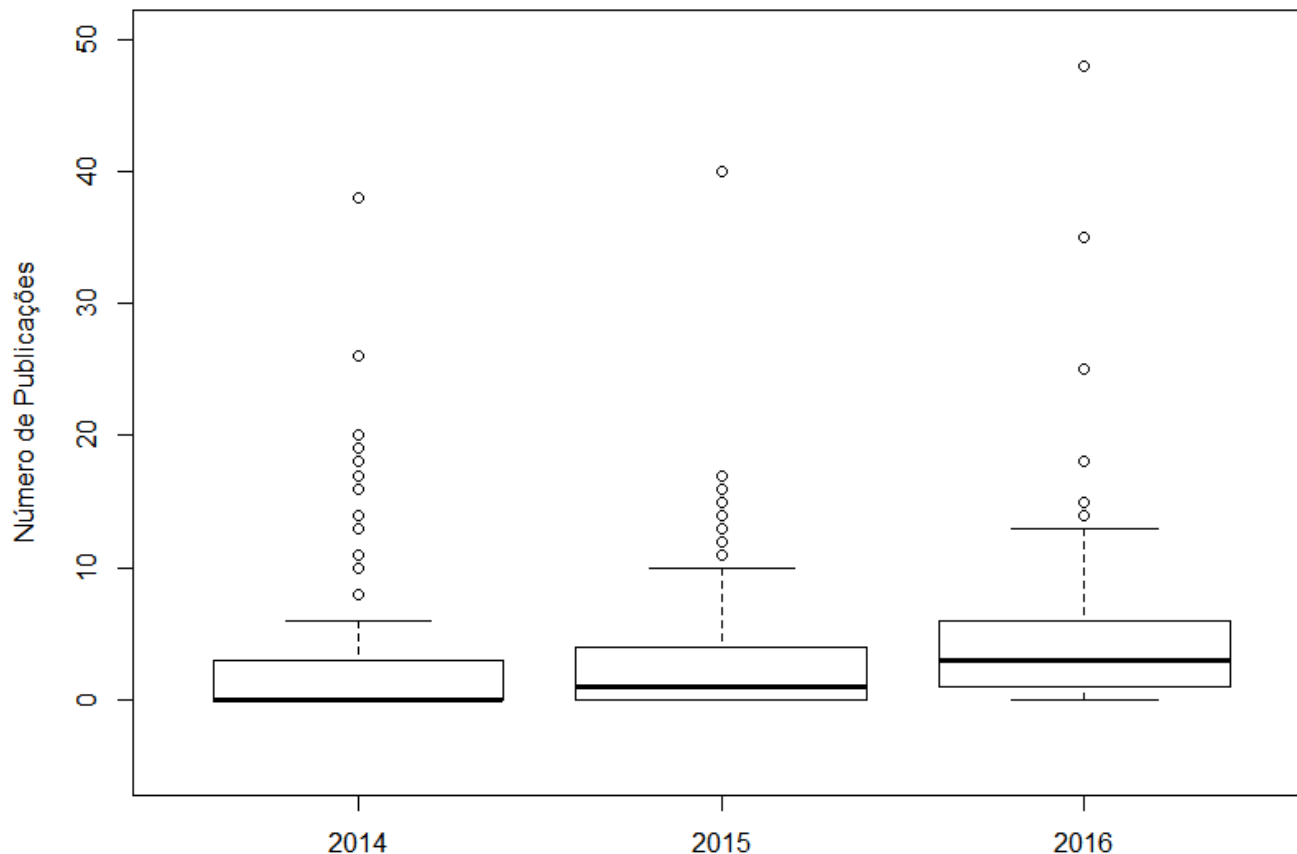
DEMONSTRAÇÃO





RESULTADOS

PC DOS ÚLTIMOS 3 ANOS, POR PESQUISADOR



MÉDIA DE PC POR ANO		
2014	2015	2016
3,23	3,46	5,34

- ❖ 15,44% dos docentes **não têm produção** nos últimos 3 anos;
- ❖ 35,29% têm de **1 a 5** produções;
- ❖ 17,64% têm de **6 a 10** produções;
- ❖ 5,88% têm de **11 a 15** produções;
- ❖ 7,35% têm de **16 a 20** produções;
- ❖ 18,38% têm **mais de 20** produções;
- ❖ Das 1251 PC, 385 possui mais de um autor

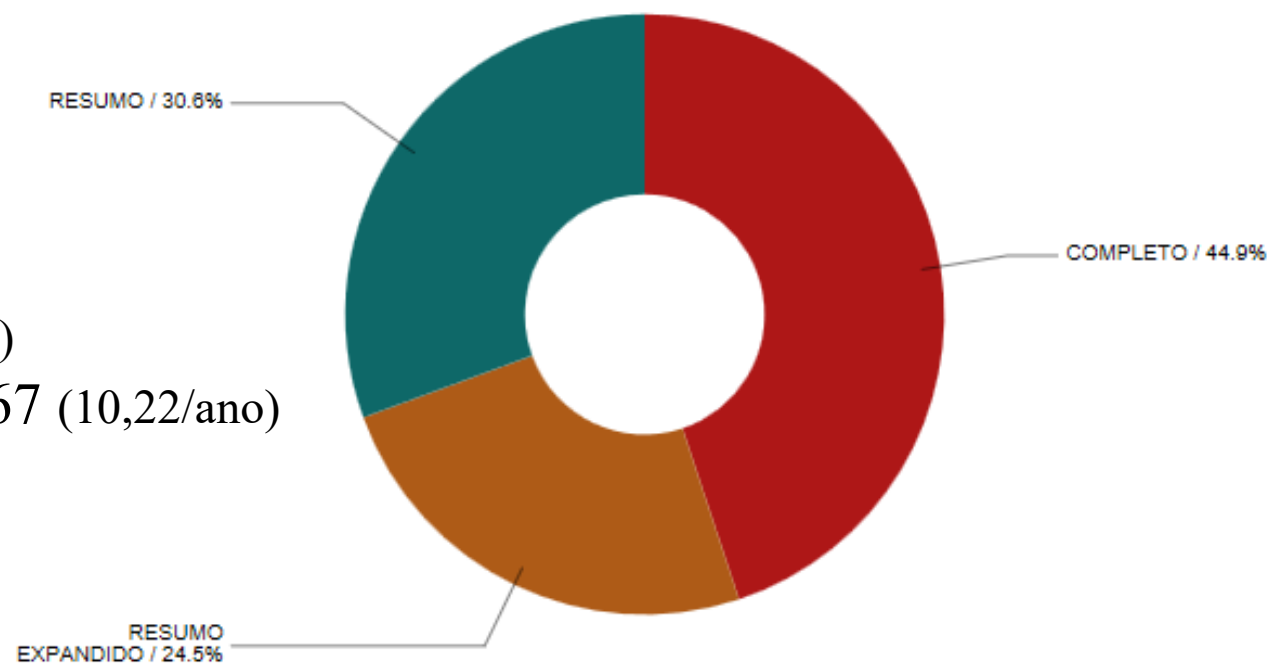


RESULTADOS

PC PUBLICADA EM ANAIS DE EVENTOS, POR COLEGIADO, NOS ÚLTIMOS 3 ANOS

■ COMPLETO ■ RESUMO EXPANDIDO ■ RESUMO

Todos os anos ~ Quantidade Publicacao



- ❖ 648 produções;
- ❖ Média/Colegiado 43,2
- ❖ Licenciatura em Química 38,33 (12,77/ano)
- ❖ Licenciatura em Ciências Biológicas 30,67 (10,22/ano)

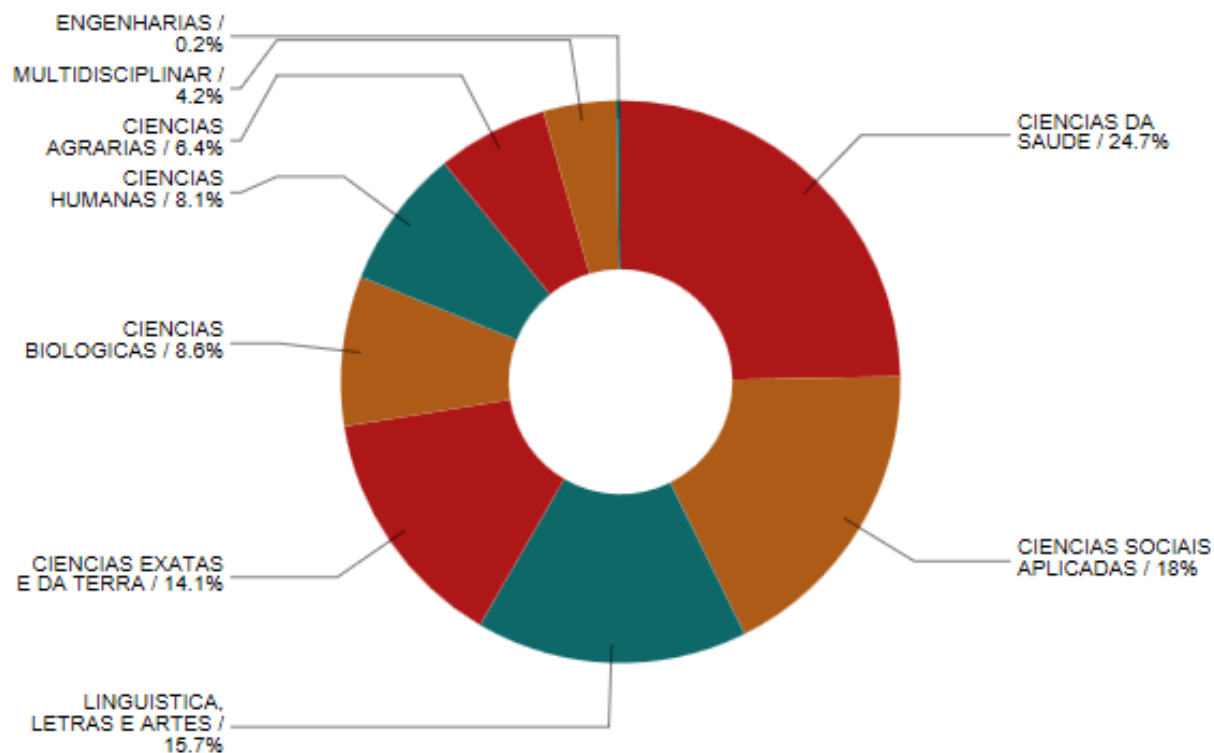


RESULTADOS

PC DOS ÚLTIMOS 3 ANOS, POR ÁREA DE CONHECIMENTO

■ CIÊNCIAS DA SAÚDE ■ CIÊNCIAS SOCIAIS APLICADAS ■ LINGÜÍSTICA, LETRAS E ARTES ■ CIÊNCIAS EXATAS E DA TERRA
■ CIÊNCIAS BIOLÓGICAS ■ CIÊNCIAS HUMANAS ■ CIÊNCIAS AGRÁRIAS ■ MULTIDISCIPLINAR ■ ENGENHARIAS

Quantidade Publicação ~ Todos os anos



Ciências da Saúde (321 - 107/ano)

Farmácia – 166 (55,33/ano)

Enfermagem – 83 (27,66/ano)

Educação Física – 72 (24/ano)

Ciências Sociais Aplicadas (234 – 78/ano)

Ciências Contábeis – 132 (44/ano)

Direito – 90 (30/ano)

Administração – 12 (4/ano)

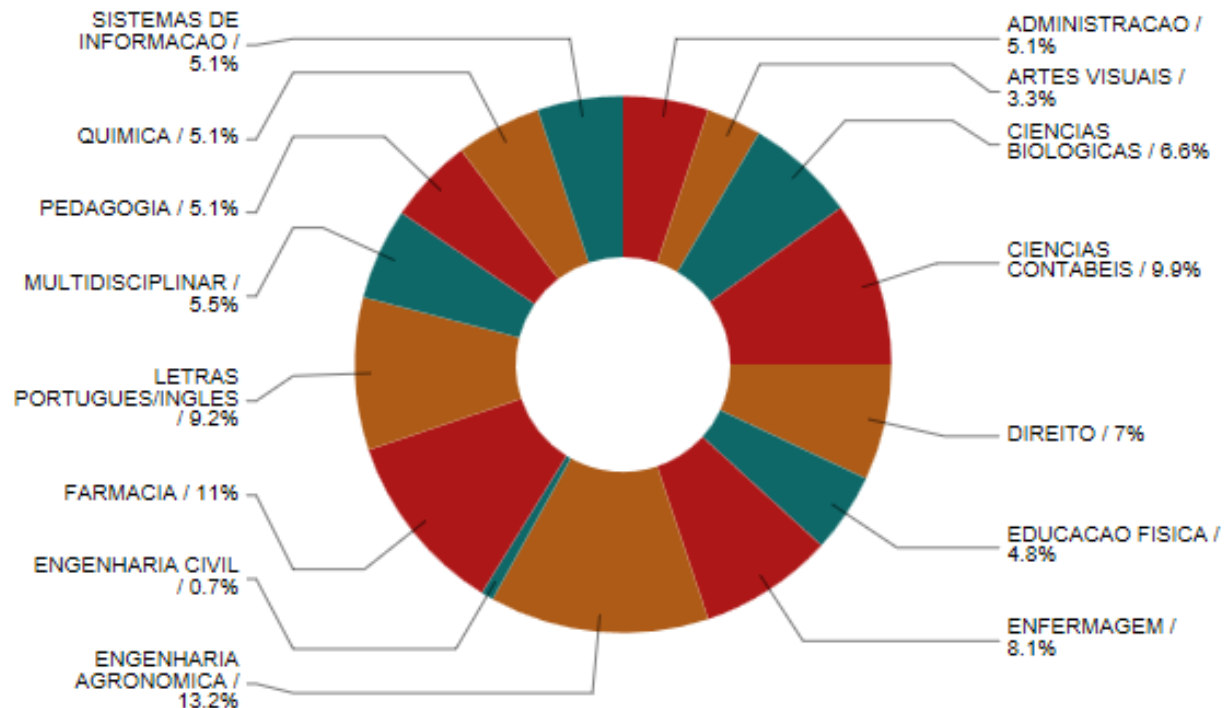


RESULTADOS

PROJETOS SUBMETIDOS AO COPE NOS ÚLTIMOS 3 ANOS, POR COLEGIADO

ADMINISTRAÇÃO ARTES VISUAIS CIÊNCIAS BIOLÓGICAS CIÊNCIAS CONTÁBEIS DIREITO EDUCAÇÃO FÍSICA
ENFERMAGEM ENGENHARIA AGRÔNOMICA ENGENHARIA CIVIL FARMÁCIA LETRAS PORTUGUÊS/INGLÊS
MULTIDISCIPLINAR PEDAGOGIA QUÍMICA SISTEMAS DE INFORMAÇÃO

Todos os anos



Engenharia Agrônômica – 36 projetos
6 extensão | 30 pesquisa
Média de 12 projetos/ano

Bacharelado em Farmácia – 30 projetos
12 extensão | 1 ensino | 17 pesquisa
Média de 10 projetos/ano

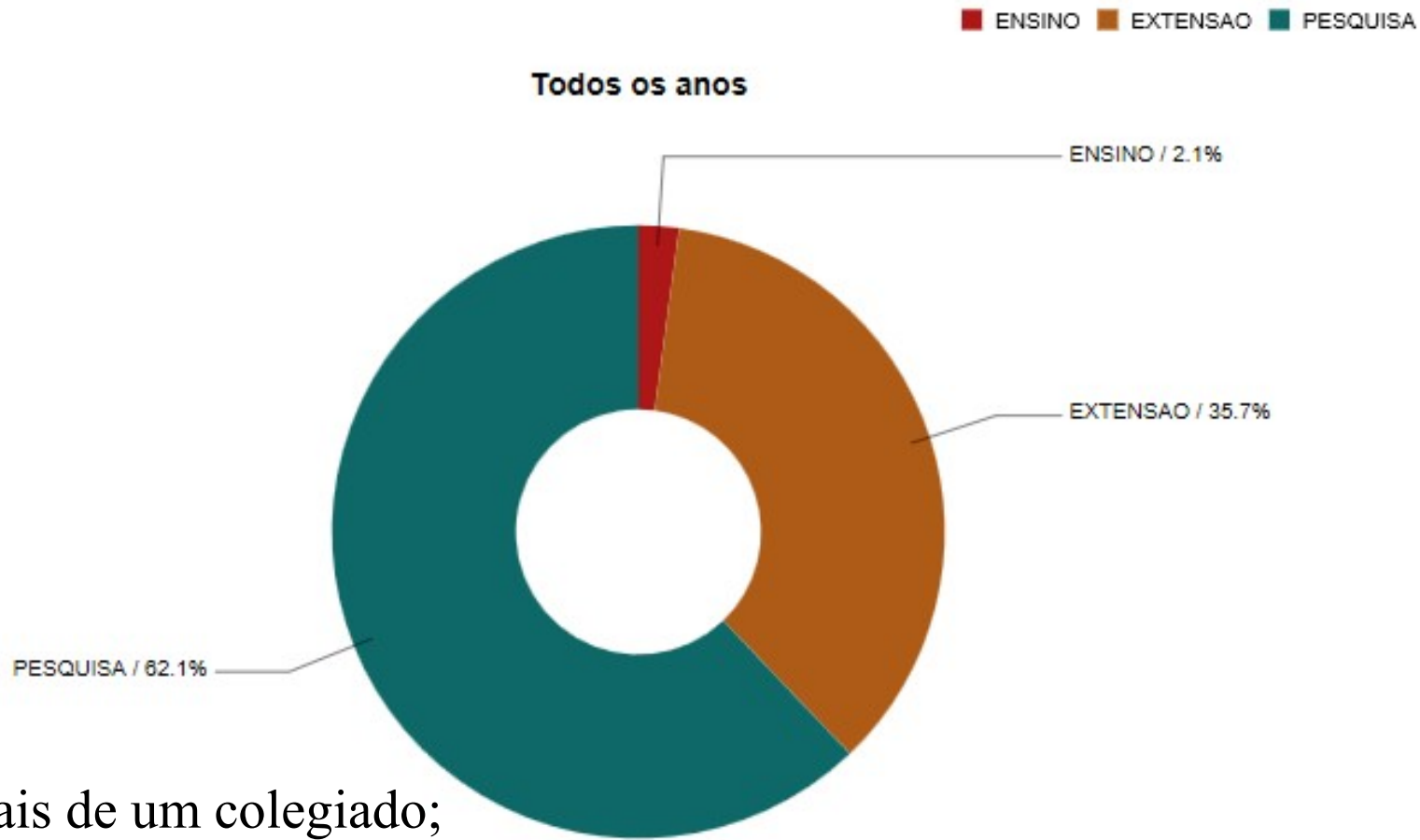


RESULTADOS

PROJETOS SUBMETIDOS AO COPE NOS ÚLTIMOS 3 ANOS, POR COLEGIADO, POR TIPO

Apenas 5 colegiados possuem pelo menos 1 projeto de ensino nos últimos 3 anos

- ❖ Farmácia;
- ❖ Ciências Contábeis;
- ❖ Educação Física,
- ❖ Letras Port./Ing.
- ❖ Pedagogia



- ❖ Total de Projetos 235;
- ❖ 36 projetos contam com docentes de mais de um colegiado;



OUTROS RESULTADOS DO TRABALHO

- ❖ Proposta de Oficina intitulada *“Integrando dados através da ferramenta gratuita Pentaho Data Integration”* para o VI SE²PIN;
- ❖ Artigo completo apresentado no VII Contextos e Conceitos intitulado *“Aplicando Técnicas de Business Intelligence sobre dados da Produção Científica – Um Estudo de Caso na Educação”*;
- ❖ Painel apresentado no VI SE²PIN intitulado *“Aplicando Técnicas de Business Intelligence sobre dados da Produção Científica – Um Estudo de Caso na Educação”* ;
- ❖ Contemplado com uma Bolsa de Iniciação em Desenvolvimento Tecnológico e Inovação (PIBITI) da Fundação Araucária;



CONSIDERAÇÕES FINAIS

- ❖ Dificuldade em encontrar ferramentas *open source* e gratuitas capazes de realizar a extração de dados e informações do Currículos cadastrados na Plataforma Lattes;

ScriptLattes (Mena-Chalco e Junior, 2013)

LattesDataXplorer (Dias, 2013)

- ❖ Qualidade dos dados – ETL

- ❖ Impossibilidade de conseguir relacionar as produções com os projetos mantidos pelo COPE;



TRABALHOS FUTUROS

- ❖ Ampliar a análise estatística sobre os dados da produção científica do campus e a partir da aplicação de técnicas de mineração de dados, identificar padrões, tendências e correlações entre os mesmos;
- ❖ Automatização do processo de ETL para realizar cargas semestrais no DW.



AGRADECIMENTOS

Prof^ª. Dra. Rosineide da Paz (UFSCar/USP)

Prof^ª. Me. Lauriana Paludo (IFPR-Campus Palmas)

Obrigada!

Talita Monteiro

tfpmonteiro@outlook.com

