Relatório do Projeto:

Árvore da produção acadêmica da Pós-Graduação da UNIFESP (até o ano de 2019)

Mauricio dos Santos Palazzuoli[1], Ângela Tavares Paes[2], Andréa Slemian[3]

Colaboradores: Alexsandro Cardoso Carvalho[4], Andreia do Carmo[5]

Maio/2020

Resumo

As Instituições de Ensino Superior (IES) públicas federais são centros de ensino e de pesquisa que tem como um dos seus papéis prioritários formar e congregar profissionais que gerem conhecimento em cada uma de suas áreas. Esse conhecimento gerado na forma de publicações em periódicos, livros e patentes movimentam a comunidade científica e pautam os caminhos das pesquisas e das suas principais áreas de concentração. No Brasil das últimas duas décadas ao menos, observou-se uma expansão do sistema universitário e da pesquisa produzida, da qual a UNIFESP, após seu processo de expansão iniciado em 2015, é um dos exemplos mais expressivos. O presente trabalho tem como objetivo mapear a produção acadêmica da UNIFESP por meio do levantamento e análise lexicométrica da produção (que concebemos aqui como "árvore"), indicando suas principais áreas de concentração e contato. Diante do volume ininteligível de dados e informações disponíveis, faz-se necessário o uso de ferramentas de Tecnologia da Informação e Comunicação (TIC) para extrair, transformar e analisar esses dados, também para gerar visualizações desse corpus textual expondo as relações entre as diversas origens do conhecimento. Desenhar esse mapa de maneira visualmente atraente, inteligível e significativa, para fins de divulgação dos produtos entregues à sociedade também foi um de nossos objetivos. Para tanto desenvolvemos uma metodologia que apresentaremos circunstanciadamente.

Palavras chave: lexicometria; cientometria; pós-graduação; TIC; IES

Introdução e objetivos

A identidade de uma universidade é essencialmente definida não somente pelo ensino, mas pela ciência que produz. Nesse sentido, a qualidade do corpo docente e sua produção científica têm papel primordial na construção da imagem de uma instituição de ensino superior (IES). Eles abrem caminhos para geração de recursos que ajudem a viabilizar novas pesquisas, promoção de eventos científicos, no estabelecimento de comunicação e parcerias entre instituições nacionais e internacionais, na construção de laboratórios, atividades de extensão, entre outras. Assim sendo, um mapeamento de sua produção pode ser valioso para estudantes, pesquisadores, agências de fomento e gerentes de pesquisa e desenvolvimento, indústria e mercado e provedores de dados.

[1] *Superintendência de Tecnologia da Informação da Unifesp, São Paulo, SP*.
[2] *Pró-Reitoria de Pós-Graduação e Pesquisa da Unifesp, São Paulo, SP*.
[3] *Departamento de História da Unifesp, Guarulhos, e Pró-Reitoria de Pós-Graduação e Pesquisa da Unifesp* .
[4] *Superintendência de Tecnologia da Informação da Unifesp, São Paulo, SP*.
[5] Coordenadoria de Rede de Bibliotecas da Unifesp (CRBU).

Sendo assim, é fundamental que uma universidade saiba o que produz. O conhecimento das pesquisas, bem como a relação entre seus temas e áreas, foi imensamente facilitado nos últimos anos por meio de ferramentas de Tecnologia da Informação e Comunicação (TIC) para visualização de dados. A importância que esta ação pode ter, tanto para que se possa pensar pontos de convergência entre campos e pesquisadores, potencialidades de pesquisa, bem como para difusão dos saberes produzidos pelas IES e, em grande, parte financiados por órgãos públicos, é imensa. Sobretudo para instituições como a UNIFESP que, além de possuir uma larga tradição de pesquisa, passou por um processo de vertiginoso crescimento após sua entrada no plano de expansão desde 2010, onde seis novos campi foram criados no Estado de São Paulo (Santos, Diadema, Guarulhos, Osasco, São José dos Campos e Zona Leste) nas mais distintas áreas. O aumento do número de docentes, departamentos e grupos de pesquisa, gerou dados em de escala ininteligível que, sem uso de programas e desenvolvimento de metodologias específicas, seria impossível captar sua produção científica como um todo.

É objetivo específico deste artigo apresentar uma experiência coletiva de construção de um caminho para visualização do que concebemos como um "mapa", ou seja, uma representação, da pesquisa feita na instituição. Nesse sentido, elaboramos uma metodologia para desenvolvimento da mesma que será apresentada a seguir. Tendo em vista as possibilidades de identificação das relações entre temas da produção bibliográfica dos docentes, também realizamos análises preliminares dos dados encontrados diante das especialidades de cada campus. Como estamos seguros que as gráficos e diagramas produzidos falam por si mesmos, não nos detemos nas suas múltiplas possibilidades de leituras e interconexões.

Como se trata de um primeiro resultado produzido por este trabalho conjunto, optamos por nos centrar nos dados extraídos a partir da produção dos docentes vinculados a Programas de Pós-Graduação das Unidades Universitárias da Unifesp (em número de oito). A ideia é que este trabalho, apresentado aqui com um caráter exploratório, possa ser ampliado para a Universidade como um todo o mais rápido possível.

Em linhas gerais, almejamos contribuir tanto para discussão sobre as formas atuais de gerenciamento de dados científicos, propondo formas inteligíveis e significativas de visualização e análise, como de divulgação dos resultados dos produtos científicos da universidade pública à sociedade.

Metodologia

Para a confecção dos mapas foram necessárias algumas escolhas. Primeiramente no tocante à origem dos dados. De imediato, a solução mais fácil seria partir das informações presentes em bases de dados disponíveis, como Web of Science e Scopus. No entanto, como é sabido, há áreas que estão muito pouco contempladas nestes indexadores que fazem uma seleção rigorosa de acordo com alguns padrões pré-estabelecidos. O que resultaria em deixar de fora grande parte dos dados de pesquisas feitas na Universidade. Como nossa preocupação é menos o fator de impacto, mas muito mais as potencialidades que podem vir a ser desenvolvidas na instituição, optamos por buscar os dados no Currículo Lattes.

Em segundo lugar, os critérios que adotamos para a extração de dados. Optamos por selecionar todos os docentes da instituição que estiveram credenciados em algum programa de pós-graduação até o ano de 2019. Nossa intenção foi mesmo a de representar um mapa que, como o das estrelas, permita que seja possível olhar o movimento do passado ao presente da pesquisa, buscando ter registro das principais áreas de concentração e sinergias que foram se acumulando aos longos dos anos, décadas. E, igualmente, suas principais lacunas. Nesse sentido, para nosso propósito, consideramos que perderíamos mais se deixássemos de lado estes dados já que os algoritmos dos programas operam de modo a não valorizar a visualização do que não teve continuidade. Para a extração das palavras, incluímos todos os títulos dos artigos científicos, livros e capítulos dos livros (incluindo em inglês, que foram traduzidos). Isso porque, a despeito de uma tendência atual para valorização da publicação de artigos científicos em lugar de livros, para algumas áreas os livros continuam sendo a maior referência, sendo não poucas vezes mais citados do que artigos.

Além disso, foi nossa opção dar um título a cada um dos agrupamentos (clusters) visualizado claramente nos Dendrogramas. Obviamente que tais títulos são tentativos, visando dar conta das áreas predominantes em cada um deles, servindo sobretudo para facilitar a apresentação de uma primeira aproximação analítica.

Processo

Utilizamos a ferramenta scriptLattes(1) para a extração de dados da plataforma Lattes do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). O scriptLattes é um software que permite a extração e compilação organizada dos itens que compõem o currículo Lattes do pesquisador. Para o presente trabalho, utilizamos parcialmente as capacidades da ferramenta, extraindo o conteúdo da produção bibliográfica, especificamente títulos de artigos, títulos de livros e títulos de capítulos de livros. Para realização de extrações através do scriptLattes é necessária uma lista contendo os códigos Lattes dos pesquisadores dos quais se deseja extrair o currículo na íntegra. A listagem foi concebida a partir da informação institucional de todos os orientadores e orientadoras que já foram em algum momento credenciad@s nos programas de pós-graduação da Unifesp e que possuíam código do currículo Lattes cadastrado no banco de dados institucional totalizando 2181 currículos. Finalizado o processo de execução do software, obtém-se, além de muitos outros arquivos padrão das análises do software, um arquivo com extensão XML contendo todos os dados de todos os currículos extraídos.

Dessa forma, as etapas do processo foram:

1. Extração em formato XML (Extensible Markup Language) dos currículos Lattes, na íntegra, dos 2181 docentes com situação de quaisquer tipos de credenciamento como orientadores na pós-graduação e pesquisa da Unifesp datada sem limitação de início até 31 de dezembro de 2019 utilizando a ferramenta scriptLattes(1). A extração dos dados foi realizada entre os dias 17 e 22 de janeiro de 2020;
2. Limpeza para reparação estrutural dos dados obtidos do Lattes em formato XML utilizando a ferramenta *Pentaho Data Integration* (PDI)(2);
3. Conversão do formato XML para formato de planilha utilizando a ferramenta EasyMorph(3);
4. Tradução automatizada via *Google Sheets* dos títulos dos artigos, livros e capítulos de livros a partir de línguas diversas para a língua portuguesa e realimentação da planilha principal com os títulos traduzidos;
5. Inserção, na planilha, das Unidades Universitárias (UU) às quais pertencem os pesquisadores de acordo com seu cadastro nos programas de pós-graduação;
6. Formatação dos dados para carga no software Iramuteq utilizando a própria planilha para concatenar os títulos com variáveis e códigos de formatação do software (4 asteriscos no início da linha + espaço + um asterisco seguido da palavra PESQ_+código Lattes do pesquisador + código de quebra de linha + título + código de quebra de linha);
7. Extração dos dados da produção bibliográfica separados por UU e um do total, 9 arquivos texto;

8. Conversão do texto concatenado para texto com quebras de linha realizada no editor de texto Notepad++;
9. Carga e subsequente análise lexicométrica do corpus textual, de cada um dos arquivos separadamente, no software IRaMuTeQ(4);
10. Produção das visualizações gráficas no software IRaMuTeQ.

Análises do software IRaMuTeQ

O IRaMuTeQ é um acrônimo de "Interface de R pour les Analyses Multidimensionnelles de Textes et de Questionnaires" (Interface do R para análises multidimensionais de textos e questionários). É um software livre distribuído sob a licença GNU GPL v2, ligado obrigatoriamente ao pacote estatístico R para análises de conteúdo, lexicometria e análise do discurso. Foi desenvolvido pelo "Laboratoire d'Études et de Recherches Appliquées en Sciences Sociales" (LERASS) da Universidade de Toulouse.

O Iramuteq permite realizar diversos tipos de análises, a mais importante é baseada no método de classificação de Max Reinert(5), porém o Iramuteq amplia as análises realizadas anteriormente pelo Alceste que é um software proprietário desenvolvido anteriormente por Reinert.

O software realiza mineração de dados em textos, permitindo a realização de várias análises quantitativas dos corpora linguísticos, são elas: estatísticas textuais clássicas (contagem de palavras); pesquisa de especificidades de grupos; classificação hierárquica descendente; análise de similitudes e nuvem de palavras.

As análises foram feitas separadamente por Unidades Universitárias e para a Unifesp como um todo.

A partir do banco de dados foram realizadas **Análise Fatorial de Correspondência (AFC) - Classificação pelo método de Reinert (gráfico cartesiano), Classificação Hierárquica Descendente – CHD (dendrograma) e Análise de Similitude (grafo)**.

Análise Fatorial de Correspondência (AFC) - Classificação pelo método de Reiner (gráfico cartesiano): Apresenta um gráfico das palavras, distribuídas em um plano cartesiano onde o tamanho da fonte é proporcional à força do elo entre a palavra e sua classe, a posição e a distância entre palavras e entre palavras e o ponto central do plano são influenciadas pela correlação entre palavras e entre palavra e sua classe.

Classificação Hierárquica Descendente – CHD (dendrograma): Apresenta as classes do gráfico anterior, AFC, em forma de dendrograma, permitindo a compreensão das relações entre as classes e o peso da classe dentro do corpus textual.

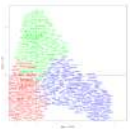
Análise de Similitude (grafo): Baseado na teoria dos grafos, apresenta as ligações entre palavras permitindo visualizar a estrutura de construção do corpus textual e a força entre palavras e temas.

Padrões de configuração das análises

Para todas as análises foram utilizadas as seguintes configurações, salvo descrito em contrário em cada análise.

Comum para todas as análises

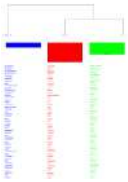
Dicionário da língua portuguesa padrão do Iramuteq contendo 142.598 palavras, conjunto de caracteres dos arquivos texto UTF-8, consideradas como formas ativas os adjetivos, advérbios, substantivos, verbos e formas não reconhecidas, lematização ativada, processo de produção de segmentos de texto por ocorrências (palavras), tamanho dos segmentos de texto de 100 caracteres.



Análise Fatorial de Correspondência (AFC) - Classificação pelo método de Reinert (Gráficos cartesianos)

Para a análise do corpus textual: Parâmetros padrão do Iramuteq.

Para o gráfico: Representação por coordenadas; Variáveis: formas ativas; Dimensões: 1000 x 1000; Evitar sobreposição: ativado; Texto proporcional à frequência: 9 a 40.



Classificação Hierárquica Descendente – CHD (Dendrogramas)

Dimensões do gráfico: altura 1100 x largura 800; tipo de dendrograma: filograma.

Análise de Similitude (Grafos)

Formas: selecionadas as 200 formas mais frequentes; Escore: co-ocorrências; Layout: fruchterman reingold (algoritmo); Árvore máxima: ativada; Limiar das bordas: específico para cada UU; Texto nos vértices: ativado; Comunidades: ativado - tipo: edge.betweenness; Halo: ativado; Dimensões do gráfico: 1000 x 1000; Tamanho do texto no vértice proporcional à frequência: efetivo 10 a 25; Largura das bordas proporcionais ao escore: ativado de 1 a 10; Tamanho do vértice fixo em 0 (zero).

Resultados

Foram realizadas análises lexicométricas sobre os títulos das produções incluindo artigos publicados em periódicos, livros e capítulos de livros constantes nos currículos Lattes de 2181 orientadores credenciados nos programas de pós-graduação stricto sensu da Unifesp, desde os primórdios da pós-graduação na instituição que remonta à década de 70 até 31 de dezembro de 2019. A distribuição dos pesquisadores por Unidade Universitária da Unifesp como encontra-se a seguir.

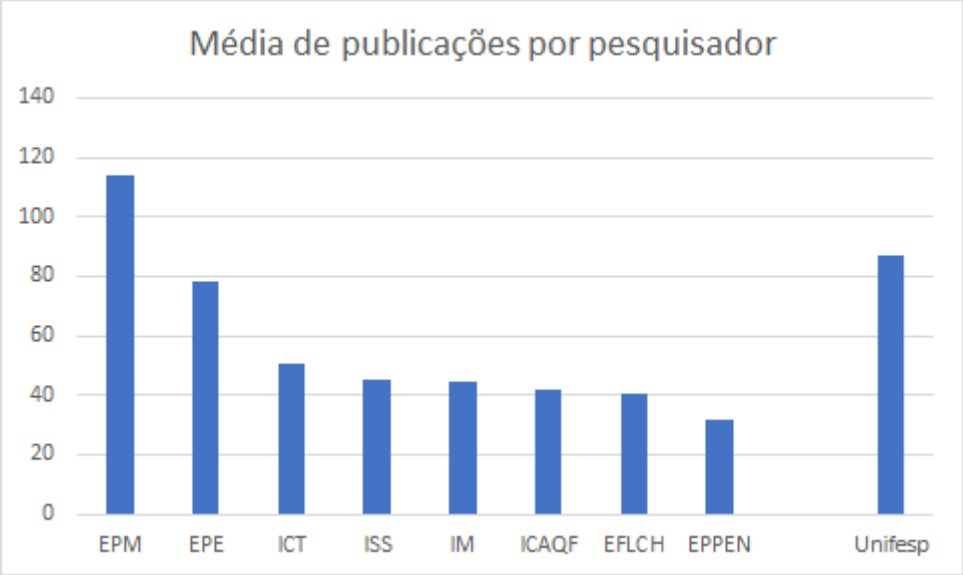
Tabela 1: Distribuição dos docentes segundo unidades universitárias da Unifesp.

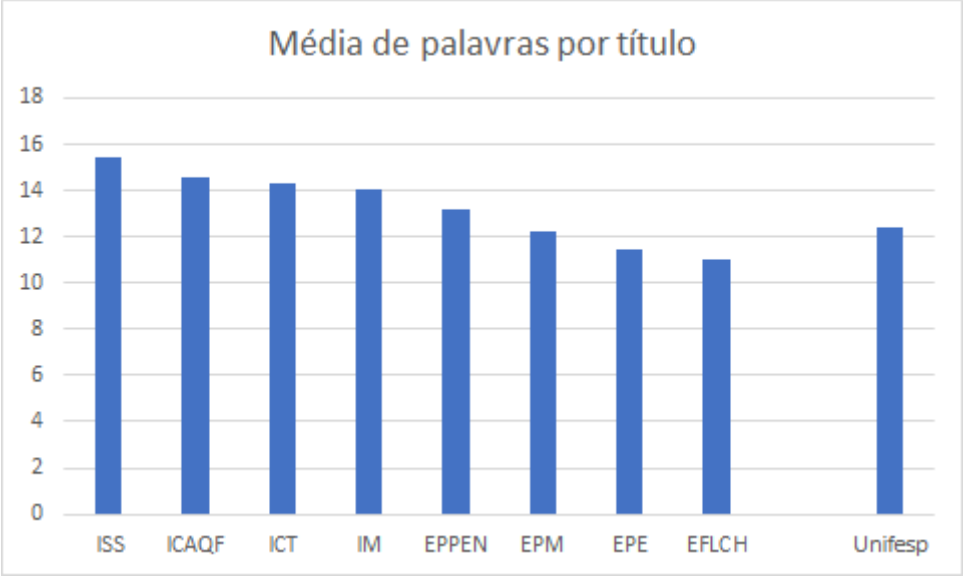
Unidade Universitária	**Sigla**	**Campus**	**Pesquisadores***	**%**
Escola de Filosofia, Letras e Ciências Humanas	EFLCH	Guarulhos	228	9,5
Escola Paulista de Enfermagem	EPE	São Paulo	164	6,9
Escola Paulista de Medicina	EPM	São Paulo	1357	56,8
Escola Paulista de Política, Economia e Negócios	EPPEN	Osasco	39	1,6
Instituto de Ciências Ambientais, Químicas e Farmacêuticas	ICAQF	Diadema	241	10,1
Instituto de Ciência e Tecnologia	ICT	São José dos Campos	179	7,5
Instituto do Mar	IM	Baixada Santista	20	0,8
Instituto de Saúde e Sociedade	ISS	Baixada Santista	161	6,7
Universidade Federal de São Paulo	Unifesp	Todos	2181 / 2389*	100%

* A soma dos pesquisadores das Unidades totaliza 2389, maior que o número da Instituição com um todo, isso se dá pela possibilidade de credenciamento do pesquisador em mais de um programa de pós-graduação distribuídos pelas Unidades Universitárias.

Tabela 2: Estatísticas do Corpus textual

	EFLCH	EPE	EPM	EPPEN	ICAQF	ICT	IM	ISS	Unifesp
Número de pesquisadores	228	164	1357	39	241	179	20	161	2181
Número de publicações não exclusivas (Corpus = total de títulos dos artigos, livros e capítulos de livros)	9263	12787	155156	1250	10051	9044	896	7300	189873
Média de publicações por pesquisador	**40,63**	**77,97**	**114,34**	**32,05**	**41,71**	**50,53**	**44,80**	**45,34**	**87,06**
Ocorrências de palavras (contagem geral)	101552	146854	1904312	16513	146515	129665	12587	112594	2350238
Média de palavras por título	**11,02**	**11,48**	**12,27**	**13,21**	**14,58**	**14,34**	**14,05**	**15,42**	**12,38**
Formas não lematizadas (palavras distintas)	12836	12787	67648	3398	19332	15564	3402	12926	85452
Formas lematizadas (palavras agrupadas pelo significante)	10465	10478	58429	2834	16381	12955	2906	10636	73716
Hápax (palavras que aparecem apenas uma vez)	5351	4910	25196	1438	8037	5974	1619	4948	32791
% hápax (nas formas lematizadas)	51,13	46,86	43,12	50,74	49,06	46,11	55,71	46,52	44,48
% hápax (no total de ocorrências)	**5,27**	**3,34**	**1,32**	**8,70**	**5,49**	**4,61**	**12,86**	**4,39**	**1,40**


Média de publicações por pesquisador
140
120
100
80
60
40
20
0
EPM
EPE
ICT
ISS
IM
ICAQF
EFLCH
EPPEN
Unifesp


Média de palavras por título
18
16
14
12
10
8
6
4
2
0
ISS
ICAQF
ICT
IM
EPPEN
EPM
EPE
EFLCH
Unifesp

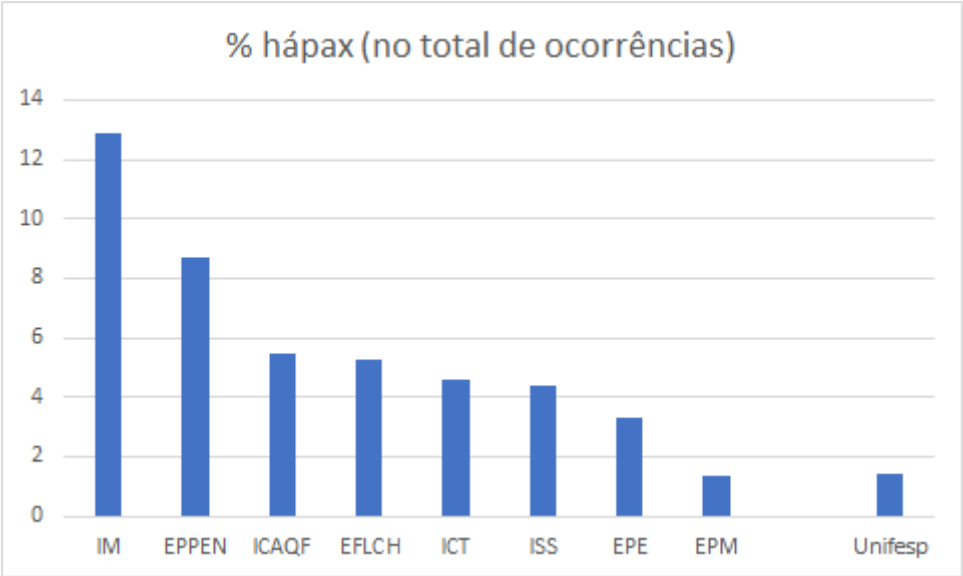
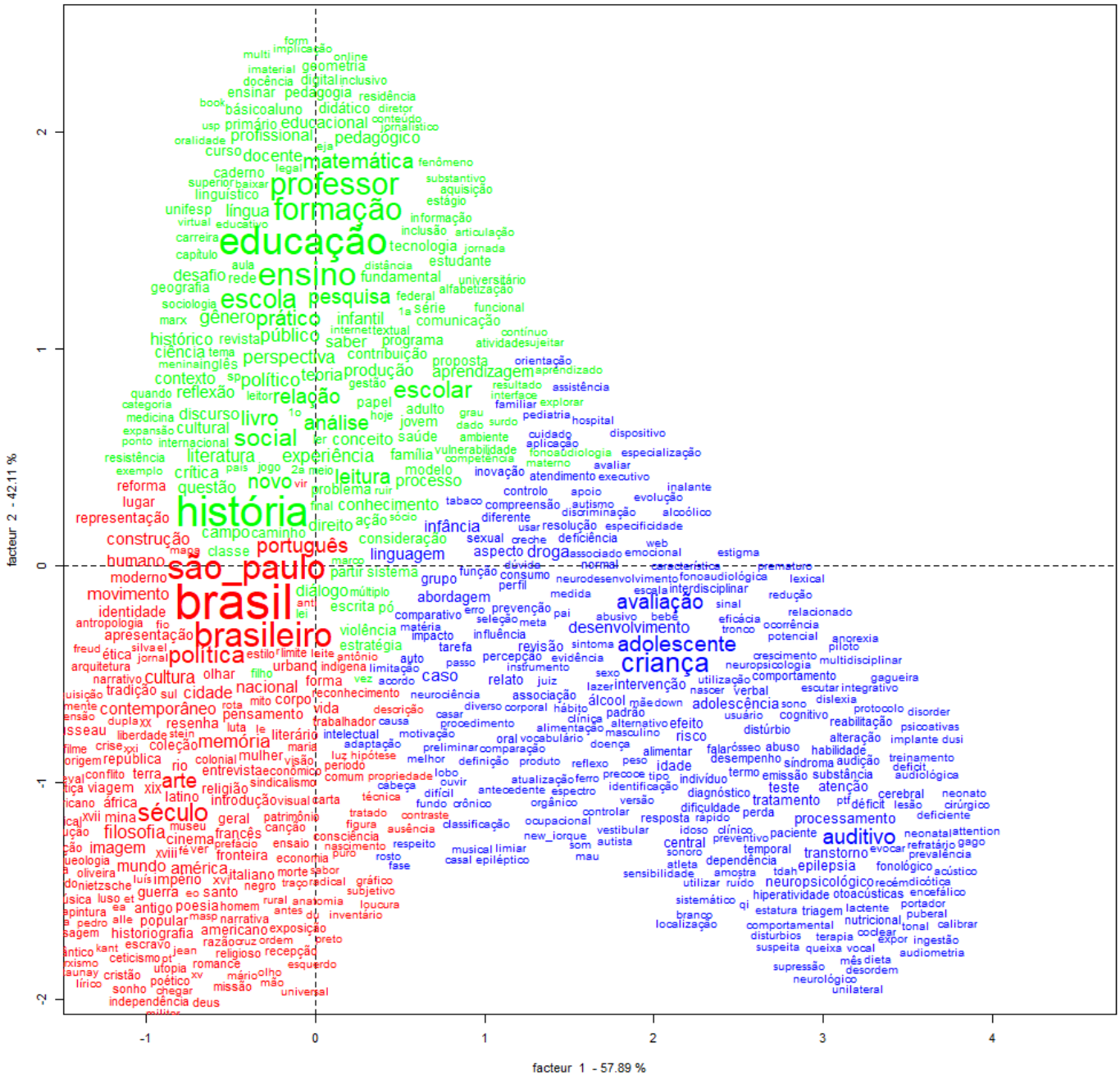

% hápax (no total de ocorrências)
14
12
10
8
6
4
2
0
IM
EPPEN
ICAQF
EFLCH
ICT
ISS
EPE
EPM
Unifesp

GRÁFICO 1 (Análise Fatorial de Correspondência)

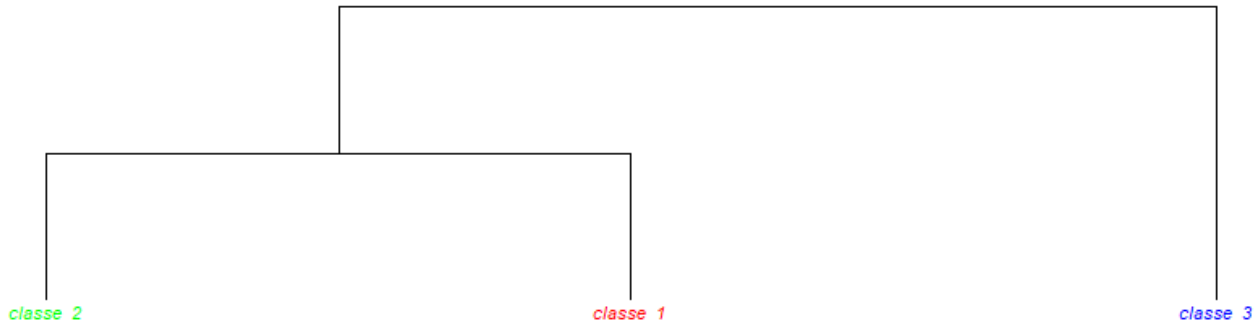
Escola de Filosofia, Letras e Ciências Humanas (EFLCH)

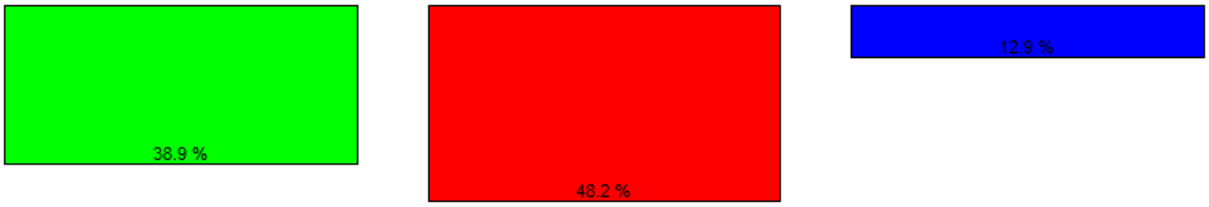


DENDROGRAMA 1 (Classificação Hierárquica Descendente)

Escola de Filosofia, Letras e Ciências Humanas (EFLCH)

Educação	Humanidades	Saúde/Educação





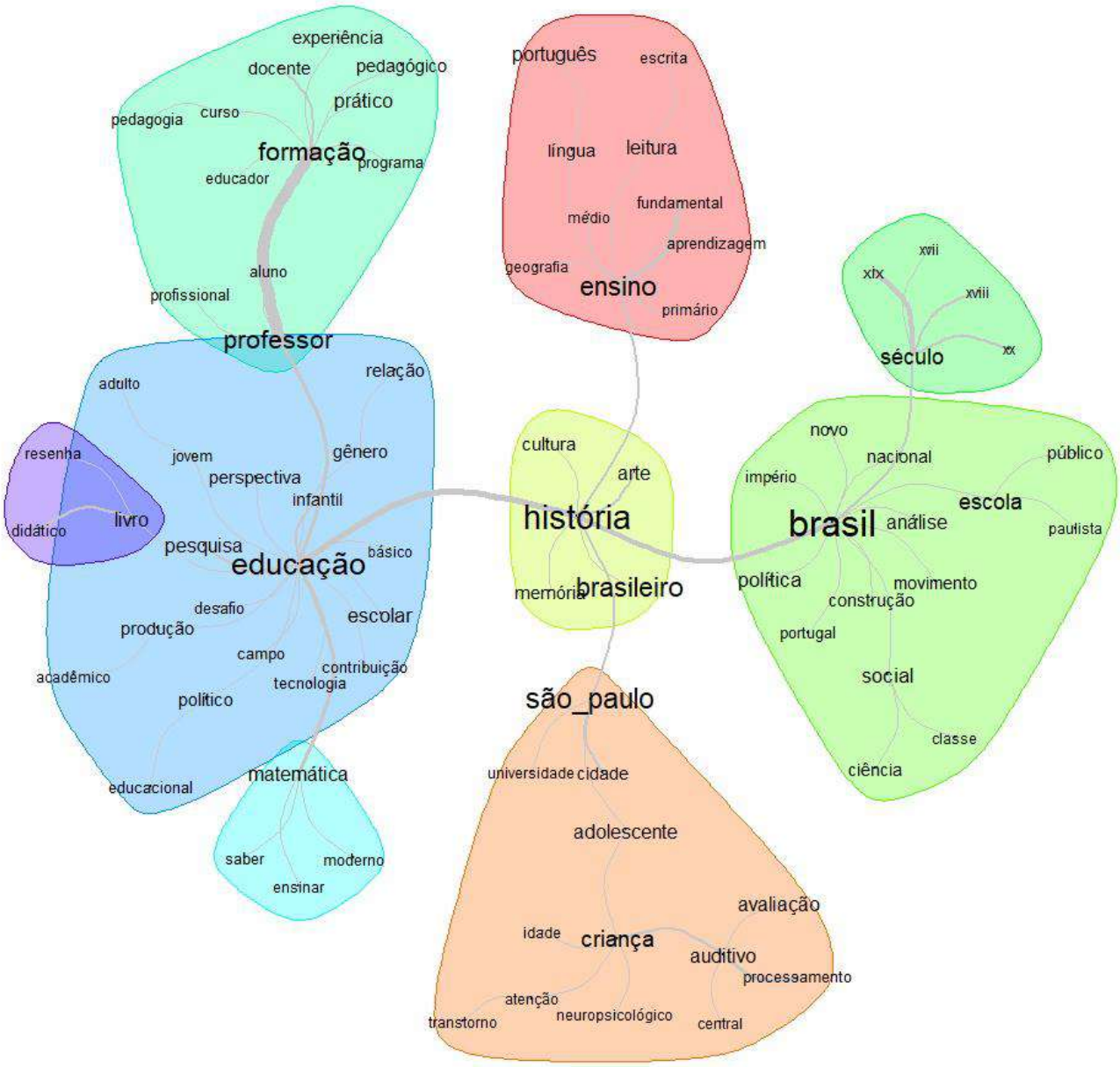
educação
professor
ensino
formação
escola
matemática
prático
docente
pesquisa
pedagógico
público
língua
curso
educacional
produção
gênero
saber
profissional
didático
infantil
escolar
desafio
tecnologia
aluno
fundamental
pedagogia
perspectiva
básico
ensinar
primário
digital
político
caderno
relação
educador
unifesp
social
teoria
geometria
linguístico
residência
livro
contexto

século
arte
filosofia
imagem
mundo
américa
império
guerra
xix
antigo
xviii
terra
memória
rio
poesia
cinema
áfrica
xvii
latino
viagem
política
popular
escravo
santo
brasil
historiografia
crise
nietzsche
colonial
americano
mina
contemporâneo
arqueologia
pintura
rousseau
ceticismo
bernardo
africano
tradução
música
resenha
paisagem
cidade

auditivo
criança
processamento
adolescente
avaliação
transtorno
epilepsia
teste
neuropsicológico
atenção
adolescência
tratamento
paciente
hiperatividade
cerebral
triagem
desenvolvimento
déficit
substância
síndroma
central
habilidade
emissão
audição
perda
abuso
efeito
otoacústicas
nutricional
alteração
distúrbio
idade
falar
desempenho
intervenção
temporal
indivíduo
evocar
comportamental
cognitivo
risco
fonológico
dependência

GRAFO 1 (Análise de Similitude)

Escola de Filosofia, Letras e Ciências Humanas (EFLCH)



Nos diagramas da Escola de Filosofia, Letras e Ciências Humanas (EFLCH) temos a formação de três agrupamentos. Um primeiro, que concentra o maior número de ocorrência, denominado por nós de **Humanidades** (48,2%), em que parece estar contemplada a produção dos Programas de Filosofia, Ciências Sociais, História, História da Arte, Letras, e em parte pelos docentes vinculados ao ProfHistória (Profissional). Seguido pelo grupo de **Educação** (38,9%), que contempla majoritariamente os docentes do Programa de Educação, com destaque para junção temática da produção do ProfHistória. E um terceiro, com menor ocorrência (12,9%), de **Saúde/Educação** que contempla a produção de parte dos docentes envolvidos com o Programa de Educação e Saúde da Criança e do Adolescente.

No tocante às áreas de maior concentração e intersecção, vemos:

- a partir do Dendrograma 1,

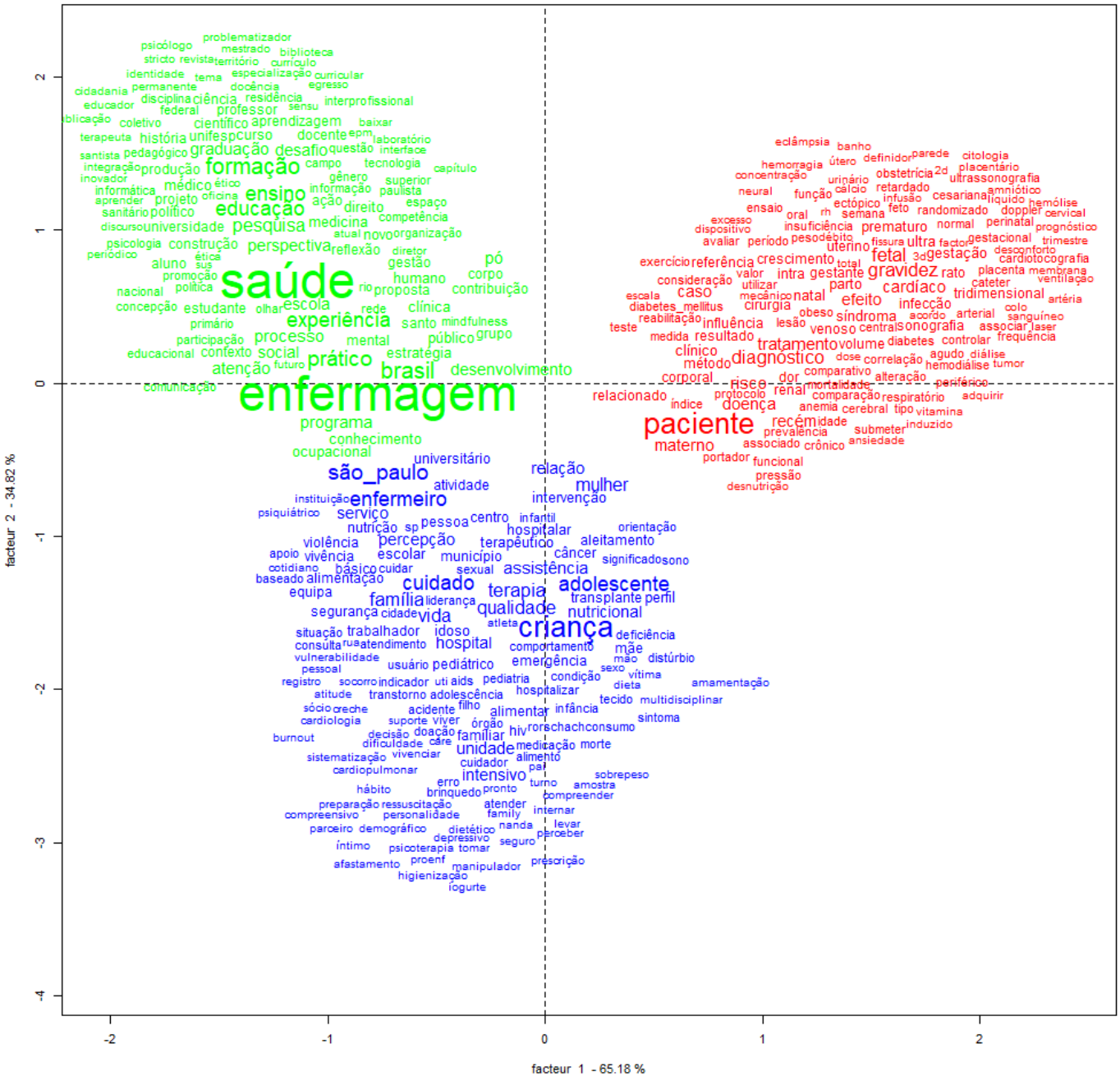
- o agrupamento denominado por nós de **Humanidades**, destacam-se as palavras "Brasil", "São Paulo", "brasileiro" e "política", que aproximam este grupo daquele de **Educação**, especialmente por meio da palavra "história". Neste sentido, vale destacar como esta palavra possui centralidade no diagrama de similitude (Grafo 1), articulando as áreas de história da educação, ensino, arte e os processos históricos propriamente, estando longe de congregar apenas a produção do Programa de História. Além disso, nos chama a atenção sua configuração mais equânime para o resto das ocorrências lexicométricas, que parecem indicar produções com uma tendência a uma menor articulação entre seus temas;

- no agrupamento de **Educação**, além de "história", as palavras mais citadas e que formam uma clara concentração são "educação", "professor", "ensino" e "formação" (seguidas por "escola" e "pesquisa"). O que indica igualmente uma tendência dos docentes envolvidos no Programa de Educação em pesquisas vinculadas à formação de professores. A aproximação com o grupo de **Saúde/Educação** se dá a partir da palavra "escolar", e mais precisamente por "estudante", "aprendizagem", "atendimento", "vulnerabilidade", "comunicação", indicando pesquisas vinculadas aos problemas de aprendizagem;

- no agrupamento **Saúde/Educação**, a palavra "criança" é a mais citada. Nele se visualiza uma concentração temática ao redor da palavra "auditivo", onde aparecem uma série de léxicos que descortinam pesquisas articuladoras das questões de aprendizagem e da saúde. Note-se que este conjunto possui muitos pontos de ligações com os outros dois, ainda que tímidos em ocorrências: do grupo de **Humanidades**, ao redor da palavra "caso", mas igualmente por "motivação", "juvenil", "descrição", "população", e mesmo "Freud"; e do grupo de **Educação** pela palavra "droga", seguida de "compreensão", "consumo", "deficiência", mas igualmente de "infância", "linguagem".

GRÁFICO 2 (Análise Fatorial de Correspondência)

Escola Paulista de Enfermagem (EPE)

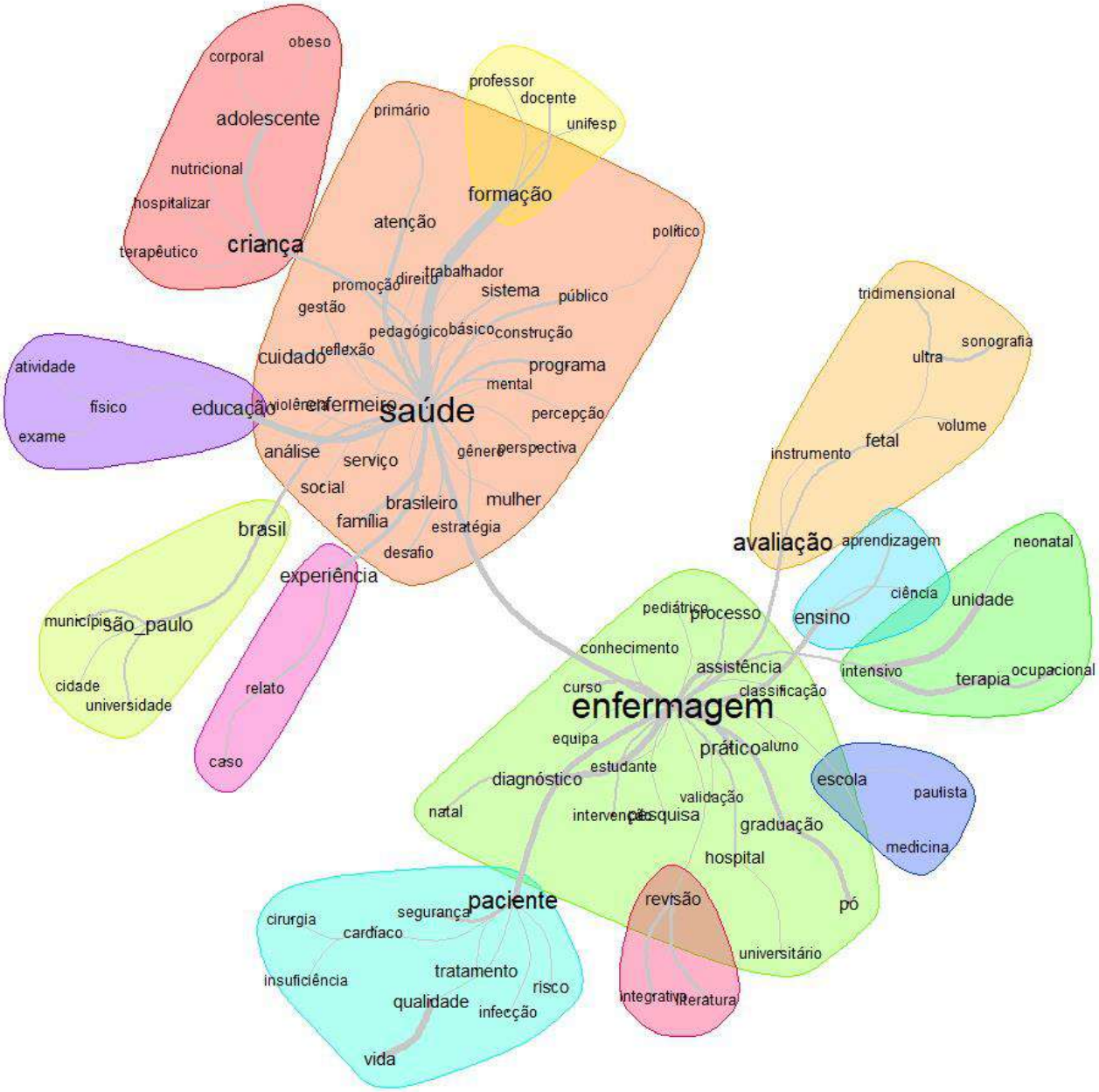


DENDROGRAMA 2 (Classificação Hierárquica Descendente)

Escola Paulista de Enfermagem (EPE)

GRAFO 2 (Análise de Similitude)

Escola Paulista de Enfermagem (EPE)



Na análise da produção da Escola Paulista de Enfermagem (EFE), o Dendrograma 2 indica a formação de 3 agrupamentos. O primeiro denominado aqui como **Enfermagem obstétrica e perinatal** envolve os cuidados com mulheres durante a gravidez e recém nascidos. A atenção à mulheres em fase gestacional aparece nas palavras "gravidez", "gestação/gestacional;gestante", "fetal/feto", "placenta", "crescimento", "semana", "intra" e "uterino". Estão presentes nesse agrupamento, exames comuns durante a gestação ("ultrassonografia", "tridimensional", "cardiotocografia"), aspectos do parto e dos neonatos ("prematuridade", "cesariana", "normal", "recém", "perinatal"). Nota-se também palavras ligadas a comorbidades ("cardíaco", "infecção", "insuficiência" e "diabetes"). A palavra mais frequente nesse grupo é "paciente" (ref mapa).

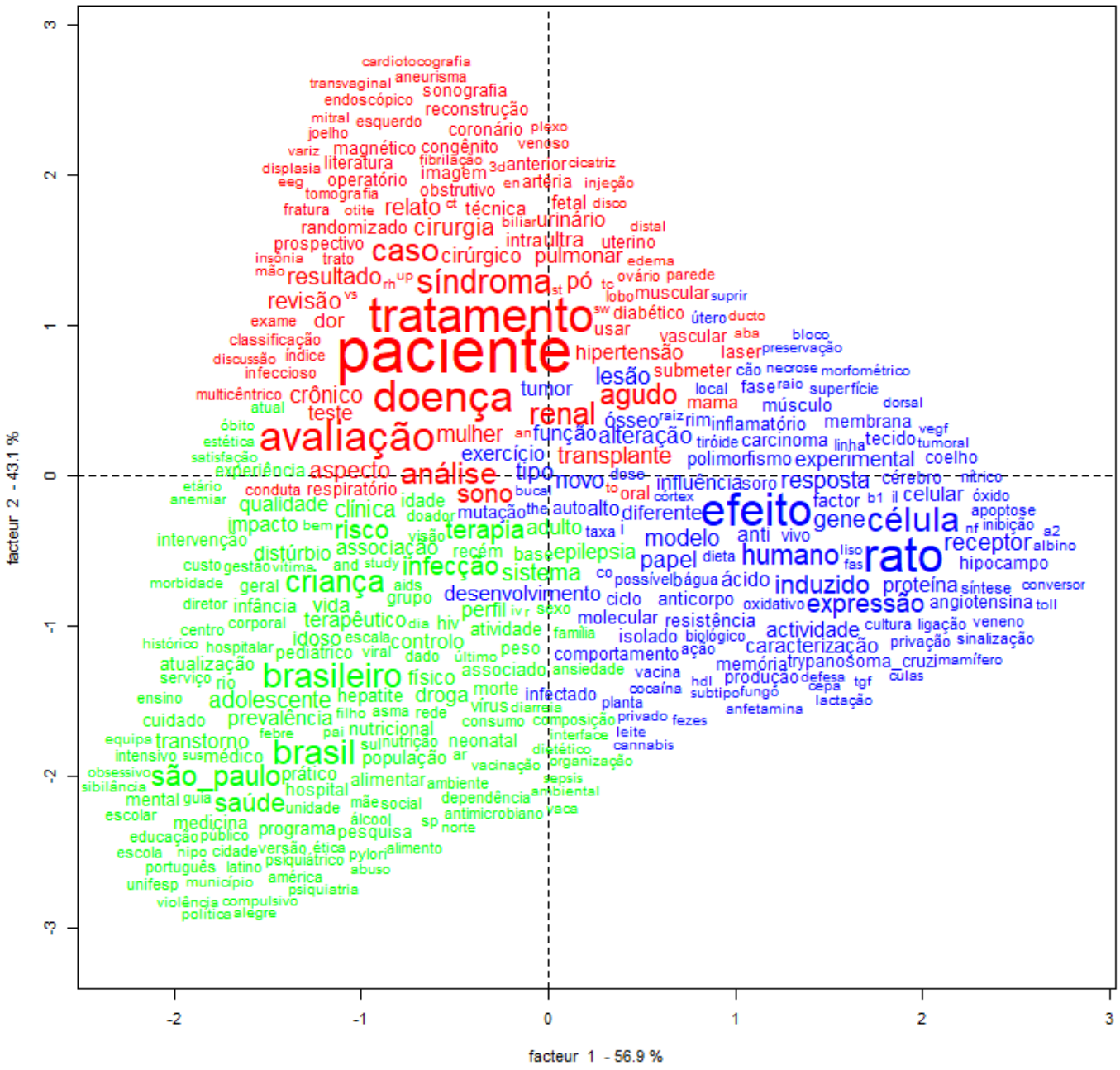
O segundo agrupamento **Educação em Saúde**, trata de aspectos relacionados à formação do profissional de saúde. A parte de ensino está clara diante das palavras "ensino", "educação", "graduação", "curso", "aprendizagem", "disciplina", "aluno", "docente", "professor", "experiência", "informação", "escola", "universidade"e "Unifesp". Também a pesquisa em educação ("pesquisa", "ciência", "científico", "projeto") e questões desafiadoras do ensino (por meio das palavras "desafio", "tecnologia", "novo", "competência", "perspectiva"). Como o esperado, a área de Enfermagem tem destaque como pode ser visualizado no Gráfico 2.

O terceiro agrupamento relaciona-se à **Assistência à saúde** de pacientes em diferentes fases da vida ("criança", "pediátrico/pediatria", "adolescente/adolescência", "idoso"), individual e coletivamente ("família/familiar", "filho", "cuidador"). Surgem nesse grupo palavras diretamente relacionadas à assistência ("intensivo", "cuidado", "terapia", "assistência", "qualidade de vida", "enfermeiro", "atendimento/atender", "percepção"). Foram identificadas algumas palavras ligadas à nutrição ("alimentar/alimento", "nutricional") e outras mais específicas como "HIV", "aids" e "brinquedo" (este provavelmente relacionado a brinquedos terapêuticos). Observando o mapa percebe-se que há maior frequência de trabalhos com crianças.

No Grafo 2, foram identificados claramente dois grandes núcleos: "saúde" e "enfermagem". Quanto ao primeiro, aparecem os temas de formação do docente da Unifesp, de nutrição/obesidade da criança e adolescente, de experiência e relato de caso; de educação/atividade física, além das regiões (Brasil, São Paulo, município, cidade, universidade). Quanto ao segundo, vinculado à Enfermagem, estão os temas de obstetrícia (avaliação, fetal, ultrassonografia, volume), revisão de literatura, paciente –segurança, risco, tratamento, unidade neonatal-terapia ocupacional – intensivo, e ensino-aprendizagem-ciência.

GRÁFICO 3 (Análise Fatorial de Correspondência)

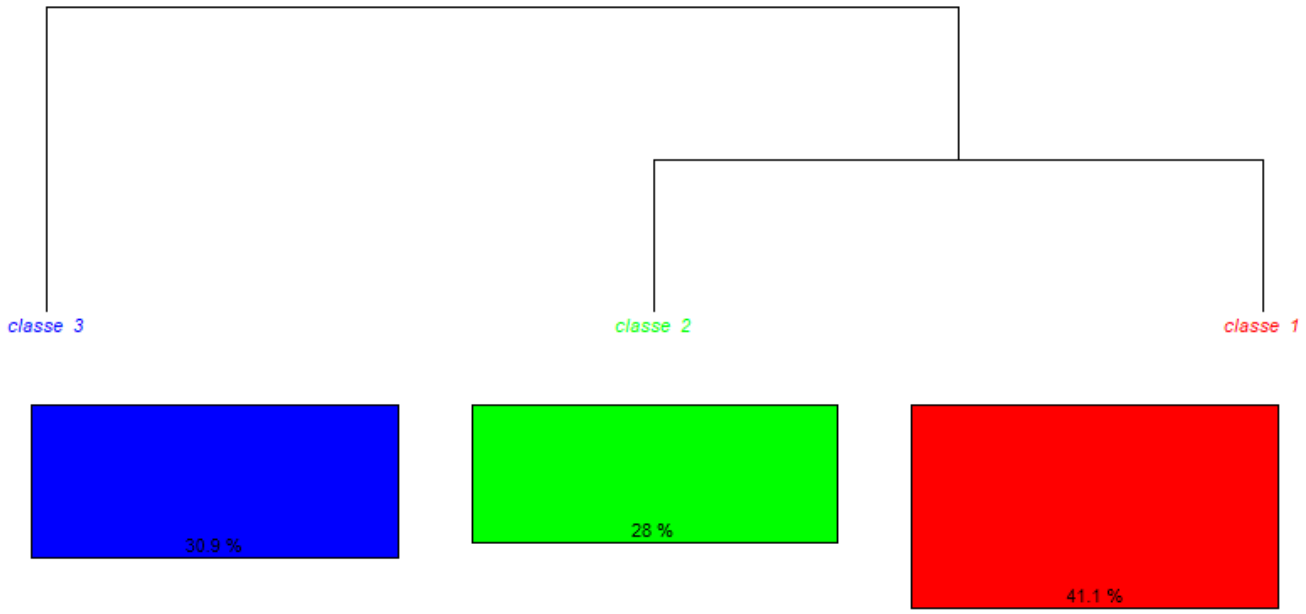
Escola Paulista de Medicina (EPM)



DENDROGRAMA 3 (Classificação Hierárquica Descendente)

Escola Paulista de Medicina (EPM)

Ciência básica	Impacto Social	Pesquisa clínica



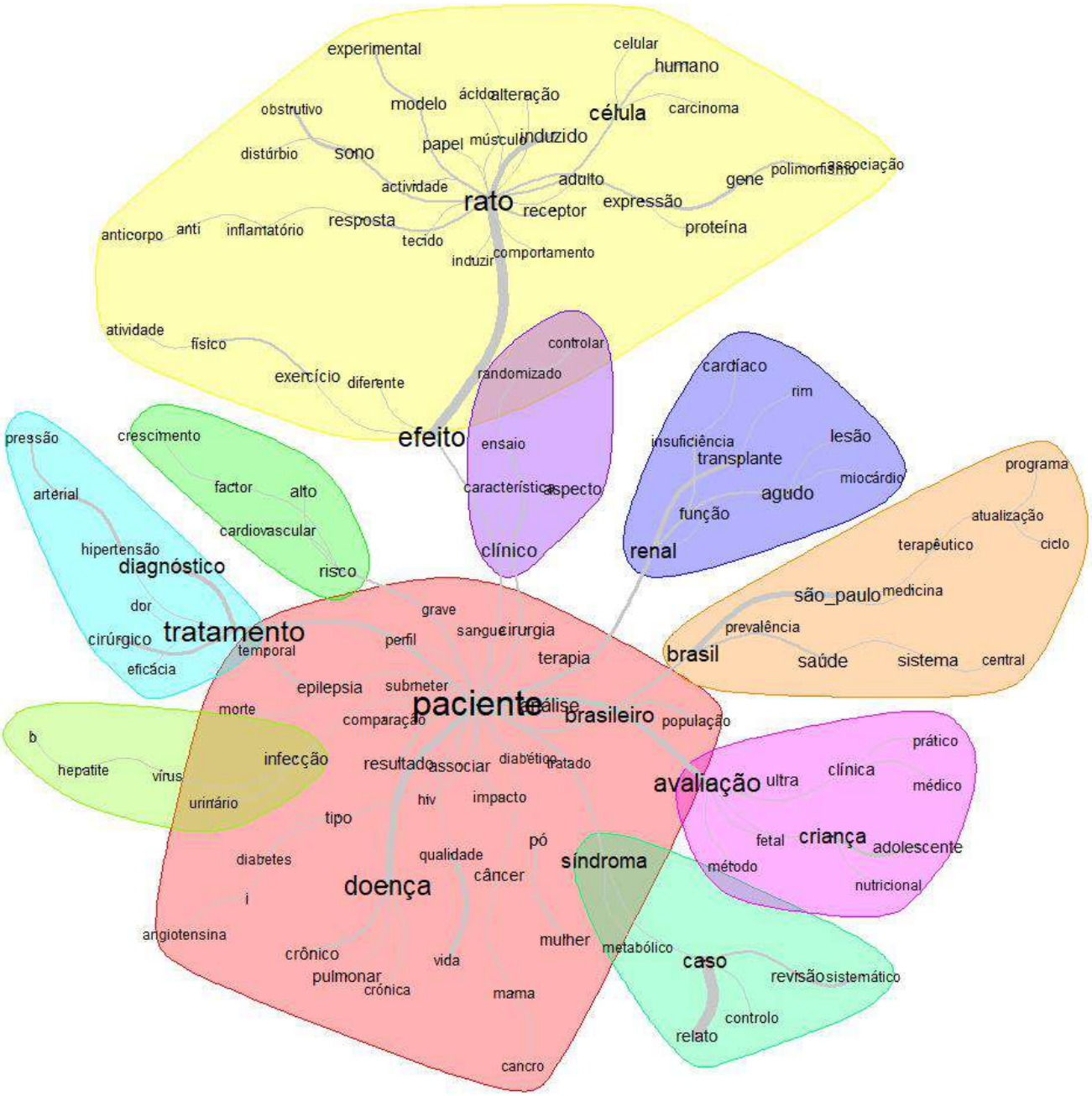
rato
célula
efeito
induzido
expressão
receptor
humano
proteína
gene
actividade
celular
modelo
experimental
angiotensina
resposta
trypanosoma_cruzi
caracterização
inibidor
induzir
ácido
enzima
papel
hipocampo
tecido
anticorpo
mecanismo
síntese
cálcio
inibição
escherichia_coli
activação
in_vitro
apoptose
sinalização
privação
anti
óxido
macrófago
mediar
oxidativo
coelho

brasil
são_paulo
saúde
brasileiro
médico
medicina
adolescente
criança
transtorno
prático
programa
prevalência
hospital
nutricional
atualização
pesquisa
público
cidade
intensivo
cuidado
educação
mental
escola
unidade
escolar
latino
população
hospitalar
alimentar
idoso
américa
vida
guia
rio
nutrição
droga
epidemiologia
emergência
atenção
manual

tratamento
caso
paciente
síndroma
relato
cirurgia
cirúrgico
diagnóstico
pó
revisão
ultra
obstrutivo
magnético
randomizado
clínico
artéria
doença
operatório
pulmonar
resultado
avaliação
sonografia
ressonância
tomografia
fetal
congênito
técnica
reconstrução
insuficiência
dor
imagem
intra
arterial
doppler
ocular
complicação
coronário
literatura
venoso
sistêmico

GRAFO 3 (Análise de Similitude)

Escola Paulista de Medicina (EPM)



Na análise da produção dos docentes da Escola Paulista de Medicina (EPM), foram identificados três grupos de palavras (Dendrograma 3). O maior deles (com 41,1% dos dados), encabeçado pelas palavras “tratamento”, “caso” e “paciente”, representa a **Pesquisa Clínica**, focada em diagnóstico de doenças e intervenções terapêuticas. Neste grupo, aparecem com frequência as palavras “síndrome”, “diagnóstico”, “relato”, “dor”, ”insuficiência”, ”crônico”, “resultado” e “avaliação”, além de nomes relacionados a exames como “ressonância magnética”, “tomografia computadorizada”, “ultra sonografia”, “imagem” e “doppler”, que mostram a descrição sobre sinais e sintomas que auxiliam no entendimento das doenças e do estado clínico de pacientes. Identifica-se também uma vertente para os tratamentos cirúrgicos, evidenciada pela força das palavras “cirurgia”, “cirúrgico”, “operatório”, “técnica”e “reconstrução”. Ainda neste grupo, o embasamento e produção do conhecimento médico por meio de estudos publicados fica claro nas palavras “revisão”, ”randomizado” e “literatura”. Com relação às áreas específicas da Medicina, há uma certa dificuldade em identificá-las, dado que tratamentos clínicos e cirúrgicos se aplicam a várias delas. No entanto, há indícios de produção em medicina fetal (palavras “fetal” e “congênita”), oftalmologia (palavra “ocular”), cardiologia (“coronário”, além de outras relacionadas), nefrologia (palavra “urinário”) e ortopedia (palavra “joelho”).

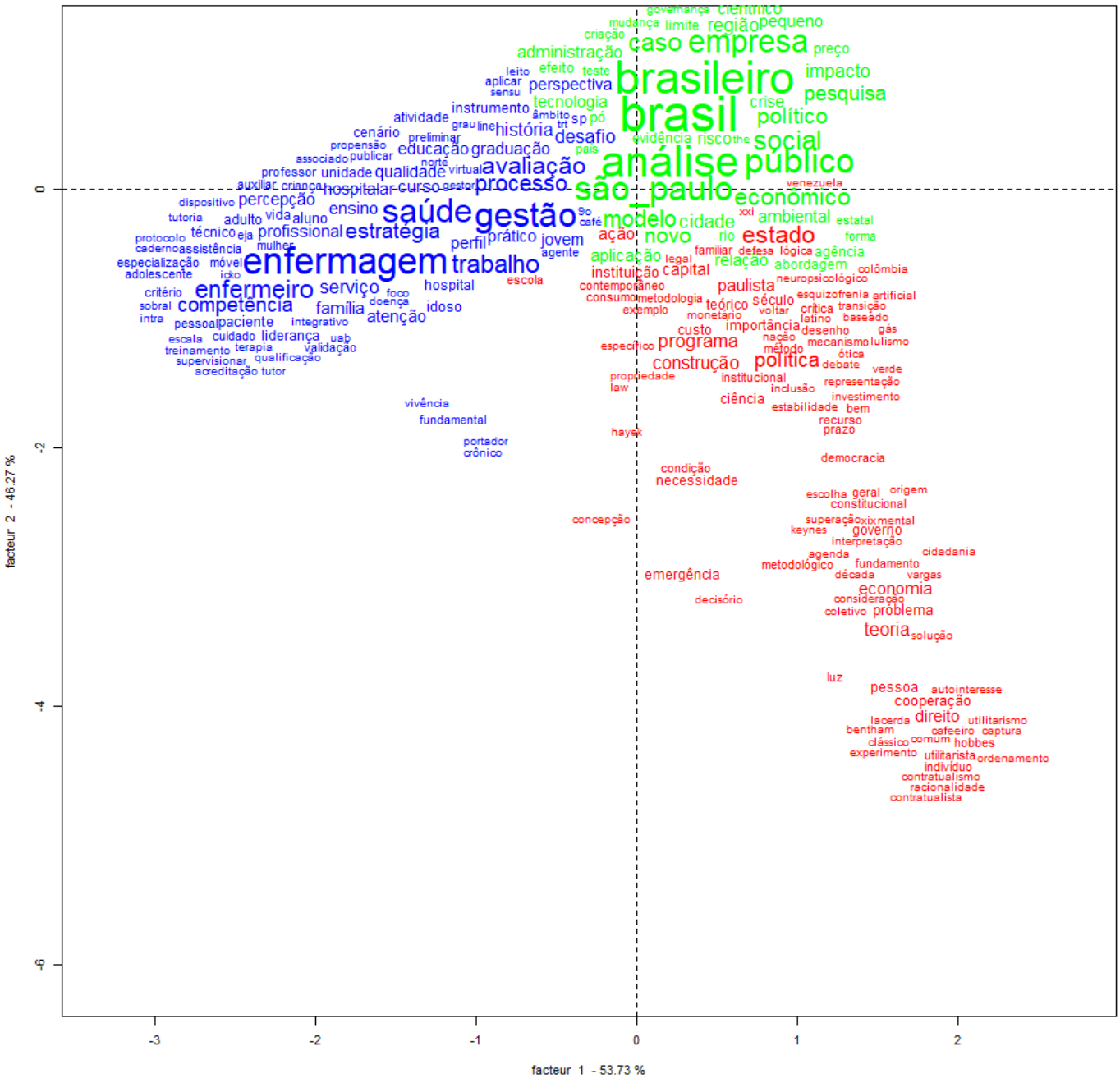
O segundo grupo de palavras mais expressivo (com 30,9% dos dados), corresponde à **Pesquisa Básica,** nitidamente identificado pelas palavras “rato”, célula”, “receptor”, “expressão” vindo de expressão gênica, “gene”, “proteína”, “atividade celular”, “apoptose”, “ácido”, “enzima”, “angiotensina”, “tecido”, “anticorpo”, “síntese”, “sinalização”, “macrófago”, “humano” e “*in vitro*”. É clara a presença de estudos experimentais em modelo animal, representados pelas palavras “induzido/induzir”, “inibidor”, “efeito” “modelo”, “experimental”, “rato” e “coelho” (sendo “rato” bem mais frequente que “coelho”). Os estudos de agentes etiológicos aparecerem com força pelas palavras “trypanossoma cruzi”, causador da doença de Chagas e “escherichia coli”, comumente responsável por doenças de trato urinário e digestivo. O interesse em fisiologia é característico nesta classe (palavras “efeito”, “atividade”, “resposta”, “caracterização”, “papel”, “mecanismo”, “mediar”, “ligação”, “modulação”). Nota-se no Gráfico 3 que esta classe se distancia um pouco mais dos outros dois grupos.

O agrupamento por nós denominado de **Impacto Social** (com 28% das palavras) está mais próximo da **Pesquisa clínica** do que da **Pesquisa Básica**, e de um modo geral está relacionada à pesquisa de impacto social e estudos de comportamento. É encabeçada pelas palavras “Brasil/brasileiro”, “São Paulo”, “saúde” e “médico/medicina”. Aparentemente é voltada para pessoas (“criança”, “adolescente/adolescência”, “público”, “população”, “idoso”), no ambiente em que se encontram (“hospital/hospitalar”, “escola/escolar”, “unidade”, “cidade”, “américa”, “latino”) com relação ao cuidado ou educação que recebem ou deveriam receber (“cuidado”, “atenção”, “educação”, “programa”, “ensino”, “serviço”, “manual”, “guia”, “prático” e “atualização”). Destaca-se o enfoque na saúde mental (“transtorno”, “mental”, “psiquiátrico”), nutrição (“nutricional”, “alimentar”, “nutrição”) e epidemiologia (“prevalência”, “pesquisa”, “epidemiologia/epidemiológico”).

No Gráfico 3, observa-se que as três classes conversam entre si e que algumas palavras ficam em posições intermediárias como, por exemplo, “desenvolvimento” e “sistema” (entre **Pesquisa Básica** e **Impacto Social**), “lesão”, “tumor”, “função” e “mama” (entre **Pesquisa Clínica e Básica**), “avaliação”, “qualidade” e “análise” (entre pesquisa clínica e social) e “terapia”, próxima das três classes. Na representação gráfica, é possível identificar algumas áreas específicas como saúde da mulher (“mulher”, “ovário”, “uterino”, “trasnvaginal”, “cardiotocografia”), doenças crônicas (“diabético”, “hipertensão”) e sono (“sono”, “insônia”, “sonografia”).

GRÁFICO 4 (Análise Fatorial de Correspondência)

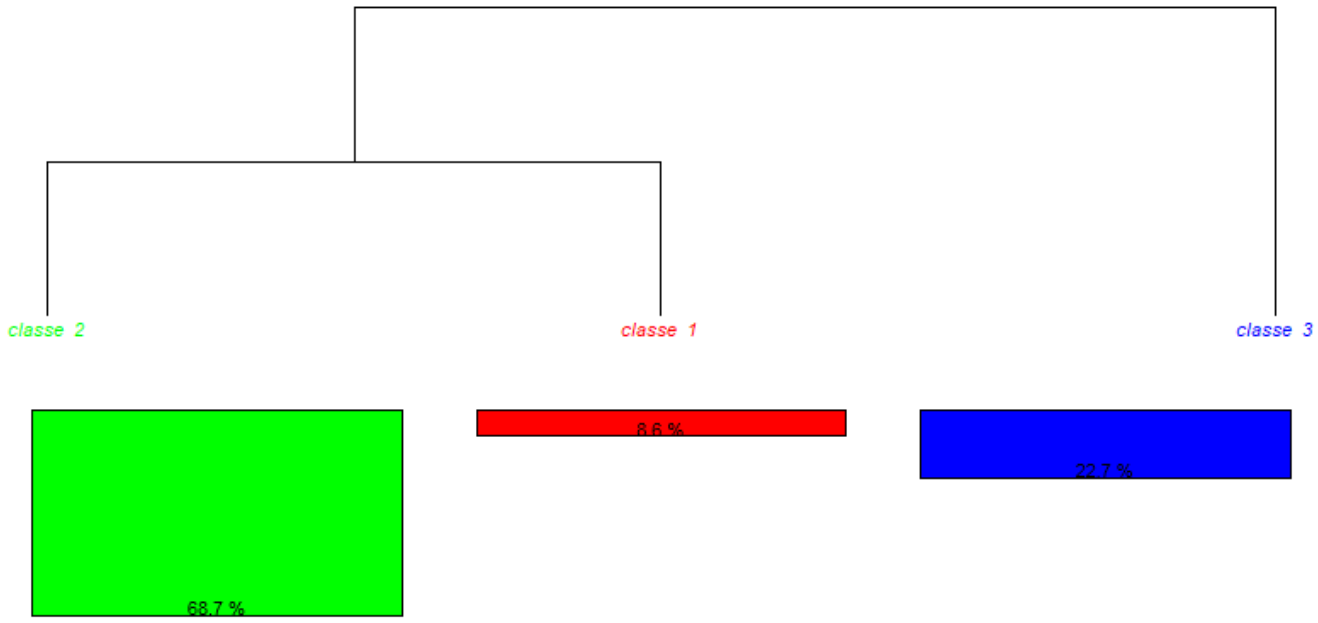
Escola Paulista de Política, Economia e Negócios, Osasco (EPPEN)



DENDROGRAMA 4 (Classificação Hierárquica Descendente)

Escola Paulista de Política, Economia e Negócios, Osasco (EPPEN)

Economia	Teoria econômica	Gestão/Organizações públicas



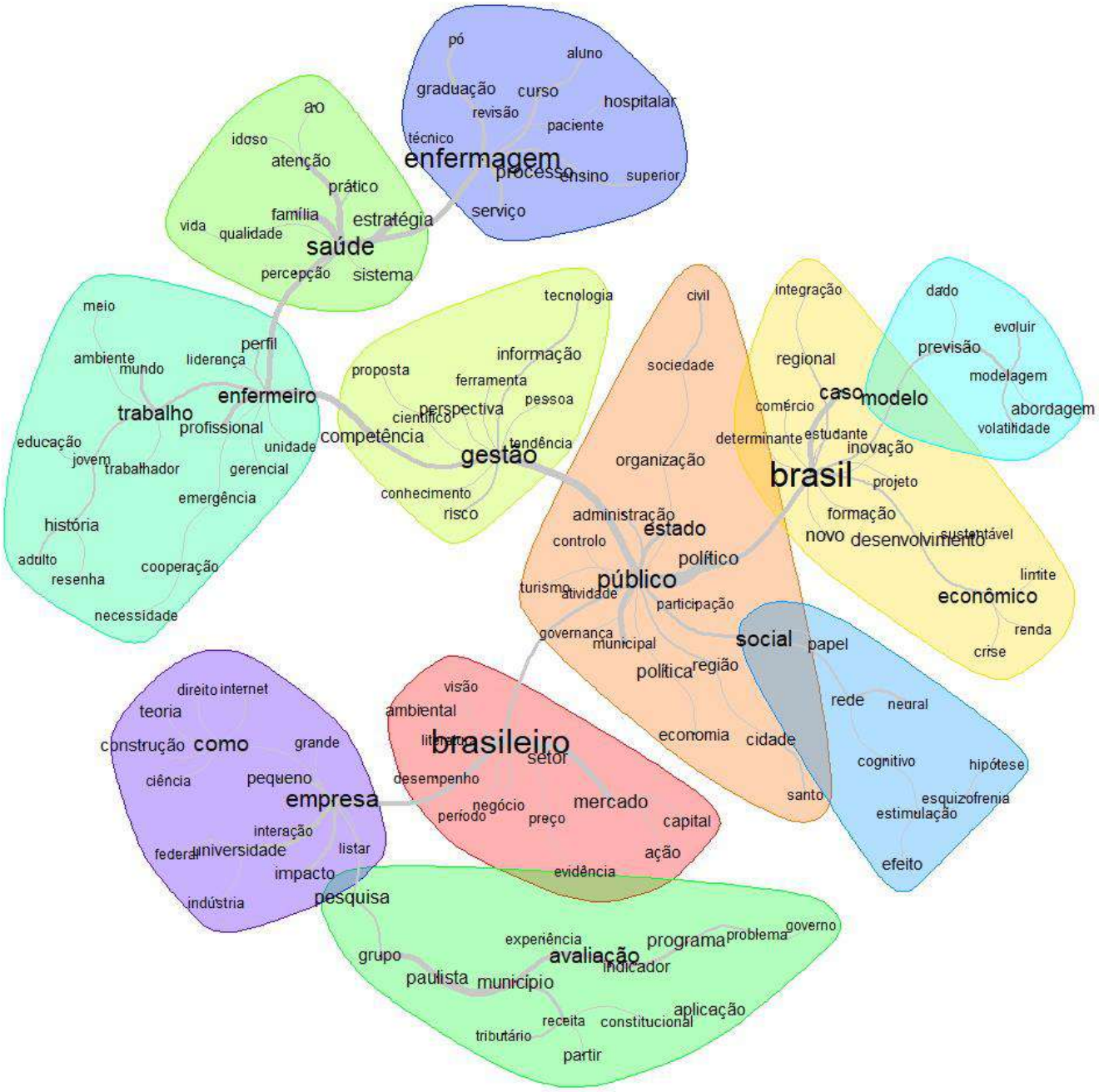
brasileiro
empresa
brasil
público
político
setor
impacto
região
social
informação
pequeno
caso
rede
pesquisa
científico
mercado
econômico
resenha
preço
controlo
são_paulo
regional
papel
grande
determinante
indústria
universidade
negócio
modelo
cognitivo
crise
inovação
limite
renda
estrutura
comércio
grupo
santo
integração
empírico
lei

teoria
direito
cooperação
economia
problema
pessoa
utilitarista
hobbes
indivíduo
comum
contratualismo
governo
emergência
luz
cafeeiro
utilitarismo
racionalidade
lacerda
contratualista
política
fundamento
solução
clássico
ordenamento
experimento
captura
bentham
autointeresse
constitucional
geral
metodológico
democracia
construção
vargas
década
consideração
coletivo
necessidade
programa
origem

enfermagem
enfermeiro
competência
saúde
profissional
serviço
percepção
paciente
estratégia
hospitalar
família
ensino
curso
adulto
técnico
perfil
atenção
especialização
adolescente
qualidade
aluno
eja
móvel
cuidado
assistência
unidade
liderança
prático
pessoal
critério
trabalho
idoso
vida
treinamento
supervisionar
sobral
qualificação
intra
integrativo
icko
dispositivo

GRAFO 4 (Análise de Similitude)

Escola Paulista de Política, Economia e Negócios, Osasco (EPPEN)



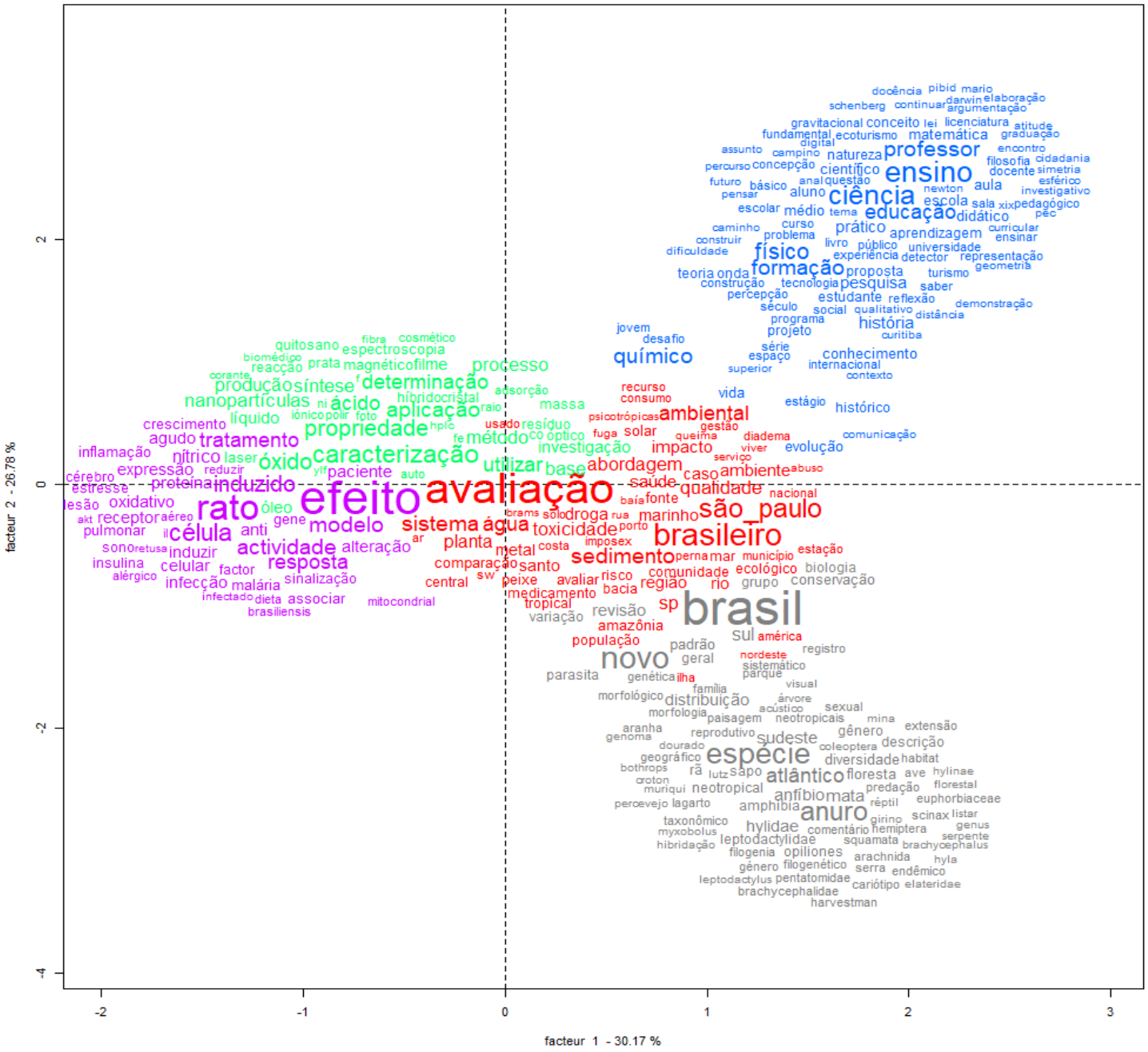
Os diagramas da EPPEN indicam que houve uma separação em 3 três agrupamentos que chamamos genericamente de **Economia, Teoria Econômica** e **Gestão/Políticas Públicas**. Pelo Dendrograma 4, podemos ver como a primeira concentra grande parte da produção do campus (68,7%), de onde o agrupamento de Teoria Econômica seria um sub-produto. Ambas identificam estar alinhadas ao Programa de Pós-Graduação de Economia e Desenvolvimento. Já a terceira classe, demonstra concentrar a produção vinculada ao Programa de Gestão de Políticas e Organizações Públicas.

No tocante às áreas de maior concentração e intersecção, vemos:

- uma concentração de pesquisas que mencionam as palavras “Brasil”, “brasileiro”, “análise”, “público”, “empresa”, “São Paulo”. O que permite descortinar, sobretudo a partir da representação do Grafo 4, dois conjuntos: um primeiro, articulado a partir da palavra “público”, e que parece ter maior centralidade, e vincular-se sobretudo com “Brasil”, “modelo”, econômico; outro não menos expressivo, vinculando palabras como “brasileiro”, “empresa”, “desempenho”, “mercado”;
- a palavra “gestão” demonstra claramente ser um ponto de conexão entre o agrupamento de **Economia** com o de **Gestão/Políticas Públicas**, bem como “avaliação”, “processo”, “desafio”, “perspectiva”;
- na concentração da classe de **Gestão**, há uma preponderância de trabalhos que mencionam “enfermagem”, “saúde”, “competência”, que parecem indicar áreas com maior ocorrência trabalhos de campo ou estudos de caso;
- no tocante, ao agrupamento que chamamos de **Teoria Econômica**, há uma maior diluição de temas expressados pelas palavras que indicam temas com abordagens teóricas, seja da história, de economia e de direito. Há um maior agrupamento em torno à “política”, e nota-se que seu ponto de maior conexão com o agrupamento de **Economia** se expressa por meio da palavra “Estado”.

GRÁFICO 5 (Análise Fatorial de Correspondência)

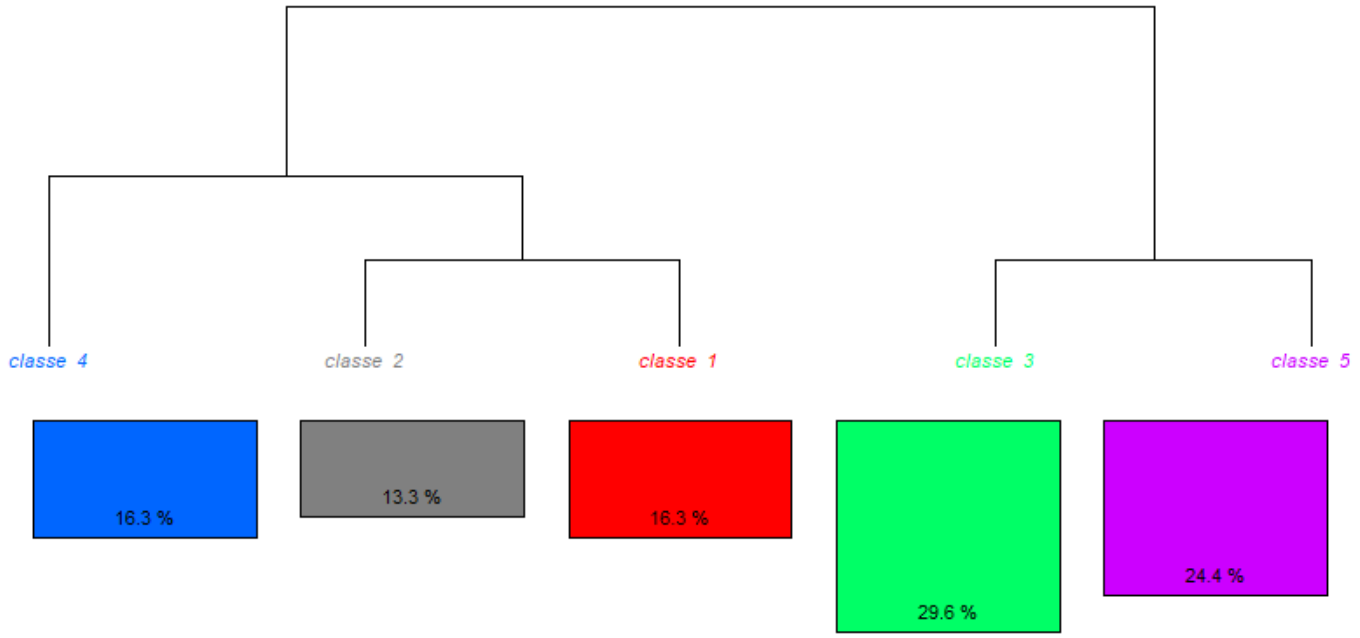
Instituto de Ciências Ambientais, Químicas e Farmacêuticas - ICAQF



DENDROGRAMA 5 (Classificação Hierárquica Descendente)

Instituto de Ciências Ambientais, Químicas e Farmacêuticas - ICAQF

Educação/Ensino	Biodiversidade/Ambiental	Biologia Química	Quimíca	Farrmácia



ensino
ciência
professor
educação
físico
formação
matemática
didático
aula
científico
prático
escola
história
médio
aprendizagem
pesquisa
conceito
docente
sala
pedagógico
aluno
natureza
gravitacional
curso
livro
fundamental
detector
proposta
licenciatura
público
filosofia
ensinar
conhecimento
estudante
reflexão
universidade
encontro
questão

anuro
espécie
atlântico
novo
mata
hylidae
anfíbio
brasil
amphibia
leptodactylidae
neotropical
sudeste
opiliones
diversidade
floresta
gênero
descrição
rã
filogenético
sapo
arachnida
squamata
filogenia
distribuição
género
pentatomidae
brachyopephalidae
hemiptera
scinax
reprodutivo
euphorbiaceae
taxonômico
comentário
harvestman
cariótipo
ave
serra
lagarto

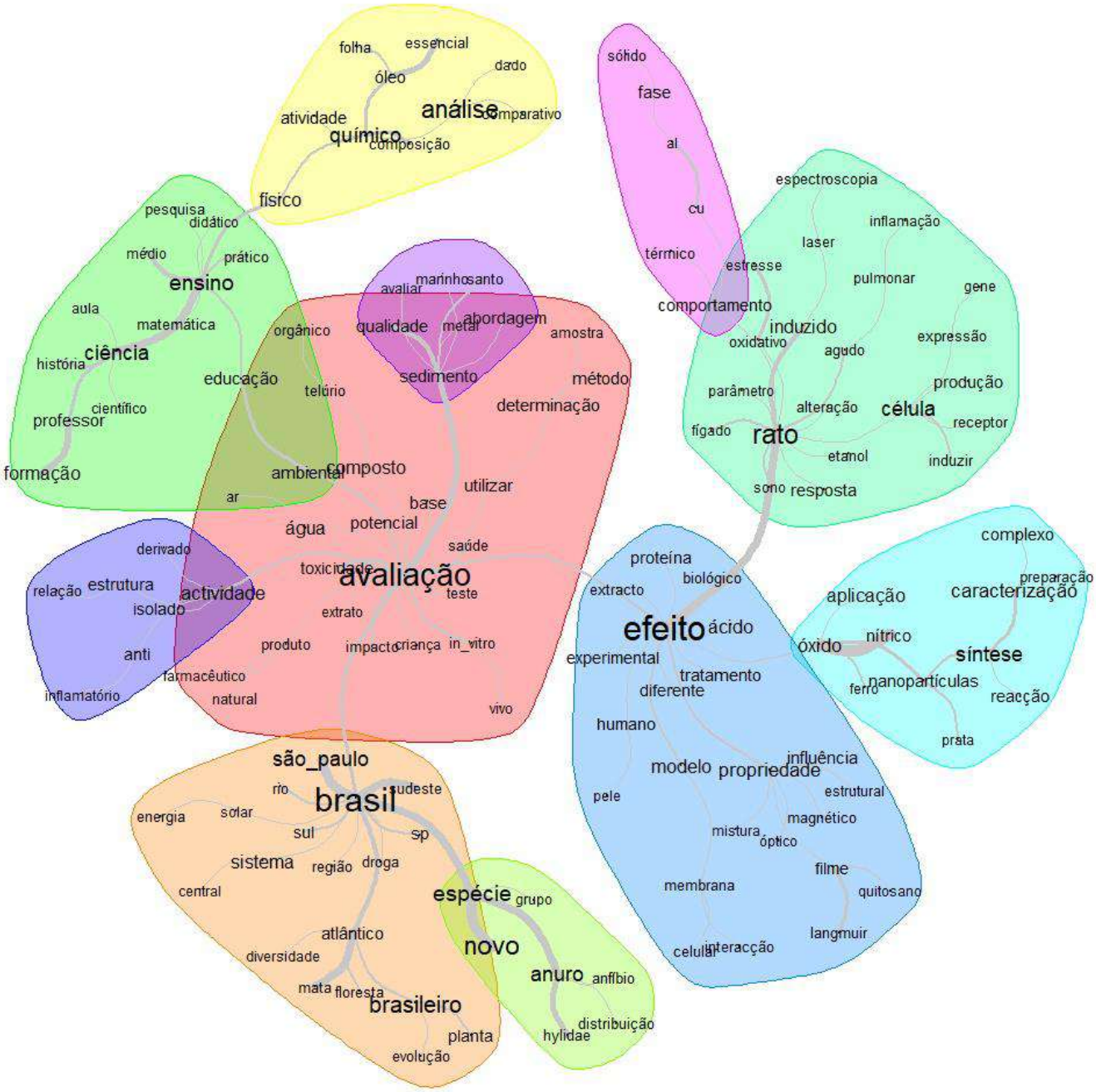
sedimento
qualidade
brasil
são_paulo
saúde
marinho
impacto
toxicidade
espanha
aerossol
avaliação
ambiental
costeiro
poluição
santo
mar
contaminação
estuário
integrar
medicamento
zona
costa
cidade
tropical
mexilhão
sp
metal
ouriço
rio
levantamento
litoral
estuarino
ecossistema
avaliar
porto
biomarcadores
risco
atmosférico

propriedade
nanopartículas
determinação
síntese
caracterização
aplicação
método
complexo
espectroscopia
filme
líquido
magnético
laser
preparação
conter
dopar
quitosano
prata
sólido
cristal
mistura
polianilina
raman
térmico
composto
processo
cerâmica
espectrometria
reacção
pele
formulação
iónico
fase
material
técnica
emulsão

rato
célula
induzido
expressão
oxidativo
induzir
receptor
inflamação
agudo
pulmonar
efeito
actividade
celular
inflamatório
malária
estresse
sono
isolado
resposta
infecção
crónica
gene
fígado
sinalização
macrófago
endotelial
stress
proteína
respiratório
insulina
leishmania
cérebro
hepático
lesão
trypanosoma_cruzi
nítrico
apoptose
anti

GRAFO 5 (Análise de Similitude)

Instituto de Ciências Ambientais, Químicas e Farmacêuticas - ICAQF



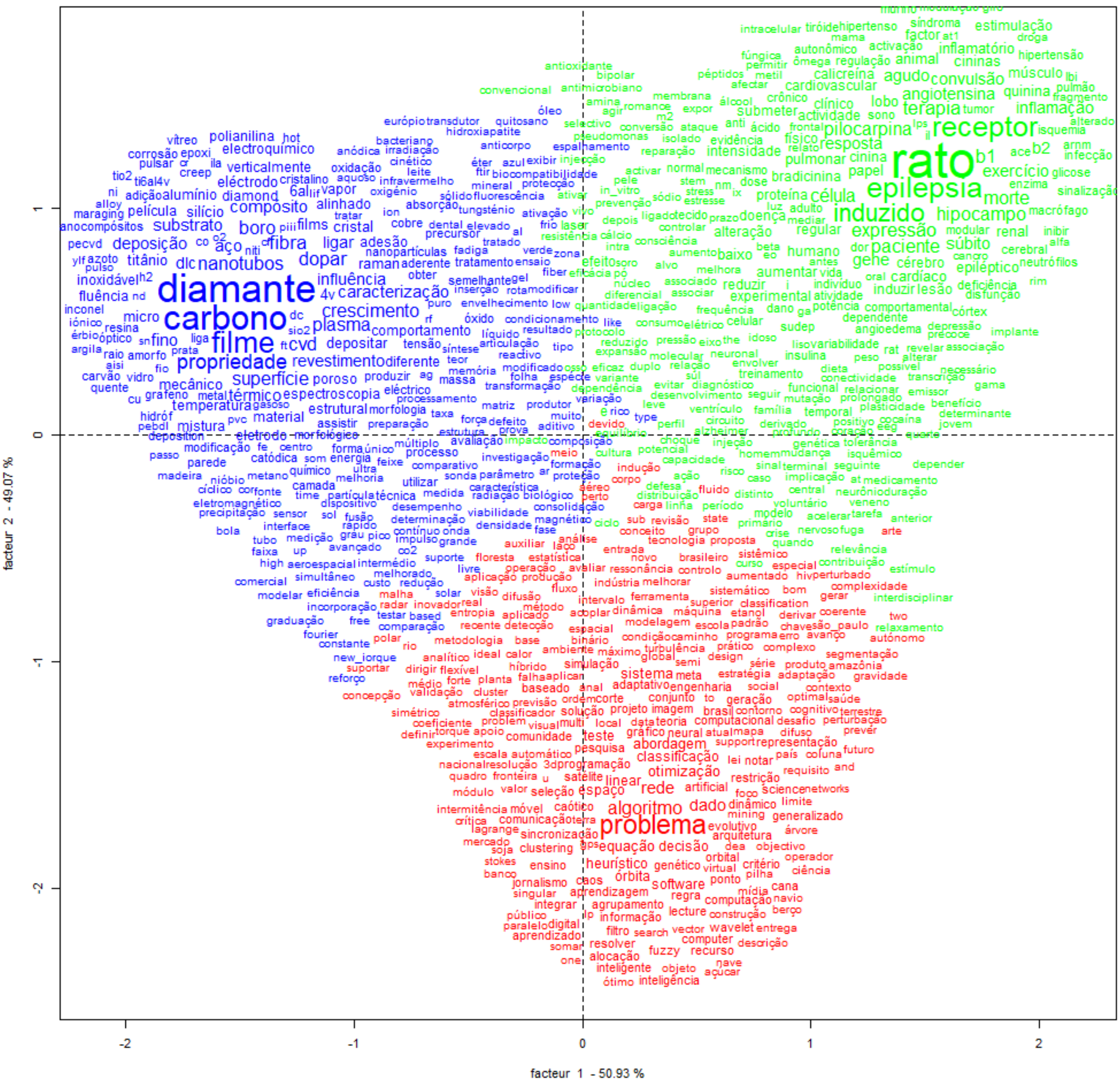
No dendrograma do Instituto de Ciências Ambientais, Químicas e Farmacêuticas (ICAQF), vemos a formação de 5 agrupamentos, cujos títulos por nós propostos já demonstram uma grande interseção entre cada um deles. O que concentra maior número de palavras é a classe que denominamos de **Química** (29,65), seguida bem de perto, e bastante articulada, com o grupo **Farmácia** (24,4%). O que se deve, majoritariamente, à produção vinculada aos Programas de Ciências Farmacêuticas e de Engenharia Química, representados aqui por palavras que nos remetem, sobretudo, à pesquisa básica. O campo que denominamos **Biologia Química** (16,3%) possui igualmente uma certa centralidade no Gráfico 5, tanto por se conectar com as ciências farmacêuticas e químicas, mas também pela sua conexão com o agrupamento de **Biodiversidade/Ambiental**; o que se deve também pela produção dos Programas de Biologia Química como de Química - Ciência e Tecnologia da Sustentabilidade. E a produção do grupo de **Biodiversidade/Ambiental** (13,3%) demonstra representar os Programas de Análise Ambiental Integrada e de Ecologia e Evolução. O agrupamento que denominamos **Educação/Ensino** (16,3%) parece representar uma parte significativa de pesquisas que se vinculam, preponderantemente, aos Programas de Ensino de Ciências e Matemática e de Matemática em Rede Nacional (Profmat-DM).

No tocante às áreas de maior concentração e intersecção, vemos:

- como a frequência das palavras “efeito” e “avaliação”, em agrupamentos diferentes, têm centralidade no Gráfico 5; a primeira agregando palavras como “rato”, “célula”, e a segunda, “água”, “planta”, “droga”, “toxicidade”, entre outras. O que pode ser explicado pelo papel com que as áreas de farmácia e química cumprem na experimentação de elementos e substâncias, não apenas para si, mas igualmente para outras pesquisas de biologia e de biodiversidade. Neste sentido, é notável como no Grafo 5, a palavra “efeito” aparece como mais conectada às pesquisas básicas, enquanto “avaliação” das área de biologia, ambiental e biodiversidade;
- há que se notar que as palavras “Brasil”, “brasileiro”, “São Paulo” aparecem integrando o agrupamento de **Biologia Química** com o da **Biodiversidade/Ambiental,** com maior frequência; em menor frequência, mas não menos significativas, estão “amazônia”, região”, ecológico”, todas representando o foco regional, nacional das pesquisas. Palavras como “anuro”, atlântico” e “espécie” como um ponto de concentração neste conjunto de Biodiversidade reforçam esta tendência;
- o agrupamento por nós intitulado de **Educação/Ensino** apresenta uma área de concentração nas palavras “ciência”, “ensino”, “educação” e “professor”, bem como “físico” e “formação”. Este agrupamento aparece, nesta escala, menos conectado com as outras áreas; no entanto, seus pontos de aproximação das áreas de biologia química, se dá pelas palavras “química”, “vida”, “estágio” e “ambiental” o que pode ser indicativo de temas para possíveis sinergias.

GRÁFICO 6 (Análise Fatorial de Correspondência)

Instituto de Ciência e Tecnologia - ICT



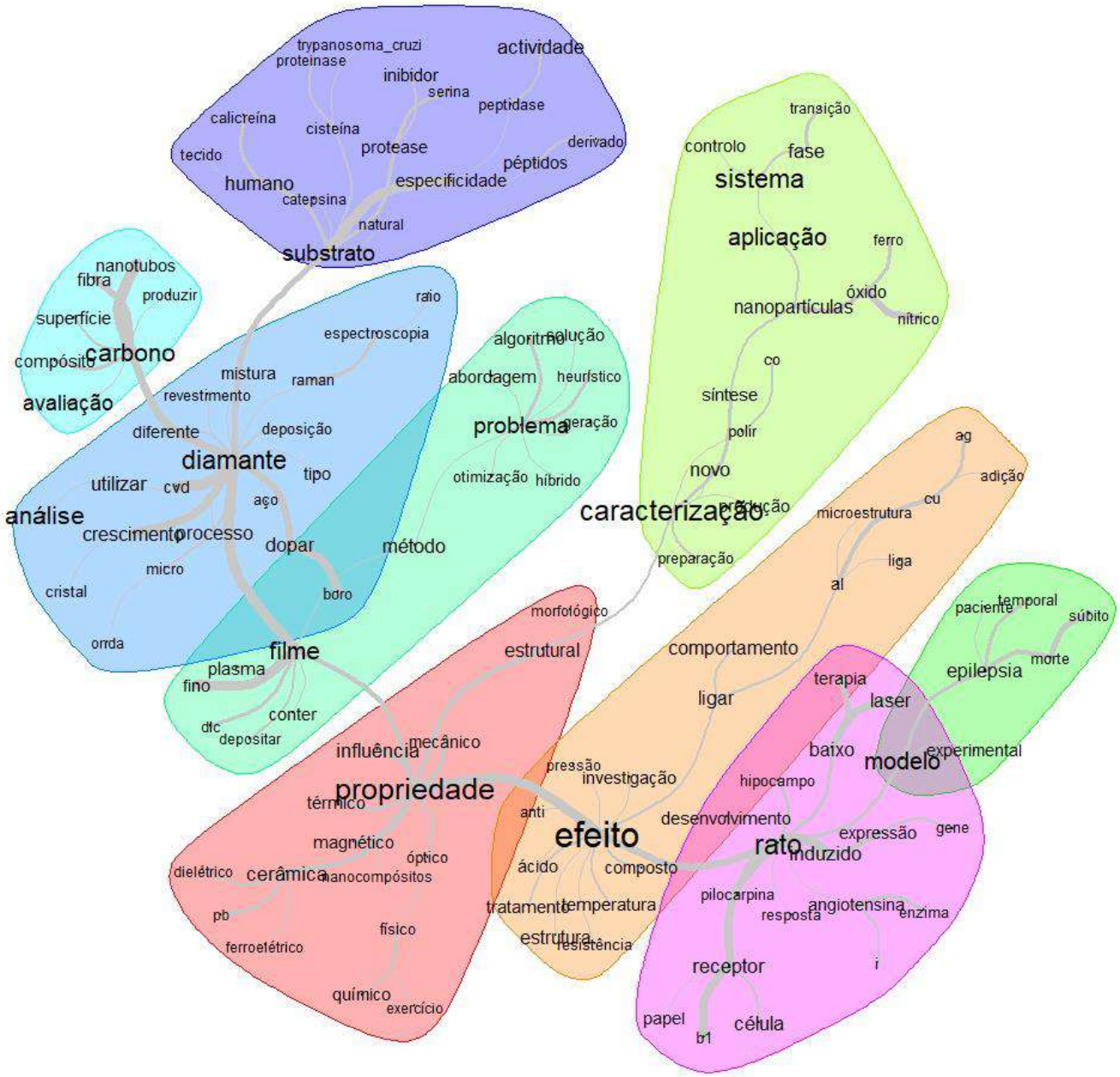
DENDROGRAMA 6 (Classificação Hierárquica Descendente)

Instituto de Ciência e Tecnologia - ICT

Bioengenharia/Interdisciplinar	Engenharia e ciência de materiais	Computação/ciência de dados/matemática

GRAFO 6 (Análise de Similitude)

Instituto de Ciência e Tecnologia - ICT



Na análise dos dados de São José dos Campos, foram identificados três agrupamentos com relativa proximidade entre si e diretamente relacionadas aos programas de pós graduação deste campus. Percebe-se claramente uma classe de palavras ligadas à área computacional, que denominamos de **Computação/ciência de dados/matemática,** na qual as palavras "problema" e "algoritmo" surgem em destaque, tanto no dendrograma quanto no mapa. O foco em programação fica claro em palavras como "sistema". Outras palavras que aparecem nessa classe remetem a tópicos de pesquisa em alta como mineração de dados e inteligência artificial, que engloba, por exemplo, aprendizagem de máquinas, algoritmos de classificação e redes neurais. Esta classe, de uma certa forma, reflete a produção dos docentes dos programas que trabalham com ciência de dados como Ciência da Computação, Pesquisa Operacional e Matemática.

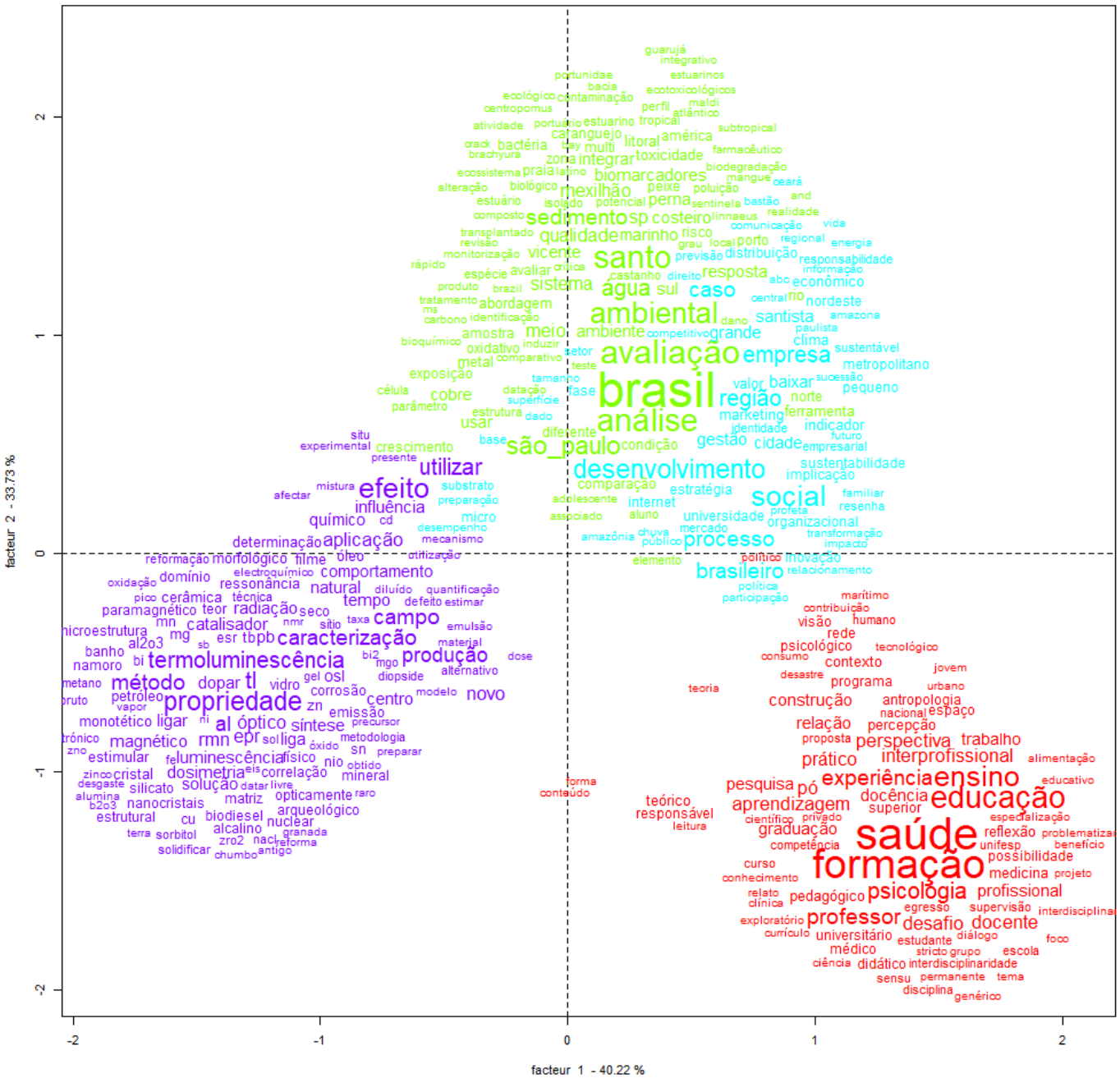
A segunda classe que concentra quase um terço dos dados refere-se ao campo de **Engenharia e Ciência de Materiais**, um dos programas mais fortes e bem avaliados do campus SJC. Dentre os vários materiais identificados - biomateriais, metais, cerâmicos e polímeros - há grande destaque para as palavras "diamante" e "carbono". Isso pode ser explicado pelo fato de o diamante ser uma forma natural de carbono que tem alta densidade atômica e, como consequência, possui muitas propriedades com grande aplicação em engenharia, tais como dureza, elasticidade e condutividade térmica.

Outro grupo de palavras identificado na análise tem uma característica mais **interdisciplinar** e se aproxima da área de **ciência básica**, com estudos experimentais em animais (em sua maioria ratos), envolvendo biologia molecular e fisiologia. Esta classe expressa em parte as áreas de concentração de outros dois programas de pós graduação do campus SJC que são Biotecnologia e Engenharia Biomédica. O denominamos de **Bioengenharia/Interdisciplinar**.

No Grafo 6, percebe-se claramente como o agrupamento de palavras vinculado a **Bioengenharia/Interdisciplinar** se vincula ao campo da Engenharia e Ciência de Materiais, por meio dos vetores que ligam as palavras "rato" e "efeito" a "propriedade", "diamante", "carbono" e "substrato". Todas estas demonstram ter muito menos pontos de ligação com o grupo que se vincula à computação.

GRÁFICO 7 (Análise Fatorial de Correspondência)

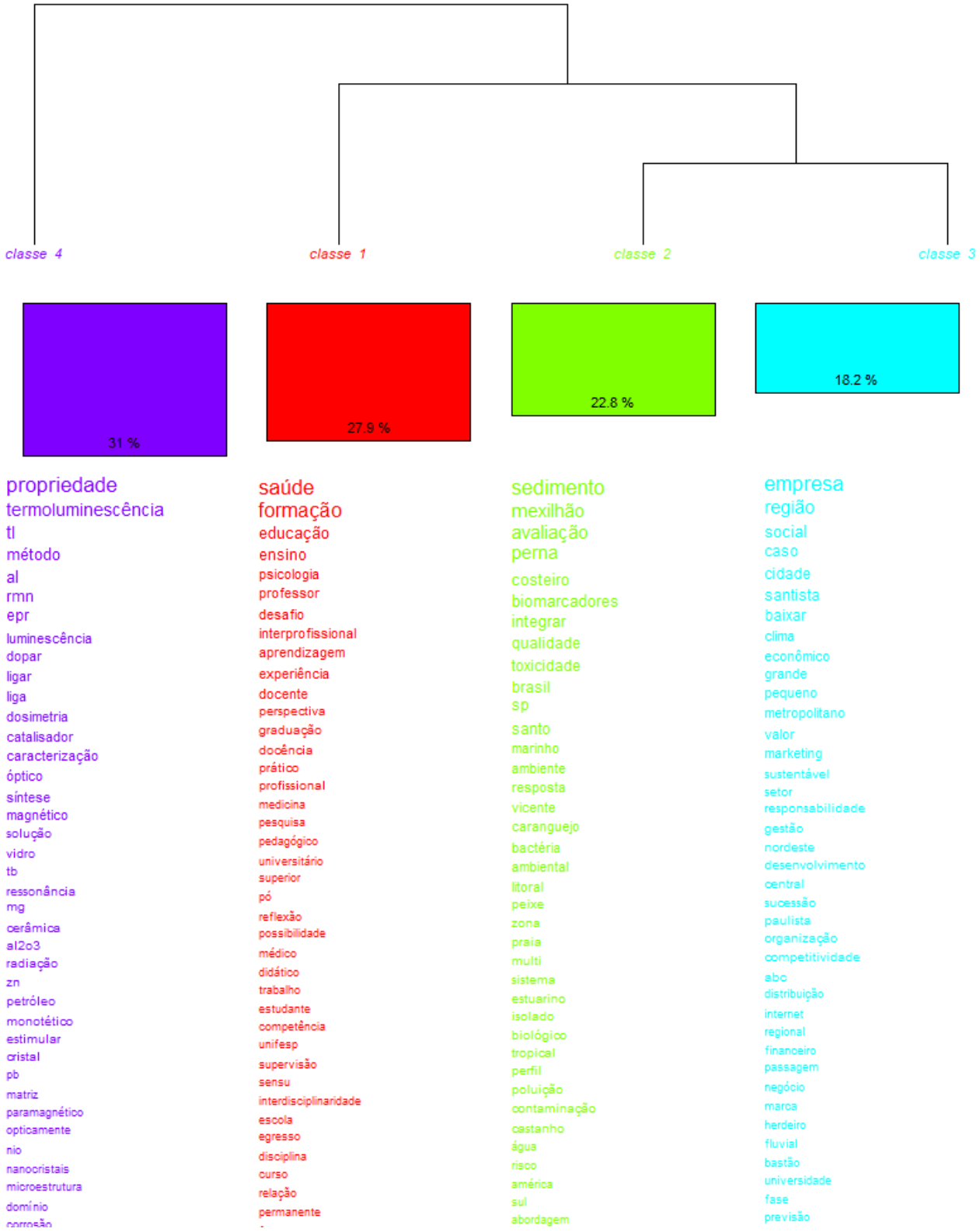
Instituto do Mar (IM)



DENDROGRAMA 7 (Classificação Hierárquica Descendente)

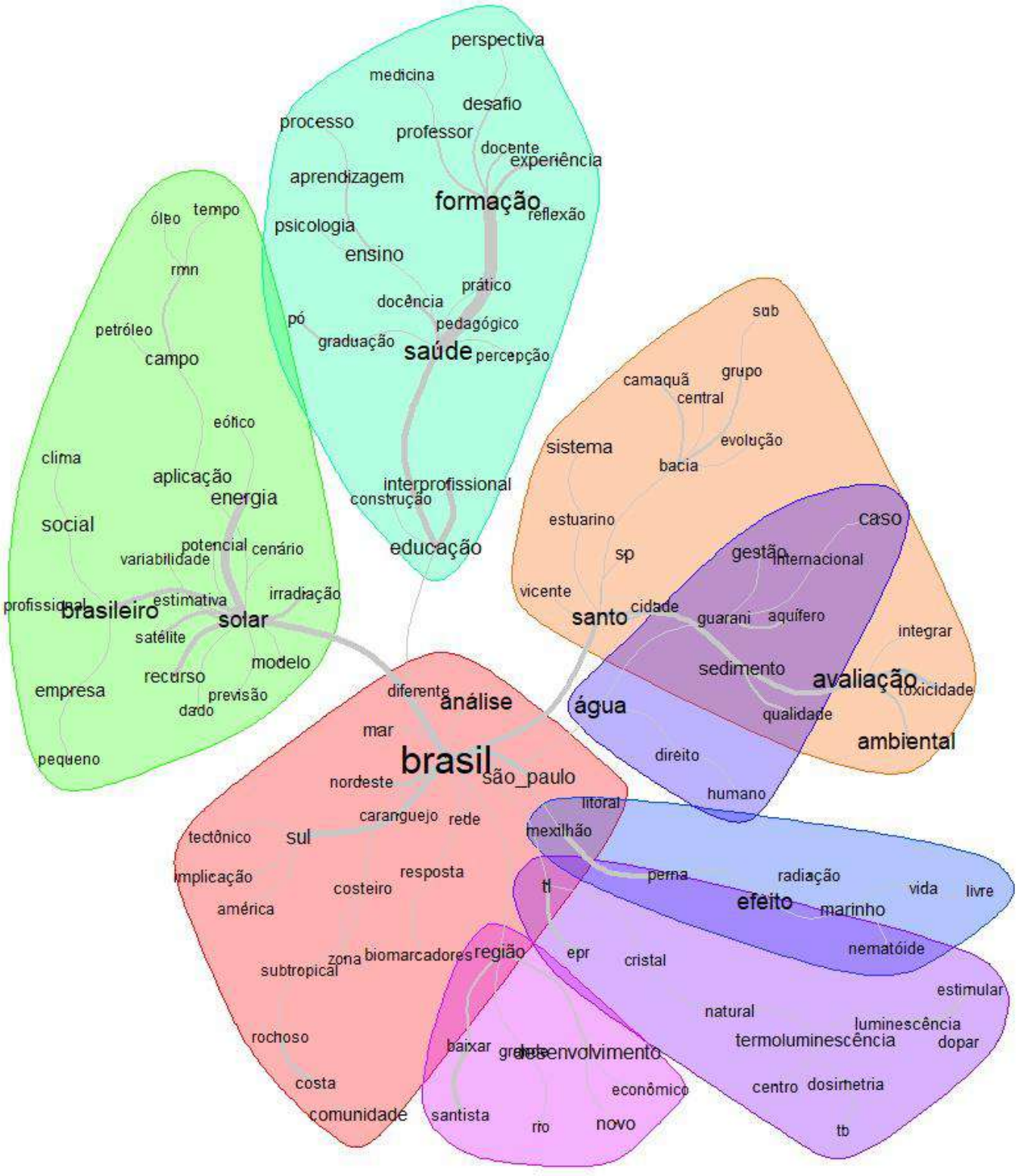
Instituto do Mar (IM)

Tecnologia/Experimentação	Saúde/Educação	Biodiversidade/Ecologia	Sustentabilidade



GRAFO 7 (Análise de Similitude)

Instituto do Mar (IM)



No Dendrograma do Instituto do Mar (IM, Campus Baixada Santista), aparecem quatro agrupamentos da produção dos docentes que estão vinculados a dois Programas: o de Biodiversidade e Ecologia Marinha e Costeira e o Interdisciplinar em Ciência e Tecnologia do Mar. Um primeiro que denominamos de **Tecnologia/Experimentação** que possui uma concentração da produção (39%), muito voltada para os experimentos de materiais com predominância de análises químicas e matemáticas. Um segundo de **Saúde/Educação (**27,9%), mesmo que ele não se refira diretamente aos temas dos Programas. E os outros dois, que chamamos de **Biodiversidade/Ecologia** e **Sustentabilidade** que demonstram ser subgrupos de uma mesma classe, razão pela qual aparecem no Gráfico 7 como bastante integrados, devido igualmente a proposta interdisciplinar de um dos Programas.

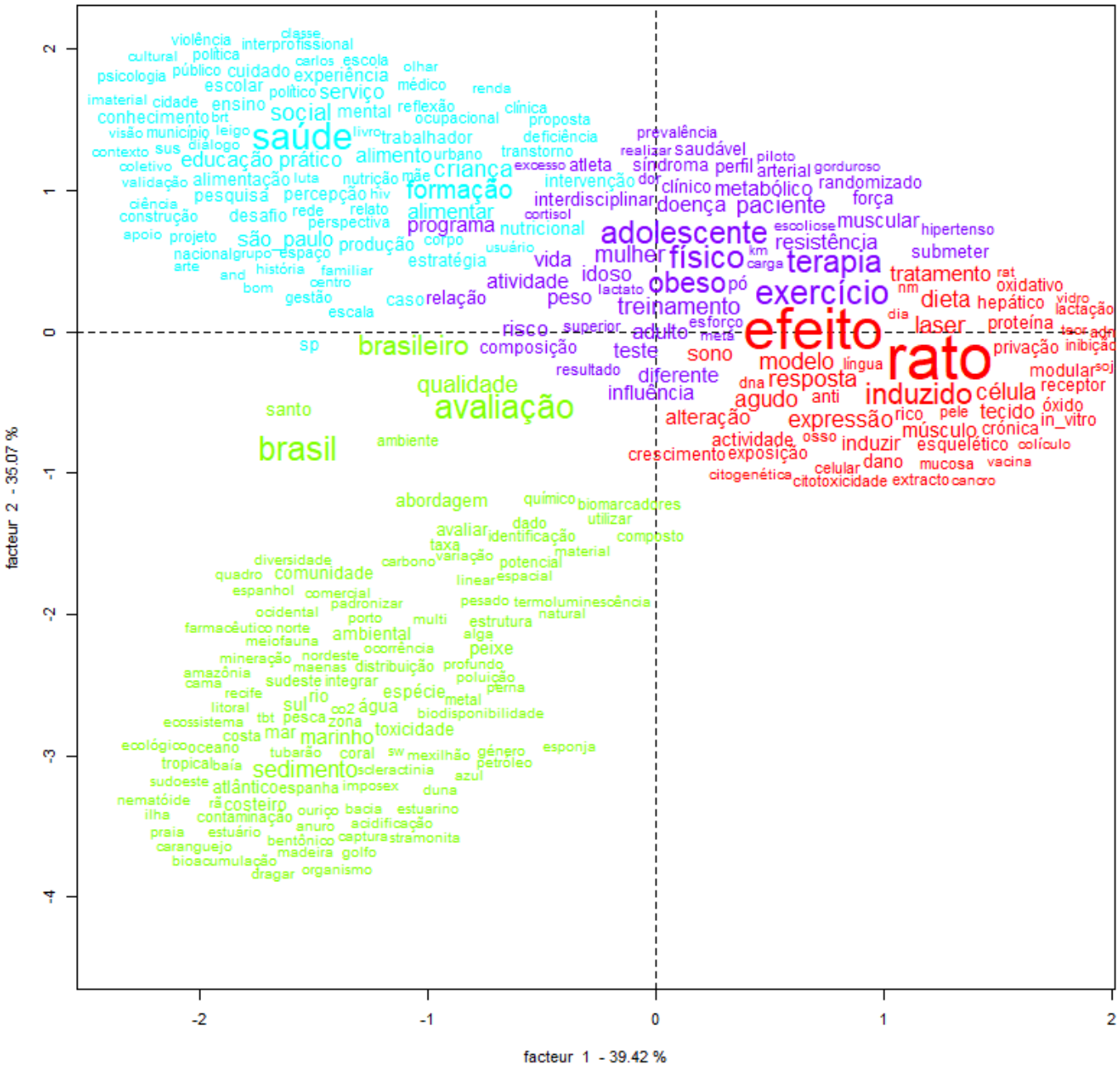
Por essa razão, o Gráfico acaba por representar três conjuntos, os quais, há que se notar, relacionam-se muito pouco entre si. O que parece estar confirmado pelo Grafo 7, em que há um conjunto praticamente sem contato com os outros vinculados pelas palavras “saúde” e “formação”; um outro, marcado pela maior frequência na articulação de “Brasil” com os conjuntos em que estão a palavras “solar” e “santos”; e vários outros conjuntos representados pelas palavras “efeito”, “radiação”, “termoluminescência”, “desenvolvimento” que mantém intersecções mais periféricas entre si, e com os outros conjuntos. Com exceção feita para a palavra “água” que aponta ser uma palavra que vincula dois conjuntos.

Voltando ao Gráfico 7, vemos que:

- no agrupamento **Tecnologia/Experimentação** há uma concentração em torno das palavras “propriedade”, “caracterização” e “termoluminescência”; no entanto, a palavra que o aproxima dos agrupamentos de **Biodiversidade/Ecologia** e **Sustentabilidade** é “efeito”. O que indica, da mesma forma que em Diadema, que as áreas mais experimentais testam elementos e substâncias de vinculados às pesquisas em biologia;
- como notado, os agrupamentos de **Biodiversidade/Ecologia** e **Sustentabilidade** demonstram estar profundamente articulados, com alta frequência para as palavras “Brasil”, “análise”, “Santos”, “ambiental” e “desenvolvimento”. Há que se notar que o que chamamos Sustentabilidade estão palavras como “região”, “empresa”, “social”, “processo”, em que estão representadas as iniciativas de pesquisa com o setor privado.
- no agrupamento de **Saúde/Educação**, há uma maior frequência exatamente das palavras “saúde”, “formação”, “educação”, “ensino” e “psicologia”, e, em menor escala, “desafio” e “interprofissional”, o que indica haver, ou ter existido, uma produção relevante dos pesquisadores nessa área, ainda que não expresso nas temáticas dos Programas.

GRÁFICO 8 (Análise Fatorial de Correspondência)

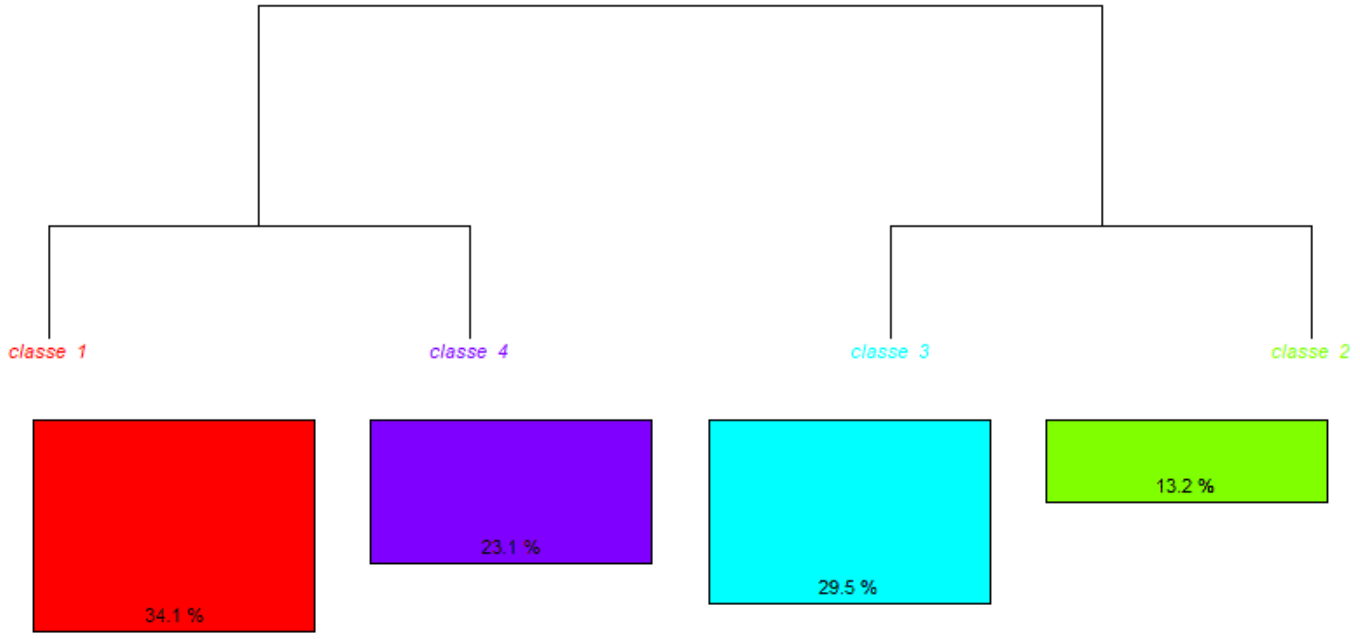
Instituto de Saúde e Sociedade (ISS)



DENDROGRAMA 8 (Classificação Hierárquica Descendente)

Instituto de Saúde e Sociedade (ISS)

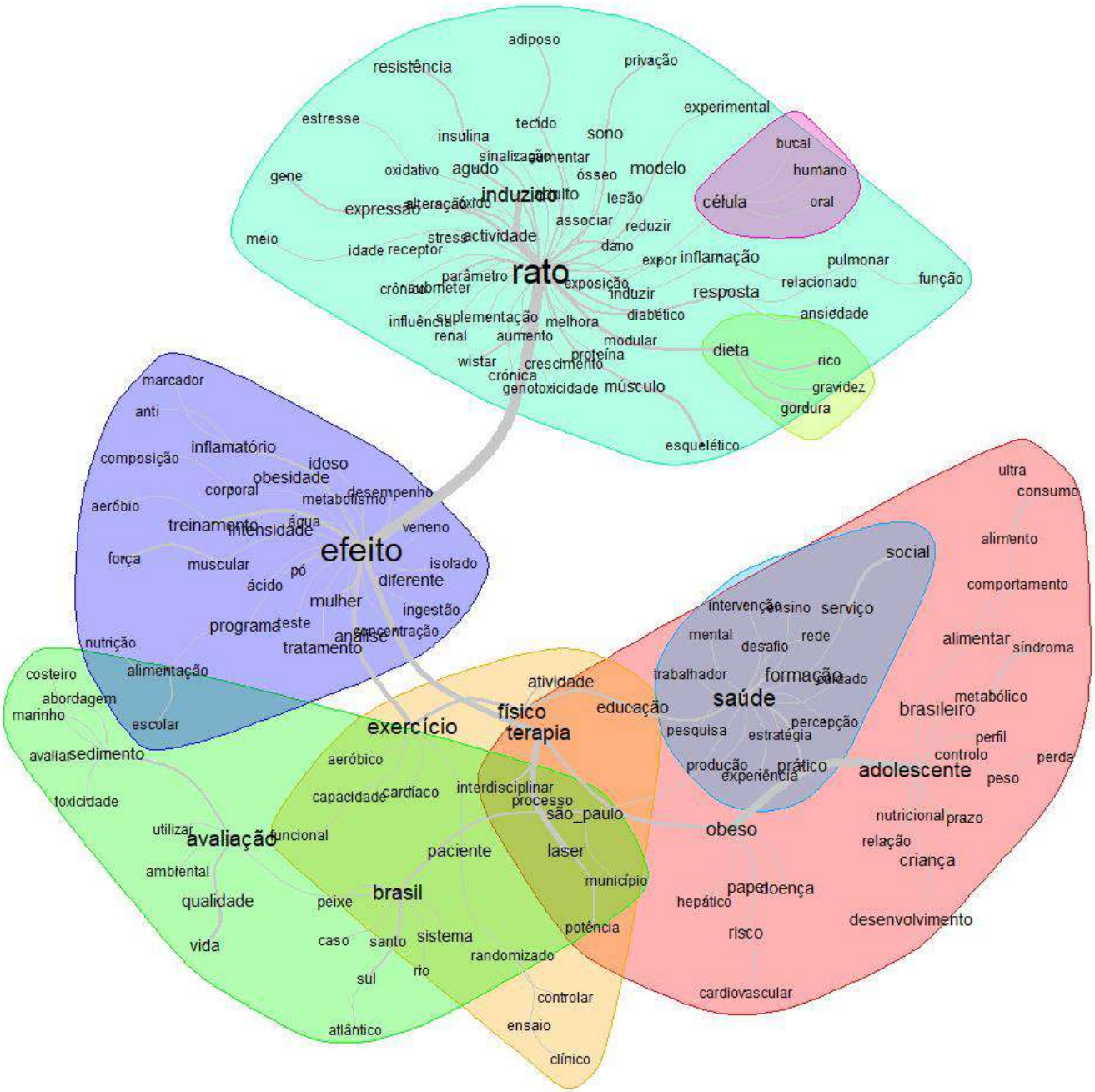
Ciências básicas	Movimento/Nutrição	Saúde/Social	Biodiversidade



classe 1	classe 4	classe 3	classe 2
rato	obeso	saúde	sedimento
induzido	adolescente	social	marinho
célula	treinamento	educação	brasil
expressão	paciente	serviço	atlântico
laser	exercício	prático	sul
induzir	mulher	experiência	mar
efeito	físico	são_paulo	costeiro
inflamação	força	ensino	toxicidade
receptor	metabólico	formação	água
tecido	randomizado	alimentação	espécie
dano	síndroma	alimento	costa
proteína	desempenho	cuidado	rio
óxido	muscular	pesquisa	espanha
crónica	interdisciplinar	mental	peixe
modelo	controlar	público	contaminação
esquelético	clínico	criança	oceano
dieta	indivíduo	trabalhador	ambiental
músculo	idoso	desafio	coral
experimental	multidisciplinar	conhecimento	metal
in_vitro	resistência	município	comunidade
ácido	pressão	atenção	tropical
rico	teste	reflexão	sudeste
sinalização	peso	rede	pesca
genotoxicidade	equilíbrio	interprofissional	sudoeste
privação	obesidade	escola	ouriço
mucosa	corporal	escolar	tubarão
stress	cardiovascular	cidade	imposex
wistar	aeróbio	nutrição	integrar
modular	prazo	universidade	bacia
cérebro	perfil	ocupacional	scleractinia
actividade	máximo	político	zona
reparação	esforço	professor	estuário
inibição	lactato	alimentar	bentônico
gene	leptina	política	captura
propriedade	alcoólico	estudante	sw
oral	saudável	família	avaliação
inibidor	homocisteína	médico	abordagem
pele	idiopática	graduação	ecossistema
adn		universitário	

GRAFO 8 (Análise de Similitude)

Instituto de Saúde e Sociedade (ISS)



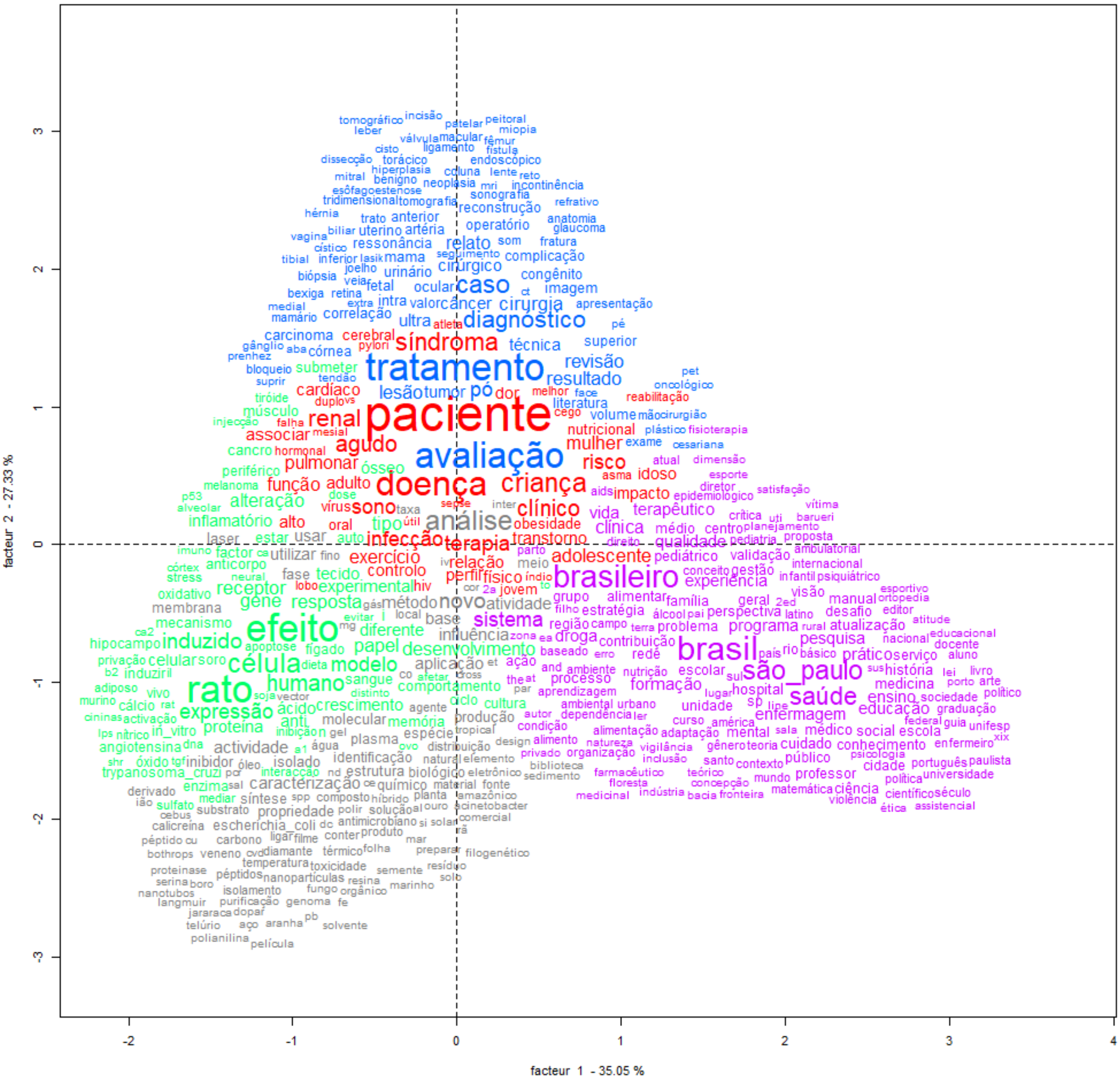
No Dendrograma 8 do Instituto de Saúde e Sociedade (ISS) temos 4 agrupamentos que demonstram uma ampla articulação na produção dos sete Programas de Pós-Graduação lá existentes: Alimentos, Nutrição e Saúde, Bioprodutos e Bioprocessos, Ciência do Movimento Humano e Reabilitação, Interdisciplinar em Ciências da Saúde, Saúde da Família e Serviços Social e Políticas Sociais. No primeiro, que chamamos de **Ciências Básicas** (34,1%), estão representadas as palavras que, em mais de um deles, referem-se às pesquisas básicas vinculadas à substâncias e de também de reações físicas. No segundo, que designamos por **Saúde/Social** (29,5%), estão representadas as pesquisas que tratam de serviço social, também da família, trabalhadores e alimentação. Seguidas por um agrupamento de **Movimento/Nutrição** (23,1%), em que estão representadas pesquisas na área de reabilitação, bem como de nutrição e saúde. E um último de **Biodiversidade** que, ainda que menos expressivo em termos quantitativos (13,2%), é representativo por se referir a pesquisas voltadas a análises regionais ambientais, de ciências do mar, que parecem aglutinar produções de docentes de maneira interdisciplinar nos programas.

No tocante às áreas de maior concentração e intersecção, vemos:

- como as análises do grupo de **Ciências Básicas,** com alta frequência e mesmo integração das palavras “rato” e “efeito”, tem um forte papel na integração de áreas experimentais com as de saúde (movimento e alimentação), como se vê na sua junção com o agrupamento de **Movimento/Nutrição** - onde predominam palavras como “adolescente”, “físico”, “obeso”, “terapia” e “exercício”. O que pode ser igualmente comprovado no diagrama de similitude, onde predominam as associações entre estes dois agrupamentos;
- no agrupamento **Saúde/Social**, a maior frequência é para a palavra “saúde”, circundada por praticamente todas as outras possíveis de ver nesta escala com pequena variação na sua ocorrência (salvo por “Criança” e “formação”). O que indica uma difusão de temas baixo sua relação com saúde (o que se vê igualmente no diagrama de similitude);
- o grupo de **Biodiversidade**, vê-se um pouco mais isolado em relação aos outros, mas cuja intersecção com o restante ocorre sugestivamente por meio das palavras “avaliação”, “Brasil”, “qualidade”, “brasileiro”. O que parece indicar sua vinculação com as outras áreas por meio de processos avaliativos de elementos e substâncias em uma perspectiva regional, vinculados especialmente com estudos do mar.

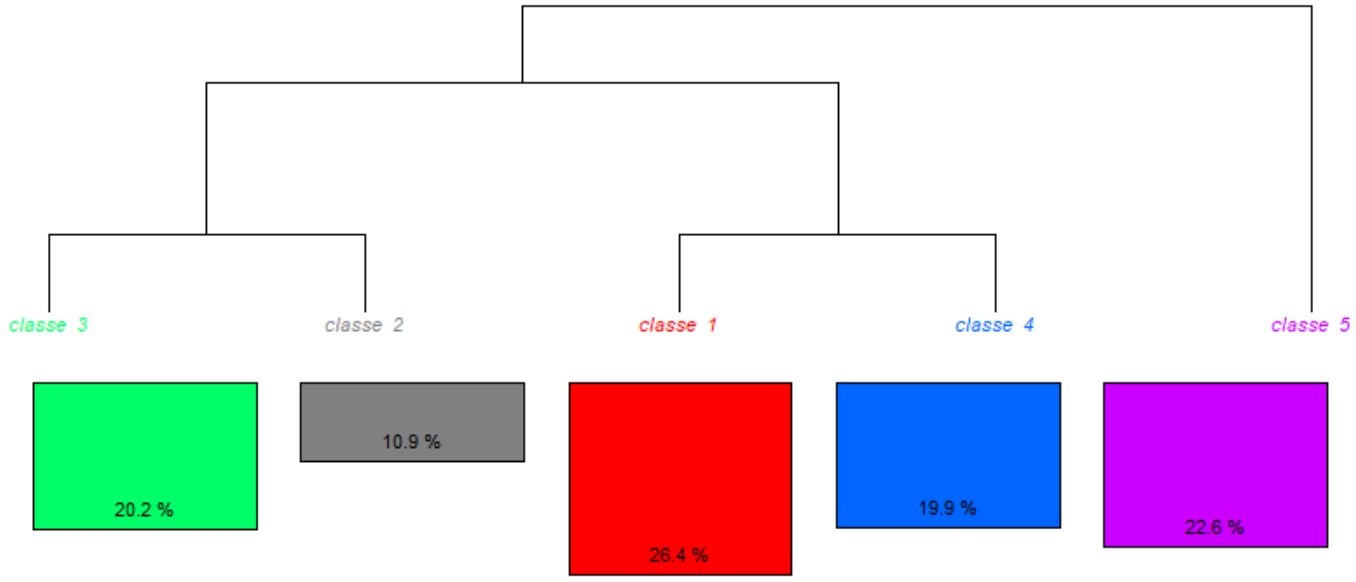
GRÁFICO 9 (Análise Fatorial de Correspondência)

Universidade Federal de São Paulo (UNIFESP)



DENDROGRAMA 9 (Classificação Hierárquica Descendente)

Universidade Federal de São Paulo (UNIFESP)



rato
célula
induzido
efeito
expressão
receptor
proteína
celular
resposta
modelo
induzir
papel
gene
humano
hipocampo
experimental
angiotensina
apoptose
trypanosoma_cruzi
privação
sinalização
oxidativo
anti
tecido
cálcio
activação
inflamação
macrófago
inflamatório
anticorpo
ácido
regulação
isquemia
metabolismo
inibição
modular
reperfusão
stress
cérebro
mecanismo
modulação
imune
nítrico
actividade
pilocarpina
proliferação

caracterização
propriedade
isolado
escherichia_coli
substrato
carbono
filme
diamante
estrutura
veneno
síntese
químico
nanopartículas
composto
determinação
microscopia
dopar
antimicrobiano
purificação
actividade
cerâmica
estrutural
conter
identificação
especificidade
péptidos
espectroscopia
bothrops
temperatura
isolamento
protease
preparação
água
molecular
aplicação
nanotubos
sequência
estirpe
serina
utilizar
térmico
reacção
mistura
proteinase
compósito

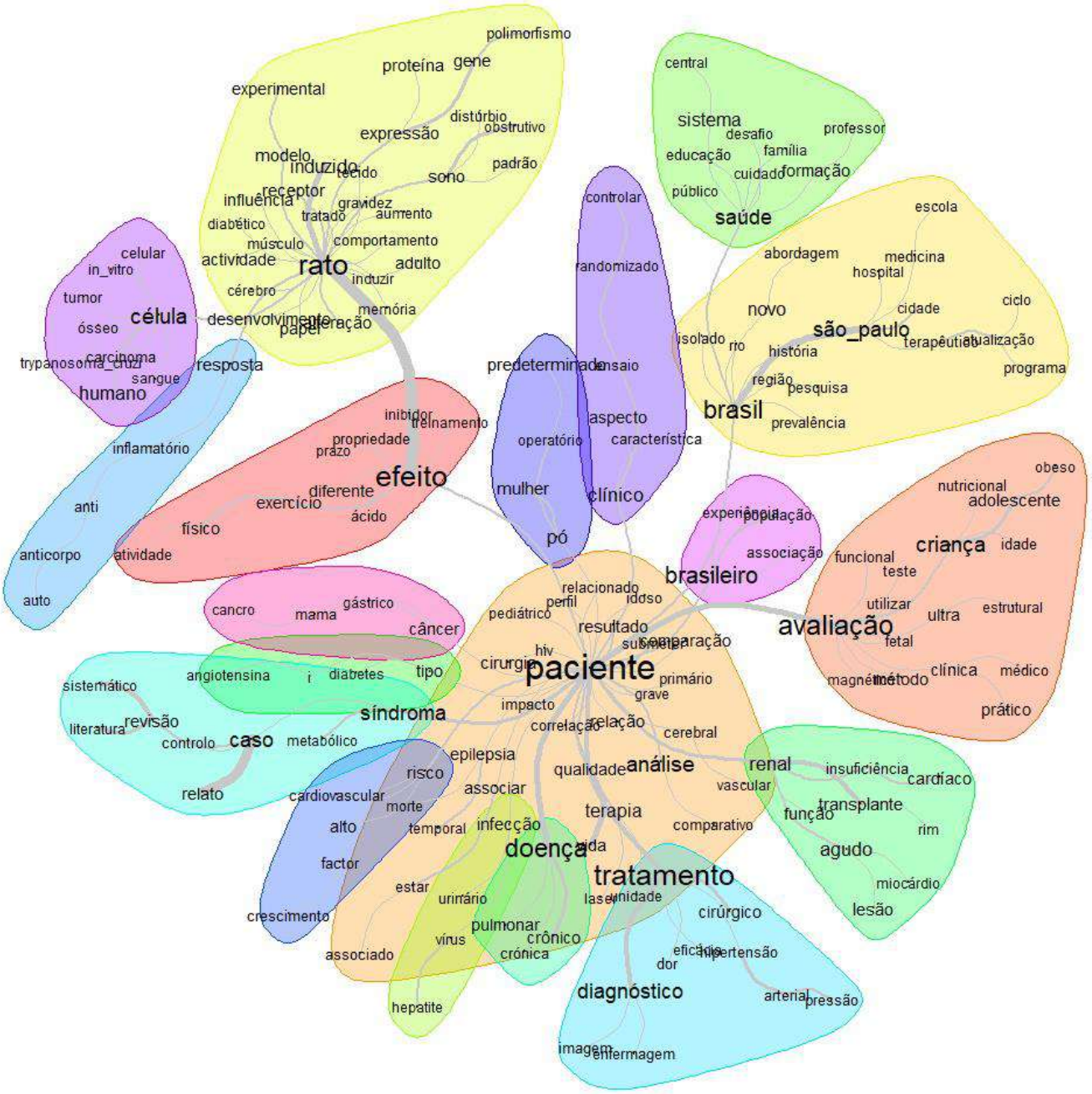
paciente
doença
síndroma
crônico
renal
risco
agudo
obstrutivo
arterial
insuficiência
sono
clínico
transplante
randomizado
mulher
hipertensão
distúrbio
metabólico
hepatite
cardíaco
cardiovascular
pulmonar
exercício
diabetes
epilepsia
sistêmico
teste
idoso
apneia
sintoma
respiratório
associação
obesidade
artrite
indivíduo
controlar
eritematoso
dor
auditivo
apnéia
juvenil
lúpus
função
esclerose
criança
cego

caso
relato
cirúrgico
cirurgia
tratamento
câncer
ultra
sonografia
técnica
intra
reconstrução
congênito
tomografia
diagnóstico
ocular
uterino
tumor
fetal
anterior
mama
urinário
operatório
artéria
ressonância
neoplasia
fratura
incontinência
ligamento
glaucoma
som
imagem
coluna
tridimensional
literatura
computadorizar
revisão
endoscópico
magnético
implante
cervical
genital
posterior
aneurisma
inferior
joelho

saúde
são_paulo
brasil
educação
ensino
enfermagem
prático
medicina
pesquisa
médico
escola
social
público
história
professor
ciência
cuidado
serviço
programa
cidade
rio
mental
formação
português
conhecimento
científico
atualização
brasileiro
unidade
política
hospital
nacional
intensivo
desafio
graduação
experiência
assistência
emergência
universidade
político
unifesp
guia
estudante
manual
atenção

GRAFO 9 (Análise de Similitude)

Universidade Federal de São Paulo (UNIFESP)



Referências

1. Mena-Chalco JP, Cesar-Jr RM. scriptLattes: An open-source knowledge extraction system from the Lattes platform. J Braz Comput Soc. 2009;15(4):31–9.
2. Pentaho Data Integration - Create Data Pipelines [Internet]. [citado 22 de outubro de 2019]. Disponível em: https://www.hitachivantara.com/en-us/products/data-management-analytics/pentaho-platform/pentaho-data-integration.html
3. Data preparation and ETL tool for non-technical people | EasyMorph [Internet]. [citado 6 de abril de 2020]. Disponível em: https://easymorph.com/
4. Ratinaud P. IRAMUTEQ: Interface de R pour les analyses multidimensionnelles de textes et de questionnaires [Internet]. [citado 22 de outubro de 2019]. Disponível em: http://www.iramuteq.org/
5. Reinert A. Une méthode de classification descendante hiérarchique : application à l'analyse lexicale par contexte. Cah Anal Données. 1983;8(2):187–98.