

# A PESQUISA CIENTÍFICA E TECNOLÓGICA SOB A ÓTICA DA BIBLIOTECA DIGITAL BRASILEIRA DE TESES E DISSERTAÇÕES

*Lillian Alvares*

Faculdade de Ciência da Informação, Universidade de Brasília

*Rafael Berlanga-Llavori*

Escola Superior de Tecnologia i Ciències Experimentals, Universitat Jaume I

# OBJETIVO

---

- ζ Extrair conhecimento da Biblioteca Digital Brasileira de Teses e Dissertações, por meio de mineração de dados, com objetivo de descrever a pesquisa científica e tecnológica desenvolvida no país nos programas de pós-graduação strictu sensu oferecido em universidades e institutos de pesquisa no período de 2006 a 2015.

# OBJETIVOS ESPECÍFICOS

---

## **Extraír informações estratégicas da BDTD:**

- ζ Análise quantitativa e distribuição por instituições das áreas de pesquisa.
- ζ Perfil da pós-graduação em relação a trajetória, temas consolidados da pesquisa e principais representantes.
- ζ Capacidade inovadora, em relação ao pioneirismo nas pesquisas realizadas.
- ζ Quantidade de trabalhos produzidos por áreas temáticas.
- ζ Número de orientadores e quantidade de trabalhos orientados.
- ζ Redes de cooperação entre áreas do conhecimento e pesquisadores .
- ζ Produtividade dos programas por comparação por área do conhecimento.

# A BIBLIOTECA DIGITAL BRASILEIRA DE TESES E DISSERTAÇÕES (BDTD)

# BDTD

---

- ζ Banco de dados que contém os resultados da pós-graduação *stricto sensu* do Brasil
- ζ Bases de dados de produção científica são fundamentais para a geração de indicadores científicos e tecnológicos.

# COMPOSIÇÃO

---

- ζ 400.000 registros
  - ζ 100.821 Teses e 276.344 Dissertações
- ζ 105 instituições brasileiras

# OAI-PMH

---

- ζ Baseados no *Open Archives Initiative Protocol for Metadata Harvesting* (OAI-PMH) ...
  - ζ ... protocolo desenvolvido pela *Open Archives Initiative* que...
    - ζ ... define um mecanismo para coleta de registros de metadados em repositórios de dados, adotando assim, o modelo baseado em padrões de interoperabilidade.

# SISTEMA DE SISTEMAS

---

- ζ A BDTD é uma...
  - ζ ... rede distribuída de sistemas de informação que...
    - ζ ... gerenciam teses e dissertações, com texto completo...
      - ζ ... sendo as instituições de ensino e pesquisa as provedoras de dados e o ...
        - ζ ... IBICT operando como agregador

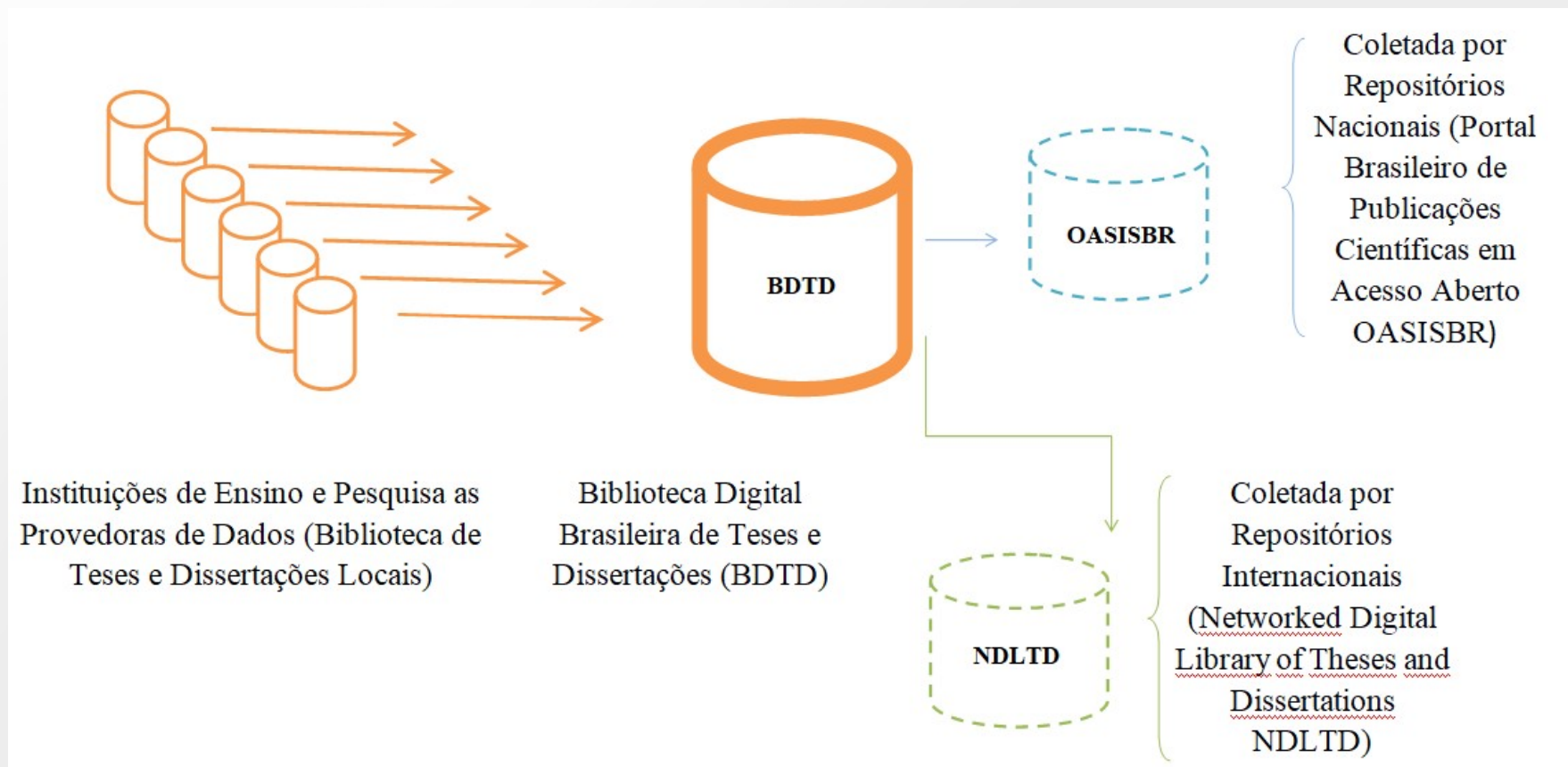


# PROVEDORA

---

- ζ O processo abrange a coleta dos metadados das teses e dissertações dos provedores (instituições de ensino e pesquisa)...
- ζ .... organiza-os em um único repositório e os expõem para coleta para outros provedores de serviços, como por exemplo a *Networked Digital Library of Theses and Dissertations NDLTD*

# ECOSSISTEMA DA BIBLIOTECA DIGITAL BRASILEIRA DE TESES E DISSERTAÇÕES



# MOTIVAÇÃO PARA CRIAÇÃO DA BDTD

---

- ζ Estudar experiências existentes no Brasil e no exterior de desenvolvimento de bibliotecas digitais de teses e dissertações,
- ζ Definir padrões de metadados e tecnologias a serem utilizadas pelo sistema e
- ζ Desenvolver um sistema de publicação eletrônica de teses e dissertações para atender àquelas instituições de ensino e pesquisa que não possuíam sistemas automatizados para implantar suas bibliotecas digitais.

# REPOSITÓRIOS INSTITUCIONAIS

---

- ζ De fato, todas avançaram e essa última foi o impulso necessário à criação dos repositórios institucionais de teses e dissertações, possibilitando, assim, o pleno cumprimento da missão da BDTD.

# ANÁLISE DA INFORMAÇÃO

---

- ζ Naquele momento de criação, no entanto, não havia a intenção de explorar o grande manancial de informação reunida sobre a pesquisa *stricto sensu* brasileira.
- ζ Ao grande volume de dados e informações científicas disponibilizados, não havia a correspondente realização da etapa de análise da informação.
- ζ Isso é, o projeto da BDTD cumpre todo o ciclo informacional, mas não considera a etapa final de análise.

# CICLO INFORMACIONAL

---

S  
E  
L  
E  
Ç  
Ã  
O

A  
Q  
U  
I  
S  
I  
Ç  
Ã  
O

T  
R  
A  
T  
A  
M  
E  
N  
T  
O

A  
R  
M  
A  
Z  
E  
N  
A  
M  
E  
N  
T  
O

D  
I  
S  
S  
E  
M  
I  
N  
A  
Ç  
Ã  
O

A  
N  
Á  
L  
I  
S  
E

# DESCOBERTA DE CONHECIMENTO

---

- ζ Sistema convencional de recuperação de informação em bases de dados não seria capaz
- ζ Mineração de dados, que se resume em uma técnica de análise automática da informação para grande volume de informações
- ζ Extrair informação, previamente desconhecida e de máxima abrangência a partir de grandes bases de dados

# MINERAÇÃO DE DADOS

---

Escolha da Fonte de Dados

Extração dos Dados de Trabalho

Preparação dos Dados

Bases de Dados de Trabalho

Processo de Data Mining

Análise dos Resultados

Disseminação dos Resultados



# METODOLOGIA

# SOFTWARES UTILIZADOS

---

- ζ Softwares de preparação (incluindo limpeza), extração, tratamento (incluindo manipulação), processamento (incluindo visualização), interpretação e análise de dados, processo conhecido como Descoberta de Conhecimento em Bases de Dados ou em inglês Knowledge Discovery in Databases (KDD).
- ζ Referencial em software livre para tratamento automático de informação:
  - ζ Open Refine para carregamento e leitura dos dados
  - ζ Tableau para limpeza, transformação e visualização dos dados.
  - ζ Knime para mineração de dados
  - ζ Jupyter para programação em Python
  - ζ Go JS para confecção dos gráficos

# PERÍODO DE AVALIAÇÃO

---

- ζ O período de avaliação: 2006 a 2015, tendo em vista a confiabilidade dos dados
  - ζ Somente a partir de 2006 as instituições que oferecem programas de pós-graduação *stricto sensu* foram obrigadas a depositar os seus resultados de pesquisa nos repositórios institucionais.
  - ζ Até 2015, uma vez que a extração do conjunto de dados de interesse foi realizado em 2016 e portanto uma parte das instituições ainda estavam alimentando seus repositórios institucionais daquele ano.
    - ζ *As instituições alimentam cotidianamente e retrospectivamente seus repositórios institucionais e a possui registros significativos desde 1997.*

# BARREIRAS

---

- ζ O processo de análise de informação: fonte de dados homogênea e sem ruídos informacionais e aí teve início a primeira barreira a ser superada.
- ζ Identificaram-se problemas na alimentação de dados pelas instituições e problemas na gestão do próprio sistema de coleta da BDTD. Falta de padronização: acentuação e palavras em inglês
  - ζ Assim, a primeira etapa foi a realização da limpeza na base de dados em uso no projeto por meio identificação de regras de correção que foram implementadas por meio de programação a partir de modelos estatísticos.
  - ζ Com essa informação, gerou-se uma relação de indicação de solução e sugestões de melhorias na entrada de dados nos repositórios de dados por instituição.

# PROBLEMAS NA ALIMENTAÇÃO DE DADOS E INDICAÇÃO DE SOLUÇÃO

## ζ Regras para Autores

- ζ Laus\ José Luiz [UNESP] (em dois campos)
- ζ Lavarda\ Carlos Eduardo Facin\ (em dois campos)
- ζ Lavinia Schuler Faccini (em um campo)
- ζ Universidade do Parana Faculdade de Direito (com a instituição)
- ζ Acentos

## ζ Regras para Palavras-chave

- ζ "Beckett\, Samuel\, " → Samuel Beckett
- ζ Crohn\, Doença de → Doença de Crohn
- ζ 8.02.10.00-7 (CDU)
- ζ Recém-nascidos - Mortalidade - Brasil, Sistemas de recuperação da informação - Saúde pública - Brasil, Sistema Único de Saúde (Brasil), SAUDE PUBLICA, SAUDE PUBLICA",

## ζ Regras para Rede de Pesquisadores

- ζ Orientador, coorientador, banca examinadora

## SUGESTÕES DE MELHORIAS NA ENTRADA DE DADOS NOS REPOSITÓRIOS DE DADOS POR INSTITUIÇÃO

---

ζ A UFMT coloca nome, CPF e Lattes:

Adad\, Maria Angela, CPF:489.444.312-01,

<http://lattes.cnpq.br/5694939964123686>.

ζ A FURB, PUC\_PR, coloca o nome completo da universidade, o programa de pós-graduação e uma data, que parece ser o nascimento do orientador.

Amal\, Julio\, 1960-, Universidade Regional de Blumenau.

Programa de Pós-Graduação em Ciências Contábeis e Administração.

- ζ As “regras para rede de pesquisadores” consideraram que os campos reservados para orientador, às vezes continham um, às vezes dois nomes e às vezes cinco nomes.
- ζ Concluiu-se (e assumiu-se) que com dois havia também o coorientador e com mais de dois, a banca examinadora.
- ζ Com isso, criaram-se regras para verificar redes de pesquisadores explícitas e as palavras-chave em torno dos orientadores.

- ζ Outro processo importante envolvido na construção do conjunto de dados para análise foi a classificação automática de palavras-chave de acordo com sua coocorrência.
- ζ Basicamente, as palavras-chave foram tratadas como um documento e as probabilidades condicionais entre elas foram calculadas, a fim de classificá-los em um conjunto predefinido de áreas de investigação, resultante da probabilidade entre palavras-chave e áreas predefinidas.
- ζ Finalmente, a cada entrada no banco de dados foi atribuído as áreas mais prováveis para a pesquisa de acordo com as probabilidades acima. Grande parte do projeto centrou-se na identificação de grandes áreas de pesquisa predefinido, de modo que todas as entradas no banco de dados estivessem completamente cobertas



- ζ O próximo passo, com a base de dados homogênea (e enriquecida), traduz as infinitas possibilidades de análise da pesquisa brasileira:
  - ζ (i) qual área do conhecimento será analisada? Por qual palavra-chave? Em uma instituição? Em várias instituições? Quais os principais orientadores? Ou,
  - ζ (ii) qual a instituição será analisada? Em várias áreas? Em algumas áreas? Ou ainda,
  - ζ (iii) qual a trajetória de pesquisa dos orientadores? Ou também
  - ζ (iv) como se classifica determinado segmento, qual a taxonomia praticada?

ζ Enfim, a opção por qual perspectiva analisar são, de fato, infinitas. E a decisão tomada para a análise aprofundada foi de algumas áreas aleatórias do conhecimento, em várias instituições, em busca de como se organiza essa área do conhecimento no Brasil, quais os principais pesquisadores e como se organizam em redes de colaboradores e quais as instituições desenvolvem pesquisa na área.

# RESULTADOS ALCANÇADOS

# PRIMEIRA FASE

---

- ζ Avaliação Global da BDTD. O campo estudado foi o de palavras-chave (PC).
- ζ O conjunto completo de todos os termos absolutos usados como palavras-chave corresponde a 3.243.871 palavras ou expressões. Depois de aplicar as regras de limpeza e padronização, obteve-se aproximadamente 700.000 palavras-chave.
- ζ As primeiras 281 ocorrências correspondem a 342.418 PC, aproximadamente 50% do total de palavras-chave. O restante 50% estão dispersos em palavras que não tem representatividade estatística, a chamada *long tail*, que apesar de poder conter temas emergentes, também se encontram ali informações com erros de preenchimento ou falta de padronização na entrada de dados.
- ζ De fato, essas situações são muitas e deverão ser abordadas em outro momento de pesquisa. Nesse momento, a opção foi considerar apenas as primeiras 281 ocorrências.

# SEGUNDA FASE

---

## Avaliação Aprofundada por Assunto ou Instituição

- ζ Educação
- ζ Biblioteconomia
- ζ Pesquisa Espacial
- ζ Violência Contra Mulher
- ζ Universidade de Brasília
- ζ Universidade de São Paulo

# **AVALIAÇÃO GLOBAL DA BDTD**

# PROPOSTA DE CLASSIFICAÇÃO DA CIÊNCIA DESENVOLVIDA NO BRASIL

---

- ζ Aproximadamente 700.000 palavras-chave
- ζ 281 primeiras correspondem a 342.418, aproximadamente 50% das palavras-chave.
- ζ O restante é “long tail”, que pode ser analisada em termos de
  - ζ Emergência do tema
  - ζ Erros de preenchimento
  - ζ Falta de padronização na entrada de dados
- ζ Opção: desconsiderar

EDUCAÇÃO	9659
ENGENHARIA DE PRODUÇÃO	5236
ESTUDO E ENSINO	4155
ADMINISTRAÇÃO	3994
HISTÓRIA	3944
EDUCATION	3789
AVALIAÇÃO	3476
ENFERMAGEM	3445
FILOSOFIA	3087
POLÍTICAS PÚBLICAS	2960
EDUCAÇÃO	2691
ESTADO	2688
QUALIDADE DE VIDA	2605
APRENDIZAGEM	2506
DIREITO	2457
FORMAÇÃO	2456
PSICOLOGIA	2344
EPIDEMIOLOGIA	2282
TRABALHO	2272
LITERATURA	2138
ENSINO SUPERIOR	2132



EDUCAÇÃO  
CIÊNCIAS DA SAÚDE  
SOCIOLOGIA  
**TECNOLOGIA**  
LETRAS  
CIÊNCIAS AMBIENTAIS  
CIÊNCIAS EXATAS  
TECNOLOGIA DA INFORMAÇÃO  
CIÊNCIAS AGRÁRIAS  
POLÍTICA  
HISTÓRIA  
INTERDISCIPLINAR  
ADMINISTRAÇÃO  
QUALIDADE  
MEDICINA VETERINÁRIA  
PSICOLOGIA  
PLANEJAMENTO  
ESTADOS BRASILEIROS  
CULTURA  
GEOGRAFIA  
DIREITO  
FILOSOFIA  
COMUNICAÇÃO  
ECONOMIA  
EDUCAÇÃO FÍSICA  
TURISMO

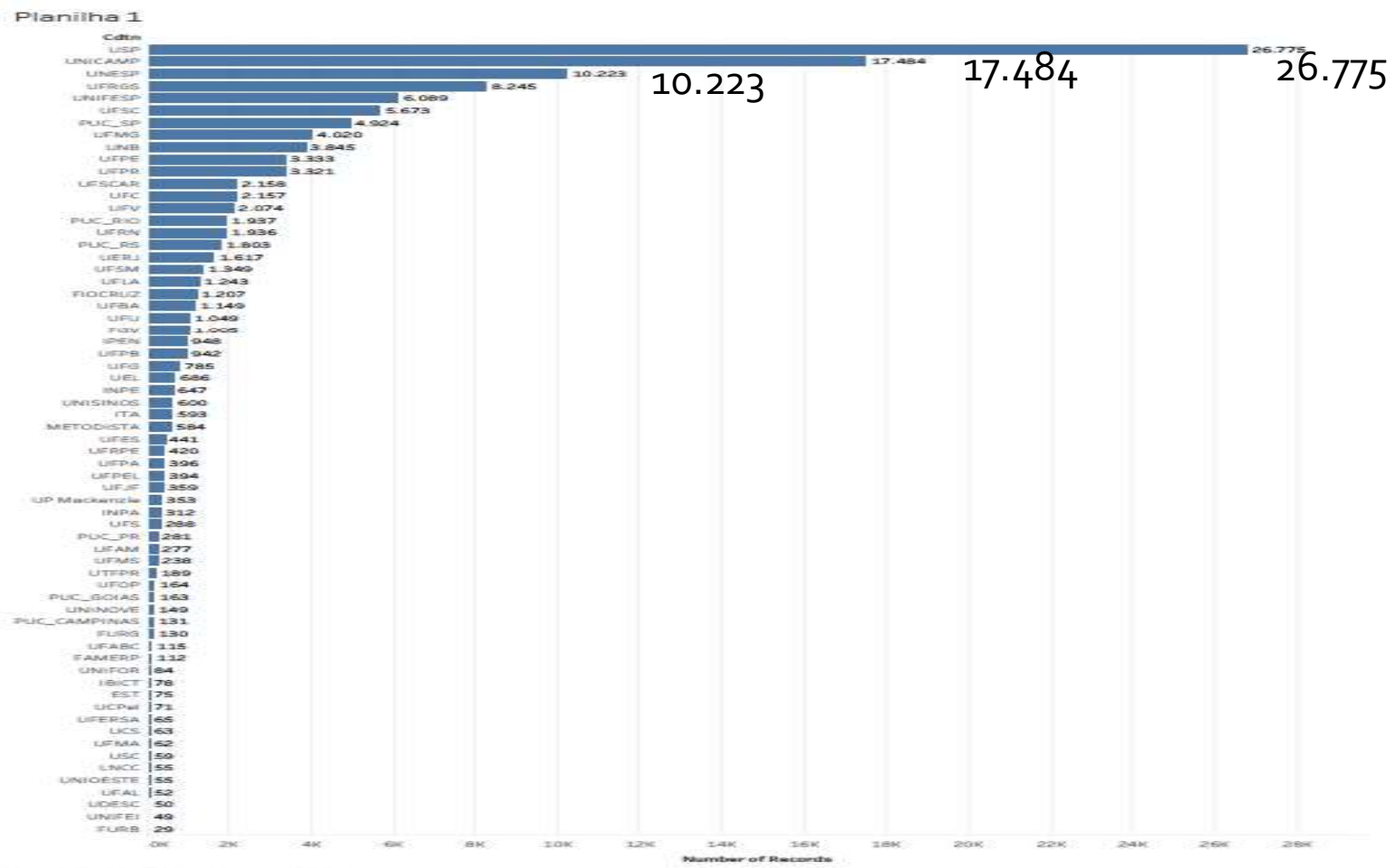
<b>TECNOLOGIA</b>	24.267	ENGENHARIA DE PRODUÇÃO	6.292		
	9,02%	ENGENHARIA MECÂNICA	1.881		
		ENGENHARIA ELÉTRICA	2.409		
		ENGENHARIA CIVIL	1.587	CONSTRUÇÃO CIVIL	815
				ESTRUTURAS	873
		BIOTECNOLOGIA	1.460		
		TECNOLOGIA NUCLEAR	1.333		
		ENERGIA	1.078		
		ENGENHARIA DE MATERIAIS	1.071	MATERIAIS	955
		ENGENHARIA DE SOFTWARE	992		
		TECNOLOGIA	986	ENGENHARIAS	766
		ENGENHARIA QUIMICA	945		
		INOVAÇÕES TECNOLÓGICAS	824		

PROPOSTA DE CLASSIFICAÇÃO DA  
CIÊNCIA DESENVOLVIDA NO BRASIL

# AVALIAÇÃO DETALHADA

UNIVERSIDADE DE SÃO PAULO

# USP EM COMPARAÇÃO



EDUCAÇÃO

# AVALIAÇÃO DA ÁREA DE EDUCAÇÃO

---

ζ 23.823 (PALAVRAS-CHAVE)

ζ 138.127 (OCORRÊNCIAS)



EDUCAÇÃO	9659
<b>EDUCATION</b>	3789
EDUCACAO	2691
EDUCAÇÃO AMBIENTAL	1269
EDUCAÇÃO FÍSICA	1069
EDUCAÇÃO INFANTIL	1011
TECNOLOGIA EDUCACIONAL	954
EDUCAÇÃO ESPECIAL	917
EDUCAÇÃO E ESTADO	795
<b>HIGHER EDUCATION</b>	786
<b>ENVIRONMENTAL EDUCATION</b>	770
EDUCAÇÃO A DISTÂNCIA	602
<b>DISTANCE EDUCATION</b>	594
EDUCAÇÃO MATEMÁTICA	589
EDUCAÇÃO EM SAÚDE	545
<b>PHYSICAL EDUCATION</b>	540
POLÍTICA EDUCACIONAL	504
<b>SPECIAL EDUCATION</b>	477
EDUCACAO ESPECIAL	463
EDUCAÇÃO DE JOVENS E ADULTOS	457
EDUCAÇÃO DE CRIANÇAS	442

UNIVERSIDADE DE BRASÍLIA

# ATUAÇÃO DOCENTE

NEVES, FRANCISCO DE ASSIS ROCHA	CIÊNCIAS DA SAÚDE
	CIÊNCIAS FARMACÊUTICAS
	PATOLOGIA MOLECULAR
TOMAZ, CARLOS ALBERTO BEZERRA	BIOLOGIA ANIMAL
	CIÊNCIAS DA SAÚDE
GARRAFA, VOLNEI	BIOÉTICA
	CIÊNCIAS DA SAÚDE
SUAIDEN, EMIR JOSÉ	CIÊNCIA DA INFORMAÇÃO
NASCIMENTO, ELIMAR PINHEIRO DO	DESENVOLVIMENTO SUSTENTÁVEL
PRATESI, RICCARDO	CIÊNCIAS DA SAÚDE
	CIÊNCIAS MÉDICAS
BARBATO, SILVIANE BONACCORSI	EDUCAÇÃO
	PROCESSOS DE DESENVOLVIMENTO HUMANO E SAÚDE
	PSICOLOGIA CLÍNICA E CULTURA
VIEIRA, JOSÊNIA ANTUNES	LINGUÍSTICA
BANDEIRA, LOURDES MARIA	CIÊNCIAS SOCIAIS
	HISTÓRIA
	SOCIOLOGIA

# ÁREAS DA CIÊNCIA DA INFORMAÇÃO

17,89%

CIÊNCIA DA INFORMAÇÃO

11,38%

ARQUITETURA DA INFORMAÇÃO

9,76%

TECNOLOGIA DA INFORMAÇÃO

9,76%

SERVIÇOS DE INFORMAÇÃO

8,94%

GESTÃO DA INFORMAÇÃO

8,13%

GESTÃO DO CONHECIMENTO

6,50%

ORGANIZAÇÃO DA INFORMAÇÃO

6,50%

INCLUSÃO DIGITAL

6,50%

RECUPERAÇÃO DA INFORMAÇÃO

6,50%

ONTOLOGIA

4,88%

SOCIEDADE DA INFORMAÇÃO

4,07%

SISTEMAS DE RECUPERAÇÃO DA INFORMAÇÃO

4,07%

COMPETÊNCIA INFORMACIONAL

4,07%

LETRAMENTO INFORMACIONAL

4,07%

BIBLIOTECAS UNIVERSITÁRIAS

4,07%

ARQUITETURA DA INFORMAÇÃO

3,25%

ARQUIVOLOGIA

# ÁREAS DE CIÊNCIAS DA SAÚDE

ABORTO LEGAL, VIOLÊNCIA CONTRA AS MULHERES, BRASIL, VIOLÊNCIA SEXUAL  
ABSENTEÍSMO, ATESTADO MÉDICO, ENFERMEIROS, MÉDICOS, GESTÃO  
HOSPITALAR, PROFISSIONAIS DE SAÚDE  
ABSENTEÍSMO, SAÚDE E TRABALHO, SERVIDORES PÚBLICOS  
ACANTHAMOEBA, NAEGLERIA  
ACIDENTE VASCULAR CEREBRAL, CÉREBRO  
ACIDENTE VASCULAR CEREBRAL, HEMIPARESIA, COORDENAÇÃO  
MOTORA, MEMBROS SUPERIORES, CAPACIDADE MOTORA  
ACIDENTE VASCULAR CEREBRAL, HEMORRAGIA, HIPERTENSÃO  
ACIDENTES DO TRABALHO, ENFERMAGEM  
ÁCIDO FÓLICO, DEFEITOS DO TUBO NEURAL (DTNS), DEFEITOS  
CONGÊNITOS, PREVENÇÃO, FARINHAS, FORTIFICAÇÃO  
ACROMEGALIA, QUALIDADE DE VIDA, APTIDÃO FÍSICA  
ACUPUNTURA, MEDICINA CHINESA, SAÚDE PÚBLICA, TRATAMENTO MÉDICO  
ADESIVOS DENTINÁRIOS, SOLVENTES, DENTINA, DENTES  
ADESIVOS DENTINÁRIOS, SOLVENTES, RESTAURAÇÃO (ODONTOLOGIA)

# BIBLIOTECONOMIA

# BIBLIOTECONOMIA

---

ζ 572 PALAVRAS-CHAVE COM "BIBLI"

ζ 1738 OCORRÊNCIAS

# BIBLIOTECONOMIA

---

- ζ BIBLIOMETRIA,149
- ζ BIBLIOTECONOMIA,125
- ζ BIBLIOMETRICS,85
- ζ BIBLIOTECAS UNIVERSITÁRIAS,63
- ζ BIBLIOTECAS DIGITAIS,59
- ζ BIBLIOTECAS,49
- ζ BIBLIOTECA UNIVERSITÁRIA,48
- ζ BIBLIOTECA,37
- ζ BIBLIOTECAS ESCOLARES,35
- ζ BIBLIOTECAS UNIVERSITARIAS,32
- ζ BIBLIOTECARIOS,25
- ζ BIBLIOTECA ESCOLAR,24
- ζ BIBLIOTECA DIGITAL,23
- ζ BIBLIOTECAS DIGITAIS TESES.,20
- ζ ANÁLISE BIBLIOMÉTRICA,19
- ζ BIBLIOGRAFIA,19
- ζ BIBLIOMETRIC ANALYSIS,17
- ζ BIBLIOTECA PÚBLICA,15
- ζ BIBLIOTECAS PÚBLICAS,15
- ζ PESQUISA BIBLIOGRÁFICA,15
- ζ PESQUISA BIBLIOGRAFICA,13



# PESQUISA ESPACIAL BRASILEIRA

# Registros Retirados

Agricultura Familiar  
Agronomia  
Alimentos  
Animais  
Arte  
Aspectos Ambientais  
Aspectos Sociais  
Alimentação  
Cidadania  
Ciencias Agrarias  
Cinema  
Comunicação  
Conservação  
Consumo  
Crianca  
Cultivo  
Cultura  
Democracia  
Desenvolvimento Sustentável  
Direito  
Doencas  
Ecologia  
Educacao

Educação Ambiental  
Educacao Física  
Enfermagem  
Engenharia Civil  
Ensino  
Epidemiologia  
Estado  
Farmacologia  
Filosofia  
Formacao  
Genero  
Geografia  
Historia  
Historia Social  
Identidade  
Idoso  
Inflamacao  
Internet  
Leitura  
Letras  
Linguagem  
Linguistica  
Literatura

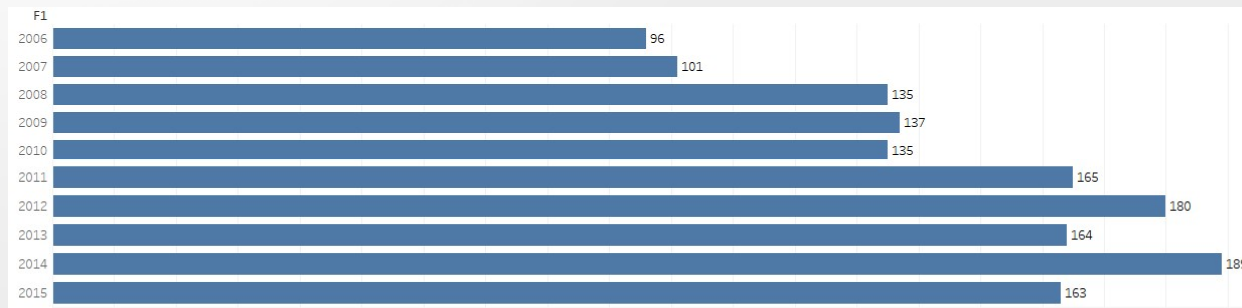
Marketing  
Medicina  
Meio Ambiente  
Memoria  
Nutricao  
Obesidade  
Odontologia  
Para  
Plantas  
Psicanalise  
Psicologia  
Ratos  
Reproducao  
Rio Grande do Sul  
Santa Catarina  
Saude Publica  
Sociologia  
Solos  
Subjetividade  
Sustentabilidade  
Trabalho  
Tratamento  
Turismo

# Registros Considerados

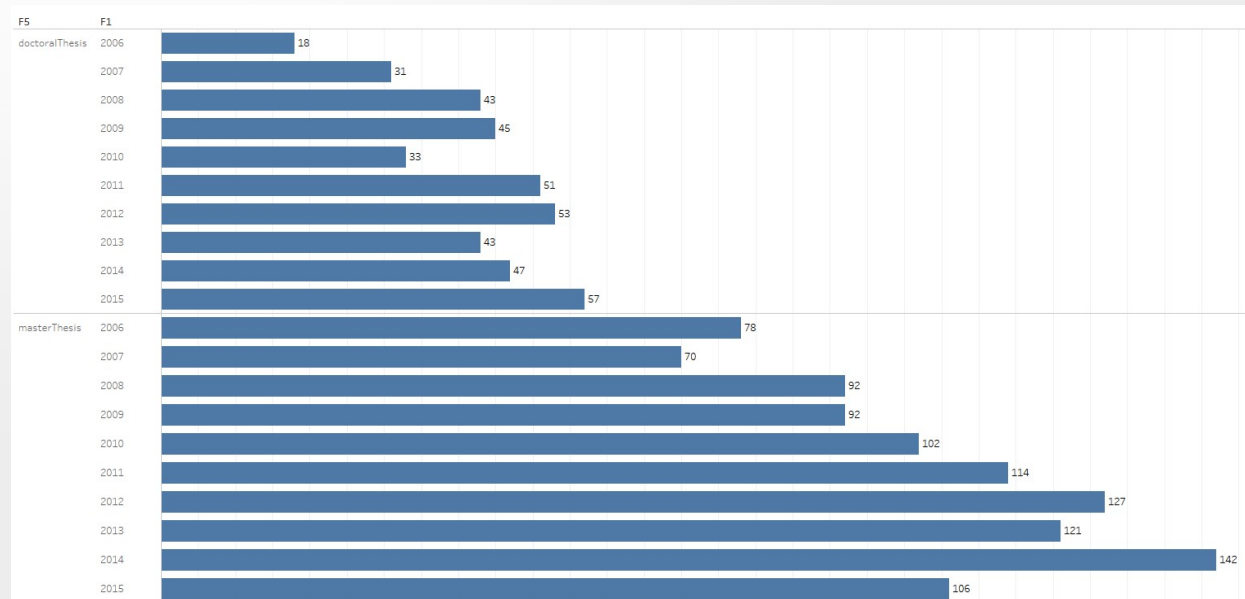
- ζ ???
- ζ Administracao
- ζ Aprendizagem
- ζ Astronomia
- ζ Biotecnologia
- ζ Ciencia da Computacao
- ζ Controle
- ζ Desenvolvimento
- ζ Economia
- ζ Energia
- ζ Engenharia de Materiais
- ζ Engenharia de Produçao
- ζ Engenharia Eletrica
- ζ Engenharia Mecanica
- ζ Fatores de Risco
- ζ Fisica
- ζ Fisica Nuclear
- ζ Matematica
- ζ Performance
- ζ Planejamento
- ζ Qualidade
- ζ Quimica
- ζ Sensoriamento Remoto
- ζ Tecnologia
- ζ Tecnologia da Informacao
- ζ Tecnologia Nuclear

# Número de pesquisas por ano

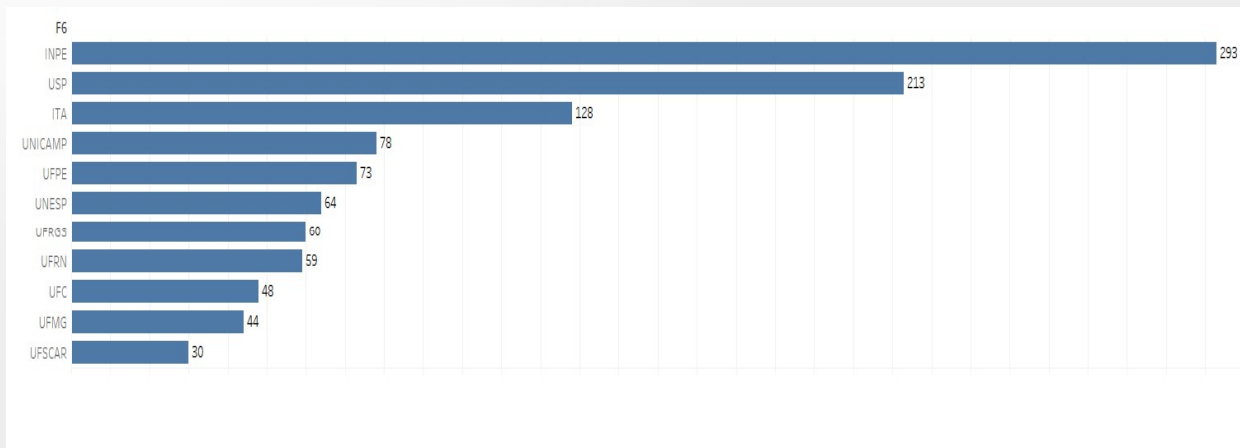
---



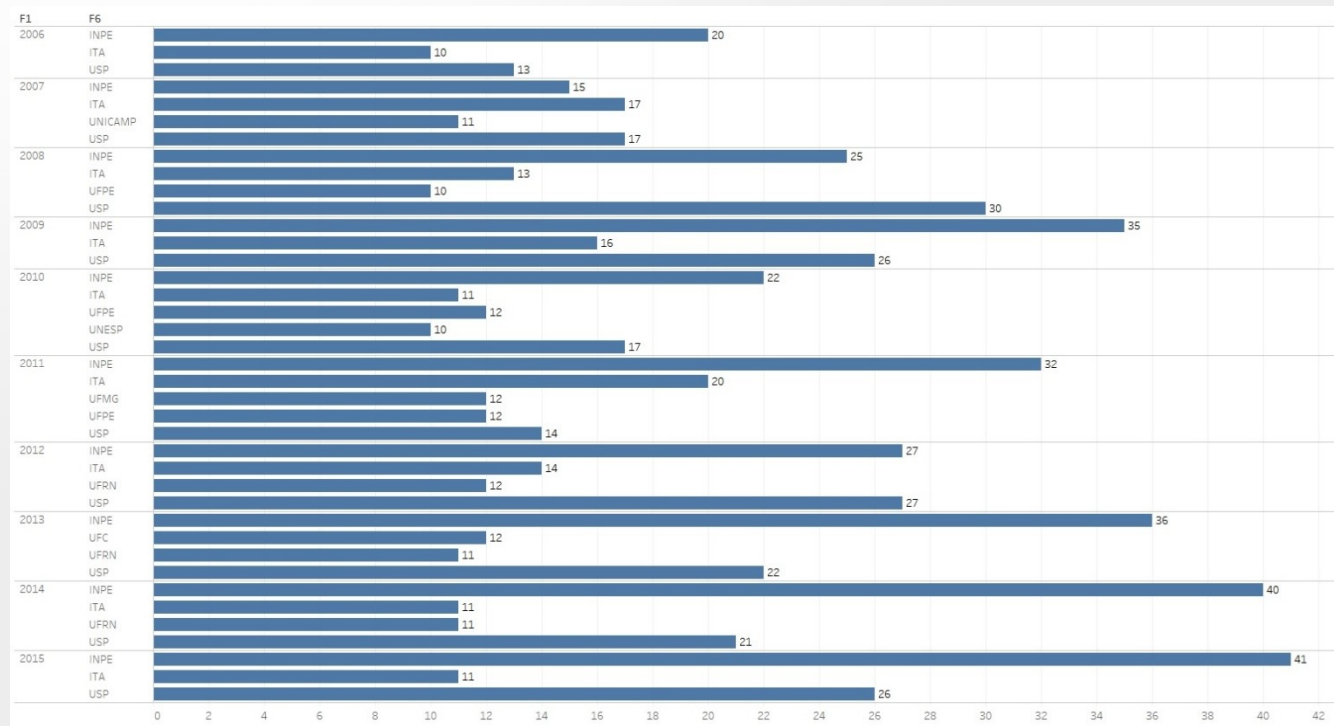
# Número de pesquisas de mestrado e doutorado por ano



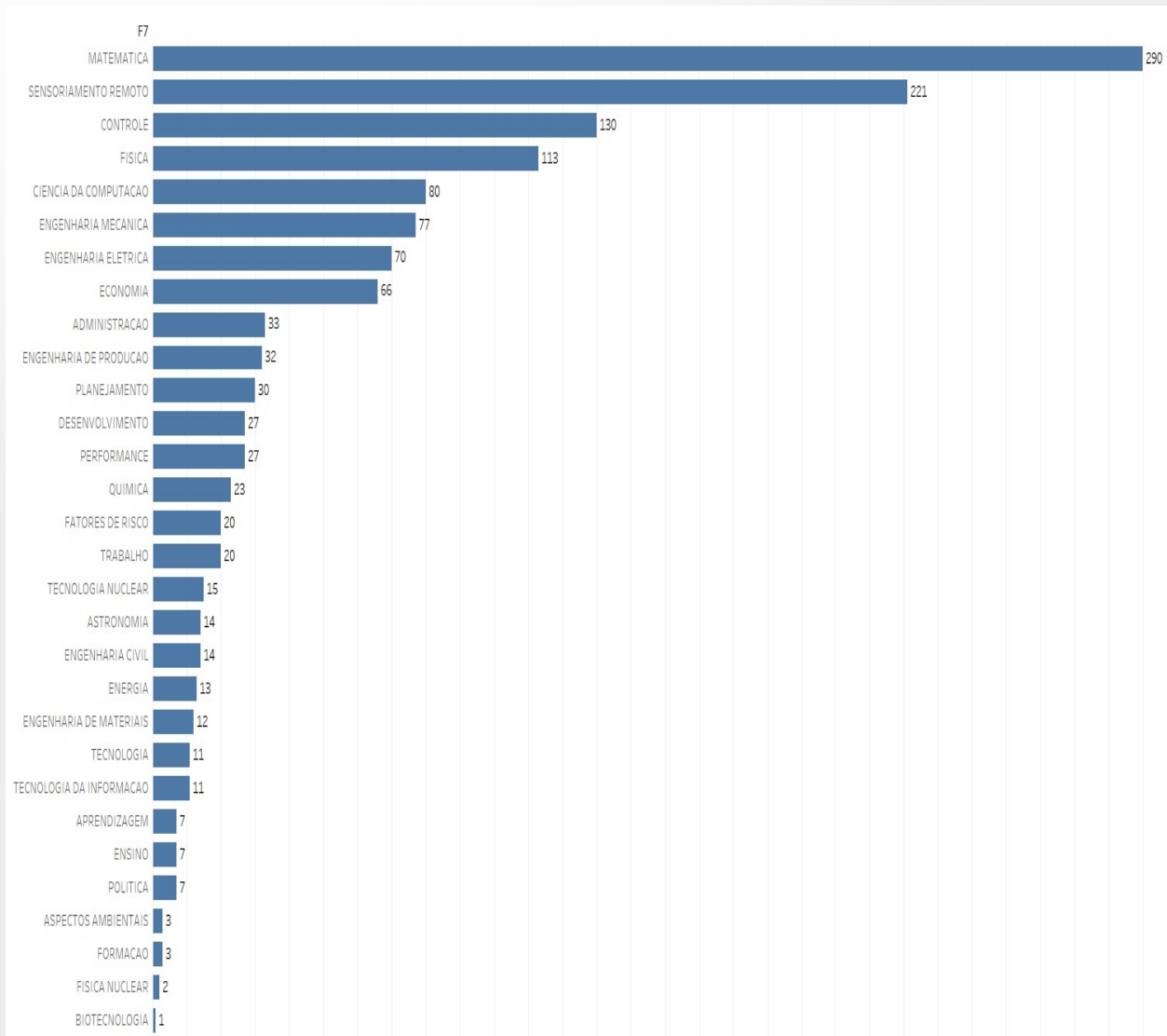
# Ranking das instituições que desenvolveram acima de 30 pesquisas por ano



# Número de pesquisas por ano e por instituição

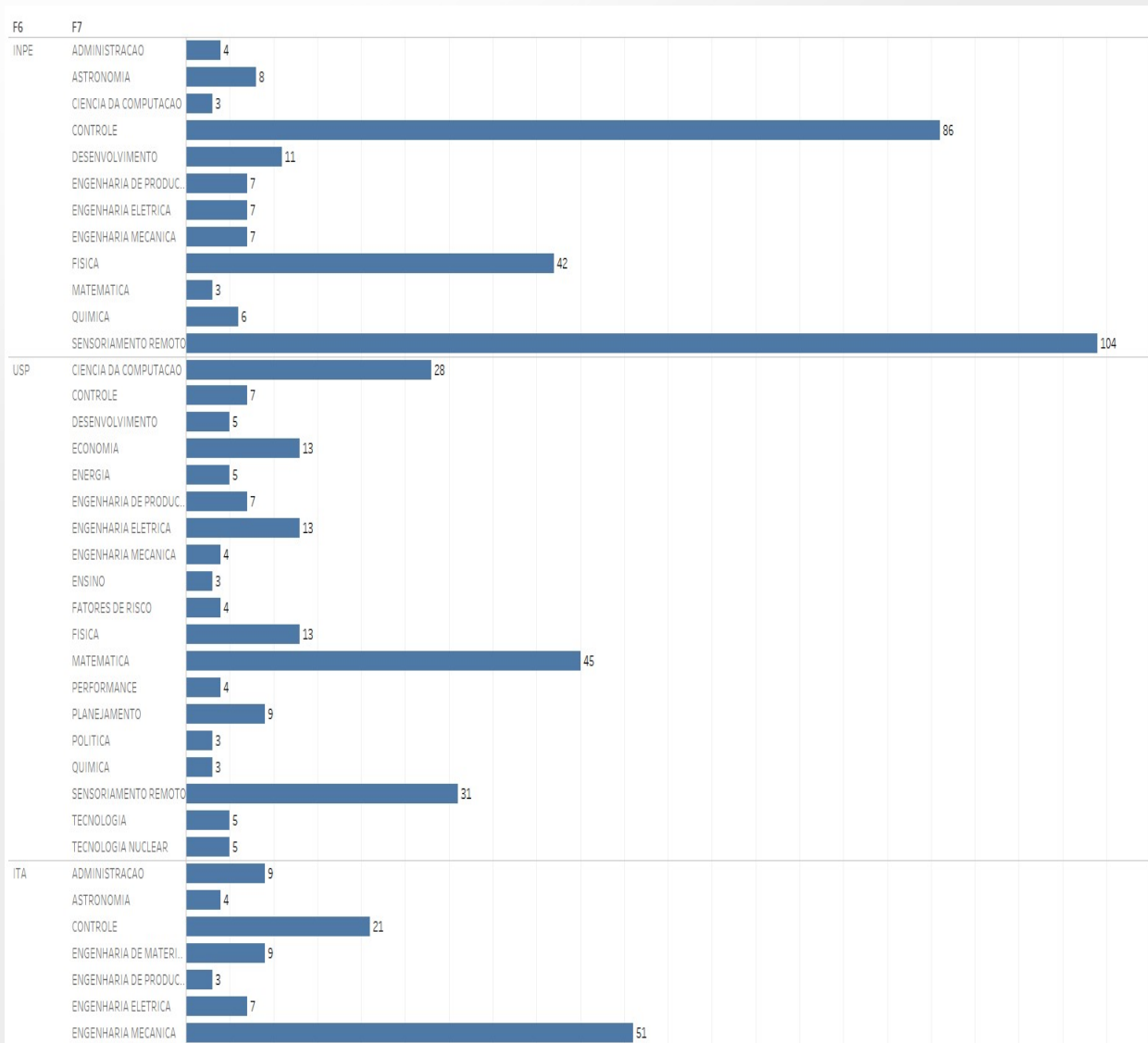


# Descritores selecionados e apresentados na ordem de ocorrência

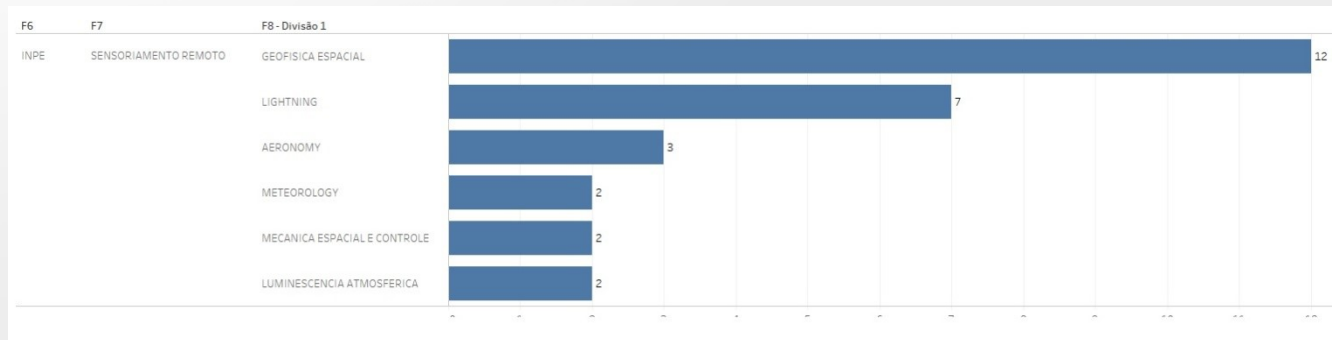




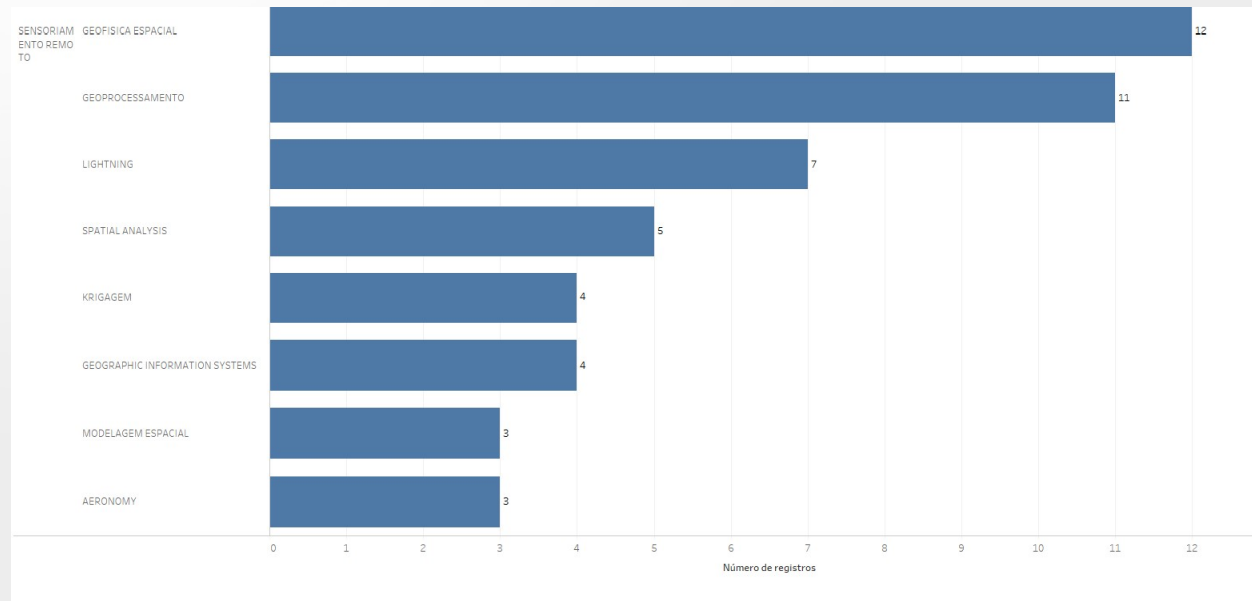
# Principais descritores por instituição



# Descritor sensoriamento remoto e sua relação com as primeiras palavras-chave mais utilizadas no INPE



# Descritores sensoriamento remoto e sua relação com as primeiras palavras-chave mais utilizadas.



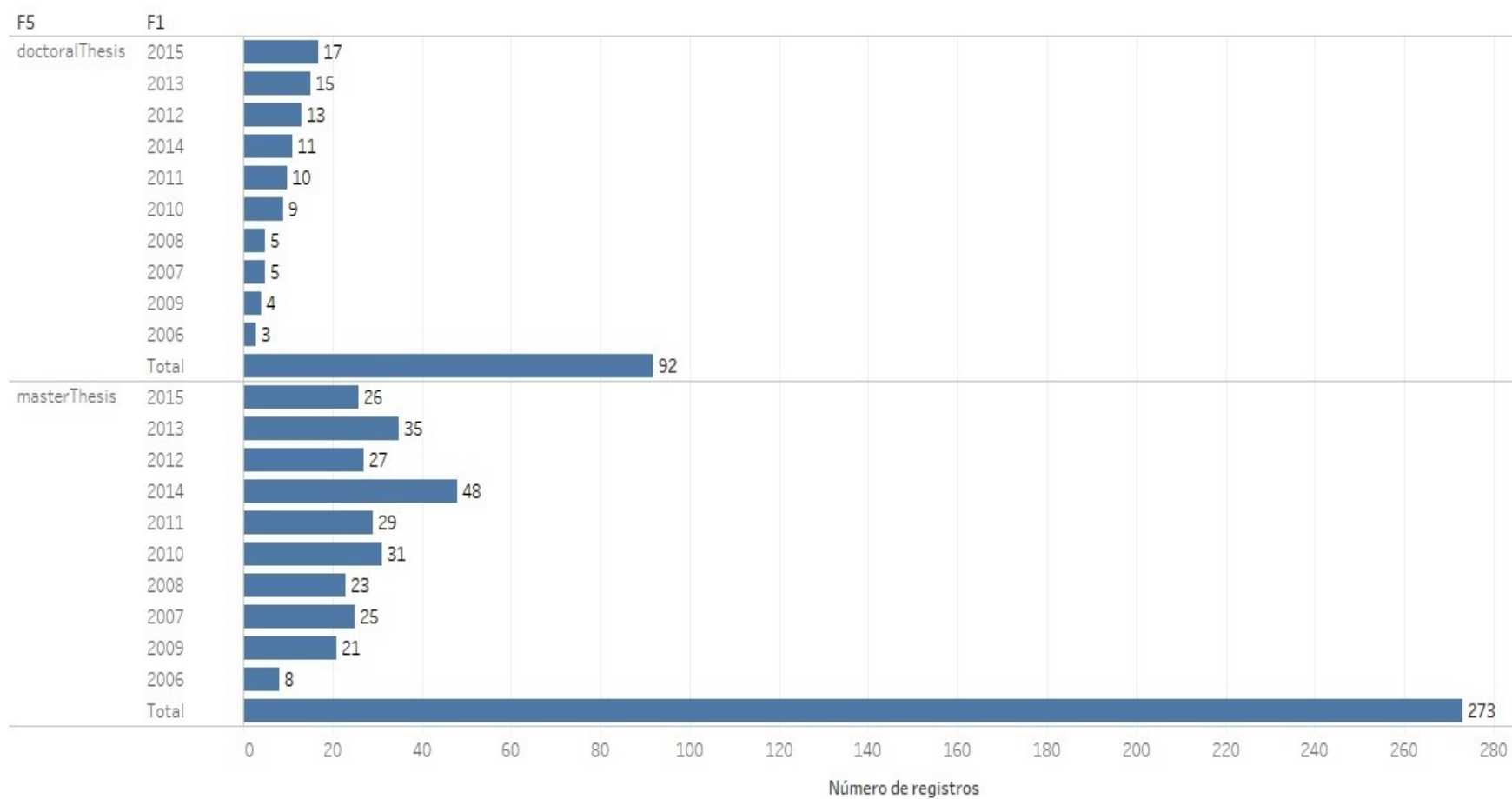
# Relacionamento de palavras-chave no INPE

F6	Palavra-Chave 1	Palavra-Chave 2	Palavra-Chave 3
INPE	MECANICA ESPACIAL E CONTROLE	ALGORITMOS DE IDENTIFICACAO DE E..	SENSOR DE ESTRELAS
		AMBIGUITY RESOLUTION	CYCLE SLIPES
		ATTITUDE ESTIMATION	UNSCENTED KALMAN FILTER
		COLISAO	INCLINATION
			SEPARACAO DE ESTAGIO
		CONTROLE DE ATITUDE	ESTIMACAO
			ESTRUTURA RIGIDO
			SIMULADOR
		CONTROLE DE PRECESSAO	SIMULACAO MATLAB SIMULINK
		DECAIMENTO ORBITAL	ATMOSPHERIC FRAGMENTATION
		DEVELOPMENT FLOW	CO SIMULATION EMULATION
		ESTABILIDADE	TRIPLE SYSTEMS
		INPUT SHAPING	STEP MOTOR
		INTERPLANETARY SPACECRAFT	TRITON
		LQG LTR CONTROL	LYAPUNOV STABILITY
		LQR CONTROL DESIGN	CONTROLE H
		PAINEL SOLAR	CONDUCTANCE
		PARAMETRIZACAO	ENCONTRO E ACOPLAMENTO ATRAC..
		PLANETARY MIGRATION	EVOLUCAO ORBITAL
		PROBLEMA BI	LIPTICO INCLINADO PRECESSANDO
		REACTION WHEEL	ATITUDES DE SATELITES
			ELECTRONIC CONTROL
			RODA DE REACAO
		REAL TIME	INTERNAL SYNCHRONIZATION
		TETRAEDRO	UMI
		THERMAL DESIGN	PROJETO TERMICO
		TOLERANCIA A FALHAS	FAULTS IN SENSORS AND ACTUATORS
		TRAJECTORY	SIMULACAO
		TRANS	NEPTUNIAN OBJECTS
		UMI	MEMS
		VELA SOLAR	SOLAR SAIL
		WELCH	LYNCH
	GEOFISICA ESPACIAL	AIRGLOW	IMAGING SPECTROMETER
		BRIGHT POINTS	PONTOS BRILHANTES
		IONOSPHERE	EQUATORIAL ANOMALY

F6	Palavra-Chave 1	Palavra-Chave 2	Palavra-Chave 3
INPE	LIGHTNING	ESPECTROSCOPIA SEM FENDA	DESCARGAS ATMOSFERICAS
		GEOFISICA ESPACIAL	LIDER BI LIGHTNING LOCATION SYSTEMS THESIS
PMM		GEOFISICA ESPACIAL CIENCIAS ATMOSFERICAS	DVAR WRF MESOSCALE MODEL
		HIGH	SPEED CAMERAS
		LIDER	ATMOSPHERIC DISCHARGE
		ANALISE DE MISSAO	PROPRIEDADES DE MASSA
SYSTEMS ENGINEERING		MECANICA ESPACIAL E CONTROLE	AVIONICA MODULAR INTEGRADA
		RECONFIGABLE CONTROL	MECANICA ESPACIAL E CONTROLE
		SISTEMAS DE CONTROLE	SUBSISTEMA DE CONTROLE DE ATIT..
		SUBSISTEMA PROPULSIVO	SIMULACAO
		AERONAUTIC NAVIGATION	NAVEGACAO AERONAUTICA
MAGNETOSFERA		ARCHITECTURE DEVELOPMENT	ENGENHARIA DE SISTEMAS ENGINEERING
		E	ENGENHARIA DE SISTEMAS
		EFEITOS DE RADIACOES IONIZANTES	ENGENHARIA DE SISTEMAS
		ENGENHARIA SIMULTANEA	ENGENHARIA DE SISTEMAS
MANOBRAS ORBITAIS		GEOFISICA ESPACIAL	SPACE ELECTRODYNAMICS
		MAGNETOPOUSE	MAGNETOPAUSA
		NUMERICAL SIMULATION	MHD
		TRANSIENT EVENT	MAGNETIC FIELD RECONNECTION
RAIOS COSMICOS		MECANICA ESPACIAL E CONTROLE	CONTROLE DE TRAJETORIA TRANSFERENCIA ORBITAL
		MINOR BODIES	MECANICA ESPACIAL E CONTROLE
		SATELITES ARTIFICIAIS	ORBIT DETERMINATION
		ANALISE ESPECTRAL	SOLAR ACTIVITY
AERONOMY		ATMOSFERA TERRESTRE	TEMPESTADES GEOMAGNETICAS
		GEOMAGNETIC ACTIVITY	ATIVIDADE SOLAR
		PREVISAUDE TEMPESTADES GEOMAG..	GEOFISICA ESPACIAL
		GEOFISICA ESPACIAL	ATMOSPHERIC ELECTRIC FIELDS ATMOSPHERIC TEMPERATURE
CONTROLE		PROCESSAMENTO DE IMAGENS	GEOFISICA ESPACIAL
		DINAMICA ORBITAL	MECANICA ESPACIAL E CONTROLE
		MECANICA ESPACIAL E CONTROLE	SYSTEM SIMULATION
		SISTEMAS DE CONTROLE	OTIMIZACAO
COHERENT STRUCTURES		EDGE OF CHAOS	GEOFISICA ESPACIAL CIENCIAS DO A..
		SYNCHRONIZATION	GEOFISICA ESPACIAL
CORRENTE CONTINUA		GEOFISICA ESPACIAL	RAIOS POSITIVOS
		LIGHTNING	GEOFISICA ESPACIAL CIENCIAS ATM..
DINAMICA ORBITAL		MECANICA ESPACIAL E CONTROLE	CONTROLE DE TRAJETORIA
		RAZAO D H	D H RATIO
LUMINESCENCIA ATMOSFERICA		BOLHAS DE PLASMA	GEOFISICA ESPACIAL
		PERKINS INSTABILITY	INSTABILIDADE PERKINS
METEOROLOGY		DEPARTMENT OF AIR SPACE CONTROL	DECEA
		DIAS DE TEMPESTADE	GEOFISICA ESPACIAL CIENCIAS ATM..
OTIMIZACAO		MECANICA ESPACIAL E CONTROLE	ASTRODINAMICA M

# Violência Contra a Mulher

## Mestrado e Doutorado de 2006 a 2015: 365



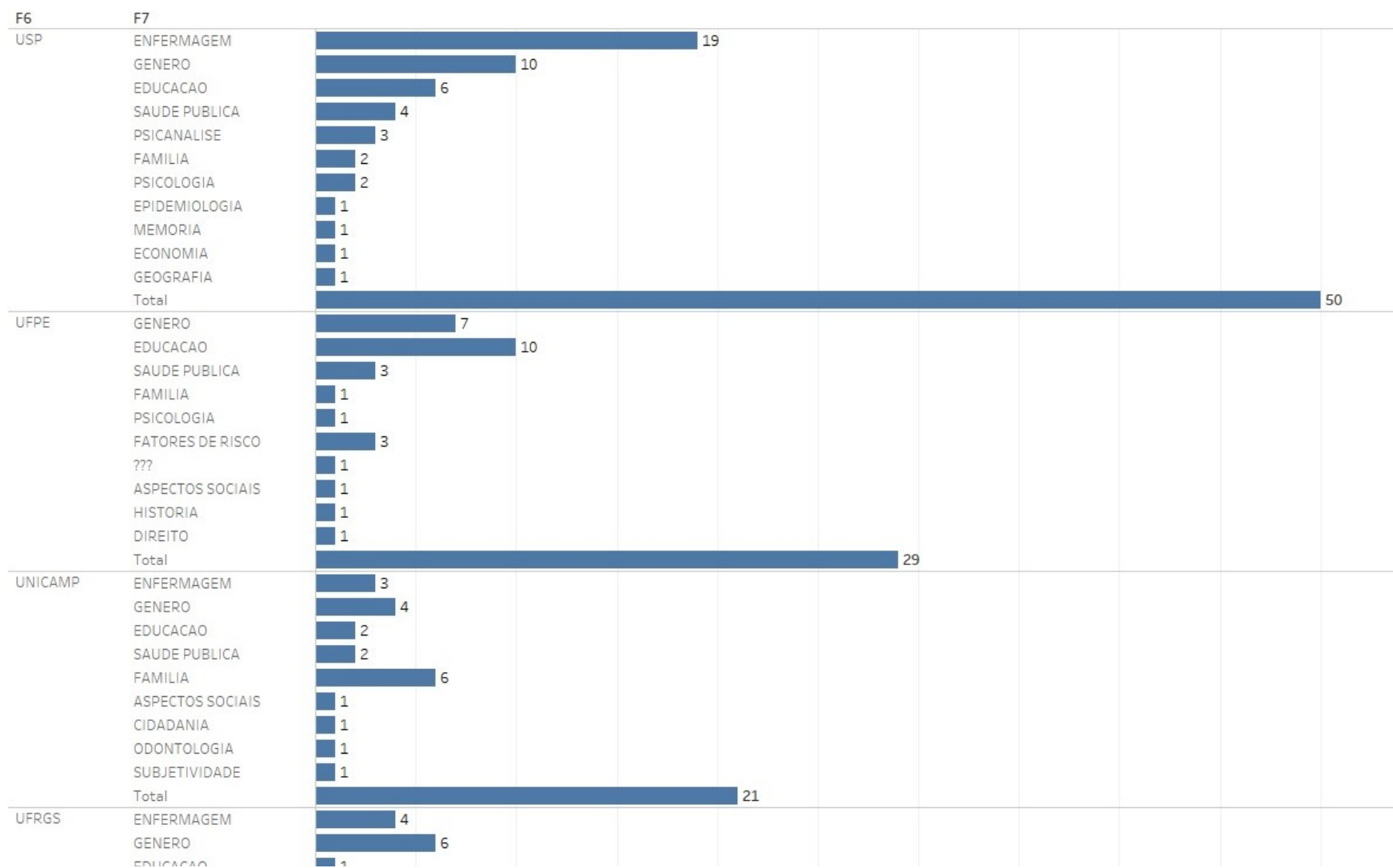
## Detalhamento das Pesquisas de Doutorado: Instituição e Área

F5	F6	F7	
doctoralThesis	USP	ENFERMAGEM	7
		EDUCACAO	4
		GENERO	3
		PSICOLOGIA	1
		EPIDEMIOLOGIA	1
		Total	16
	UFRGS	ENFERMAGEM	2
		GENERO	2
		FAMILIA	1
		EPIDEMIOLOGIA	1
HISTORIA SOCIAL		1	
Total		7	
UNICAMP	ENFERMAGEM	1	
	EDUCACAO	1	
	SAUDE PUBLICA	2	
	FAMILIA	2	
	CIDADANIA	1	
	Total	7	
UNIFESP	ENFERMAGEM	1	
	SAUDE PUBLICA	3	
	FAMILIA	1	
	EPIDEMIOLOGIA	1	
	Total	6	
UNESP	EDUCACAO	4	
	GENERO	1	
	SAUDE PUBLICA	1	
	Total	6	
UFBA	ENFERMAGEM	3	
	EDUCACAO	1	
	CIENCIAS DA SAUDE	1	
	ESTADO	1	
	Total	6	
FIOCRUZ	ENFERMAGEM	2	
	TRABALHO	3	
	Total	5	
UFPB	GENERO	1	





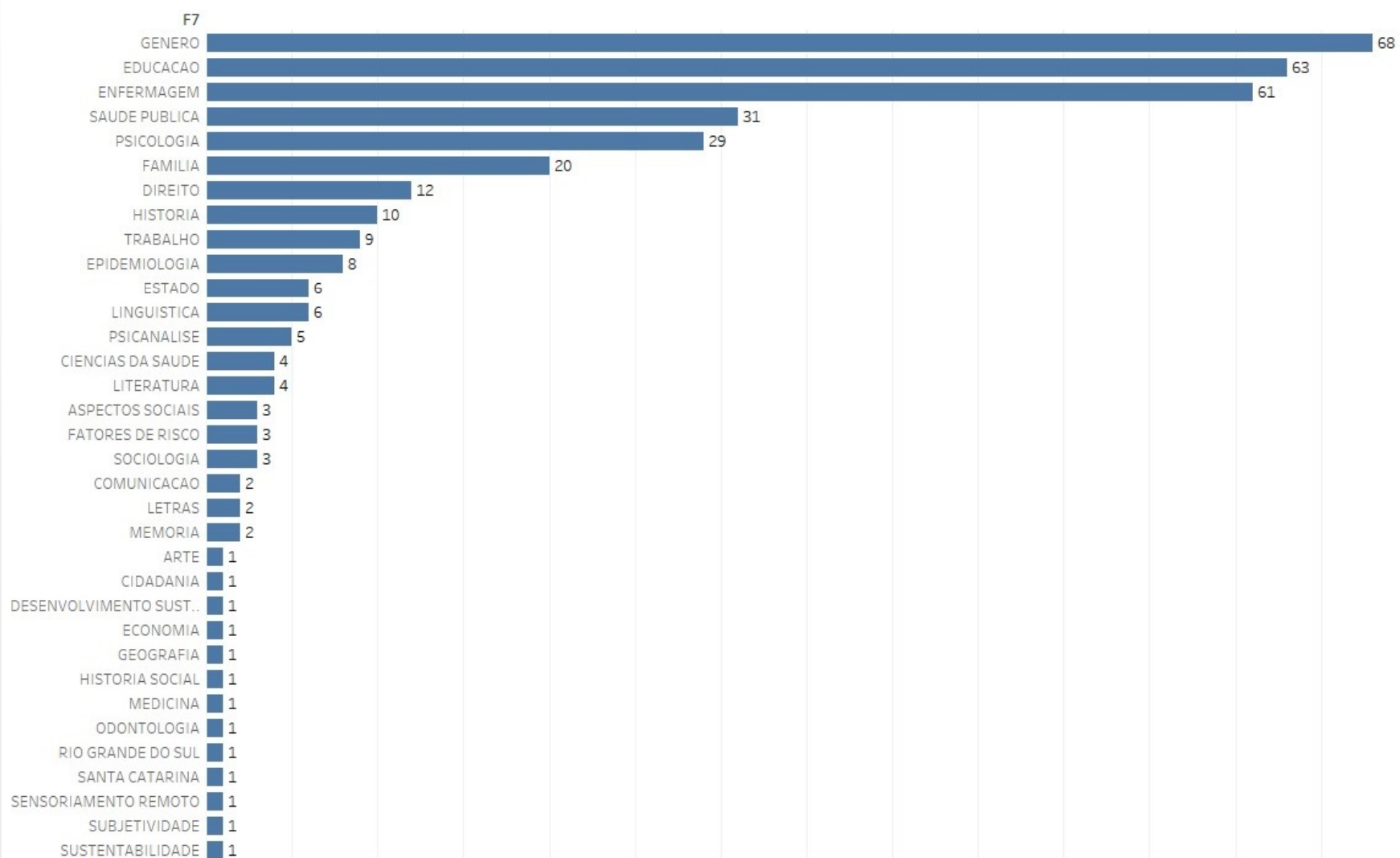
## Instituições e Áreas de Pesquisa



## Principais Orientadores e Instituições

F4 - Divisão 1	F6	
Elza Berger Salema Coelho	UFSC	6
Normelia Maria Freire Diniz	UFBA	6
Ana Flavia Pires Lucas D Oliveira	USP	5
Marta Julia Marques Lopes	UFRGS	5
Rita de Lourdes de Lima	UFRN	5
Rosa Maria Godoy Serpa da Fonseca	USP	5
Ana Marcia Spano Nakano	USP	4
Augusta Thereza de Alvarenga	USP	4
Benedito MedradoDantas	UFPE	4
Guita Grin Debert	UNICAMP	4
Lucia Helena Garcia Penna	UERJ	4
Marlene Neves Strey	PUC_RS	4
Stela Nazareth Meneghel	UFRGS	2
	UNISINOS	2
Lourdes de Maria Leitao Nunes Rocha	UFMA	3
Luiza Jane Eyre de Souza Vieira	UNIFOR	3
Marilia Martins Vizzotto	METODISTA	3
Raimunda Magalhaes da Silva	UNIFOR	3
Alessandro Leite Cavalcanti	UEPB	2
Aloisio Jose Bedone	UNICAMP	2
Ana Bernarda Ludermir	UFPE	2
Ana Maria de Almeida	USP	2
Anibal Faundes	UNICAMP	2
Antonio Augusto Moura da Silva	UFMA	2
Arlete Maria dos Santos Fernandes	UNICAMP	2
Blanca Susana Guevara Werlang	PUC_RS	2
Claudia Maria de Mattos Penna	UFMG	2
Cynthia de Carvalho Lins Hamlin	UFPE	2
Elisa Medici Pizao Yoshida	PUC_CAMPINAS	2
Elisabeth Meloni Vieira	USP	2
Heloisa Buarque de Almeida	USP	2
Henrique Caetano Nardi	UFRGS	2
Joseli Bastos da Costa	UFPB	2
Karla Galvao ADRIAO	UFPE	2
Lindioneza Adriano Ribeiro	UFU	2

## Grandes Áreas



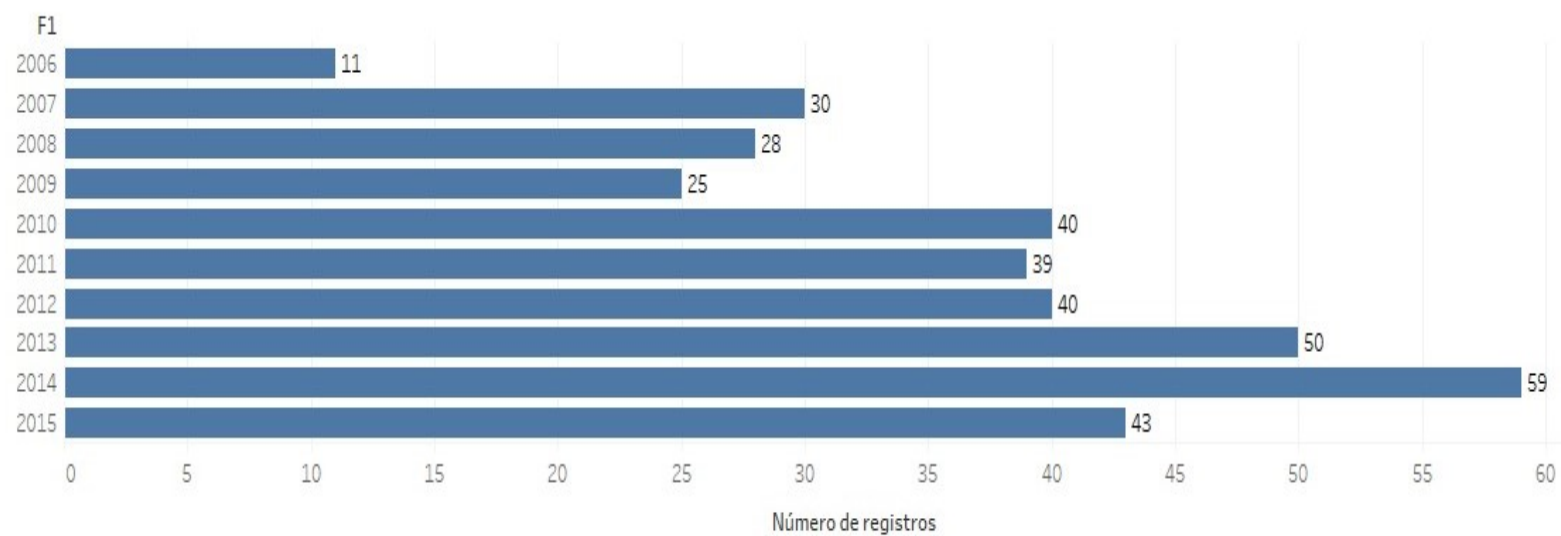
## Primeira Divisão das Grandes Áreas

F7	F8 - Divisão 1	
GENERO	VIOLENCIA DOMESTICA	5
	PATRIARCADO	5
	CICLOS DE VIDA E SOCIEDADE	3
	WOMEN	2
	VIOLENCIA CONTRA A MULHER	2
	VIOLENCIA CONJUGAL	2
	MOVIMENTOS SOCIAIS	2
	LEI	2
	DIREITOS HUMANOS	2
	WOMEN VIOLENCE AGAINST	1
	VULNERABILITY	1
	VIOLENCIA NAS RELACOES INTIMAS	1
	VIOLENCIA DOMESTICA CONTRA A MULH..	1
	VIOLENCIA CONTRA MULHERES	1
	VIOLENCIA CONTRA LA MUJER	1
	VIOLENCIA CONTRA AS MULHERES	1
	VIOLENCE DE GENRE	1
	VIOLENCE AGAINST WOMEN	1
	VIOLENCE	1
	SOCIOLOGIA	1
	SOCIAL SERVICE	1
	SOCIAL POLICY	1
	SAUDE DA MULHER	1
	RS	1
	REPRESENTACOES SOCIAIS	1
	RELACIONAMENTO CONJUGAL	1
	REFLECTION GROUP OF WOMEN VICTIMS...	1
	PUBLIC POLICIES FOR WOMEN	1
	PUBLIC HEALTH SERVICE	1
	PROFISSIONAIS DE SAUDE	1
	PROFESIONAL DE LA SALUD	1
	PRATICAS SOCIAIS	1
	POLITICAS PUBLICAS PARA AS MULHERES	1
	POLITICAS DE ENFRENTAMENTO A VIOLE..	1
	MUTILATION	1
	MULHERES DA PAZ	1

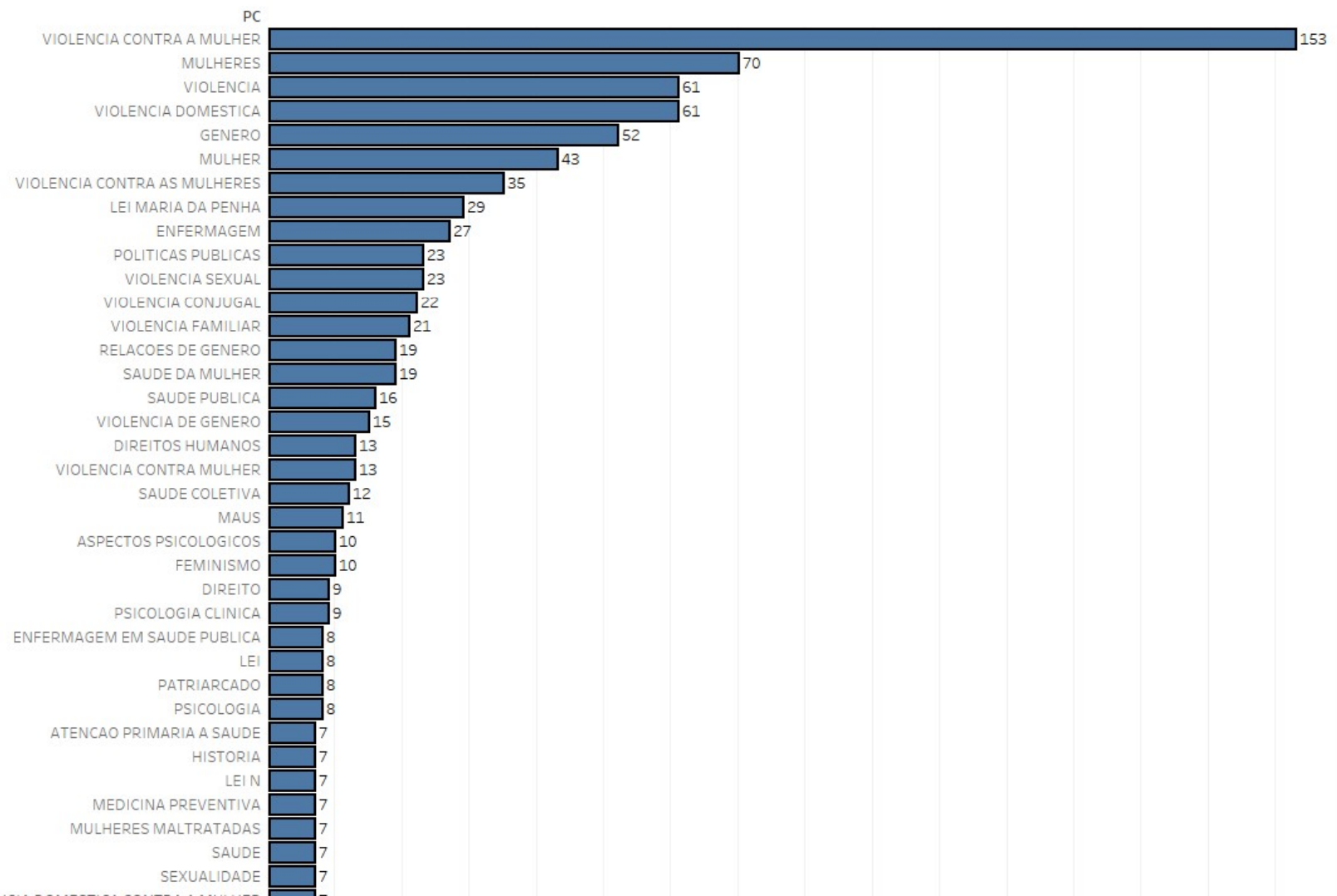
## Segundo detalhamento de áreas

F8 - Divisão 1	F8 - Divisão 2			
VIOLENCIA DOMESTICA	WOMEN		3	
	LEI MARIA DA PENHA		2	
	GENERO		2	
	VIOLENCIA CONTRA LA MUJER		1	
	TESTE DAS RELACOES OBJETAIS		1	
	NURSE CLINICIANS		1	
	MULHER		1	
	MEDICINA PREVENTIVA		1	
	LEGAL FIELD		1	
	ATENCAO A SAUDE		1	
	ASSISTENCIA SOCIAL		1	
	ASPECTOS PSICOSSOCIAIS		1	
VIOLENCIA CONTRA A MULHER	GENERO		1	
	CUIDADO DE ENFERMAGEM		1	
	DESENVOLVIMENTO SUSTENTAVEL		1	
	DIREITOS DA MULHER		1	
	FEMICIDIO		1	
	MEMORIA		1	
	MULHERES MALTRATADAS		1	
	SAUDE PUBLICA		2	
	VIOLENCIA POR PARCEIRO INTIMO		1	
	ASPECTOS PSICOLOGICOS	WOMEN		1
CLINICAL PSYCHOLOGY			2	
ESTUDANTES UNIVERSITARIOS			1	
PSICOLOGIA			1	
PSICOLOGIA DA MULHER			1	
RELACOES HOMEM			1	
VIOLENCIA CONTRA A MULHER			1	
WOMEN AND VIOLENCE			1	
HISTORIA	DIREITOS DAS MULHERES		1	
	DITADURA E DITADORES		1	
	IDENTIDADE		1	
	INCLUSAO SOCIAL		1	
	RIO GRANDE DO SUL		1	
	TITULO		1	
	PATRIARCADO	CONCEICAO DO COITE		1
CULTURA PATRIARCAL			1	
DIREITOS REPRODUTIVOS			1	
PATRIARCADO			1	

## Evolução da quantidade de pesquisa por ano



## Principais Descritores da Pesquisa





## Direito

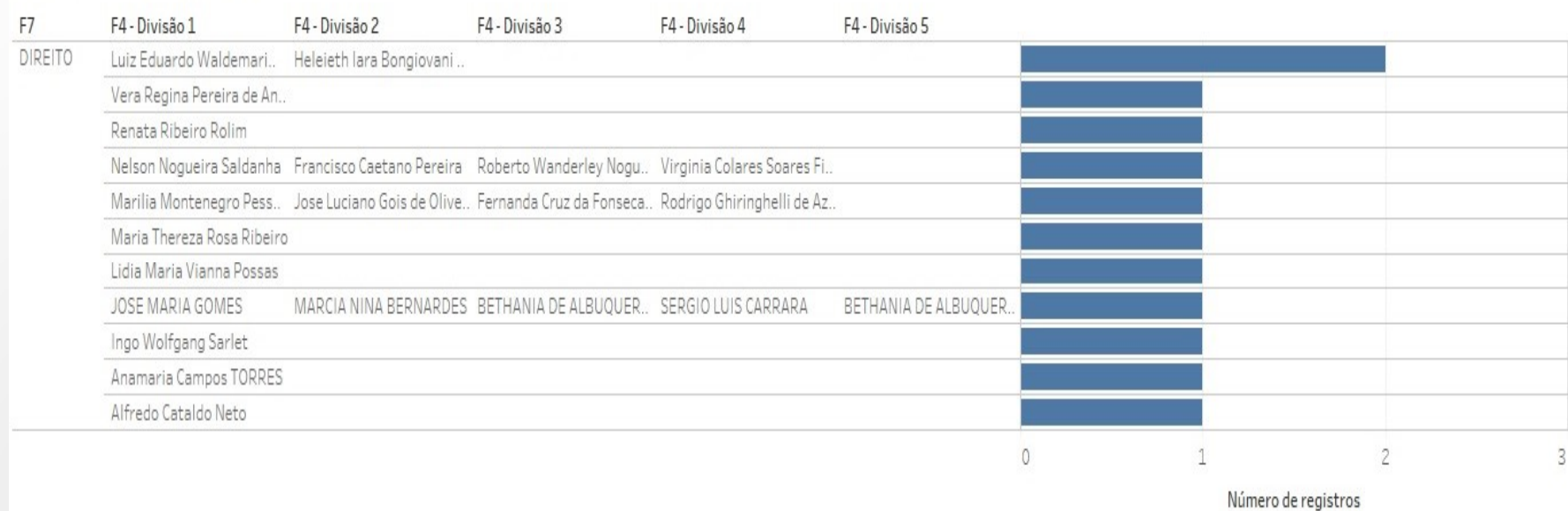
F7	F8 - Divisão 1	F6	F4 - Divisão 1		
DIREITO	VIOLENCIA FAMILIAR	PUC_RIO	JOSE MARIA GOMES	1	
		UNESP	Lidia Maria Vianna Possas	1	
	VIOLENCIA DOMESTICA CONTRA MULHERES	PUC_SP	Luiz Eduardo Waldemarim Wanderley	2	
	DIREITO PENAL	UFPE	Anamaria Campos TORRES	1	
		UFSC	Vera Regina Pereira de Andrade	1	
	VIOLENCIA DOMESTICA	UFPEL	Maria Thereza Rosa Ribeiro	1	
	JUIZADOS ESPECIAIS	PUC_RS	Alfredo Cataldo Neto	1	
	HISTORIA	UNICAP	Nelson Nogueira Saldanha	1	
	DIREITOS FUNDAMENTAIS	PUC_RS	Ingo Wolfgang Sarlet	1	
	DIREITO	UNICAP	Marilia Montenegro Pessoa de Mello	1	
	ACCESS TO JUSTICE	UFPB	Renata Ribeiro Rolim	1	

Número de registros

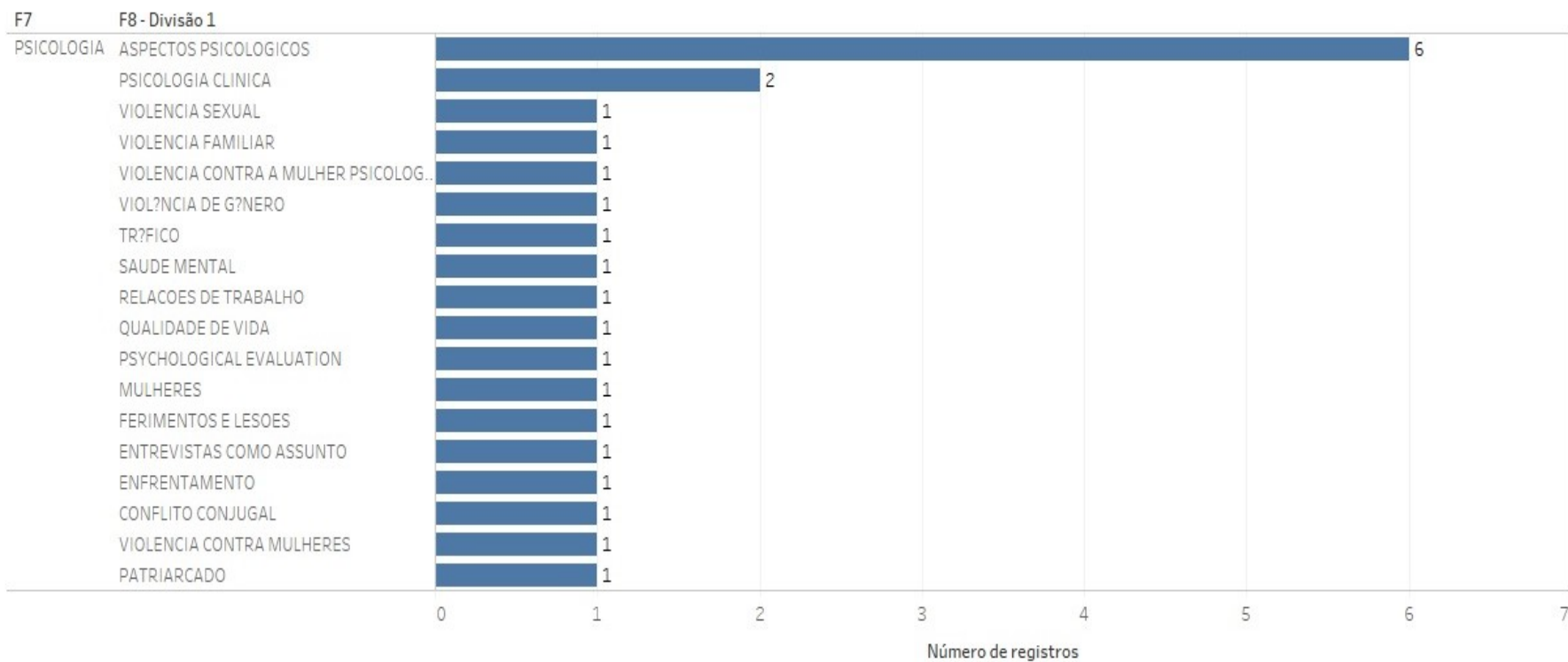
## Palavras-Chave Associadas



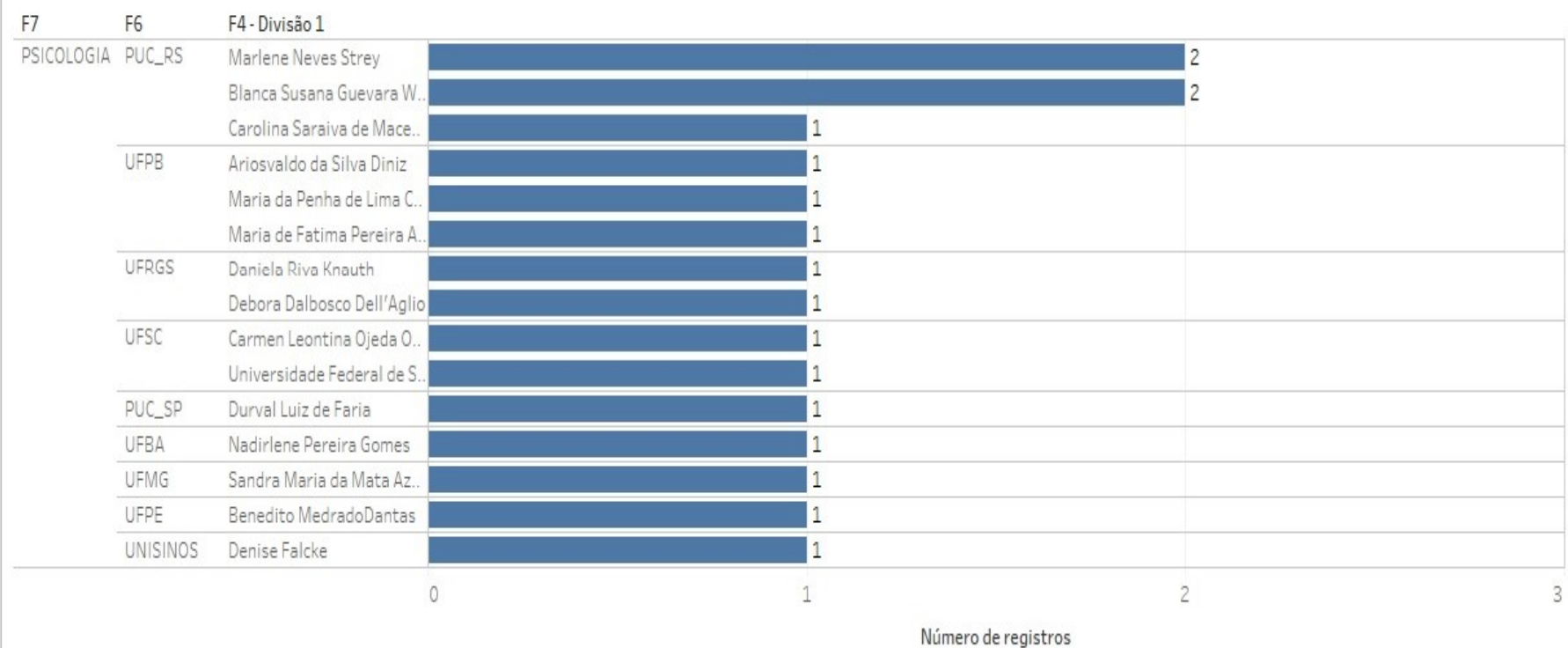
## Direito: Pesquisadores



## Primeira Divisão das Grandes Áreas: Psicologia



## Primeira Divisão das Grandes Áreas: Detalhamento Psicologia



# CONCLUSÃO

- ζ A obtenção de conhecimento sobre a pesquisa científica e tecnológica brasileira a partir da análise automática de informação na Biblioteca Digital Brasileira de Teses e Dissertações é de expressiva importância para toda a cadeia de produção do conhecimento científico e tecnológico.

ζ Pode contribuir tanto para o planejamento estratégico de políticas de ensino e pesquisa no Brasil como para a organização de bancas examinadoras. Pode tanto oferecer o perfil de programas de pós-graduação como o perfil de determinada área de pesquisa. Numericamente, apresenta os docentes e pesquisadores mais produtivos e as relações entre temas mais significativas. Enfim, é um manancial inesgotável de fonte de informação sobre a pesquisa brasileira.



ζ A criação de conjuntos de dados com vistas à análise de informação foi possível graças a identificação de softwares livres responsáveis pela preparação (incluindo limpeza), extração, tratamento (incluindo manipulação), processamento (incluindo visualização), interpretação e análise de dados, processo conhecido como Descoberta de Conhecimento em Bases de Dados ou em inglês Knowledge Discovery in Databases (KDD).

ζ Optou-se por excluir a *long tail*, isso é, os inúmeros registros que: (i) por trazerem temas emergentes, (ii) erros de digitação ou (iii) falta de padronização na entrada de dados, geram uma relação significativamente extensa, porém impossível de figurar estatisticamente.

ζ Outro resultado importante foi a geração da lista de palavras-chave, 700.000 itens aproximadamente, capaz de motivar uma proposta de classificação da pesquisa brasileira, por meio da criação de grandes grupos de conhecimento que se destacam numericamente na análise e como eles se segmentam. Essa possibilidade enriquece o conhecimento da realidade da pesquisa brasileira, auxiliando na produção de indicadores da ciência e tecnologia do país.

ζ Como subproduto da pesquisa, foi possível despertar a consciência das inúmeras possibilidades de análise da BDTD com o conhecimento das ferramentas para realizá-la, a urgência da necessidade de organizar diretrizes e recursos para o correto preenchimento da BDTD e a certeza de que o aprendizado deve ser expandido para as demais bases do conhecimento científico e tecnológico.

- ζ Necessidade premente de **organizar diretrizes** para o correto preenchimento da BDTD e **recursos para facilitar o correto** preenchimento
- ζ Expansão da metodologia para **outras bases do conhecimento**

# Agradecimentos

---

- ζ Agradecimentos à **Fundação de Apoio à Pesquisa do Distrito Federal (FAP DF)** que possibilitou minha participação no EDICIC 2017 por meio da concessão de bolsa para participação de evento científico internacional procedente do Edital 01/2017 - *Seleção pública de propostas apoio à participação em eventos, cursos de curta duração e visitas técnicas de natureza científica, tecnológica e de inovação*

# REFERÊNCIAS

---

- ζ Alvares, L. (2017). *Descoberta de conhecimento na Biblioteca Digital Brasileira de Teses e Dissertações* (Relatório de Estágio Pós-Doutoral), Castellón de la Plana, Comunidade Valenciana, Espanha. Escola Superior de Tecnologia i Ciències Experimentals do Departament de Llenguatges i Sistemes InformàticsUniversitat Jaume I, UJI.
- ζ Anaya-Sánchez, H., Pons-Porrata, A. & Berlanga, R. (2011). A document clustering algorithm for discovering and describing topics. *Pattern Recognition Letters* 31(6): 502-510 (2010)
- ζ Han, J., Kamber, M., & Pei, J. (2011). *Data Mining: Concepts and Techniques*. Morgan Kaufmann Series in Data Management Systems.
- ζ Sancho, R. (1990). Indicadores bibliometricosutilizados en la evaluación de la ciência yla tecnologia: revision bibliográfica. *Revista Española de Documentación Científica*,13(3-4), 842-65.
- ζ Triska, R.,& Café, L. (2001). Arquivos abertos: subprojeto da Biblioteca Digital Brasileira. *Revista Ciência da Informação*, 30(3), 92-96.
- ζ Vanz, S. A. de S.,& Stumpf, I. R. C. (2010). Procedimentos e ferramentas aplicados aos estudos bibliométricos. *Informação & Sociedade: Estudos*, 20(2), 67-75.
- ζ Velho, L. (1986). A avaliação do desempenho científico. *Cadernos USP*, (1), 22-40.