

UNIVERSIDADE NOVE DE JULHO

PROGRAMA DE PÓS-GRADUAÇÃO EM ADMINISTRAÇÃO - PPGA

JOSÉ EDUARDO STOROPOLI

**O USO DO *KNOWLEDGE DISCOVERY IN DATABASE (KDD)* DE INFORMAÇÕES
PATENTÁRIAS SOBRE ENSINO A DISTÂNCIA: CONTRIBUIÇÕES PARA
INSTITUIÇÕES DE ENSINO SUPERIOR**

Orientador: Prof. Dr. Emerson Antônio Maccari

Co-Orientador: Fernando Antônio Ribeiro Serra

SÃO PAULO

2016

JOSÉ EDUARDO STOROPOLI

**O USO DO *KNOWLEDGE DISCOVERY IN DATABASE (KDD)* DE INFORMAÇÕES
PATENTÁRIAS SOBRE ENSINO A DISTÂNCIA: CONTRIBUIÇÕES PARA
INSTITUIÇÕES DE ENSINO SUPERIOR**

Tese apresentada ao Programa de Mestrado e
Doutorado em Administração da Universidade
Nove de Julho – UNINOVE, para obtenção do
título de Doutor em Administração.

Prof. Émerson Antônio Maccari, Dr. – Orientador

Prof. Fernando Antônio Ribeiro Serra, Dr. – Co-Orientador

SÃO PAULO

2016

FICHA CATALOGRÁFICA

Storopoli, José Eduardo.

O uso do Knowledge discovery in database(KDD) de informações patentárias sobre ensino a distância: contribuições para instituições de ensino superior./ José Eduardo Storopoli. 2016.

193 f.

Tese (doutorado) – Universidade Nove de Julho - UNINOVE, São Paulo, 2016.

Orientador (a): Prof. Dr. Émerson Antônio Maccari.

1. Educação a distância. 2. Ensino superior. 3. Patentes. 4. Estratégia.

I. Maccari, Émerson Antônio. II. Título

CDU 658

DEDICATÓRIA

Dedico esta tese aos meus pais, Eduardo Storopoli e Maria Cristina Barbosa Storopoli, que sempre me incentivaram a perseguir meus estudos, que estiveram ao meu lado e foram cruciais na minha formação como acadêmico, como pessoa e como cidadão.

AGRADECIMENTOS

Gostaria de agradecer à Caroline Cabral Kanaan, pelo apoio e compreensão durante este árduo trabalho que muitas vezes aparentou infinito e herculano.

Ao prof. Dr. Émerson Antônio Maccari, pela orientação nos momentos difíceis.

Ao prof. Dr. Fernando Antônio Ribeiro Serra, pelas discussões e reflexões.

Ao prof. Dr. Luc Marie Quoniam, pela excentricidade e genialidade.

À profa. Cláudia Kniess, pelo apoio no conhecimento patentários.

Ao prof. Leonel Cézar Rodrigues, pela tutela e dedicação.

Ao prof. Martinho Isnard Ribeiro, pela amizade e parceria.

Ao prof. Edson Luíz Riccio, pela contribuição na definição do escopo da tese.

Ao prof. David Reymond, pela contribuição com a ferramenta *Patent2Net*.

Ao Marcos Mazzieri, pelas contribuições metodológicas.

Ao prof. Renato Ferraz, pelo apoio durante a análise de dados.

Ao André Santos, pela contribuição no processo de busca na base de dados ESPACENET.

Ao Celso Vanderlei, pelas contribuições na realização dos mapas mentais.

Ao prof. Manuel Portugal, pelo apoio e direcionamento.

Ao Grupo de Pesquisa de Gestão Estratégia de Ensino Superior, na figura do Fabrício Imbrizi.

À secretaria de pós-graduação, na figura da Marcela Ágatha de Almeida.

À UNINOVE, por me proporcionar o apoio institucional.

SUMÁRIO

LISTA DE FIGURAS	9
LISTA DE TABELAS	15
RESUMO	16
ABSTRACT.....	17
1 INTRODUÇÃO.....	18
1.1 Problematização.....	21
1.2 Objetivos do Estudo	22
1.2.1 Objetivo geral	23
1.2.2 Objetivos específicos.....	23
1.3 Justificativas.....	23
1.4 Estrutura do Trabalho	26
2 REVISÃO DA LITERATURA	27
2.1 Instituições de Ensino Superior.....	27
2.1.1 <i>Trade-offs</i> entre Ensino, Pesquisa e Extensão	27
2.1.2 Gestão em Instituições de Ensino Superior	29
2.1.3 Vantagem competitiva em Instituições de Ensino Superior.....	30
2.2 O Uso de EAD em Instituições de Ensino Superior.....	31
2.2.1 Novas Tecnologias de Ensino para o EAD	35
2.3 Estrutura e Evolução da Estratégia	37
2.3.1 Visão Baseada em Recursos	38
2.3.1.1 Origens da RBV.....	38
2.3.2 Teoria Institucional.....	40
2.3.2.1 Teoria Institucional Aplicada à Instituições de Ensino Superior	42
2.4 Relação entre RBV e Teoria Institucional.....	43
2.5 Informações Patentárias	45
2.5.1 Base de dados de Patentes	47

2.5.2	A Classificação Internacional de Patentes.....	48
3	METODOLOGIA DA PESQUISA	50
3.1	Framewok Patent Analyze <i>versus</i> KDD.....	50
3.2	As bases de patentes e suas interfaces de busca.....	52
3.2.1	Os <i>crawlers</i> de patentes e o processo <i>data mining</i>	54
3.2.2	A ferramenta computacional <i>Patent2net</i>	55
3.3	Construindo palavras-chave para EAD	56
3.3.1	Relação com a ambiguidade na expressão e busca (silêncio e ruído)	59
3.4	Análise das patenes em EAD	68
4	APRESENTAÇÃO DOS RESULTADOS	69
4.1	Mapeamento das patentes em EAD no mundo	69
4.1.1	Disponibilização pública das informações sobre as patentes em EAD	69
4.1.2	Mapeamento geral dos depósitos de patentes em EAD.....	73
4.1.3	Possibilidades de cruzamento de dados com a utilização de tabelas dinâmicas	75
4.1.4	Mapa de geolocalização dos países (inventores, das empresas solicitantes, onde a proteção da tecnologia foi solicitada)	88
4.1.5	Mapeamento das patentes concedidas em EAD	91
4.1.6	Possibilidades de cruzamento relacionados às patentes concedidas ou em processo de concessão sobre o tema EAD.....	91
4.1.7	Possibilidades de seleção de patentes com base na combinação de indicadores...	101
4.1.8	Redes relacionadas às patentes em EAD	104
4.1.9	Redes de equivalência entre as tecnologias, entre os documentos citados por outras patentes ou por fontes externas, e entre as citações entre patentes.....	110
4.1.10	Análises relacionadas às famílias de patentes	113
4.1.11	<i>Mind map</i> entre patentes	115
4.2	Metodologia de Busca de Patentes em EAD por Análise de Conteúdo	117
4.2.1	Resumos das patentes por Classificação Hierárquica Descendente (CHD) e Análise Fatorial de Correspondência (AFC)	120

4.2.2	Resumos das patentes por similaridade	133
4.2.3	Resumos das patentes por AFC (análise de 5 modalidades)	134
4.2.4	Descrições das invenções nas patentes sobre EAD por CHD e AFC.....	135
4.2.5	Descrições das invenções das patentes por similaridade.....	147
4.2.6	Reivindicações das patentes de EAD por CHD e AFC	150
4.2.7	Reivindicações das patentes por similaridade	156
4.2.8	Consolidação das análises de conteúdo das patentes sobre EAD.....	158
4.2.9	Patentes EAD por análise de agrupamento: <i>clustering</i>	160
4.3	Mapeamento das patentes em EAD em Universidades.....	165
4.4	Uso de informações patentárias na definição de estratégia em IES À luz do RBV	170
5	CONCLUSÕES E RECOMENDAÇÕES	174
5.1	Conclusões	175
5.2	Recomendações	179
	REFERÊNCIAS	180

LISTA DE FIGURAS

Figura 1 – Exemplo de curva de utilidade de uma escolha de investimentos de uma universidade	28
Figura 2 – Comparação entre procedimento metodológico KDD e Patent2Net.....	52
Figura 3 – Mapa Mental das palavras-chave de educação a distância.....	57
Figura 4 – Potencial de recuperação das palavras-chave.	58
Figura 5 – Contribuição das palavras-chaves na expressão de busca.	59
Figura 6 – Potencial da expressão de busca anterior ao refinamento.	67
Figura 7 – Potencial da expressão de busca após o refinamento.	67
Figura 8 – Tela inicial com os temas sobre os quais já foi realizada a mineração de patentes, com destaque para as patentes em EAD.	71
Figura 9 – Tela apresentando os 7 <i>links</i> de acesso aos diversos resultados relacionados à análise das 3090 patentes em EAD.....	73
Figura 10 – Montagem da curva de estabilidade para definição da estratégica de pesquisa (rodapé) das patentes em EAD com a utilização da ferramenta Patent2net....	74
Figura 11 – Cabeçalho da <i>interface</i> que permite realizar cruzamentos entre os 13 indicadores disponíveis e relacionados às patentes	76
Figura 12 – Parte superior e inferior da tela que avalia a relação entre o inventor e o seu país de origem	77
Figura 13 – Parte superior e inferior da tela que avalia a relação entre a empresa depositante e o seu país de origem	78
Figura 14 – Tela que demonstra a relação entre o país da empresa depositante e o estatuto jurídico da patente	79
Figura 15 – Tela que demonstra a relação entre a data de depósito e o país onde a patente foi depositada.....	80
Figura 16 – Tela que demonstra a relação entre o ano de depósito e o país onde a patente foi depositada	81
Figura 17 – Representação gráfica da evolução quantitativa dos depósitos das patentes em EAD.....	82
Figura 18 – Parte superior (figura à esquerda) e inferior (figura à direita) da tela que avalia a relação entre o IPCR4 e o número de citações recebidas pela patente.	83

Figura 19 – Parte superior e inferior da tela que demonstra a relação entre o IPCR4 e o país do inventor.....	84
Figura 20 – Parte superior e inferior da tela que demonstra a relação entre o IPCR4 e o país da empresa depositante.	85
Figura 21 – Parte superior e inferior da tela que demonstra a relação entre o IPCR4 e o país de depósito.	86
Figura 22 – Parte superior e inferior da tela que demonstra a relação entre o IPCR4 e o estatuto jurídico da patente.	87
Figura 23 – Parte superior e inferior da tela que demonstra a relação entre o IPCR4 e a data de depósito da patente.	88
Figura 24 – Mapa de geolocalização levando em consideração os países declarados como sendo de origem pelos inventores das tecnologias em EAD.....	89
Figura 25 – Mapa de geolocalização levando em consideração os países declarados pelas empresas detentoras da tecnologia cuja patente foi solicitada	90
Figura 26 – Mapa de geolocalização levando em consideração os países onde foi solicitada a proteção da tecnologia por meio do registro das patentes.....	91
Figura 27 – Tela que demonstra a relação entre as patentes concedidas e o país do inventor.....	92
Figura 28 – Parte superior e inferior da tela que demonstra a relação entre as patentes concedidas ou em processo de concessão, a empresa depositante e o seu país de origem.	93
Figura 29 – Tela que demonstra a relação entre as patentes concedidas ou em processo de concessão e o país de depósito.....	94
Figura 30 – Tela que demonstra a relação entre as patentes concedidas ou em processo de concessão e a data de depósito.	95
Figura 31 – Gráfico de linhas que demonstra quais foram as tecnologias mais citadas nas patentes concedidas em cada ano.	96
Figura 32 – Parte superior e inferior da tela que demonstra a relação entre o IPCR7 e o país da empresa solicitante	97
Figura 33 – Parte superior e inferior da tela que demonstra a relação entre o IPCR7 e o país da empresa depositante	98

Figura 34 – Parte superior e inferior da tela que demonstra a relação entre o IPCR7 e o país de depósito (apenas patentes concedidas ou em processo de concessão).....	99
Figura 35 – Parte superior e inferior da tela que demonstra a relação entre o IPCR7 e a data de depósito	100
Figura 36 – Parte superior e inferior da tela que demonstra a relação entre o IPCR7 e o número de citações recebidas pela patente.....	101
Figura 37 – Partes superior e inferior da tela com a visão geral das informações relacionadas às 3090 patentes em EAD.	102
Figura 38 – Recorte da tela com a visão geral das informações relacionadas à única patente em EAD depositada por uma empresa brasileira	103
Figura 39 – Rede entre as empresas depositantes das 3.090 patentes em EAD.....	105
Figura 40 – Rede entre as patentes depositadas em EAD com base no IPCR7.	106
Figura 41 – Rede entre os países e a tecnologia da patente depositada (com base no IPCR7).	107
Figura 42 – Rede mista entre os países e a tecnologia cuja patente foi solicitada (com base no IPCR7).	108
Figura 43 – Rede mista entre os inventores e a tecnologia cuja patente foi solicitada (com base no IPCR7).	109
Figura 44 – Rede mista entre as empresas e a tecnologia cuja patente foi solicitada (com base no IPCR7).	110
Figura 45 – Rede simples destacando as relações entre tecnologias equivalentes voltadas ao EAD (com base no IPCR11).	111
Figura 46 – Rede simples destacando as relações entre tecnologias que citaram os mesmos artigos científicos em suas listas de referências (com base no IPCR11).	112
Figura 47 – Rede simples destacando as relações entre tecnologias que citaram a tecnologia central, em destaque na Figura (com base no IPCR11).....	113
Figura 48 – <i>Interface</i> para seleção das famílias tecnológicas das patentes.....	114
Figura 49 – <i>Interface</i> para cruzamento das famílias tecnológicas das patentes.	114
Figura 50 – <i>Mind map</i> com os principais assuntos extraídos pelo Patent2net das patentes em EAD depositadas na Espacenet.	115

Figura 51 – Mind map com os principais assuntos extraídos pelo Patent2net das patentes em EAD depositadas na Espacenet (foco no tema principal “G: Physics (2503 documentos).	116
Figura 52 – Mind map com os principais assuntos extraídos pelo Patent2net das patentes em EAD depositadas na Espacenet (foco no tema central “G: Physics → G06 Computing: Calculatint: Counting”, com 1251 documentos)	117
Figura 53 - Dendograma da CHD dos resumos das patentes sobre EAD	122
Figura 54 – AFC dos resumos das patentes sobre EAD	123
Figura 55 – AFC dos resumos das patentes sobre EAD com as formas ativas que os compõe.	124
Figura 56 – AFC dos resumos das patentes sobre EAD com as formas suplementares que os compõe.	126
Figura 57 – AFC dos resumos das patentes sobre EAD com os cabeçalhos, indicando países depositantes e o código IPC	127
Figura 58 – Dendograma e sub-contesto do contexto 1 com as formas componentes.	128
Figura 59 – Dendograma e sub-contesto do contexto 2 com as formas componentes.	128
Figura 60 – Dendograma e sub-contesto do contexto 3 com as formas componentes.	129
Figura 61 – Análise do Sub-contesto 1, frequência de ocorrência mínima 25 e visualização da propagação do <i>label</i>	130
Figura 62 – Análise do Sub-contesto 2, frequência de ocorrência mínima 25 e visualização da propagação do <i>label</i>	131
Figura 63 – Análise do Sub-contesto 3, frequência de ocorrência mínima 25 e visualização da propagação do <i>label</i>	131
Figura 64 – Análise de similaridade de formas ativas de 2.870 resumos de patentes sobre EAD	134
Figura 65 – Dendograma da CHD das descrições das invenções nas patentes sobre EAD	137
Figura 66 – AFC das descrições das invenções das patentes sobre EAD	139
Figura 67 – AFC das descrições das invenções das patentes sobre EAD com as formas ativas que as compõe	140
Figura 68 – AFC das descrições das invenções das patentes sobre EAD com as formas suplementares que as compõe	140

Figura 69 – AFC das descrições das invenções das patentes sobre EAD com os cabeçalhos, indicando países depositantes, depositantes e o código IPC.....	141
Figura 70 – Dendograma e sub-contexto 1 do contexto 1 com as formas componentes	142
Figura 71 – Dendograma e sub-contexto 2 do contexto 2 com as formas componentes	142
Figura 72 – Dendograma e sub-contexto 3 do contexto 3 com as formas componentes	143
Figura 73 – Análise do Sub-contexto 1, frequência de ocorrência mínima 25 e visualização do tipo propagação do <i>label</i>	144
Figura 74 – Análise do Sub-contexto 2, frequência de ocorrência mínima 25 e visualização do tipo propagação do <i>label</i>	145
Figura 75 – Análise do Sub-contexto 3, frequência de ocorrência mínima 25 e visualização do tipo propagação do <i>label</i>	146
Figura 76 – Análise de similaridade de formas ativas de 344 descrições das invenções das patentes sobre EAD	148
Figura 77 – Dendograma e sub-contexto 1 do contexto 1 com as formas componentes	152
Figura 78 – Dendograma e sub-contexto 2 do contexto 2 com as formas componentes	152
Figura 79 – Análise do Sub-contexto 1, frequência de ocorrência mínima 25 e visualização do tipo propagação do <i>label</i>	154
Figura 80 – Análise do Sub-contexto 2, frequência de ocorrência mínima 25 e visualização do tipo propagação do <i>label</i>	155
Figura 81 – Análise de similaridade de formas ativas de 322 reivindicações presentes nas patentes sobre EAD	157
Figura 82 – <i>Clusters</i> dos resumos das patentes sobre EAD usando o algoritmo <i>Lingo</i> e a visualização <i>Foam tree</i> (88 clusters)	162
Figura 83 – <i>Clusters</i> dos resumos das patentes sobre EAD usando o algoritmo <i>K-means</i> e a visualização <i>Foam tree</i> (25 clusters)	163
Figura 84 – <i>Clusters</i> dos resumos das patentes sobre EAD usando o algoritmo <i>K-means</i> e a visualização Aduna (25 Clusters).....	164
Figura 85 – Parte superior e inferior da tela que demonstra a relação entre as universidades solicitantes de patentes em EAD e o país de origem dessas instituições .	166
Figura 86 – Parte superior e inferior da tela que demonstra a relação entre as universidades solicitantes de patentes em EAD e o país onde a patente foi solicitada...	167

Figura 87 – Parte superior e inferior da tela que demonstra a relação entre as universidades solicitantes de patentes em EAD e o país onde a patente foi solicitada...168
Figura 88 – Gráfico de linhas representando a evolução temporal das solicitações de patenteamento de tecnologias voltadas ao EAD por universidades.....168
Figura 89 – Gráfico de área representando a tecnologia em EAD com os maiores números de patentes solicitadas por universidades, com base no indicador IPCR7.....169
Figura 90 – Parte superior e inferior da tela que demonstra a relação entre as universidades solicitantes de patentes em EAD, o número de citações recebidas pelos documentos, e a tecnologia presente nas solicitações de patenteamento.170

LISTA DE TABELAS

Tabela 1 – Categorias das Fontes de Vantagem Competitiva em IES.....	31
Tabela 2 – Fatores críticos de Sucesso EAD.....	33
Tabela 3 - Antecedentes Institucionais e Respostas Institucionais Previstas	44
Tabela 4 – Principais desinências da língua portuguesa.....	62
Tabela 5 – Saturação e nível de melhoria das buscas.....	68
Tabela 6 - Resultados da CHD dos resumos das patentes sobre EAD.	120
Tabela 7 – Síntese de análise de conteúdo dos resumos de patentes sobre EAD.	135
Tabela 8 – Resultados da análise de CHD das descrições das invenções nas patentes sobre EAD.	136
Tabela 9 – Síntese de análise de conteúdo das descrições das invenções de patentes sobre EAD.....	149
Tabela 10 – Resultados da CHD das reivindicações das patentes sobre EAD.....	150
Tabela 11 – Síntese de análise de conteúdo dos resumos de patentes sobre EAD	158
Tabela 12 – Consolidação das análises de conteúdo das patentes de EAD: Resumo, Descrições e Reivindicações	159

RESUMO

O ensino a distância tem uma longa história de sucessos e fracassos, existe pelo menos desde o final do século XVIII. O ensino superior a distância iniciou no Brasil em meados de 1994, tendo como principal fator a expansão da internet. A busca de inovações e novos modelos relacionados ao processo de ensino a distância (EAD) torna-se importante, tanto do aspecto operacional como estratégico. Face a estes desafios, as informações existentes nos bancos de dados de patentes podem contribuir de forma significativa para a definição de estratégias de EAD em instituições de ensino superior (IES), portanto, o objetivo da tese foi analisar o uso do *Knowledge Discovery in Database* (KDD) de informações patentárias e suas possíveis contribuições para o EAD em IES. A metodologia utilizada foi a estrutura do KDD para exploração, análise, seleção, pré-processamento, limpeza, transformação, *data mining*, interpretação e avaliação de dados de informações patentárias sobre EAD da base do *European Patent Office* (EPO) que possui aproximadamente 90 milhões de documentos. A coleta dos dados utilizou-se o emprego de *data mining* por meio do *software crawler* Patent2Netv.2. A amostra de patentes adquiridas com o uso de aprimoradas expressões de busca resultou em 3.090 patentes, que foram analisadas por meio de tabelas dinâmicas, análises de rede, mapas mentais, análise de conteúdo e *clustering*. Os principais resultados: (1) possibilitaram apresentar o diagnóstico sobre as patentes relacionadas a EAD no mundo; (2) o desenvolvimento de uma metodologia de uso do KDD para análise de conteúdo de informações patentárias em EAD para IES; (3) o mapeamento das patentes em EAD em Universidades; e, finalmente, (4) a avaliação do uso de informações patentárias e sua utilização na definição de estratégias de adoção de EAD em IES, à luz do Visão Baseada em Recursos.

Palavras-chave: Educação a distância, Ensino Superior, Patentes, Estratégia.

ABSTRACT

Distance learning (DL) has a long history of success and failures, and has existed for at least since the end of the XVIII century. Higher education DL began in Brazil during 1994, having the expansion of the internet as the main factor. The search of innovations and new models related to the process of DL has become critical, both from the operational and strategic aspect. Regarding those challenges, the available information in patent databases can contrive add to, in an important manner, the design of DL strategies in higher education institutions (HIE), therefore, the thesis' objective is: to analyze the employment of Knowledge Discovery in Database (KDD) in patent information and its main contributions to DL in HIE. The method employed was the KDD structure to discovery, analysis, selection, pre-processing, filtering, transformation, data mining, interpretation and assessment of patent information data from the European Patent Office's (EPO) database, composed of 90 million documents. The data collection was based on a sample of patents acquired through enhanced search expressions, by the crawler software Patent2Netv.2. The data of 3.090 patents were analyzed by dynamic tables, network analysis, mindmaps, content analysis and clustering. The main results: (1) provided the diagnosis of patents related to DL in a global perspective; (2) developed a methodology for the use of KDD to analyze the content of DL patent information to HIE; (3) mapping of DL patents from HIE; and, ultimately, (4) assessment the use of patent information in order to formulate strategies on adopting DL in HIE, in the light of the Resource-based View.

Keywords: Distance learning, Higher Education, Patents, Strategy.

1 INTRODUÇÃO

Os novos modelos educacionais com seus processos de mediações procuram atender a espaços e tempos diferentes, submetidos também a contextos diversos, como é o caso da educação ou ensino a distância (EAD). Segundo Moore, Kearsley, Galman, & Mello (2007), a educação a distância é um aprendizado planejado que ocorre, na maioria das vezes, com dissociação da dicotomia espaço-tempo. Neste modelo professores e alunos podem estar em lugares e meios diferentes do local do ensino, o que exige técnicas especiais de educação com uso de Tecnologias de Informação e Comunicação e disposições organizacionais e administrativas especiais.

O ensino a distância tem uma longa história de sucessos e fracassos; existe pelo menos desde o final do século XVIII, com um expressivo desenvolvimento a partir de meados do século XIX, quando foi criado o primeiro curso por correspondência (*Correspondence Colleges* – Reino Unido). Atualmente utiliza várias mídias, desde o material impresso até os simuladores *online*, com grande interação entre o aluno e o centro produtor, quer fazendo uso de inteligência artificial quer por comunicação síncrona entre professores e alunos (Duarte, 2011).

No Brasil, o ensino a distância surgiu com a criação do Instituto Rádio Monitor, em 1939 e com o Instituto Universal Brasileiro, a partir de 1941. Já o Ensino Superior a distância iniciou no Brasil em meados de 1994 e dentre os principais fatores que contribuíram para a sua difusão estão a expansão da Internet junto às Instituições de Ensino Superior (IES) e a publicação da Lei de Diretrizes e Bases para a Educação Nacional de 1996 (Brasil, 1996), que possibilitou que o EAD se estabelecesse no cenário nacional. Isso motivou as Instituições de Ensino Superior a iniciar experiências educacionais com a oferta de cursos livres a distância, sobretudo com a aplicação das Tecnologias de Informação e Comunicação (TICs).

O artigo 80 da LDB lei n. 9.394/96 1996 (Brasil, 1996), regulamenta que a educação a distância é uma forma de ensino que possibilita a autoaprendizagem, com a mediação de recursos didáticos sistematicamente organizados, apresentados em diferentes suportes de informação, utilizados isoladamente ou combinados, veiculados pelos diversos meios de comunicação.

A Secretaria de Educação a Distância - SEED, criada pelo MEC, define no Decreto 5.622, de 19.12.2005 (que revoga o Decreto 2.494/98), que regulamenta o Art. 80 da Lei 9394/96 (LDB) Educação a Distância “ como modalidade educacional na qual a mediação didático-pedagógica nos processos de ensino e aprendizagem ocorre com a utilização de meios e tecnologias de informação e comunicação, com estudantes e professores desenvolvendo atividades educativas em lugares ou tempos diversos.”

Diferentemente dos cursos da modalidade presencial, em que o estudante precisa comparecer em local e horário determinados, na Educação a Distância é o estudante quem administra os horários de estudo e o ritmo de aprendizagem. A EAD busca constantemente estreitar as relações entre professores e alunos e tornar mais efetivo o processo de ensino e de aprendizagem por meio dos Ambientes Virtuais de Aprendizagem (AVA).

O EAD contribui para a difusão e de democratização da educação e se caracteriza como uma das melhores opções para a inclusão social e para a melhoria quantitativa e qualitativa do processo educacional, face à limitação do sistema educativo convencional (presencial) de responder às demandas pleiteadas pela evolução da sociedade e dos processos de comunicação (Michael G. Moore & Kearsley, 2011; Michael G. Moore et al., 2007; Michael Grahame Moore, 2013).

Acredita-se que Educação a Distância (EAD) promoveu uma revolução na educação nos últimos anos. No entanto, esta não é uma mudança fácil e nem uniforme - especialmente nos modelos de gestão (Moran, 2007). Essa problemática de gestão da EAD não está equilibrada nas instituições, ainda sem um modelo de gestão definido e claro para a modalidade no Brasil, apresentando significativas rupturas (problemas) de gestão (Roesler, 2008).

O modelo de (Sartori & Roesler, 2005) prediz a estrutura organizacional de um programa de Educação a Distância composta por unidades responsáveis pela administração financeira e acadêmica, pela produção e entrega de materiais didáticos, pelo atendimento pedagógico aos alunos, pelo suporte técnico e informacional, pela pesquisa e avaliação e pela elaboração de novos projetos pedagógicos, entre outros. Os autores discutem que a partir da análise, diagnóstico e prognóstico das ações, estratégias e processos pode ser possível incorporar melhorias e inovação na área de atuação do programa de EAD. Isto contribuiria para gerar um banco de conhecimento que propicie gestão do conhecimento. Isso exige uma nova

postura com relação aos desafios organizacionais contemporâneos, imprescindível para o desenvolvimento das competências organizacionais.

O termo competências organizacionais, mais popular com o desenvolvimento do conceito de *core competences* (competências essenciais), tem sua origem na evolução da chamada visão baseada em recursos (VBR - do inglês, *Resource-based View of the Firm*, ou RBV), cuja perspectiva assume que é a partir da articulação dos recursos que a competência se constitui (Takahashi & Fischer, 2009).

Segundo Crubellate, Pascucci e Grave (2008): “A *Resourced-based View* (RBV) é uma perspectiva explicativa do comportamento estratégico, fundamentada na ideia de que seleção, obtenção e disposição de recursos e desenvolvimento de competências únicas ou de difícil imitação resultam em diferenciação e vantagem competitiva”.

A RBV lança seu foco principalmente sobre suas condições internas, analisando como são obtidos, combinados e aplicados os recursos. Estes são definidos como bens ou inputs tangíveis ou intangíveis que uma organização possui, controla ou a que tem acesso em base semipermanente (Helfat & Peteraf, 2003).

De acordo (Retamal, Behar, & Maçada, 2009), as organizações que optaram pelo desenvolvimento de ações em EAD devem procurar estabelecer o equilíbrio da utilização dos recursos e potencializá-los visando a sua maximização, com base na RBV. A utilização destes recursos de forma diferenciada por uma determinada empresa pode fazer com que ela obtenha vantagem competitiva em relação a seus concorrentes.

Nesse sentido, as patentes podem ser consideradas fontes de vantagem competitiva em vários setores. Os dados de patentes podem mostrar alterações na estrutura e no desenvolvimento de atividades criativas de um país nas tecnologias, na indústria e nas empresas. As informações patentárias também podem indicar as mudanças de dependência de determinadas tecnologias, além de sua disseminação e penetração científica, técnica e, em última instância, mercadológica (OECD, 2002).

De acordo com (Quoniam, Kniess, & Mazzieri, 2014), as patentes podem ser consideradas como valiosas ferramentas para a disseminação da informação, podendo ser

utilizada como: (a) fonte de dados para os indicadores do grau de desenvolvimento tecnológico e econômico; (b) fonte de acompanhamento da evolução tecnológica; (c) fonte de identificação detentores de tecnologias concorrentes, tendências tecnológicas e mercados potenciais.

Neste contexto, face aos desafios da implementação do EAD em IES, esta pesquisa tem como fio condutor a análise e o uso de informações patentárias para definição de estratégias de adoção de tecnologias/metodologias para o EAD em Instituições de Ensino Superior.

1.1 PROBLEMATIZAÇÃO

A globalização proporcionou diversos efeitos no cenário competitivo de organizações na indústria de ensino superior. Houve um aumento significativo do acesso, mais pessoas ao redor do mundo estão ingressando no grau superior. Isso gerou um acirramento da competição, que precisou atender a novos níveis de demanda, não só no quesito acesso, mas também na qualidade do serviço (Van der Wende, 2003).

A demanda por maiores padrões de qualidade no serviço prestado no setor do ensino superior fez com que as organizações desse setor reavaliassem as bases de sustentação de suas vantagens competitivas, trazendo à tona novos diferenciais e o aprimoramento dos já existentes. Posicionamentos em conceitos defasados e ultrapassados não garantirão a continuidade do diferencial competitivo que muitas instituições de ensino se embasavam, sendo que “a crescente demanda por educação superior não é apenas uma demanda por mais do mesmo” (Van der Wende, 2003, p. 203). Foi necessário que a estrutura do setor de ensino superior se modificasse para dar espaço às novas maneiras e posicionamentos de competição.

As novas bases da competição, tais como educação a distância e *e-learning* (Salmon, 2005), internacionalização (Altbach & Knight, 2007), inovações de currículo (Fallows & Steven, 2000) etc. permitiram que instituições de ensino oferecessem novos padrões de ensino, proporcionando uma melhor educação e aprendizagem. O advento da tecnologia aplicada ao ensino também permitiu a criação de materiais didáticos dinâmicos e interativos trazendo novas maneiras de como o professor pode se relacionar com o aluno, maximizando o potencial dessa relação (Sharpe, Benfield, & Francis, 2006). A tecnologia, adicionalmente, trouxe novas maneiras de geração de informação para o apoio à tomada de decisão que faz com que seja

possível identificar padrões que demonstrem deficiências tanto no processo de educação (atividade primária) quanto em outras atividades de apoio das instituições de ensino (Fatimah, Gazi, & Saedah, 2010).

A educação a distância apresenta-se como uma forma de adaptação das instituições de ensino superior às exigências do mercado. Pode ser considerada uma estratégia competitiva utilizada pela IES para atender a demanda das classes de baixa renda, já que possibilita a oferta de cursos a preços inferiores aos praticados na modalidade presencial. Outra vantagem dessa modalidade é expandir geograficamente a atuação da IES e desta forma que os estudantes das mais diversas regiões (Duarte, 2011).

O uso educativo das tecnologias pode ser considerado uma perspectiva conceitual que orienta a análise da educação a distância. Se por um lado EAD é uma modalidade de ensino que responde aos desafios das sociedades contemporâneas, por outro, deve ser considerada, do ponto de vista conceitual, parte de um processo educacional mais amplo (Belloni, 2002).

Segundo Belloni, 2005, a integração das TIC à educação exige considerar dois aspectos: a dupla dimensão de tal integração, ou seja, que é preciso considerar as mídias como objetos de estudo (aspectos éticos e estéticos) e como ferramentas pedagógicas (aspectos pedagógicos e didáticos). O segundo aspecto é que o uso adequado das TICs promove e exige uma abordagem interdisciplinar da educação, pois requer que o professor domine o uso das ferramentas e trabalhe coletivamente com profissionais de diferentes áreas.

A busca de inovações e novos modelos relacionados ao processo de ensino a distância torna-se cada vez mais importante, tanto do aspecto operacional como estratégico. Face a estes desafios, as informações existentes nos bancos de dados de patentes podem contribuir de forma significativa para a definição de estratégias de EAD em IES. Neste sentido, este trabalho busca responder a seguinte questão de pesquisa: **“Como o uso do KDD de informações patentárias pode contribuir com o ensino a distância em instituições de ensino superior?”**

1.2 OBJETIVOS DO ESTUDO

Os objetivos geral e específicos da tese são descritos a seguir.

1.2.1 Objetivo geral

O objetivo geral desta tese é analisar o uso do KDD de informações patentárias e suas possíveis contribuições para o Ensino a Distância em instituições de ensino superior.

1.2.2 Objetivos específicos

Os objetivos específicos do trabalho são:

- ✓ Apresentar o diagnóstico sobre as patentes relacionadas ao Ensino a Distância no mundo;
- ✓ Desenvolver uma metodologia, com o uso do KDD, para análise de conteúdo de informações patentárias em EAD para IES;
- ✓ Descrever o mapeamento das patentes em EAD em Universidades;
- ✓ Avaliar como o uso de informações patentárias podem ser utilizadas na definição de estratégias de adoção de Ensino a Distância em IES, à luz do RBV.

1.3 JUSTIFICATIVAS

O mercado global de ensino a distância foi estimado em 165 bilhões de dólares em 2014 e deve manter uma taxa de crescimento na ordem de 5,08% anual e alcançar o valor de 243 bilhões de dólares em 2022 (Statistics MRC, 2015).

Segundo dados de 2012, os Estados Unidos estão na liderança no número de alunos matriculados em cursos em EAD, com 7,1 milhões que cursaram alguma disciplina na modalidade a distância em 2012 (Straumsheim, 2014). Dentre as regiões do mundo que mais crescem, se destacam Ásia (17,3%), Europa Oriental (16,9%), África (15,2%) e América Latina (14,6%). Especificamente na América Latina o faturamento do ensino a distância deve dobrar em um período de 5 anos, alcançando patamares de 2,29 bilhões de dólares em 2016, considerando que em 2011 esse número era de apenas 1,16 bilhão (Docebo, 2014). Dentre os países que compõem a América Latina, o Brasil possui o maior crescimento projetado (21,5%), seguido pela Colômbia (18,6%), Bolívia (17,8%) e Chile (14,4%).

No Brasil, os dados do Censo da Educação Superior de 2012 demonstraram que o EAD no Brasil encerrou o ano com 1,2 milhão de alunos matriculados, ante um total de 7 milhões de matrículas, perfazendo cerca de 17% das matrículas. O crescimento de ingressantes na modalidade a distância foi de 12,2%, enquanto na educação presencial o crescimento médio foi de 4,4% comparando o ano 2012 com o ano de 2011 (Ministério da Educação, 2013). Dentre as escolhas dos ingressantes no ensino superior que optaram pela modalidade a distância, os cinco cursos que mais se destacam, em ordem de procura, foram: (1) pedagogia; (2) administração; (3) serviço social; (4) competências gerenciais e (5) ciências contábeis. A demanda também pode ser representada por curso, neste caso atingindo o número de 390,67 alunos por curso a distância no Brasil em um total de 1.773 cursos. (ABED, 2013)

Tal crescimento e destaque da educação a distância fez com que o Brasil se voltasse para a modalidade como ponto de apoio para o cumprimento das metas do Plano Nacional de Educação - PNE (Ministério da Educação, 2014). O PNE foi confeccionado no começo da década de 2010, porém apenas publicado oficialmente em 2014. É um plano composto por vinte metas que devem ser cumpridas até 2020 para o aprimoramento da nação e é um dos pilares fundamentais do conjunto de políticas públicas denominado de “Pátria Educadora”. Em especial, é importante destacar a meta 12: “elevar a taxa bruta de matrícula na educação superior para 50% (cinquenta por cento) e a taxa líquida para 33% (trinta e três por cento) da população de 18 (dezoito) a 24 (vinte e quatro) anos, assegurada a qualidade da oferta e expansão para, pelo menos, 40% (quarenta por cento) das novas matrículas, no segmento público”. A taxa de matrícula bruta mencionada na meta é a razão do número de alunos matriculados em relação à população que se encontra na faixa etária de 18 a 24 anos (faixa etária relativa ao ensino superior). Já a taxa líquida indica a razão do número de alunos matriculados pertencentes à faixa etária (18 a 24 anos) em relação ao total da população na faixa etária de 18 a 24 anos. Ou seja, a taxa bruta mensura o número de alunos matriculados no ensino superior e a taxa líquida mensura o número de alunos matriculados que pertencem à faixa etária considerada ideal do ensino superior (18 a 24 anos). Para o cumprimento desta meta, o governo brasileiro já declarou que se apoiará fortemente na modalidade a distância.

Além da sua relevância como política pública, a educação superior também possui destaque como setor produtivo na economia brasileira. Mais de 80% dos cursos a distância incidem direta ou indiretamente sobre a formação de profissionais para o mercado de trabalho

(ABED, 2013). A educação superior privada movimentou, em 2012, um total de 28,3 bilhões de reais e tem um faturamento estimado de 32 bilhões e a expectativa é que este faturamento aumente para 35,9 bilhões em 2014 e 40,1 em 2015 (HOPER Educação, 2015). No Brasil, há 2.391 instituições de ensino superior, sendo 2.090 privadas e 301 públicas responsáveis por boa parte dos 1.790.442 profissionais no mercado formal de trabalho do setor de ensino como um todo (HOPER Educação, 2015).

A utilização de plataformas virtuais vem mudando o panorama das salas de aula. Neste sentido, a interação entre aluno e professores aumentou significativamente. Em uma sala de aula presencial, em média, um professor faz 3 perguntas por hora; porém, em uma sala de aula virtual, esse número aumenta para 21 perguntas por hora, gerando 117 respostas individuais de alunos por hora (Capper, 2001).

A educação superior está ainda em expansão no Brasil e grande parte deste crescimento é devido à influência positiva da educação a distância. A EAD é um fator estratégico tanto para o governo quanto para o setor privado de educação superior. O uso de tecnologias permite com que as interações entre alunos e professores sejam mais intensas e produtivas, eliminando eventuais barreiras de comunicação e fazendo com que estejam mais próximos do que em uma sala de aula presencial. A educação a distância possui, além de sua relevância econômica, relevância estratégica para a formação de profissionais de trabalho mais capacitados e de cidadãos questionadores e conscientes. O foco deste trabalho é colocar em discussão a educação a distância como fator econômico, social e tecnológico no contexto global e analisá-la como um fenômeno amplo e complexo, porém de potencial analítico, e sua aplicabilidade e pertinência ao mercado brasileiro de ensino superior.

É sabido que organizações buscam fontes de vantagem competitiva dentro de seu ambiente interno para adquirirem uma vantagem significativa perante seus concorrentes, com o objetivo da maximização de receita e redução do custo gerando o máximo de valor percebido ao seu cliente. Portanto, instituições de ensino superior buscam novos diferenciais oriundos da mobilização de recursos e capacidades de seus ambientes internos para atrair mais alunos, receitas e prestígio. Neste contexto, parte-se do pressuposto que informações patentárias, aliadas ao *Resource-based View* (RBV), podem ser utilizadas na definição de estratégias de Ensino a Distância em instituições de ensino superior.

Neste sentido, este trabalho apresenta como contribuição a relação entre os três elementos de pesquisa: “ensino a distância”, “*Resource-based View*” e “patentes” em prol da melhoria da tomada de decisão e competitividade das instituições de ensino superior.

1.4 ESTRUTURA DO TRABALHO

Quanto à forma, esta tese está estruturada em cinco capítulos: introdução, revisão da literatura, metodologia da pesquisa, apresentação dos resultados e conclusões.

O capítulo I contempla a introdução, problemática de pesquisa, justificativa e objetivos do trabalho. O capítulo II é composto pela revisão teórica sobre os principais tópicos relacionados ao trabalho, dentre os quais se destacam: instituições de ensino superior, *Resource-based View* e informações patentárias. O capítulo III detalha os procedimentos metodológicos utilizados durante a pesquisa, como foco na obtenção e análise das informações disponíveis no banco de dados de patentes. O capítulo IV apresenta os resultados obtidos. O capítulo V expressa as conclusões da pesquisa embasadas pela teoria correlata e as recomendações.

2 REVISÃO DA LITERATURA

A revisão de literatura desta tese foi estruturada em torno de três lentes teóricas. A primeira diz respeito às Instituições de Ensino Superior – IES e o uso do EAD por elas; A segunda descreve a estrutura e evolução da estratégia, juntamente com a RBV, Teoria institucional e a relação entre RBV e Teoria Institucional e, a terceira, versa sobre as Informações Patentárias.

2.1 INSTITUIÇÕES DE ENSINO SUPERIOR

O ensino e aprendizagem nos dias atuais contam com o suporte de tecnologias e ambientes virtuais que podem auxiliar o aluno. Isso implica que o gestor universitário deve estar aberto ao pensar do ensino, o que inclui considerar a possibilidade de uso de tecnologias (Laurillard, 2013), além de estar atento a escolher, dentro de seus preceitos didáticos da instituição de ensino, qual tecnologia adotar. Além da escolha sobre qual tecnologia seria a mais adequada, há ainda a governança sobre os aparatos de tecnologia de informação na tomada de decisões de um gestor de universidade (Wilmore, 2014).

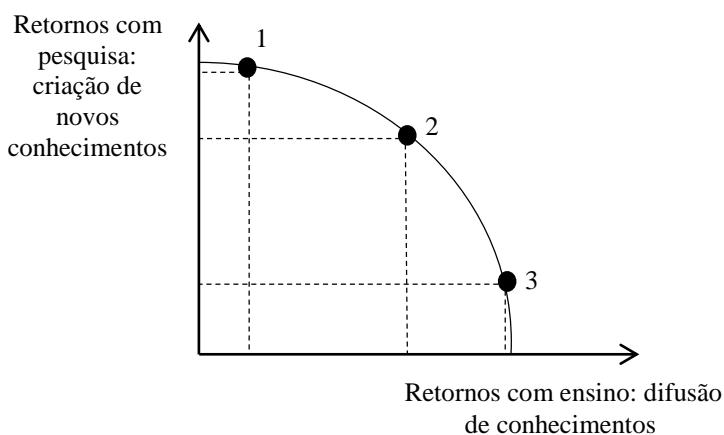
É importante que o gestor tenha conhecimento e habilidades para lidar com tecnologias contemporâneas capazes de oferecer produtividade e ações de vanguarda para as atividades de ensino e de trocas e geração de conhecimentos. Ou seja, deve-se ter cuidado para escolher ferramentas não só de ensino, mas também de pesquisa, que costumam ter nível técnico bastante específico e custos elevados.

2.1.1 *Trade-offs* entre Ensino, Pesquisa e Extensão

A gestão complexa da universidade requer ainda que existam escolhas sobre a alocação de recursos. O gestor universitário deve articular essas escolhas sob diversos interesses distintos (Éster & Melo, 2008). A grande questão é que algumas dessas escolhas trazem *trade-offs* ao gestor de uma universidade, no qual sua escolha deve se pautar conforme seus objetivos e prioridades.

Ou seja, ainda que exista uma curva de utilidade sobre a melhor maneira de se investir os recursos da uma organização universitária, as relações estipuladas que sustentam essa curva por muitas vezes dependem dos critérios pessoais do gestor. Mesmo que se considere que ele deva tomar suas decisões nos limiares de uma curva de utilidade, cabe a ele escolher onde investir ao longo dessa curva. A figura a seguir exemplifica uma curva hipotética a partir do trecho exposto do texto de Santos (1989), demonstrando duas variáveis para a escolha de investimentos na gestão universitária: o ensino e a pesquisa.

Figura 1 – Exemplo de curva de utilidade de uma escolha de investimentos de uma universidade



Fonte: Adaptado de Santos (1989)

Percebe-se, portanto, que existe uma tomada de decisões do gestor da universidade que pode ser influenciada por diversos interesses (seus ou de outros *stakeholders*). Na situação 1, o gestor prefere investir mais em pesquisa, favorecendo seus programas de pós-graduação na geração de novos conhecimentos; na situação 3, ao contrário, o gestor prefere investir mais em ensino de graduação para difusão de conhecimentos já consolidados; na situação 2, o gestor busca um equilíbrio maior entre os retornos obtidos com a pesquisa e ensino.

Alinhar os interesses de diversos interessados em um único ponto dessa curva de utilidade não é uma tarefa fácil. Enquanto pressões de professores podem se direcionar a criação de novos conhecimentos, por outro lado, pressões de estudantes de graduação tencionem a escolha de maiores retornos para a difusão do conhecimento. Isso se agrava ainda mais quando a realidade da universidade envolve também a escolha de investimentos em outras variáveis, tal como as atividades de extensão, por exemplo.

Esse tipo de tomada de decisão é especialmente comum na gestão de pós-graduação das universidades. O foco em pesquisa geradora de conhecimento, função típica dos programas de pós-graduação *stricto-sensu*, está comumente em embate dentre os objetivos e prioridades das universidades. Cabe então ao gestor de organizações universitárias ter a habilidade em escolher a alternativa que possa manter um alto nível de produtividade e geração de conhecimento sem comprometer as funções da universidade ou em detrimento destas, lembrando que suas decisões devem respeitar preceitos morais e éticos (Eyal, Berkovich, & Schwartz, 2011), de forma que se mantenha uma escolha dentro dos objetivos da universidade.

2.1.2 Gestão em Instituições de Ensino Superior

Diante desse ambiente contemporâneo das instituições de ensino superior, cabe aos acadêmicos e aos gestores desenvolverem formas efetivas de gerir as organizações universitárias. Isto é, administrá-la na direção de evitar que o ensino seja colocado de forma subalterna ao mercantilismo do ensino e valorizar a difusão e criação do conhecimento de qualidade como elementos centrais e indissociáveis à sua existência. É um grande desafio, portanto, gerir organizações universitárias dentro do atual contexto brasileiro. Se por um lado, é um grande desafio tornar a universidade atualizada a seu tempo, já que ela é enrijecida e não interveniente (Sampaio & Laniado, 2009; Santos, 1989) em suas decisões gerenciais; por outro lado, não se pode cair nos modismos mercantilistas das universidades, no qual as tradições retrógradas interferem na obtenção do seu objetivo central da universidade que é transferir e gerar conhecimento. Por isso a importância em se ter uma gestão universitária capaz de manter os objetivos centrais da universidade mediante a crescente complexidade de demandas e requerimentos do ambiente externo à instituição de ensino, independentemente de ser uma organização privada ou pública.

Importante ressaltar que não se trata de algo contra o modelo de lucro das universidades, mas sim da busca do lucro em detrimento da qualidade do ensino e do comprometimento da geração de conhecimento em pesquisa. Daí a importância em se ter uma profissionalização da gestão universitária.

2.1.3 Vantagem competitiva em Instituições de Ensino Superior

As fontes da vantagem competitiva em instituições de ensino superior podem ser dividida em 12 categorias: Aprendizagem Organizacional; Tecnologia da Informação; Marca; Formação de Alianças; Cultura Organizacional; Inovação; Currículo; Desenvolvimento de Docentes; Infraestrutura; Qualidade do Corpo Discente; Pesquisa e Doações.

A primeira categoria de fontes de vantagem competitiva, aprendizagem organizacional, refere-se ao uso do conhecimento obtido pela organização para aprimoramentos que tragam vantagem competitiva para a IES (Mazzarol & Soutar, 1999; Secundo, Margherita, Elia, & Passiante, 2010). Tecnologia da Informação (TI) envolve tanto o seu uso como ferramenta de suporte para atividades organizacionais, como instrumento no fluxo de informações e meio de navegação na internet (Clarke, 1997; Dyson, 2004; Mazzarol & Soutar, 1999). A marca de uma instituição de ensino superior é uma fonte importante de vantagem competitiva, pois é um tipo de recurso que possui lastro à história e posicionamento da organização e, portanto, possui alto potencial de difícil imitação (Dyson, 2004; Hemsley-Brown & Oplatka, 2006; Ivy, 2001; Mazzarol & Soutar, 1999). IES também utilizam como fonte de vantagem competitiva a formação de alianças estratégicas com outras organizações (Dyson, 2004; Mazzarol & Soutar, 1999) e o uso de doações (Dyson, 2004).

Outros fatores internos também são utilizados como fontes de vantagem competitiva: cultura organizacional (Dyson, 2004; Mazzarol & Soutar, 1999; Secundo et al., 2010), inovação (Mazzarol & Soutar, 1999), currículo (Clarke, 1997), desenvolvimento de docentes (Dyson, 2004; Secundo et al., 2010), infraestrutura (Dyson, 2004), qualidade do corpo docente (Dyson, 2004; Secundo et al., 2010) e pesquisa (Dyson, 2004; Marginson, 2006). A Tabela 1 sintetiza os estudos que evidenciam fontes de vantagem competitiva em IES.

Tabela 1 – Categorias das Fontes de Vantagem Competitiva em IES

	Mazzarol & Soutar (1999)	Clarke (1997)	Dyson (2004)	Hemsley-Brown & Oplatka (2006)	Ivy (2001)	Marginson (2006)	Secundo et al. (2010)
Aprendizagem Organizacional	X						X
TI	X	X	X				
Marca	X		X	X	X		
Formação de Alianças	X		X				
Cultura Organizacional	X		X				X
Inovação	X						
Curriculum		X					
Desenvolvimento de Docentes			X				X
Infraestrutura			X				
Qualidade do Corpo Discente			X				X
Pesquisa			X			X	
Doações			X				

Fonte: Elaborado pelo autor

Para a elaboração da Tabela 1, foram definidas, com base na análise dos estudos, 12 categorias de fontes de vantagem competitiva. Para cada estudo, é elencado quais fatores de vantagem competitiva foi evidenciado em relação às categorias comuns identificadas.

2.2 O USO DE EAD EM INSTITUIÇÕES DE ENSINO SUPERIOR

Um dos pontos importantes no uso do ensino a distância em instituições de ensino superior é a maximização do uso da tecnologia a favor da funcionalidade da educação (Moore, 2013). Para tanto, além de demandar uma abordagem sistêmica (Moore & Kearsley, 2011), alguns acadêmicos defendem que EAD deve ser vista também como um sinônimo do emprego de tecnologia educacional para o ensino e que o ensino deve ser também entregue pelos canais eletrônicos: *e-learning* (Guri-Rosenblit, 2005).

O perfil dos alunos engajados em cursos de educação a distância difere dos perfis dos demais alunos matriculados em cursos presenciais. Alunos EAD possuem uma maior proficiência de leitura e escrita, aprendem de maneira independente e possuem domínio de funções básicas de computação (Kerr, Rynearson, & Kerr, 2006). Tais alunos demandam um ritmo próprio de estudo e respondem melhor à orientações frequentes dos professores e às

instruções multimidiáticas, ou seja, que possuam diversos meios (vídeo, texto, imagens *etc.*) (Liaw, Huang, & Chen, 2007).

Os alunos EAD podem estar sujeitos a problemas inerentes da sua relação entre professores e tecnologias. É documentado que a experiência de ensino dos alunos foi impactada de maneira grave por falhas de comunicações entre seus professores e por dificuldades técnicas na plataforma de ensino (Hara, 2000). A severa ruptura da experiência de ensino dos alunos em plataformas EAD está diretamente ligada à evasão. Os alunos evadidos apresentam níveis de satisfação menores do que alunos que concluíram o curso e obtiveram seus diplomas (Levy, 2007).

Instituições de ensino devem estar cientes de que para serem bem sucedidas em suas empreitadas na educação a distância, devem estar atentas a certos fatores que se demonstraram cruciais. Os fatores críticos de sucesso expostos na Tabela 2 foram selecionados na literatura científica especializada, sendo que a maioria foi obtida por meios quantitativos estatísticos de amostras pertinentes.

Tabela 2 – Fatores críticos de Sucesso EAD

Fator	Fonte(s)
Competência Docente	(Soong et al., 2001); (Selim, 2007)
Atitude Docente	(Soong et al., 2001); (Sun et al., 2008); (Selim, 2007)
Colaboração	(Soong et al., 2001); (Selim, 2007)
Suporte de plataforma	(Soong et al., 2001); (Selim, 2007)
Expectativa de Aprendizagem	(Sun et al., 2008)
Flexibilidade	(Sun et al., 2008)
Qualidade	(Sun et al., 2008); (Liaw, 2008)
Pertinência	(Sun et al., 2008); (Liaw, 2008)
Facilidade de Uso	(Sun et al., 2008); (Liaw, 2008)
Diversidade de Conteúdos e Atividades	(Sun et al., 2008); (Liaw, 2008)
Satisfação	(Liaw, 2008)

Fonte: Elaborado pelo autor

Os fatores críticos de sucesso foram resumidos e enquadrados em categorias amplas para melhor comparação e sintetização. Competência docente pode ser explicada como a habilidade que os professores atuantes na plataforma de ensino possuem em promover uma experiência de ensino para seus alunos (Soong, Chan, Chua, & Loh, 2001). O nível de titulação do corpo docente dedicado a alunos EAD também influencia o nível de competência (Selim, 2007).

Um outro fator relacionado aos docentes EAD, atitude docente, é definido como a atitude dos professores atuantes na plataforma de ensino com relação ao processo de ensino que ocorre na plataforma (Soong et al., 2001). A atitude está inteiramente ligada às crenças que o docente possui na educação a distância (Selim, 2007). Professores que ainda são céticos em relação a essa modalidade de ensino tendem a possuir atitudes negativas com relação a EAD.

A colaboração deve ser vista como um pré-requisito da plataforma EAD. Alunos, ao serem estimulados a interagirem uns com os outros e compartilharem experiências, atingem

níveis superiores de engajamento (Soong et al., 2001). Além de se dispor de uma plataforma com funcionalidades colaborativas, IES devem estimular seus docentes para que permitam que um ambiente colaborativo emerja.

Muitos alunos possuem dificuldades ao acessarem e utilizarem a plataforma para seus estudos. Para tanto, é crítico que IES providenciem suporte técnico para seus alunos e professores (Selim, 2007). O suporte à plataforma deve permitir com que alunos e professores desfrutem das funcionalidades e benefícios da educação a distância de maneira íntegra e funcional (Soong et al., 2001).

IES devem ajustar sua plataforma de ensino para que comportem a expectativa dos alunos EAD com relação ao processo de aprendizagem. A jornada do aluno pela plataforma, desde seus conteúdos e diversas atividades, devem estar claras e coerentes com o que os alunos esperam. Assimetrias tanto no processo de ensino quanto na execução da jornada do aluno na plataforma durante seus estudos impactam na experiência EAD tanto de alunos quanto de professores (Sun, Tsai, Finger, Chen, & Yeh, 2008).

Plataformas de ensino devem também comportar a flexibilidade. Prazos, tipos de atividades e critérios devem ser flexíveis. IES que mantém uma rigidez excessiva no processo de aprendizagem em suas plataformas podem estar em desvantagem e proporcionar problemas na experiência de aprendizagem de seus alunos, impactando também a atitude docente (Sun et al., 2008).

A qualidade de uma plataforma de ensino é um dos fatores mais abstratos, pois qualidade é um conceito subjetivo. A qualidade pode ser analisada pelo ponto de vista funcional (Sun et al., 2008) ou de experiência do usuário (Liaw, 2008). Uma plataforma que tenha todas as suas funcionalidades em perfeito estado possui qualidade, assim como uma plataforma que consiga promover uma boa experiência de consumo também possui qualidade.

Para que alunos possam desfrutar de maneira eficiente e mantenham sua motivação durante o processos de aprendizagem, eles devem estar com níveis elevados de percepção da pertinência (ou utilidade) do curso ou disciplina que estão cursando em EAD (Liaw, 2008).

Plataformas de ensino também devem ser de fácil uso para que professores e alunos possam interagir sem grande esforço ou sobrecarga cognitiva (Sun et al., 2008). Ao permitir com que a interação flua sem grandes barreiras, plataformas conseguem capitalizar todas suas funcionalidades e, consequentemente, obtém níveis elevados de engajamento (Sun et al., 2008).

Alunos e professores conseguem obter uma melhor interação e engajamento quando existem conteúdos disponíveis com uma grande diversidade (Sun et al., 2008). A diversidade de conteúdos abrange não só a natureza midiática (vídeo, áudio, texto, imagem *etc.*), mas também a variedade dos objetivos de aprendizagem e também a contemporaneidade (Liaw, 2008).

Por fim, a satisfação, tanto de professores quanto de alunos, é fundamental para o sucesso de IES engajadas no ensino a distância (Liaw, 2008). Assim como qualidade, satisfação também é um conceito abstrato, pois deriva de subjetividade.

2.2.1 Novas Tecnologias de Ensino para o EAD

Ao se aplicar os recursos das Tecnologias de Informação e Comunicação (TICs) na educação percebe-se um rompimento das barreiras de tempo e espaço (Moran, 2007). Os autores dizem que para aprender não é preciso estar em uma sala de aula presencial, haja vista que estamos vivendo em uma era da educação sem fronteiras. Essa mudança impacta tanto no que se refere às modernas TICs que facilitam a educação a distância, como novas formas de considerar a experiência ao longo da vida como créditos acadêmicos (Maccari & Quoniam, 2008). Dessa forma, percebe-se que o mundo está se adaptando a estas novas formas de se estudar, aprender e de valorizar o que foi aprendido.

Nessa linha, Itacaramby Pardim & Maccari (2014) afirmam que para que todos tenham a possibilidade de optar pela educação ao longo da vida ou que, antes disso, tenham acesso à educação, é preciso que o sistema educacional esteja preparado e que propicie as condições necessárias para ingresso, permanência e conclusão dos cursos nos diversos níveis de ensino. Desta forma, com a significativa ampliação do acesso à educação ocorrida nos últimos anos, isso não significa apenas criar novos prédios escolares e disponibilizar mais vagas, mas, também, flexibilizar a forma de oportunizar o acesso às aulas para atender a um público que não consegue frequentar os cursos regulares. As Instituições de ensino superior (IES) estão

adotando e implementando de forma intensiva as tecnologias de informação e comunicação (TIC's) em seus processos e rotinas. A essa característica soma-se a crescente demanda por educação a distância – EAD e faz com que as IES coloquem esta nova dinâmica em seus processos decisórios e estratégicos.

Ao se falar especificamente da educação a distância (EAD), o ceticismo que havia em relação ao seu uso está se dissipando rapidamente. Davies, Howell, & Petrie, (2010) afirmam que estudos que comparam diretamente salas de aulas presenciais e a distância para determinar o mérito específico das modalidades tiveram um decréscimo significativo em 2007 comparado a 1998, o que pode indicar que EAD está se tornando aceito como uma experiência educacional viável e importante.

Houve também um decréscimo da ênfase em estudos que tiveram foco em problemas de tecnologia, tais como analisar a qualidade da tecnologia da educação a distância e questionamentos da habilidade de educadores em prover uma experiência de aprendizagem aceitável do ponto de vista tecnológico.

As novas TIC's permitem que alunos e professores interajam em plataformas virtuais de aprendizagem num conceito novo denominado de web 2.0. Recentemente há dois temas importantes nas TIC's no cenário de plataformas e IES: criatividade e participação do aluno; e formação de uma identidade *online* (Greenhow, Robelia, & Hughes, 2009).

De acordo com Fiorillo, Maccari, & Martins (2015) quando se trabalha no processo decisório com esta nova dinâmica sobre o uso de EAD pelas IES, para o reconhecimento dos cursos nesta modalidade de educação, é importante cumprir os requisitos legais. Os autores destacam que, das doze competências essenciais requeridas pelo instrumento do MEC para obtenção do reconhecimento de curso superior a distância levantadas na análise do marco regulatório, as quatro principais são: 1) Gestão de Avaliação de Curso, que está diretamente relacionada ao reconhecimento do MEC, sem o qual, o curso não tem validade nenhuma; 2) Gestão Estratégica e de Políticas para a EAD, por sua abrangência e seu caráter dinâmico, inovador e motivador da aprendizagem e de outras competências; 3) Gestão de Recursos Humanos e das Equipes Multidisciplinares, a qual refere-se à manutenção e ao aprimoramento constante de professores, tutores e funcionários técnico-administrativos capacitados com a finalidade principal de manter o aluno nas instituições de ensino, inseridas em um cenário de

competitividade acirrada; 4) Gestão e o Planejamento da Infraestrutura Física e Tecnológica, considerando-se o avanço das Tecnologias de Informação e Comunicação, responsáveis pela disseminação da EAD por todo o país.

2.3 ESTRUTURA E EVOLUÇÃO DA ESTRATÉGIA

Uma maneira de se examinar e compreender como um campo da ciência se desenvolveu ao longo do tempo é aplicando análises e métodos bibliométricos. Estudos bibliométricos são utilizados para mapear a ciência e elucidar uma maior compreensão sobre a literatura científica. Dentre os estudos notórios da administração, seis estudos foram mapeados e são apresentados a seguir. Ressalta-se que os estudos analisam o campo da estratégia, porém destaque também é dado para o papel da RBV.

No primeiro estudo bibliométrico aplicado à área de estratégia, evidencia-se a necessidade de maturidade do campo teórico (Ramos-Rodríguez & Ruíz-Navarro, 2004). A maioria dos documentos citados pelos artigos analisados pelo estudo são livros que podem ser considerados clássicos da literatura de administração com destaque para a confirmação de Michael E. Porter como o autor mais influente e que a RBV é a contribuição mais importante para o campo da estratégia. Por ser um marco na literatura, o estudo de (Ramos-Rodríguez & Ruíz-Navarro, 2004) impulsionou outros estudos bibliométricos. Dentre os quais, destaca-se como continuação da análise o estudo de Serra, Ferreira, de Almeida, & de Souza Vanz (2012), que expandiu o escopo temporal, porém constatou quase as mesmas conclusões que o estudo inicial. Os autores encontraram também fortes influências da teoria dos custos de transação sob a literatura de estratégia. Por fim, um estudo amplo de 26 anos de literatura de estratégia conclui que a evolução do pensamento acadêmico da estratégia se manifesta como uma integração dos níveis estratégicos corporativos e competitivos, transcendendo quaisquer noções de hierarquias estratégicas (Furrer, Thomas, & Goussevskaia, 2008).

A evolução do pensamento estratégico migra para uma maior integração entre as diferentes escolas teóricas, mas também dá um maior foco às questões internas das organizações. A pesquisa em estratégia, na sua conjuntura atual, é denominada metaforicamente como um pêndulo. Nos primórdios do campo, os trabalhos seminais focaram em uma perspectiva contingencial, na qual era necessário que a organização encontrasse um “*fit*” ideal

entre seu ambiente interno e externo. Os trabalhos seminais posteriores foram fortemente influenciados pela escola da organização industrial, ramo de estudo da economia, quantificando e objetivando o posicionamento estratégico das organizações em seu ambiente externo, dando prioridade para este último. Nos anos mais recentes, os trabalhos seminais voltaram-se novamente para o ambiente interno das organizações, mais especificamente os seus recursos e capacidades; ou seja, a metáfora do pêndulo explica tal retorno ao foco inicial como o movimento de um pêndulo que vai e volta (Hoskisson, Hitt, Wan, & Yiu, 1999).

2.3.1 Visão Baseada em Recursos

A literatura sobre visão baseada em recursos é extensa e se encontra na segunda década de sua existência. É importante ressaltar a relevância dos estudos bibliométricos para a análise e compreensão de um determinado campo da ciência ao longo de um determinado período. O primeiro estudo bibliométrico relevante sobre estratégia foi conduzido por (Ramos-Rodríguez & Ruíz-Navarro, 2004) que selecionou como amostra todos os artigos publicados no *Strategic Management Journal* desde 1980, o ano de sua fundação. Seguindo a mesma abordagem, (Serra et al., 2012), continuaram a análise dos artigos publicados no *SMJ* entre 2001 e 2007. À respeito da visão baseada em recursos (RBV), Ferreira et al. (2014) analisou bibliometricamente a influência da RBV na literatura de negócios internacionais entre 1991 e 2010. Há também estudos críticos que reveem o impacto e analisam o futuro da RBV na ciência da estratégia. Barney, Wright, & Ketchen, (2001) e Barney, Ketchen, & Wright (2011) lançam tais olhares dez e vinte anos após a publicação do artigo seminal de Barney (1991). Por fim, Hoskisson, Hitt, Wan, & Yiu(1999) examina o desenvolvimento do campo da estratégia e a influência das correntes teóricas ao longo dos anos. Estes seis trabalhos são considerados o ponto de partida para analisar a literatura e ciência da estratégia pela lente da RBV, suas contribuições, impacto e futuro.

2.3.1.1 Origens da RBV

A visão baseada em recursos (VBR) também chamada de *resource-based view* (RBV), em inglês, é uma abordagem que categoriza e analisa uma empresa como proprietária de recursos que podem gerar vantagem competitiva. O objetivo de observar uma empresa por meio de seus recursos é compreender os elos complexos entre a vantagem competitiva e suas fontes.

A RBV também é uma das teorias mais tangíveis e objetivas da estratégia, pois visa destacar quais recursos são fundamentais para a estratégia e posicionamento da empresa no mercado perante seus clientes e concorrentes.

O primeiro trabalho que utilizou o termo *resource-based view* foi o de (Wernerfelt, 1984), o qual defende que é mais proveitoso analisar empresas pelos recursos do que pelos produtos. Tomando-se por base a análise de recursos, é possível encontrar a otimização das atividades produto-mercado. O autor argumenta que não é nova a perspectiva de abordar uma empresa por meio de seus recursos, tendo suas raízes no trabalho de economistas como Penrose, (1959). Para Wernerfelt (1984), recursos são ativos tangíveis ou intangíveis lastreados semi-permanentemente à uma empresa.

Os recursos se originam como fontes de vantagem competitiva desde que sejam atrativos e a atratividade é o resultado de seu potencial fornecido à organização para criação de barreiras ou mecanismos de isolamento à imitação. Recursos devem ser a base de um *portfolio* de produtos mas também necessitam ser de difícil imitação para que o posicionamento da organização, por meio de seus recursos, seja sustentável a longo prazo (Dierickx & Cool, 1989; Wernerfelt, 1984). Portanto, a RBV traz uma nova maneira de enxergar como organizações obtém vantagem competitiva, não por meio de seu portfólio de posições de produtos no mercado, mas sim pelo portfólio de recursos que dispõem.

A abordagem de analisar recursos como fontes de vantagem competitiva em organizações possui as suas origens no trabalho de (Barney, 1991), no qual o autor estende a definição clássica de recursos dando uma função específica dos recursos: habilitam a empresa a conceber e implementar estratégias que aprimorem sua eficiência e efetividade. Todos os recursos devem ser usados no processo estratégico ou não são considerados recursos, apenas meros ativos da organização. O autor se destacou por seu modelo de análise de empresas: o VRIO, que baseia-se na capacidade de reconhecer e analisar os recursos que podem ser caracterizados como fontes de vantagem competitiva sustentável. Para isso, é necessário que a empresa verifique o potencial desses recursos, levando em consideração quatro atributos essenciais de um recurso a saber: (V) ser valioso para que a empresa possa explorar uma oportunidade ou uma ameaça ambiental; (R) ser raro entre os concorrentes da empresa; (I) ser de difícil imitação pelos concorrentes; e (O) a empresa deve estar organizada para aproveitar

ao máximo, o potencial desse recurso (Barney & Hesterly, 2008). O modelo VRIO se tornou conhecido por seu potencial gerencial de análise e universalidade de aplicação empresarial. Atualmente, a RBV é componente curricular de qualquer curso ou disciplina de estratégia empresarial.

A visão baseada em recursos (RBV) possui um atenção especial para as capacidades de uma organização. As capacidades não são apenas as que as empresas buscam na condição de aprimoramento competitivo, mas também aquelas desenvolvidas por novas empresas que atuam em novos negócios, sem ter como preocupação principal a competitividade. Capacidades podem ser definidas como “recursos acumulados com o tempo e não adquiridos, portanto agregando uma condição de inimitabilidade” (Collis, 1994, p. 146).

Há diferentes naturezas de capacidade, que podem ser definidas como dinâmicas ou substanciais (Sapienza, Autio, George, & Zahra, 2006), mas alguns autores as classificam de maneira similar porém com nomenclatura diferente: nível zero e nível primário (Winter, 2003), simples e dinâmicas (Collis, 1994). Essas distinções geram discussões na própria definição de cada uma delas e na relação entre ambas, provocando debates sobre os antecedentes e os resultados de cada uma das capacidades e sobre sua forma de relacionamento com os diferentes tipos de ambientes competitivos.

Dentre os tipos de capacidade, as dinâmicas devem ser destacadas. Originalmente concebidas por (Teece, Pisano, & Shuen, 1997), são definidas como capacidades que gerenciam a habilidade de uma organização de aprender, adaptar, mudar ou renovar ao longo do tempo. De uma maneira mais tangível, capacidades dinâmicas são um conjunto específico de processos identificados como desenvolvimento de produto, tomada de decisão estratégica e formação de alianças (Eisenhardt & Martin, 2000).

2.3.2 Teoria Institucional

A teoria institucional busca compreender como o contexto institucional, definido como “regras, normas e ideologia da sociedade” (Meyer & Rowan, 1983, p. 84), influencia a racionalização e difusão de burocracias formais na sociedade moderna. Sob este olhar, as regras emergentes do contexto institucional atuam como “mitos” que organizações incorporam para adquirir legitimidade, recursos, estabilidade e uma maior sobrevivência (Meyer & Rowan,

1977). Os mitos rationalizados são aceitos como prescrições de conduta apropriada e, portanto, institucionalizam-se nas organizações, gerando isomorfismo, ou seja, elas assumem a mesma “forma” e estrutura que as demais. Organizações, ao buscarem legitimidade perante o contexto institucional, adaptam suas estruturas de maneira similar, fazendo com que o campo organizacional regido pelo contexto institucional se torne isomórfico (DiMaggio & Powell, 1983).

A legitimação que é buscada por organizações se refere ao “nível que os relatos culturais estabelecidos provem explanações sobre sua existência, funcionamento, jurisdição, e a escassez ou negação de alternativas” (Meyer & Scott, 1983, p. 201). Uma organização completa seria aquela que nenhuma dúvida pode ser levantada. A legitimação de uma organização pode ser subcategorizada em três dimensões, de acordo com (Scott, 1995) (a) regulativa - ocorre perante os órgãos e instituições que possuem poder de regulação e o exercem no mercado em que a organização atua, (b) normativa - refere a uma organização estar em conformidade com as normas e leis de uma sociedade e (c) cognitiva – que é uma aceitação por parte da sociedade e comunidade que reconhecem a organização cognitivamente. . . .

A legitimidade de organizações também é analisada conforme as suas fontes que podem ser definidas como “aqueles que possuem a capacidade de confrontar a organização, não em termos de poder, mas em questão de autoridade cultural” (Meyer & Scott, 1983, p. 202). A importância do estudo da legitimação de organizações é que seu processo geralmente ocorre em paralelo com o processo de institucionalização, fazendo com que elas busquem a legitimação ao se adaptarem e moldarem aos padrões esperados do contexto institucional (Lawrence, Winn, & Jennings, 2001).

Uma das questões mais importantes da teoria institucional é o que faz organizações que compartilham o mesmo ambiente adotarem estruturas e “formas” similares ao buscarem um ajuste (*fit*) com o seu ambiente. Isto leva a dois argumentos chave do isomorfismo. O primeiro diz respeito às organizações se conformarem a mitos rationalizados oriundos da sociedade sobre o que se constitui uma organização. Tais mitos emergem como soluções para amplos problemas organizacionais e são devidamente rationalizados quando se acredita que eles são as soluções apropriadas. Conforme as organizações se adequam a esses mitos, mais se tornam institucionalizadas, levando a um isomorfismo institucional (John W. Meyer & Rowan,

1977). De acordo com (DiMaggio & Powell, 1983), o isomorfismo institucional é amplificado por processos e práticas dentre as organizações que aumentam a difusão dos mitos e dita também qual estrutura organizacional será adotada. O segundo argumento do isomorfismo condiciona que, quando adaptações às pressões institucionais contradizem necessidades internas de ganho de eficiência, as organizações geralmente anunciam que se adaptaram, mas, na realidade, não se adaptaram; elas desvinculam a ação da estrutura para preservar a eficiência organizacional. Organizações desvinculam a sua estrutura formal das atividades de produção quando há um conflito entre elas ou quando sofrem de pressões institucionais divergentes. Tal desvinculação permite com que organizações alcancem a legitimidade que a conformidade aos mitos racionalizados possibilita, mas também se mantenham eficientes tecnicamente nas suas atividades de negócio.

O isomorfismo institucional é decorrente de três tipos de pressões que organizações são submetidas no ambiente institucional: (a) coercitiva, (b) mimética e (b) normativa (DiMaggio & Powell, 1983). Pressões coercitivas são resultantes de relações de poder geralmente entre o estado por meio de seus órgãos regulamentadores, ou entre organizações certificadores e empresas. Pressões miméticas surgem primariamente durante momentos de incerteza, pois elas costumam imitar os concorrentes que são percebidos como influentes ou bem sucedidos. As pressões normativas estão relacionadas com forças que levam à organizações adotarem certas práticas ou estruturas como a escolha moral correta; e, geralmente, são associadas com pressões das classes profissionais, como consequência dos valores educacionais e de treinamento relacionados à atividade profissional. As três pressões podem ser também pensadas em uma única perspectiva conforme a empresa é impactada. Pressões coercitivas são oriundas de “cima” (do estado), pressões miméticas se originam horizontalmente à organização (dos seus pares), e pressões normativas são provenientes de “baixo” (classes e organizações profissionais).

2.3.2.1 Teoria Institucional Aplicada a Instituições de Ensino Superior

Instituições de ensino superior são organizações que sofrem pressões do ambiente institucional. A sociedade atual emerge de um modelo institucionalizado da qual uma grande variedade de posições sociais são apropriadas em níveis requeridos de conhecimento formal,

fazendo com que as taxas de ingresso no ensino superior ao redor do mundo adquiram uma tendência forte de crescimento (Schofer & Meyer, 2005).

Tais pressões fazem com que instituições de ensino superior se reestruitem como uma resposta adaptativa. O sucesso dessa resposta pode ser visto em termos de desempenho organizacional, IES que se predispõem à reestruturação são mais suscetíveis a um aumento de desempenho e melhores taxas de sobrevivência do que IES que são incapazes ou involuntárias a mudarem (Zajac & Kraatz, 1993). Inclusive há incentivos de que IES dinamizem e adquiram novas maneiras de elaborar o seu planejamento estratégico, como é o caso recente das ações do ministério de educação holandês para combater o isomorfismo de IES holandesas (Maassen & Potman, 1990).

A reestruturação como resposta às mudanças ambientais apresenta evidências divergentes. Há tanto processos convergentes de adaptação, como principal evidência a homogeneização; como processos divergentes, evidenciado por respostas diferentes e regionalizadas (Vaira, 2004). Para isso, é necessário a integração de duas perspectivas teóricas: dependência de recursos e institucionalização (Tolbert, 1985).

2.4 RELAÇÃO ENTRE RBV E TEORIA INSTITUCIONAL

A RBV e a Teoria Institucional podem ser aplicadas de maneira conjunta para a observação e compreensão de diversos fenômenos e situações que as organizações experienciam e podem explicar melhor a vantagem competitiva (Barney et al., 2011). Utilizando como tema principal a Estratégia, o elo entre recursos organizacionais e o contexto institucional pode ser analisado sob a ótica de como que organizações respondem aos contextos organizacionais alocando e remanejando seus recursos. É possível construir teorias sobre os processos subexplorados (a *caixa preta*), do ponto de vista científico, que permeiam entre os recursos de um lado e lucratividade do outro.

Pressões institucionais são exercidas sobre organizações para que a conformidade seja instaurada. Oliver (1991) propõe uma tipologia de respostas estratégicas que variam de acordo com a agência ativa que a organização exerce. Em outras palavras, as respostas organizacionais são elaboradas conforme a liberdade de decisão. As respostas variam entre aquiescer,

comprometer, evitar, desafiar e manipular. A Tabela 3 **Erro! Fonte de referência não encontrada.** mostra as diversas respostas estratégicas e eventuais antecedentes institucionais juntamente com o poder de influência que os antecedentes possuem sobre certas respostas.

Tabela 3 - Antecedentes Institucionais e Respostas Institucionais Previstas

Fator Preditivo	Resposta Estratégicas				
	Aquiescer	Comprometer	Evitar	Desafiar	Manipular
Causa					
Legitimização	Alto	Baixo	Baixo	Baixo	Baixo
Eficiência	Alto	Baixo	Baixo	Baixo	Baixo
Constituintes					
Multiplicidade	Baixo	Alto	Alto	Alto	Alto
Dependência	Alto	Alto	Moderado	Baixo	Baixo
Conteúdo					
Consistência	Alto	Moderado	Moderado	Baixo	Baixo
Restrição	Baixo	Moderado	Alto	Alto	Alto
Controle					
Coerção	Alto	Moderado	Moderado	Baixo	Baixo
Difusão	Alto	Alto	Moderado	Baixo	Baixo
Contexto					
Incerteza	Alto	Alto	Alto	Baixo	Baixo
Interconectividade	Alto	Alto	Moderado	Baixo	Baixo

Fonte: Adaptado de Oliver (1991)

Recursos, num ambiente institucional, podem também ser vistos pela dualidade do ato de exploração. A ação de explorar um recurso pode ser tanto no sentido de aproveitar ao máximo a sua eficiência e otimizá-lo (do inglês *exploitation*), quanto no sentido de descobrir novos recursos ou novas aplicações para recursos já existentes (do inglês *exploration*). Essa dualidade entre *exploitation/exploration* é caracterizada como um contexto no qual organizações se submetem ou são submetidas pelo ambiente institucional (March, 1991). Em ambientes institucionais que favorecem a *exploitation*, há o aproveitamento máximo do potencial que um recurso pode gerar e é caracterizado pelos conceitos de refinamento, escolha, produção, eficiência, seleção, implementação e execução de recursos.

A seguir serão apresentadas as informações patentárias que podem ser utilizadas como recurso estratégico para IES.

2.5 INFORMAÇÕES PATENTÁRIAS

Os ativos intangíveis são todos os que não são físicos ou objetos financeiros. Alguns exemplos são as marcas, a propriedade intelectual, know-how ou segredos industriais. No contexto atual, de negócios baseados no conhecimento, esses ativos têm papel fundamental para as organizações, uma vez que podem gerar riqueza ao ultrapassar o nível de ganho econômico e financeiro em relação aos negócios baseados na força da mão de obra ou em outras formas de trabalho não baseadas no conhecimento (Wiederhold, Tessler, Gupta, & Smith, 2009).

A Propriedade Intelectual engloba o campo de Propriedade Industrial, os Direitos Autorais e outros Direitos sobre bens materiais de vários gêneros, tais como os Direitos Conexos e as Proteções *Sui Generis* (INPI, 2015). Os direitos de Propriedade Industrial medem a produção da atividade inovadora de um país, a partir das invenções. Apesar da invenção em si não ser considerada inovação, existe estreita relação entre patentes e atividades inovadoras. As patentes podem indicar as mudanças de dependência de determinadas tecnologias, além de sua disseminação e penetração científica, técnica e, em última instância, mercadológica (OECD, 2002).

A Propriedade Industrial pode ser considerada como o conjunto de direitos que compreende as patentes de invenção e de modelo de utilidade, os registros de desenho industrial, as marcas e as indicações geográficas (INPI, 2015). Nesse sentido, a Patente, conforme definição de Puhlmann & Moreira (2004), é um título de propriedade temporário outorgado pelo estado, por força de lei, ao inventor/autor (ou pessoas cujo direito dele decorram) para que este possa excluir terceiros, sem sua prévia autorização, da fabricação, comercialização, importação, uso, venda, dentre outras atividades que se utilizem da matéria protegida (exceto para fins de ensino e pesquisa). A mesma definição se aplica ao Registro de Desenho Industrial.

Conforme citado anteriormente, as patentes podem ser classificadas em Patentes de Invenção (PI) ou Patentes de Modelo de Utilidade (MU). As patentes de invenção têm como requisitos a novidade, a aplicação industrial, a atividade inventiva e a suficiência descritiva. O prazo de vigência da Patente de Invenção é de 20 anos a partir da data do depósito. As patentes de modelo de utilidade são objetos de uso prático, ou parte desses, suscetível de aplicação

industrial, que apresente nova forma ou disposição, envolvendo o ato inventivo, que resulte em melhoria funcional no seu uso ou fabricação. As patentes classificadas como Modelo de Utilidade têm o prazo de vigência de 15 anos a partir da data do depósito. (INPI, 2015).

Nesse contexto, para garantir a apropriação dos resultados obtidos a partir do processo constituído de inovações tecnológicas das empresas que investiram em Pesquisa e Desenvolvimento, a sociedade concede, por meio da Patente, a essas organizações um método de proteção legal temporária. A patente permite a exploração da sua inovação em troca da informação detalhada sobre parte substancial do conteúdo técnico contido naquela matéria protegida por lei.

De acordo com Fujino, Stal, & Plonski (1999), o objetivo da propriedade industrial é garantir o direito de exploração comercial desta aos titulares por período determinado de tempo, restringindo o uso não-autorizado por terceiros. Assim, pelo lado da empresa, a Patente consiste em uma reserva de mercado garantida pela mesma durante sua vigência, sobre uma novidade sob o ponto de vista técnico-científico. Para a universidade, o patenteamento e a exploração comercial de determinada tecnologia garante recursos à universidade para o financiamento de novas pesquisas, além da divulgação e aproximação da pesquisa acadêmica com as necessidades de mercado. As Patentes têm ganhado grande notoriedade para a configuração de políticas públicas nas discussões sobre o papel das universidades em um sistema caracterizado por inovações (Haase, de Araújo, & Dias, 2009).

As patentes podem ser consideradas como fontes importantes de informações, uma vez que os resultados tecnológicos apresentados para realização do pedido de patenteamento raramente são replicados em outras publicações.

Os escritórios responsáveis pelos registros de patentes mantêm bancos de dados de livre acesso para a comunidade em geral com descrições completas das invenções apresentadas para depósito. Esses pedidos de patente, aprovados ou não, bem como a descrição das invenções em termos funcionais e de aplicação, permanecem disponíveis nas bases de dados, formando uma valiosa biblioteca tecnológica (Dou, 2010). Desta forma, nessas bases de dados não estão apenas as invenções que foram validadas como título digno, mas também pedidos de registro de patentes de invenções que não foram bem sucedidas. Cada aplicação deverá indicar também

referências ao que já existe, tornando o pedido de registro de patente uma revisão de toda a literatura sobre o assunto reclamado.

Segundo (Quoniam et al., 2014), a patente é vista como uma forma de expressar ao mercado a investigação puramente técnica e tecnológica. A OECD (2002) considera que os dados de patentes podem mostrar alterações na estrutura e no desenvolvimento de atividades criativas de um país na indústria, nas empresas e tecnologias. (Puhlmann & Moreira, 2004) discutem que as patentes podem ser utilizadas como fonte de informação para diversas finalidades, dentre as quais se destacam: (a) identificação de desenvolvimentos tecnológicos já realizados; (b) identificação de alternativas tecnológicas; (c) identificação de tecnologias emergentes, de modo a caracterizar as tendências do desenvolvimento tecnológico de determinada área do conhecimento; (d) avaliação de mercados futuros, uma vez que o patenteamento costuma preceder a comercialização em alguns anos; (e) avaliação das atividades de Pesquisa e Desenvolvimento e detecção de mudanças estratégicas de instituições e empresas.

A documentação de patentes apresenta vantagens consideráveis em comparação com outras fontes de informação tecnológica, como por exemplo: divulgar informação mais rapidamente do que outras fontes porque na maioria dos países os documentos são publicados antes de sua concessão e possuir uma estrutura uniforme relativa ao “*layout*” do documento e aos dados bibliográficos, que são identificados por códigos utilizados por todos os países por meio da Classificação Internacional de Patentes – IPC (WIPO, 2013).

2.5.1 Base de dados de Patentes

Dentre as bases de dados internacionais mais importantes sobre Patentes, destaca-se o *Espacenet - European Patente Office - EPO* (<http://www.espacenet.com>) e o *Patentscope - World Intellectual Property Organization - WIPO* (<http://www.wipo.int/portal/en/index.html>)

O EPO, *European Patente Office*, disponibiliza um sistema de acesso a estas informações (Ferramenta *Espacenet*) envolvendo três bases (um agrupamento tem como base diferentes escritórios nacionais de patentes - hoje com 80 países, outro agrupa as patentes europeias e o terceiro as patentes mundiais). Todas estas três bases, portanto, contém mais de

70 milhões de documentos de patentes, contendo informações sobre invenções e desenvolvimentos técnicos, cada um com 20 páginas, em média, que representam 1,4 mil milhões de páginas, que constituem uma verdadeira "enciclopédia técnico-tecnológica", que abrangem o período de 1836 até os dias atuais. (<http://www.epo.org/searching/free/espacenet.html>) (Quoniam et al., 2014)

A base *Espacenet* libera uma interface (*Open Patent Services – OPS*) para que seja possível utilizar as ferramentas de tecnologia da informação - TI para automatizar a coleta das informações referente as patentes. Por meio da OPS, pode-se acessar dados mundiais de patentes do EPO, incrementar as suas próprias aplicações de software e integrar seus próprios dados com os dados disponíveis via OPS.

A **WIPO**, *World Intellectual Property Organization*, é um órgão da ONU e tem como objetivo manter e aprimorar o respeito pela propriedade intelectual (marcas, patentes, registro geográfico), ou seja, defende o conhecimento em sua utilização global (venda, transferência, cessão *etc*), buscando a estabilidade nos negócios e a supressão de eventuais usurpações, abusos ou distorções. Atualmente, é composta de 187 Estados-membros e administra 27 tratados internacionais. A WIPO disponibiliza um sistema de acesso livre as informações patentárias (Ferramenta *Patentscope*). Por meio da base *Patentscope* é possível pesquisar 18.780.896 documentos de patentes, incluindo 2.212.089 pedidos publicados de patentes internacionais – PCT (<http://patentscope.wipo.int/search/en/search.jsf>).

2.5.2 A Classificação Internacional de Patentes

A pesquisa documental sobre as informações patentárias pode ser realizada nas bases de dados em qualquer um dos campos documentais, incluindo o título e o resumo em forma de pesquisa "texto completo". Observa-se a ausência de campo documentário "palavra-chave". A *World Intellectual Property Office* utiliza como técnica de descrição, não as palavras-chave, mas um sistema de classificação, independente da língua, denominado Classificação Internacional de Patentes (IPC) (Quoniam et al., 2014).

A Classificação Internacional de Patentes, conhecida pela sigla IPC – *International Patent Classification* – foi estabelecida pelo Acordo de Estrasburgo em 1971, alterado em 28 de Setembro de 1979, e prevê um sistema hierárquico de símbolos para a classificação de

Patentes de Invenção e de Modelo de Utilidade, de acordo com as diferentes áreas tecnológicas a que pertencem (WIPO, 2013). Neste sentido, a Classificação Internacional de Patentes (IPC) é uma descrição tipificadora da tecnologia.

Para o IPC, desenvolvido em duas línguas "oficiais" (Francês e Inglês), a tecnologia divide-se em oito seções com aproximadamente 70.000 subdivisões descritas por um símbolo. A busca e identificação da IPC nos documentos está relacionado com a facilidade de reconhecer a área de aplicação tecnológica destes documentos em nível internacional, independente da língua que o documento de patente foi escrito e depositado. O uso do IPC permite uma recuperação fácil da tecnologia desejada, bem como fornece uma base para determinação de dados estatísticos de certos parâmetros tecnológicos (Hirata, Kniess, Cortese, & Quoniam, 2015).

3 METODOLOGIA DA PESQUISA

O capítulo de metodologia da pesquisa deste trabalho possui como intuito apresentar o funcionamento das bases de patentes e suas interfaces de buscas, como também explicitar o processo de *data mining* com a utilização de *crawlers*, em especial o *Patente2Net*. O processo de construção da amostra de patentes inicia-se com os procedimentos de construção de palavras-chave para a busca de patentes sobre educação a distância, e demanda ajustes de ambiguidade na expressão e na busca.

3.1 FRAMEWOK PATENT ANALYZE VERSUS KDD

A abordagem de um conjunto de dados, normalmente requer planejamento prévio. Tal planejamento representa o conjunto das ações que contemplam, desde a escolha do universo de dados a serem explorados, até os resultados que se pretendem. Há algumas formas de realizar tal planejamento e essas técnicas estão amplamente divulgadas na literatura, especialmente nas áreas de ciência da informação e engenharia da computação (Boyd & Crawford, 2012; Buhl, Röglinger, Moser, Heidemann, & others, 2013; Chen, Chiang, & Storey, 2012; Cukier & Mayer-Schoenberger, 2013; Dean & Ghemawat, 2008; George, Haas, & Pentland, 2014; LaValle, Lesser, Shockley, Hopkins, & Kruschwitz, 2011; Manyika et al., 2011).

A abordagem das informações presentes nas patentes requer procedimentos específicos que não foram identificados nessas literaturas. Por isso, investiu-se uma parte do tempo da pesquisa para identificar e propor procedimentos mais adequados ao tipo de dados que serão analisado (informações de patentes). Para cumprir tal intento, usou-se a estrutura básica do KDD (*Knowledge Discovery in Database*) por ser a estrutura comum entre as abordagens usadas pelos autores citados. O KDD consiste no conjunto de ações de exploração e análise de dados, envolvendo a seleção, pré-processamento, limpeza, transformação, *data mining*, interpretação e avaliação.

Constatou-se que o KDD tem espectro amplo e isso se deve ao fato de sua natureza genérica (CEINE, 2015). Como exemplo, na abordagem pelo KDD, há etapas de limpeza e transformação de dados. Isso significa que, após ter acesso aos dados, algoritmos específicos serão usados para preparar a remoção de duplicidades, enriquecer os dados pela complementação de partes ausentes e ainda realizar a transformação de formatos. Por

transformação, pode-se citar o preenchimento de campos que estejam vazios, por exemplo. No caso da análise de patentes, o *software Crawler* (no caso o Patent2Netv.2) tem a capacidade de fornecer os dados no formato necessário para a análise, no entanto, é necessário construir corretamente a expressão de busca.

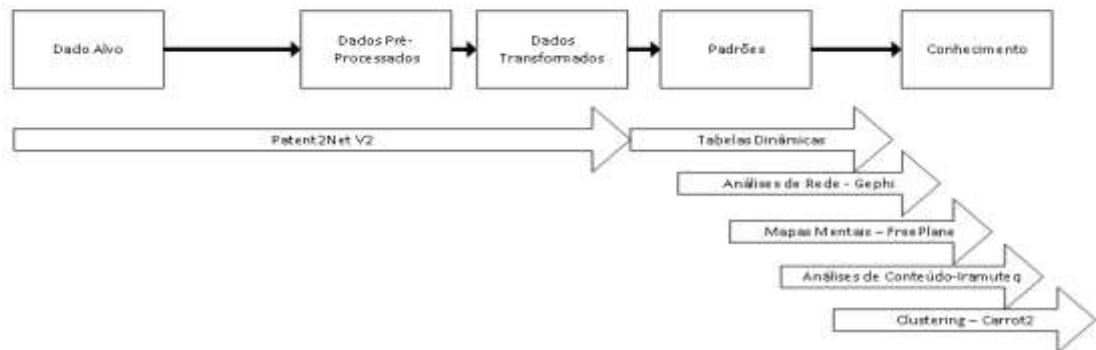
A expressão de busca precisa ser específica o suficiente para recuperar apenas os conteúdos pertinentes (baixo ruído), mas, genérica o suficiente para recuperar uma amostra expressiva (baixo silêncio). O que argumenta-se, é que as etapas iniciais que envolvem a seleção, limpeza e transformação dos dados na abordagem do KDD tradicional, no caso em estudo, das informações das patentes, são compostos pelo processo interativo da construção da expressão de busca e o ajuste da capacidade de recuperação: redução de ambiguidades, ruídos e silêncios. Dessa forma, o KDD tradicional não parece ser a melhor escolha para explorar as informações disponíveis nas patentes, por serem essas mais organizadas do que dados mais genéricos e desconhecidos (Quoniam et al., 2014).

Já a fase de *Data Mining* do KDD, que consiste em identificar padrões, regras e tendências são parecidas para o caso das informações das patentes e, portanto, entende-se que devem fazer parte do *Framework* de análise de informações de patentes proposto por nós. A última fase do KDD é a fase da interpretação e avaliação e entende-se que tal etapa, no caso da análise das informações de patentes, merece algumas ligeiras alterações e comentários. No sentido de otimizar o processo de análise e por contar com recursos automatizados no caso, toda a avaliação e interpretação foram deslocadas para a fase anterior de *Data Mining*, no caso do *Framework* de análise de patentes que está se propondo.

Nesse caso, todas as análises e identificações de tendências, regras e padrões são realizadas na fase de *Data Mining*, restando para a última fase a ligação entre os resultados das análises das informações das patentes e o campo de estudo de interesse. Isso quer dizer que pode-se realizar as análises de forma prática, verificar seus resultados e promover o debate com fundamentos científicos, teóricos ou outros estudos empíricos, interligando de certa forma as informações tecnológicas das patentes com outros campos de conhecimento: ciências sociais aplicadas, sociologia e filosofia não exaustivamente. Para sintetizar a proposta de *Framework* para análise de informações de patentes sem perder de vista as etapas de um KDD tradicional, permitindo a comparação gráfica, apresenta-se a seguir o diagrama que sintetiza as ideias descritas.

Figura 2 – Comparação entre procedimento metodológico KDD e Patent2Net

Procedimentos Metodológicos de Análise de Patentes					
Etapas do Processo Patent2Net Ver. 2	Fase 1A - Construindo a Expressão de Busca	Fase 1B - Reduzindo Ambiguidades, Ruidos e Silêncios	Fase 2 - Analizando as Patentes	Fase 3 - Modelando os achados (Informação)	
	 1 Thesaurus 2 Palavras Chave 3 Sintaxe do Banco de Dados 4 Define a Expressão de Busca	 1 Leitura dos conteúdos recuperados 2 Estudo dos Morfemas (Afícos e Desinências) das palavras chave da expressão de busca 3 Estudo dos Radicais e Identificação de Cognatas nas palavras chave da expressão de busca 4 Refaz a expressão de busca	 1 Tabelas Dinâmicas 2 Análises de Rede 3 Mapas Mentais 3 Análises de Conteúdo 4 Clustering	1 Modelos 2 Construtos 3 Protótipos 4 Proposições 5 Hipóteses 6 Artefatos 7 Métodos 8 Instanciações	
Alinhamento com os princípios do processo de KDD					
Etapas do Processo de KDD	Seleção Objetivo – Identificar quais os conjuntos de dados, atributos, características (variáveis) e os casos, observações (registros), farão parte da análise	Pré-processamento e Limpeza Objetivo – Eliminar dados redundantes, inconsistente (Ruidos), discrepantes do conjunto (Ambiguidade) e ausência de respostas para constituição da amostra (Silêncio)	Transformação Objetivo – Chegar ao formato e armazenamento ideal para posterior mineração	Data Mining Objetivo – Identificar ou descobrir padrões e regras nos dados analisados, desconhecidas a priori, pelo sistema e pelo analista	Interpretação e Avaliação Objetivo - Ideação, predição, identificação de tendências e construção da informação



3.2 AS BASES DE PATENTES E SUAS INTERFACES DE BUSCA

Para países como o Brasil, que ainda se encontram em processo de desenvolvimento, a realização de pesquisas cujo foco central esteja voltado à avaliação dos textos das patentes

são de grande valia. Em torno de apenas 5% das invenções protegidas por patentes nos países considerados desenvolvidos são estendidas para os demais países, o que torna os 95% restantes de domínio público, permitindo que sejam livremente replicados inclusive com o respaldo de tratados e legislações nacionais e internacionais (Corrêa & Gomes, 2007). Dessa forma, após a lacuna de sigilo de 18 meses, período este onde os textos das patentes ficam indisponíveis para consulta, qualquer interessado em identificar patentes livres em seu país poderá, por meio de pesquisas em diferentes bases, obter os textos dessas patentes na íntegra, avaliá-los e, se viável, replicá-los. Todavia, como os documentos de patentes se encontram alocados na *deep web*, os buscadores comumente utilizados para pesquisa na *web* de superfície não são capazes de encontrá-los adequadamente e, tão pouco, disponibilizá-los de maneira organizada e que facilite sua avaliação. Para viabilizar tais pesquisas é necessário utilizar buscadores específicos para tal, como por exemplo, o *Google Patent Search* (www.google.com/patents), o *Patentscope* (<http://www.wipo.int/patentscope/en/>) ou o *Espacenet* (<http://worldwide.espacenet.com/>).

Disponível desde 2006, o *Google Patent Search* é um motor de busca de patentes tanto concedidas quanto negadas, além dos pedidos de patentes, inicialmente depositados no *United States Patent and Trademark Office* (USPTO), datados a partir do ano de 1790, e que atualmente engloba documentos depositados também em outras bases como o *European Patent Office* (EPO) e o *World Intellectual Property Organization* (WIPO), datados a partir de 1978.

Independente de algumas críticas quanto ao desempenho da ferramenta na busca efetiva pelos documentos, seu desempenho é considerado bom em termos de tempo de resposta (European Patent Office, 2007). Estima-se que este buscador tenha acesso ao texto integral de aproximadamente 8 milhões de patentes. O *Patentscope* é uma ferramenta de busca disponibilizada pelo próprio WIPO, que permite acesso a uma base de patentes de grande porte, com cerca de 45 milhões de documentos de variados países. Já o *Espacenet*, disponibilizado pelo EPO desde 1998, considerado à época uma ferramenta que revolucionaria os mecanismos de acesso ao conteúdo das patentes (White, 2006), permite acesso ao texto integral de aproximadamente 90 milhões de documentos. Em 2013, o desempenho do aplicativo foi considerado superior aos demais quando levados em consideração indicadores relacionados especialmente à amplitude de cobertura de dados (European Patent Office, 2013).

3.2.1 Os *crawlers* de patentes e o processo *data mining*

A despeito da existência de uma série de bases onde seja possível pesquisar os textos integrais dos milhões de documentos patentários, bem como de uma série de ferramentas de busca que se dispõem a facilitar este trabalho, ainda assim a realização de uma avaliação direcionada e organizada dos conteúdos desses documentos se mostra uma tarefa bastante árdua.

Em meados da década de 1940, o soviético Genrich Altshuller iniciou a análise manual do texto das cerca de 400 mil patentes existentes à época. O referido trabalho, que durou aproximadamente 25 anos para ser concluído (cerca de 32 minutos para analisar cada patente, considerando 24 horas diárias de trabalho), deu origem à metodologia TRIZ (“teoria da resolução inventiva de problemas”, em português), que por sua vez é até hoje utilizada na busca por soluções para problemas existentes em muitas corporações ao redor do mundo (Webb, 2002).

Tomando o exemplo da metodologia empregada por Altshuller, fazendo-se uma conta simples do tempo que se gastaria para analisar os textos das cerca de 90 milhões de patentes disponibilizadas pelo EPO, apenas na *Espacenet*, seriam necessários aproximadamente 5.625 anos.

Dessa forma, a existência de ferramentas computacionais que se proponham a realizar uma análise prévia dos conteúdos das patentes, organizando-as com base em informações que permitam identificar de forma rápida e direta o ponto em que se deseja chegar, são essenciais para viabilizar a realização de estudos envolvendo estes documentos. Como exemplos de tais ferramentas, que a partir deste ponto serão chamadas de *crawlers*, destacam-se o *Intellixir* (<http://www.intellixir.com/>), o *Matheo Patent* (<http://www.matheo-software.com/en/>), o *Patent Integration* (<https://patent-i.com/>), e o *Patent Inspiration* (<http://www.patentinspiration.com/>), todos softwares pagos, além dos softwares livres: *Lens* (<https://www.lens.org/lens/>) e *Patent2net* (<http://patent2net.vlab4u.info/>).

Um *web crawler* se constitui em uma ferramenta computacional que navega pela *internet* de maneira automática e bastante metódica, cuja principal funcionalidade consiste em obter da *web* uma série de informações que não podem ser localizadas com o uso de

navegadores comuns. O referido processo, também conhecido como *data mining*, ou mineração de dados, identifica correlações significativas, padrões e tendências que se repitam em grandes conjuntos de dados alocados nos mais diversos tipos de repositórios como, por exemplo, as bases de patentes, utilizando complexas técnicas estatísticas e matemáticas (Larose, 2014).

O uso de *crawlers* para mineração de dados nas bases de patentes não é raro. Diversos trabalhos propondo metodologias com a referida função podem ser encontrados na literatura (Ernst, 2003; Fattori, Pedrazzi, & Turra, 2003), assim como a descrição dos processos necessários para sua execução (Ernst, 2003; Kim, Suh, & Park, 2008; Tseng, Lin, & Lin, 2007).

3.2.2 A ferramenta computacional *Patent2net*

O *Patent2net* é um *crawler* de uso livre cuja principal aplicabilidade é a mineração de dados na base de patentes mantida pelo EPO. A escolha pela utilização da *Espacenet*, ao invés, por exemplo, da *Patentscope* como base para as buscas, se resume ao fato de que a *Espacenet* permite o acesso à sua *Application Programming Interface* (API), cuja tradução em português é “Interface de Programação de Aplicações”, e que de um modo bastante generalista se refere ao conjunto de padrões e rotinas computacionais que permitem que outros aplicativos, como por exemplo os *crawlers*, utilizem diversas de suas funcionalidades sem necessariamente se envolver com complexos detalhes de ordem computacional, mas apenas se utilizar das funções para as quais o *software* tenha sido inicialmente concebido. E, quando esta função se refere à mineração de dados, a liberação da API é essencial (Imielinski, Virmani, & Abdulghani, 1996).

O *Patent2net* otimiza o *data mining* e se mostra bastante eficiente quando levados em consideração o volume e a variedade de informações, bem como a velocidade com que estas são processadas. Sendo assim, a utilização do *Patent2net*, que por sua vez interage diretamente com a *Espacenet* na busca de informações relacionadas às patentes disponíveis na base do EPO, é absolutamente legal e transparente, podendo inclusive ser citada como um exemplo de uso gratuito do *Open Patent Service* (OPS) para educação e informação em patentes. Para Kallas (2006), o OPS é uma metodologia real e estratégica para a divulgação de informações, contribuindo para tornar transparentes os resultados das pesquisas científicas sobre quaisquer temas, especialmente aqueles prioritários para uma nação como o Brasil, como é o caso da nanotecnologia, do manejo de terras-raras, da eficiência energética, da biodiversidade e das doenças negligenciadas.

3.3 CONSTRUINDO PALAVRAS-CHAVE PARA EAD

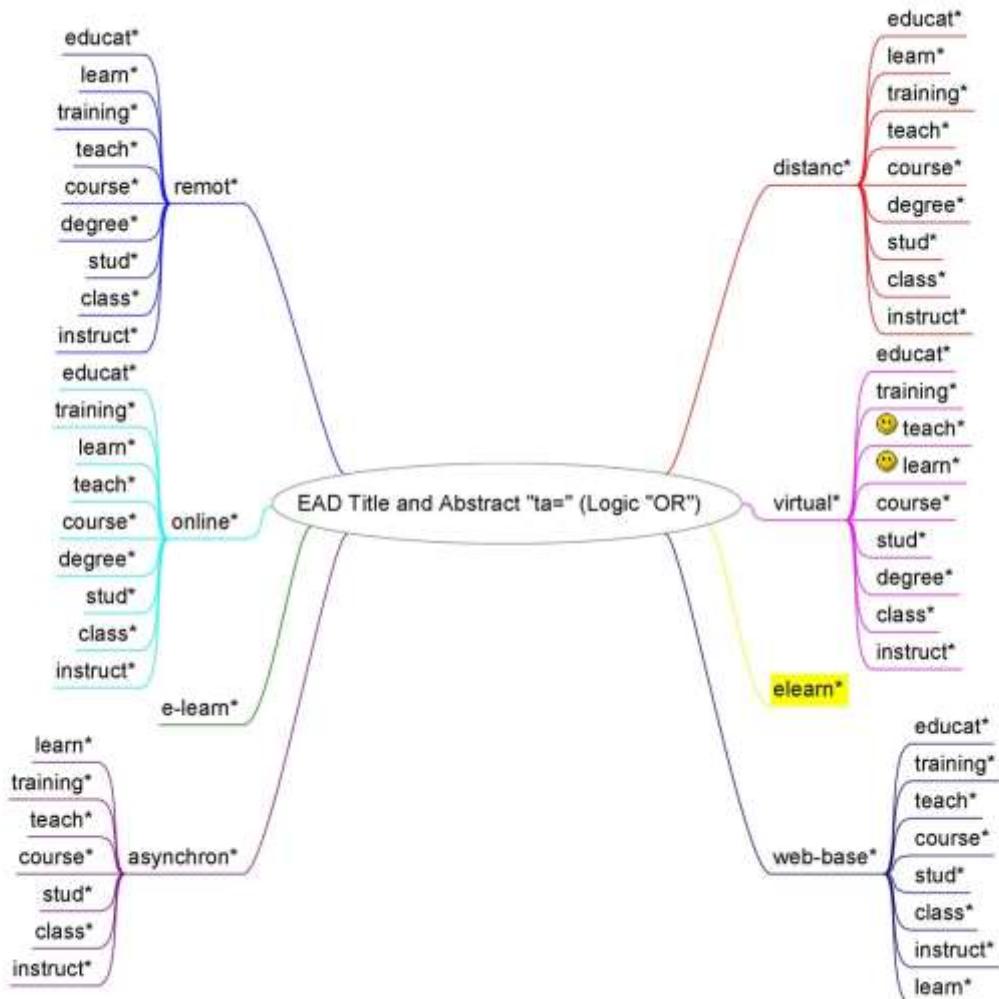
Com o intuito de formar um conjunto de palavras-chave que permita elucidar sobre ensino a distância, primeiramente foi buscado um *thesaurus* (estrutura de sinônimos) sobre *distance learning* no site Thesaurus.com (<http://www.thesaurus.com>). Além disso, o site de dicionários (<http://dictionary.reference.com/browse/distance%20learning>), mostrou mais termos, como sinônimos de ensino a distância. Em inglês, o termo que melhor representa ensino a distância é *distance learning*, conforme ambas informações dos dicionários e do Thesaurus citado. Por isso, a palavra "ensino a distância" foi definida como palavra-chave inicial para iniciar as buscas em bases de dados escritas na língua portuguesa e *distance learning* para buscas em bases de dados escritas na língua inglesa. Com o objetivo de acessar as bases de dados de patentes, no escritório Europeu de Patentes (http://worldwide.espacenet.com/?locale=en_EP) o termo em inglês pode ser suficiente, uma vez que os depósitos de patentes são obrigatoriamente classificados em inglês e em francês, em termos de conteúdo e linguagem de resumo e de palavras-chave.

Com base nos resultados do primeiro levantamento, o termo de busca *distance learning* começou a refletir o tema de interesse. Como segundo passo, empregou-se o software *Publish or Perish* idealizado pela pesquisadora Anne-Wil Harzing (www.harzing.com) para compreender as publicações científicas e os livros sobre ensino a distância. O objetivo foi encontrar outros possíveis sinônimos e termos que possam estar relacionados ao tema, mas que por questões de diferenças culturais ou escolas de pensamento científico, pudessem ser representadas com outras palavras no campo científico. Nesse caso, como o software *Publish or Perish* usa como algoritmo de busca a base de dados acadêmica *Google Scholar*, o termo de busca usado foi *distance learning*. Como modo padrão, o software *Publish or Perish* retorna até 1.000 publicações em livros, artigos científicos e citações que podem ser analisadas para compreender os principais autores e trabalhos relacionados ao tema de interesse, nesse caso, o ensino a distância.

Os documentos publicados reportados pelo software *Publish or Perish* são sempre dispostos de maneira que os mais citados são os primeiros da lista. O documento mais citado (citado por 4.580 documentos publicados), no tema de interesse *distance learning* é um livro escrito por Michael Moore et al., com o título Educação a Distância: Uma visão sistêmica da

aprendizagem on-line (tradução do título pelo autor) de 2011 (Moore & Kearsley, 2011), e se encontra na sua terceira edição. Com a mesmo procedimento, encontrou-se também, o livro de Moore: *Handbook of Distance Learning*; de 2013 (Moore, 2013): citado por 549 documentos publicados. Nesses documentos, foi possível observar os termos utilizados a respeito de EAD, visando aumentar a quantidade dos termos de busca nas bases de patentes. Com base nisso, houve o desenvolvimento de palavras-chave, que foram organizadas em formato de mapa mental, conforme Figura 3.

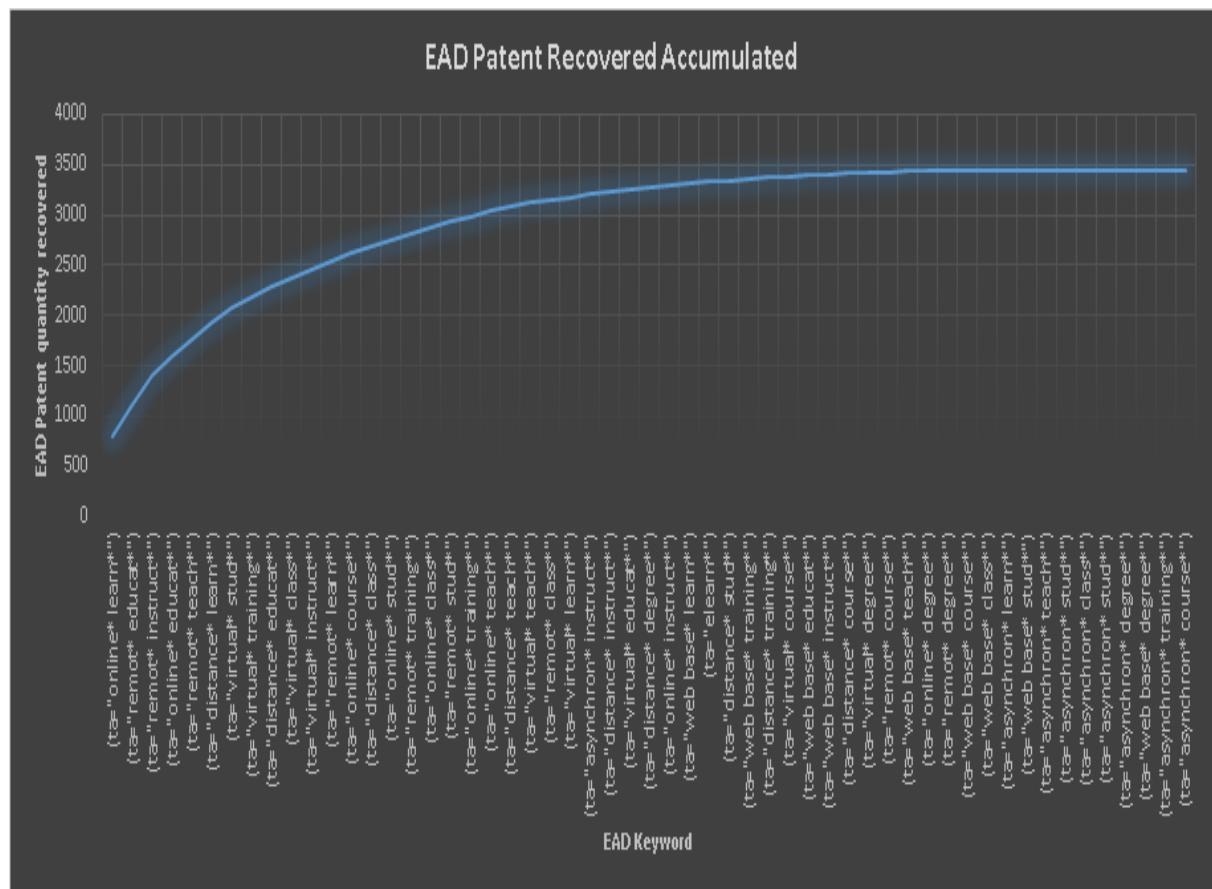
Figura 3 – Mapa Mental das palavras-chave de educação a distância



O Escritório Europeu de Patentes (EPO), tem como limitação no processo de busca, o uso de 10 termos no máximo, para cada busca. Dessa forma, por meio de um processo de lógica

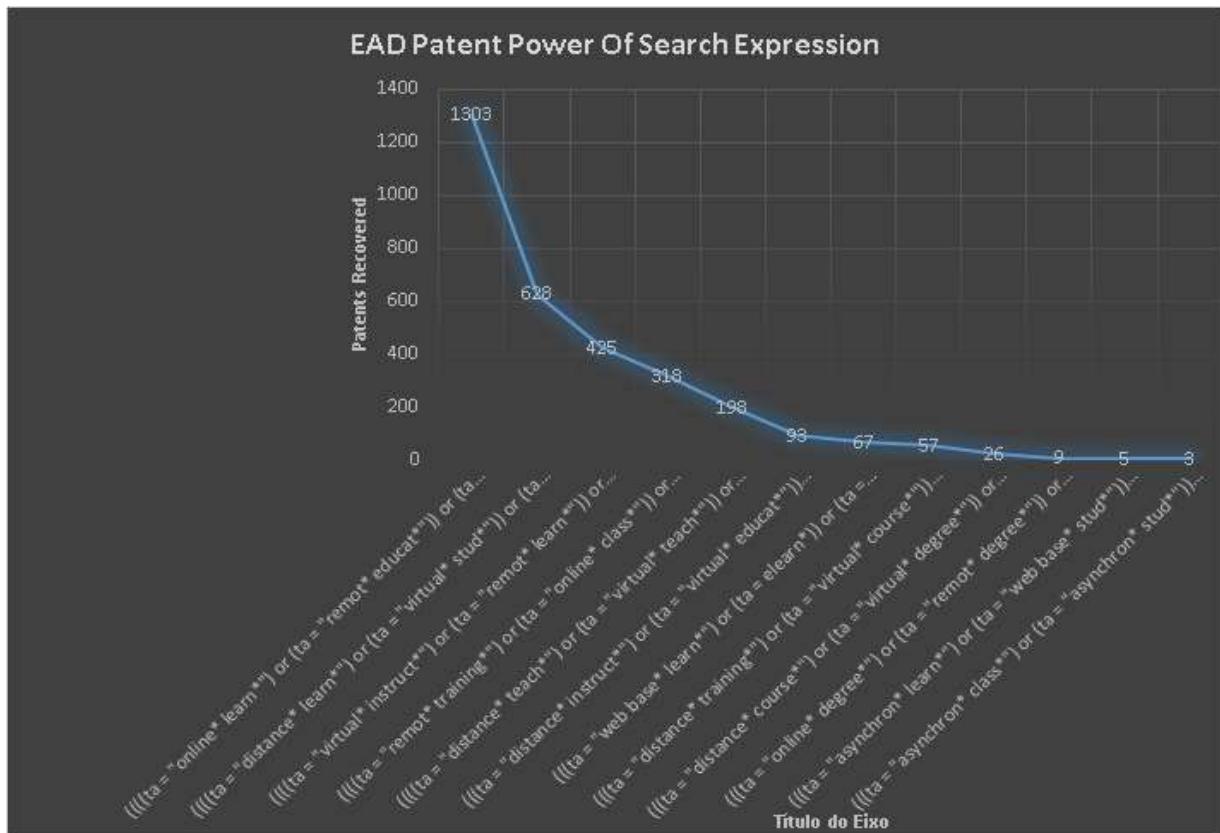
combinatória e agrupamento de 10 em 10 termos, foram construídas 12 expressões de busca. Para reduzir as sobreposições e definir a ordem em que as buscas ocorreriam, testou-se cada termo no EPO, registrando a quantidade de patentes recuperadas para cada termo isoladamente. Verificou-se que as 52 palavras-chave apresentavam um potencial de recuperação de 3.620 patentes. Na Figura 4, é apresentada a relação de palavras-chave no eixo X e a quantidade de patentes recuperadas no eixo Y.

Figura 4 – Potencial de recuperação das palavras-chave.



É importante destacar, além do potencial de recuperação das palavras-chave, a real contribuição de cada palavra-chave empregada na expressão de busca. A figura exibe quanto cada palavra-chave contribui no retorno de patentes pelo emprego da expressão de busca. Destaca-se que no eixo X se encontra a expressão de busca e no eixo Y a quantidade de patentes retornadas na base pela expressão de busca.

Figura 5 – Contribuição das palavras-chaves na expressão de busca.



3.3.1 Relação com a ambiguidade na expressão e busca (silêncio e ruído)

Quando se prepara as expressões de busca, há como base as palavras-chave (*keywords*). Palavras-chave são, por sua vez, compostas por uma ou mais palavras únicas que dão significado a um conceito ou construto de interesse. No caso de “ensino a distância”, caso separe-se a palavra “ensino” da palavra “distância”, não seria possível obter o mesmo significado. O mesmo aconteceria ao usarmos apenas a palavra “distância”. Isoladamente, nenhum das palavras tem o significado do conceito ou construto de interesse, que em linhas gerais consiste no ensino realizado sem que aluno os recursos didáticos e o professor estejam no mesmo local físico.

A busca de informações por meio de palavras-chave, está sujeita geralmente a dois efeitos indesejáveis, nomeados no campo da ciência da informação como silêncio e ruído. Pensando como um contínuo que inicia no silêncio e termina no ruído, os dois extremos são indesejáveis. O silêncio é caracterizado pela falta de retorno ou retorno pequeno em termos de tamanho da amostra, alcançada na execução de determinada palavra-chave ou expressão de

busca. Já o ruído, caracteriza-se pelo retorno de informações que respondem pela mesma palavra-chave, mas não tem relação com o conceito ou construto de interesse.

Normalmente os bancos de dados estruturados oferecem alguns operadores booleanos e funções de truncagem para que os usuários desses bancos de dados possam operar suas expressões de forma a alcançar os resultados desejados em algum ponto entre o silêncio e o ruído, que seria o ponto ótimo para cada pesquisa ou interesse. Operadores booleanos são aqueles que realizam funções lógicas como “ou”, “e” - *or, and* - e truncagem são variáveis colocadas numa determinada posição da palavra, que mostra ao sistema de busca, que esse deve retornar todos os resultados que consistem na parte da palavra escrita até a posição da variável de truncagem, independentemente das letras que vierem depois.

No caso do EPO, a variável de truncagem é o asterisco - * -. Obviamente que o ponto ótimo entre silêncio e ruído, varia de pesquisa para pesquisa e ainda em função de cada sintaxe usada nas expressões de busca construídas. Após compreendidas as bases de construção das palavras-chave e as operações lógicas envolvidas, decide-se qual o ponto exato onde as buscas serão disparadas. Isso significa que, os bancos de dados normalmente são estruturados em partes, como por exemplo, título, resumo e método não exaustivamente. No caso do EPO, o título e o resumo de uma patente foram os pontos escolhidos porque normalmente carregam por definição a temática da patente, além de haver a obrigatoriedade de depositar o resumo em Inglês e Francês. A sintaxe específica para o banco de dados EPO é “(ta=)”, sendo *t* para título - do inglês *title* - e o *a* de resumo - do inglês *abstract* -. Essas são especificidades do banco de dados EPO, no entanto, a lógica de construção da busca, guardadas as especificidades de sintaxe de cada banco de dados podem ser, em grande parte, aproveitadas.

Na primeira etapa da busca, foram utilizadas palavras-chave compostas por duas palavras e truncagem em cada uma delas, para reduzir o efeito do silêncio. Por exemplo, usando-se a palavra-chave (*ta="distance learning"*), tem-se como resposta, apenas as informações que tem no título ou no resumo as palavras em sua forma exata, como informado. Nesse caso, se por hipótese, houvesse uma informação descrita como “*distance learned*”, essa não faria parte do resultado, caracterizando o efeito conhecido como silêncio, levando a falsa conclusão quanto a quantidade e mesmo a qualidade das informações a respeito do conceito ou construto de interesse. Para buscar reduzir o efeito do silêncio, usou-se o recurso de truncagem.

No mesmo exemplo usou-se a palavra-chave (*ta= "distance* learn*"*). Nessa última expressão, os resultados receberiam a adição das informações cujas terminações após a letra *e* da palavra *distance* e também a adição das terminações após a letra *n* de *learn*.

O problema é que nesse caso, pode-se introduzir ambiguidades ou ruídos, originadas pelas palavras que tem o mesmo radical, mas com desinências diferentes que podem modificar o significado da palavra, deslocando a mesma do conceito ou construto de interesse. Mesmo fora do conceito ou construto de interesse, a informação estará presente nos resultados de busca, porque o efeito da truncagem privilegiou o radical da palavra. A explicação pode ser obtida observando as estruturas gramaticais e as raízes das palavras. As palavras são constituídas de morfemas e os morfemas são compostos principalmente por radicais, afixos e desinências (outras partes da estrutura como infixos, vogais temáticas e tema, foram desconsideradas por não se aplicarem em princípio ao caso). Radical é o elemento comum das palavras cognatas, também chamadas de palavras da mesma família e traz o significado básico da palavra. Afixos são partículas anexadas ao radical para formar outras palavras e podem ser do tipo prefixo, quando vem antes do radical ou sufixo, quando vem depois do radical.

As desinências são morfemas colocados no final das palavras para indicar flexões verbais ou nominais. As principais desinências foram consultadas em “Estrutura das palavras”, (2016), representadas na Tabela 4. Como exemplo, no caso das buscas, se usar-se a palavra-chave (*ta= "online* stud*"*), os resultados incluirão todas as informações onde há “*online study, studies, studied*”, o que é desejado, pois reduz o efeito silêncio, no entanto também retornará “*online studio*”. *Studio* é um tipo de local, sala ou escritório e não é normalmente atribuído ao conceito ou construto de interesse “*distance learning*”, caracterizando o problema da ambiguidade ou ruído. Pode-se deduzir que essa questão, impõe a necessidade de se desenvolver palavras-chave e expressões para cada língua que se pretende realizar a busca, ou seja, a língua, cuja base de dados é construída, ou ainda, a língua cujo ponto de busca - por exemplo título e resumo - é construído.

Tabela 4 – Principais desinências da língua portuguesa.

NOMINAIS	Gênero	masculino (-o) feminino (-a)
	Número	singular (não há) plural (-s)
	de tempo e modo	<p>-va, -ve: imperfeito do indicativo, 1ª conjugação</p> <p>-ia, -ie: imperfeito do indicativo, 2ª e 3ª conjugações</p> <p>-ra, -re: mais-que-perfeito do indicativo (átono)</p> <p>-sse: imperfeito do subjuntivo</p> <p>-ra, -re: futuro do presente do indicativo (tônico)</p> <p>-ria, -rie: futuro do pretérito do indicativo</p> <p>-r: futuro do subjuntivo</p> <p>-e: presente do subjuntivo, 1ª conjugação</p> <p>-a: presente do subjuntivo, 2ª e 3ª conjugações</p>
VERBAIS	de pessoa e número	<p>-o: 1ª pessoa do singular, presente do indicativo</p> <p>-s: 2ª pessoa do singular</p> <p>-mos: 1ª pessoa do plural</p> <p>-is-, -des: 2ª pessoa do plural</p> <p>-m: 3ª pessoa do plural</p>
VERBO-NOMINAIS		<p>-r: infinitivo</p> <p>-ndo: gerúndio</p> <p>-do: particípio regular</p>

Fonte: (“Estrutura das palavras”, 2016)

No caso do EPO, como descrito anteriormente, os títulos e os resumos das patentes para serem depositadas são exigidos em Inglês e Francês. O foco inicial é alcançar portanto as informações patentárias sobre “ensino a distância”, cujo ponto alvo de busca está escrito em Inglês. Para o Inglês, a estrutura das palavras são compostas de radical, afixo e desinência, que obviamente variam em relação a outras línguas. No site da organização sem fins lucrativos *Learn That Foundations*, uma organização que fornece dicionários abertos da língua inglesa, estão disponíveis mais de 180 mil palavras em Inglês e também a relação dos principais sufixos: (“LearnThat Foundation”, 2013).

A redução da ambiguidade inserida pelo uso da truncagem, passa por um ajuste fino, a partir da análise do radical, do afixo e das desinências das palavras únicas envolvidas na composição das palavras-chave, visando preparar expressões de busca que sejam mais precisas. A precisão, no estudo, ainda que relativa, será aquela que retornar maior quantidade de informações com baixa ambiguidade (baixo silêncio e baixo ruído). Os sufixos em inglês estão disponíveis em (“Suffix mania! Long list of suffixes, incl. meaning, word lists, examples”, 2016), ou ainda em (“English Grammar - English Suffixes - Learn English”, 2016) e também em (“What Are the Most Common Suffixes in English?”, 2016). O processo de construção das palavras únicas que irão compor as palavras-chave, deverão considerar a decomposição de cada palavra única, a princípio em morfema raiz (radical), o morfema derivado (prefixo e sufixo) e a desinência, nesse caso especificamente para a língua inglesa.

Sufixos inflexionais são aqueles que não mudam a classe da palavra, após a inflexão. Os sufixos inflexionais, também chamados de desinências, na língua inglesa, estão apresentados a seguir, conforme consultado em (“Suffix”, 2016):

- **-s** third person singular present
- **-ed** past tense
- **-t** past tense
- **-ing** [progressive](#)/continuous
- **-en** [past participle](#)
- **-s** plural
- **-en** plural (irregular)
- **-er** [comparative](#)
- **-est** [superlative](#)
- **-n't** [negative](#)

Como as desinências na língua inglesa, são compostas por uma a três letras, para reduzir a quantidade de informações não pertinentes como resultado das buscas, chamadas de ruídos e assim reduzir a ambiguidade das expressões de busca desenvolvidas com o recurso “*wildcard*” da truncagem “*”, optou-se por incluir recursos disponíveis na base EPO, denominados “*wildcard*” “?” e “#”. O “*wildcard*” “?” (interrogação), quando incluído numa palavra-chave, retorna as informações que tem as mesma raiz (entre o início da palavra até o *wildcard* “?”). A quantidade de “*wildcard*” “?” colocadas no final da palavra, define a quantidade de letras que as buscas irão considerar após o raiz da palavra. Por exemplo, quando uso a expressão (*ta*=“onlin?") o retorno será tudo que começa com “*onlin.....*” e tem mais uma letra para terminar a palavra, já que há apenas um “*wildcard*” “?”, como “*onlin*” e “*online*”, mas não retornará “*onlining*”. Se a expressão for (*ta*=“onlin???”), as respostas incluirão, “*onlin*”, “*online*” e “*onlining*”, ou seja, incluirá desde apenas a raiz da palavra até três letras a mais. Considerando que as desinências tem entre uma e três letras, o refinamento proposto e chamado de fase 2, foi a substituição do “*wildcard*” truncagem “*” para o “*wildcard*” “???”.

O efeito dessa substituição de “*wildcards*” nas expressões de busca, reduz o ruído, deixando de recuperar patentes que não tem relação com o conceito ou construto de interesse. Essa redução da quantidade de letras após a palavra raiz, remove também alguns efeitos indesejáveis de palavras que ocorrem com palavras mais longas. Como exemplo, pode-se usar a palavra “*class*”. Na fase 1, utilizou-se o “*wildcard*” de truncagem (*ta*=“*class**”) e o retorno incluiu todas as patentes que continham a palavra “*class*” mas também, “*classification*”, que não tem relação provável com o conceito ou construto de interesse “educação a distância”. Na fase 2, usou-se o “*wildcard*” “???", cuja expressão de busca tornou-se (*ta*=“*class*???”). Os

resultados obtidos consideraram “*class*” e todas as desinências possíveis na língua inglesa, limitado a três letras excluindo, portanto, “*classification*” não será buscada na consulta.

Outro aspecto usado para refinar as expressões de busca, que precisa ser observado são os casos das palavras cuja sufixo é derivacional. Essas palavras mudam de sentido quando recebem esses sufixos. Como exemplo, pode-se citar a palavra “*training*” que traduzido para a língua portuguesa significa treinamento. A palavra pertence ao conceito ou construto de interesse “ensino a distância”, portanto, precisa estar presente nas patentes recuperadas. A raiz dessa palavra é “*train*”, que traduzindo para a língua portuguesa significa trem, que não pertence ao conceito ou construto de interesse “ensino a distância”. Nesse caso, o uso do “*wildcard*” “#” (*hash*), pode auxiliar, porque tem um efeito similar ao “*wildcard*” “?””, mas não inclui na resposta a parte raiz da palavra. Dessa forma, a expressão (*ta=*”*train###*”) retornará todas as patentes que tem a raiz “*train*” adicionada a uma, duas ou três letras, cobrindo portanto as desinências da língua inglesa, mas não incluirá a própria raiz isoladamente, eliminando as patentes sobre “*train*”, mas incluindo “*training*”, “*trained*”, “*trainee*” e demais combinações gramaticais com até três letras depois da raiz “*train*”.

Um último aspecto, que deve ser levado em conta para o refinamento das palavras-chave e expressões de busca por consequência. Um segundo exemplo é a palavra “*study*”. O potencial da palavra “*study*” para o conceito ou construto do “ensino a distância” parece ser alto devido a frequência com que a palavra aparece na análise de conteúdo do *Hand Book Of Distance Learning* de Morre et al. (2010). Nesse caso, tem dois efeitos que precisam ser observados. A raiz da palavra é “*stud*”, mas assim, como no caso da palavra “*train*”, o significado de “*stud*” não tem relação com a palavra “*study*”. Nesse ponto, o “*wildcard*” “#” poderia resolver, pois trata-se de um problema de diferentes significados da raiz da palavra. A excessão acontece, pois há nesse caso, sufixos que também mudam o significado, como por exemplo “*studio*”. Percebo que há “*stud*” e “*studio*”, tanto a raiz como um dos sufixos não estão dentro do conceito ou construto de interesse “ensino a distância”. Nesse caso, a proposição é resolver a parte da mudança de significado da raiz da palavra pelo uso do “*wildcard*” “#” e para excluir a palavra “*studio*” dos resultados, usar a expressão boleana “*NOT*”, também disponível para a base de dados EPO. Nesse caso a expressão refinada nessa fase 2 tornou-se ((*ta=*”*stud###*”) *not studio*). Os resultados dessa expressão retornaram patentes com todas as

desinências da raiz “stud” como “study”, “studying”, “studied”, “studies”, não incluirá a raiz “stud” isoladamente e não incluirá a palavra “studio”.

A seguir, é descrita a lista de sufixos derivacionais (disponíveis em <https://en.wikipedia.org/wiki/Suffix>) da língua inglesa, que são aqueles que podem mudar as classes das palavras e que foram observadas para propor os refinamentos descritos:

- **-ise/-ize** (usually changes nouns into verbs)
- **-fy** (usually changes nouns into verbs)
- **-ly** (usually changes adjectives into adverbs)
- **-ful** (usually changes nouns into adjectives)
- **-able/-ible** (usually changes verbs into adjectives)
- **-hood** (usually class-maintaining, with the word class remaining a noun)
- **-ess** (usually class-maintaining, with the word class remaining a noun)
- **-ness** (usually changes adjectives into nouns)
- **-less** (usually changes nouns into adjectives)
- **-ism** (usually class-maintaining, with the word class remaining a noun)
- **-ment** (usually changes verbs into nouns)
- **-ist** (usually class-maintaining, with the word class remaining a noun)
- **-al** (usually changes nouns into adjectives)
- **-ish** (usually changes nouns into adjectives/ class-maintaining, with the word class remaining an adjective)
- **-tion** (usually changes verbs into noun)
- **-logy/-ology** (usually class-maintaining, with the word class remaining a noun)

Os efeitos do refinamento das expressões descritas e chamadas de fase 2, removeram dos resultados, aproximadamente 530 patentes. O termo “aproximadamente” deve ser usado porque o algoritmo de busca do EPO assim descreve os resultados, sendo essa uma possível limitação do uso do sistema de busca “smart search” do banco de dados EPO. A amostragem exploratória da fase 1, de aproximadamente 3.620 patentes foi refinada para 3.090 patentes, conforme apresentados nas Figura 6 e 6.

Figura 6 – Potencial da expressão de busca anterior ao refinamento.

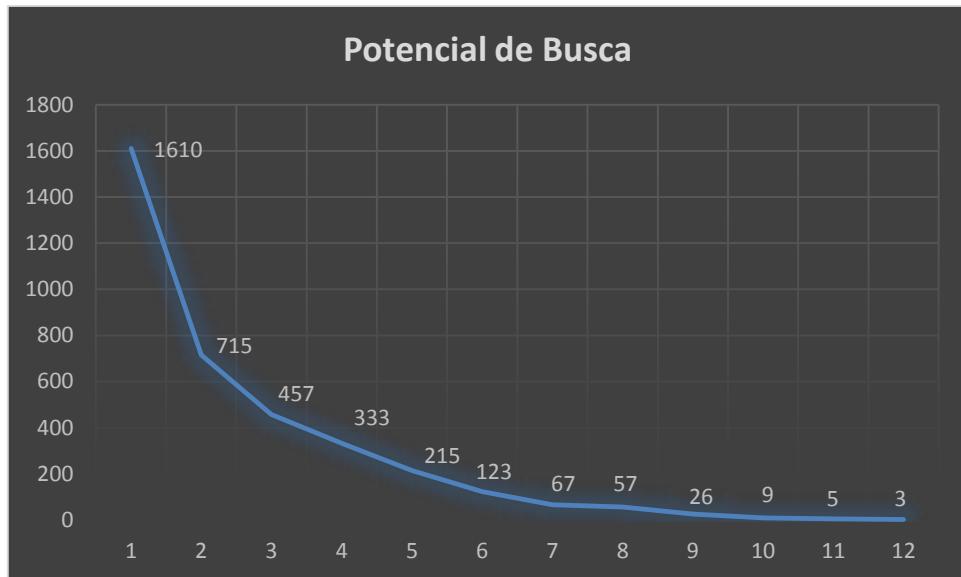


Figura 7 – Potencial da expressão de busca após o refinamento.



Na Tabela 5 são disponibilizadas as informações com os níveis de saturação para cada busca (*saturation curve*) e também o nível de melhoria busca após busca em percentuais (*increase level by search*).

Tabela 5 – Saturação e nível de melhoria das buscas.

Curva de Saturação (Patentes)	Curva de Saturação (%)	Poder Potencial da busca da expressão Patentes EAD	Nível de Aprimoramento por Busca	Número da Busca
1.607	52.02%	1.607		1
2.225	72.42%	618	39.21%	2
2.607	84.36%	381	16.49%	3
2.778	89.90%	171	6.57%	4
2.897	93.77%	120	4.30%	5
2.968	96.06%	71	2.45%	6
3.019	97.71%	51	1.71%	7
3.045	98.53%	25	0.84%	8
3.069	99.32%	24	0.80%	9
3.082	99.73%	14	0.41%	10
3.087	99.90%	5	0.17%	11
3.090	100.00%	3	0.10%	12

3.4 ANÁLISE DAS PATENTES EM EAD

Nessa etapa, conhecida como análise ou *data mining*, foram usadas técnicas para simplificação e organização dos dados recuperados, com o objetivo de obter informações desconhecidas a priori. Tais informações, obtidas por meio da descoberta de padrões dentro da massa de dados, que são gravados nas patentes depositadas e que foram recuperadas pela expressão de busca descrita em detalhes na fase anterior. Para realizar tais análises, serão consideradas as funcionalidades disponibilizadas no Patent2Net v.2 – tabelas dinâmicas, análises de rede, mapas mentais, análises de conteúdo e análise de agrupamento ou *clustering*.

4 APRESENTAÇÃO DOS RESULTADOS

Conforme definido na metodologia, utilizou-se como instrumento de pesquisa o KDD para responder a questão de pesquisa: Como o uso do KDD em informações patentárias pode contribuir com o ensino a distância em instituições de ensino superior?

Neste sentido, para melhor organização deste capítulo, abordam-se as quatro dimensões da pesquisa de forma separada de modo a possibilitar análise: a) Mapeamento das patentes em EAD no mundo; b) Metodologia de busca de Patentes em EAD por análise de Conteúdo; c) Mapeamento das patentes em EAD em Universidades; e, finalmente, e) Uso de informações patentárias na definição de estratégia em IES à luz do RBV.

4.1 MAPEAMENTO DAS PATENTES EM EAD NO MUNDO

Para representar o diagnóstico das patentes em EAD no mundo, optou-se pela abordagem da disponibilização pública das informações sobre as patentes em EAD, do diagnóstico geral dos depósitos de patentes em EAD, e da possibilidades de cruzamento utilizando tabelas dinâmicas. Adicionalmente, utilizou-se um mapa de geolocalização dos países para mostrar visualmente a localização geográfica dos inventores, das empresas solicitantes, e onde a proteção da tecnologia foi solicitada.

Abordou-se também o diagnóstico das patentes concedidas em EAD, juntamente com as possibilidades de cruzamento relacionados às patentes concedidas ou em processo de concessão sobre o tema, além da possibilidades de seleção de patentes com base na combinação de indicadores.

Finalmente, demonstrou-se visualmente as redes relacionadas às patentes em EAD e as de equivalência entre as tecnologias e documentos citados, as análises relacionadas às famílias de patentes, e o *Mind map* entre patentes.

4.1.1 Disponibilização pública das informações sobre as patentes em EAD

A Figura 8 traz a tela inicial que contém os resultados de todas as extrações realizadas até o momento com o Patent2net. Nela é possível verificar que assuntos que já foram realizados a

mineração de patentes nos diversos temas como “drone”, “terrás raras”, “impressoras 3D”, além do tema “EAD” (apontado pela seta azul).

Resumidamente, durante o seu funcionamento a ferramenta computacional *Patent2net* executa em sequência alguns módulos computacionais bastante específicos. De início, o usuário deve criar *login* na base de patentes *Espacenet* para que esta possibilite o trabalho de mineração de dados. Logo após o registro, o *crawler* extrai os textos integrais das patentes diretamente da base, convertendo estes textos de forma que possam ser analisados. Em sequência, o *Patent2net* monta as diversas redes propostas com base nas patentes, em conjunto com o *software* livre *Gephi*. Ainda, faz uma análise *full text* dos documentos patentários, em conjunto com o *software* *Iramuteq*, cria diversas tabelas de acesso dinâmico com o auxílio do *freeware Pivottable*, e, ao final, gera arquivos no formato *html* que poderão ser consultados localmente ou enviados à *web* para consulta de qualquer local.

Para se ter acesso aos resultados das extrações e tratamentos sobre as patentes em EAD, foco do presente trabalho, basta clicar sobre a palavra representativa do tema que se deseja consultar, neste caso “EAD” (seta azul na Figura 1), que na realidade é um *link* de acesso aos resultados estáticos e às tabelas dinâmicas que permitirão avaliar em profundidade as extrações realizadas.

Figura 8 – Tela inicial com os temas sobre os quais já foi realizada a mineração de patentes, com destaque para as patentes em EAD.



Fonte: <http://200.136.214.89/~patent/>

Na Figura 9 se encontram os *links* de acesso às sete modalidades de avaliação das 3090 patentes em EAD obtidas pelo processo de mineração de dados disponíveis na Espacenet.

O primeiro deles, "Patents datatable, Pivot table ", permite acesso à *interface* dinâmica para consultas direcionadas às patentes em EAD e também à tabela dinâmica para realização de cruzamentos entre diversos indicadores relacionados a essas patentes. Tais cruzamentos são apresentados mais à frente neste mesmo capítulo. Com isso, é possível delimitar a pesquisa de forma a permitir identificar, com base em uma série de critérios, as patentes que se deseja selecionar, como por exemplo, realizar cruzamentos entre os países das empresas depositantes e a tecnologia que se buscou proteger, entre a data de depósito da patente e o país de depósito, entre o país do inventor e o estatuto jurídico da patente, dentre uma série de outras possibilidades.

O segundo *link*, “Attractivity: Geolocalisation of patent covering (without EP, WO), Applicants, Inventors (when available)”, permite identificar geograficamente quais são os países com o maior número de empresas financeiras das tecnologias presentes nos

documentos patentários, quais são os países com o maior número de inventores, além de em quais países existe o maior número de solicitações de proteção de invenções.

O terceiro *link*, "Networks (Inventor, Applicant, Technology)", fornece acesso às redes de colaboração entre inventores, entre empresas, e entre as tecnologias mais estudadas sobre o tema foco da presente experimentação.

O quarto *link*, "Mixed Networks (Country-Technology, Inventor-Technology, Applicant-Technology, Applicant-Inventor)", disponibiliza redes mais complexas, baseadas nas relações existentes entre a tecnologia e o país depositante, entre a tecnologia e o inventor e entre a tecnologia e a empresa depositária. Adicionalmente, é possível ter acesso a rede entre os inventores e as empresas depositárias das patentes em EAD.

O quinto *link*, "Equivalents, Reference (References to other patents or External references), Patents citations networks", fornece acesso à rede em que é possível avaliar as relações entre as tecnologias equivalentes àquelas descritas nos documentos de patentes, à rede entre as patentes que citam os mesmos artigos científicos em sua lista de referências, assim como à rede de citações entre as diversas patentes.

O sexto *link*, "Patent family & Pivot table", possibilita acessar uma *interface* dinâmica onde podem ser realizadas uma série de consultas direcionadas às famílias das patentes, assim como a uma tabela dinâmica onde diversos cruzamentos relacionados a essas famílias podem ser realizados de maneira semelhante às possibilidades oferecidas pelo primeiro *link* já descrito.

Por fim, o sétimo *link*, "IPC's Mind-Map (FreePlane PlugIn)", possibilita avaliar as categorias das patentes de maneira organizada e em forma de fluxograma, tomando por base os títulos dos documentos.

Figura 9 – Tela apresentando os 7 links de acesso aos diversos resultados relacionados à análise das 3090 patentes em EAD.

On-line analysis tools:

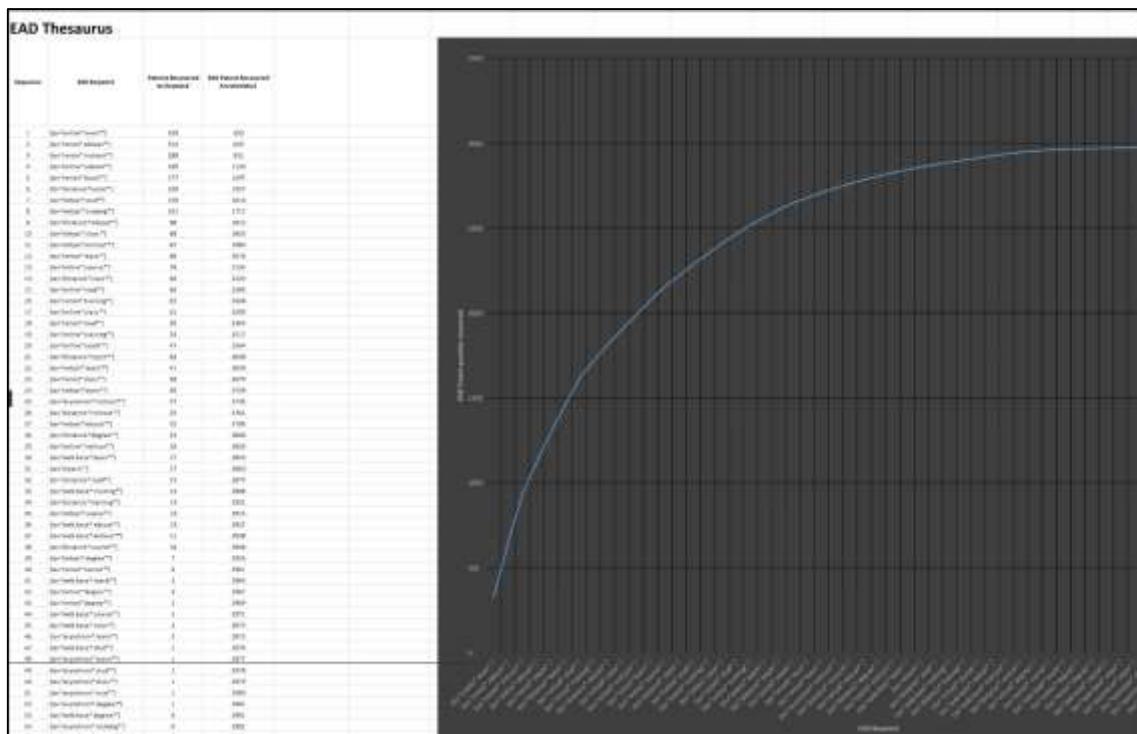
- Patents datable, Pivot table
- Attractivity: Geolocalisation of patent covering (without EP, WO), Applicants, Inventors (when available)
- Networks (Inventor, Applicant, Technology)
- Mixed Networks (Country-Technology, Inventor-Technology, Applicant-Technology, Applicant-Inventor)
- Equivalents, Reference (References to other patents or External references), Patents citations networks
- Patent family & Pivot table
- IPC's Mind-Map (FreePlane Plugin)

Fonte: <http://200.136.214.89/~patent/DONNEES/Eadv4.html>

4.1.2 Mapeamento geral dos depósitos de patentes em EAD

Os dados apresentados a seguir são oriundos da experimentação que tomou por base uma estratégia de pesquisa pré-definida, delimitada de acordo com os vários sinônimos do termo “EAD”. Esta estratégia, apresentada no rodapé da Figura 10, foi definida com base no momento em que se atingiu a estabilidade da curva de sinônimos, apresentada na mesma Figura 10. Essa estabilidade significa que, quando foram inseridos novos sinônimos do termo “EAD” na estratégia de pesquisa e havia baixo ou nenhum incremento de novas patentes consultadas na *Espacenet*. Dessa forma, considerou-se esse conjunto de palavras-chave como os mais representativos para se proceder o processo de mineração de dados e geração dos resultados pelo Patent2net.

Figura 10 – Montagem da curva de estabilidade para definição da estratégia de pesquisa (rodapé) das patentes em EAD com a utilização da ferramenta Patent2net



Na Figura 10 observa-se que está ordenada do maior para o menor a frequência de patentes que retornam da busca, dependendo da palavra-chave. As palavras-chave que tiveram o maior retorno foram ta= “online* learn*” com 329 patentes e “remot* or educat*” com 314. Já as palavras com menor retorno de patentes foram “asynchron* stud*”, com 1 patente, e “asynchron* degree*”, também com uma patente localizada na base de 90 milhões de documentos.

Note que o asterisco “*” foi utilizado ao final de cada palavra para poder capturar toda extensão da palavra após o asterisco. Exemplo learn*, trouxe patentes relacionadas com: learn, learning etc. Outro ponto importante desta pesquisa é que foram usadas 72 duas palavras chave, que retornaram 3090 patentes. Destas, os resumos são em sua grande maioria em inglês (2.870, EN), 256 em francês (FR), 10 em Alemão (DE) e 31 em outras línguas (OL). Todavia, nesta análise deve ser levado em consideração que uma patente pode apresentar mais de um *abstract* e em diversos idiomas.

O levantamento de um conjunto de patentes que se refiram a um determinado tema nunca será uma ciência exata e que leve a um número absoluto, já que podem existir patentes que falam de um tema específico, porém sem utilizar as palavras eleitas na estratégia de pesquisa. Dessa forma, sempre será possível encontrar novas palavras-chave que identifiquem outras patentes sobre o mesmo tema, fazendo com que este universo seja considerado como uma amostra, e não como a população completa das patentes sobre o tema EAD. Tal amostra é formada pela junção de patentes oriundas de fontes heterógenas, o que implica que os tratamentos realizados nem sempre representarão a totalidade dos documentos. Os números indicados em cada experimento têm o sentido apenas de “validar” a significatividade da amostra, muito mais do que representar valores absolutos. São observadas tendências relativas e não contagens absolutas.

4.1.3 Possibilidades de cruzamento de dados com a utilização de tabelas dinâmicas

A Figura 11 apresenta em sua totalidade a interface em formato de tabela dinâmica, disponível no Patent2net, onde quaisquer indicadores relacionados às patentes selecionadas podem ser conduzidos tanto para o eixo X quanto para o eixo Y, isoladamente ou em conjunto. A interface também permite escolher a forma de apresentação dos dados, seja em formato de mapas de calor, de gráfico de linhas, gráfico de barras, gráfico de área, dentre outras opções. Os 13 indicadores disponíveis para combinação são "year" (ano de depósito), "kind" (estatuto jurídico da patente), "Inventor-Country" (país do inventor), "prior-Date" (data exata de depósito), "country" (país onde a proteção foi solicitada), "applicant" (empresa financiadora da patente), "IPCR4" (código de classificação internacional com 4 dígitos), "IPCR7" (código de classificação internacional com 7 dígitos), "label" (rótulo da patente), "representative" (representatividade do documento), "inventor" (nome do inventor), e "Cita" (número de citações recebidas por outros documentos patentários). O Acesso a essa *interface* ocorre ao clicar sobre o primeiro link disponível na Figura 9 ("Patents datatable, Pivot table").

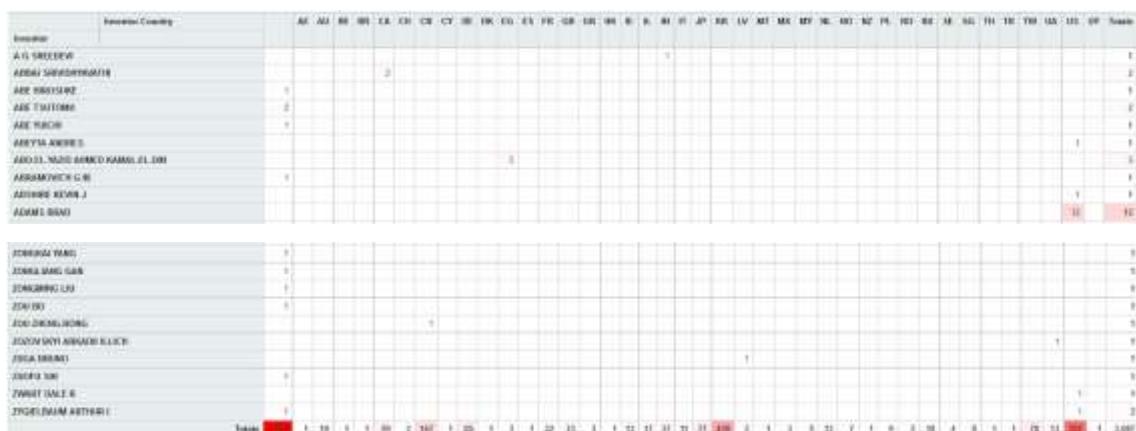
Figura 11 – Cabeçalho da interface que permite realizar cruzamentos entre os 13 indicadores disponíveis e relacionados às patentes



Fonte: <http://200.136.214.89/~patent/DONNEES/Eadv4/Eadv4Pivot.html>

Já na Figura 12, como exemplo, o termo “inventor”, que se refere ao inventor (pessoa física responsável pelo depósito da patente), está alocado no eixo X, enquanto o termo está alocado no eixo Y. Este exemplo de cruzamento, a despeito do fato de que muitos documentos de patentes não possuem o registro do país do inventor, permite avaliar quem são as pessoas que efetivamente produzem o conhecimento com relação ao tema EAD. Neste exemplo, embora parcial, destaque pode ser dado ao inventor “Adams Brad”, com 12 solicitações de patentes. Este inventor informou os Estados Unidos (US) como país de origem, que por sua vez também se destacaram como o país de origem mais informado pelos inventores, com 761 patentes depositadas sobre EAD, seguidos da Coreia do Sul (KR), com 439 solicitações, da China (CH), que totalizou 167 e do Japão (JP) com 77 pedidos de depósito de patentes. Como visto na Figura 12, e também na maioria das Figuras apresentadas em sequência, foram disponibilizadas apenas as partes superior e inferior da tela onde os referidos resultados se encontram disponibilizados, visto que a apresentação da tela em sua totalidade tornaria impossível a visualização dos dados com clareza.

Figura 12 – Parte superior e inferior da tela que avalia a relação entre o inventor e o seu país de origem



Fonte: <http://200.136.214.89/~patent/DONNEES/Eadv4/Eadv4Pivot.html>

Na Figura 13, que também demonstra apenas as partes superior e inferior de uma das telas representativas da análise de resultados, o termo “applicant”, que se refere à empresa responsável pelo depósito da patente, foi alocado no eixo X, enquanto o termo “Applicant-Country”, que define o país de origem da empresa depositante, foi posicionado no eixo Y. A análise dos resultados oriundos deste cruzamento permite, por exemplo, identificar quais são as empresas que efetivamente investem no patenteamento de tecnologias voltadas ao EAD, bem como qual é o país de origem declarado por estas corporações. Neste exemplo, nota-se que a empresa “ZTE Corp”, depositou 7 patentes. Em 6 delas, a empresa citou a China como país de origem, enquanto em 1 documento, o país de origem não foi declarado no momento da solicitação de patenteamento.

Figura 13 – Parte superior e inferior da tela que avalia a relação entre a empresa depositante e o seu país de origem



Fonte: <http://200.136.214.89/~patent/DONNEES/Eadv4/Eadv4Pivot.html>

A Figura 14 traz os resultados do cruzamento entre os termos “Applicant-Country” (país da empresa depositante) e “kind” (estatuto jurídico da patente). Entenda-se por estatuto jurídico da patente a situação do documento depositado perante aos órgãos avaliadores da proposta de patenteamento. Esta classificação se baseia em uma série de diferentes códigos (A, A0, A1, A2, A3, A4, A8, A9, B, B1, B2, C, C1, C2, U, U1 e Y), que determinam, por exemplo, se a patente se encontra em avaliação, se foi negada, se foi concedida, dentre outras situações mais específicas.

Nesta Figura 14 é possível ainda identificar, por exemplo, quais são os países que detém tanto o maior número de solicitações, quanto o maior número de patentes concedidas voltadas ao EAD. Neste caso, a despeito do fato de que em 1.501 patentes o país de origem da empresa solicitante não tenha sido declarado, destaque pode ser dado às empresas Americanas, com 611 solicitações, seguidas das empresas Sul Coreanas, com 427 pedidos, e das empresas Chinesas, com 148 depósitos. Já em relação à distribuição das patentes de acordo com o estatuto jurídico, o código A1 (“European patent application published with European search report”) lidera a lista com 949 solicitações, seguido do código U (“first publication level”) com 242, e do código A2 (“European patent application published without European search report (search report not available at publication date)”), com 132 pedidos.

Figura 14 – Tela que demonstra a relação entre o país da empresa depositante e o estatuto jurídico da patente

Applicant Country	Kind	A	AB	AF	A2	AA	B	B1	B2	C	C1	C2	E1	U	M1	Y	T1	Total
AE	1																	1
AU	1	7																9
BR		1																1
CA	2	37	2															43
CH		7																8
CN	74	29	1															143
DE	12	27	19	9	1													81
ES		2	1															3
FR		16	1															16
GB	3	11	2															16
GR	1	2																3
HK		1	1															2
IE	1	6	7															14
IL		10	1															11
IN	2	13	3															28
IT	2	6	2															10
JP	23	91	8	2	1													179
KR	200	17	8															227
LV	1	1																2
MT		1																1
MX	1		1															2
MY	1	2																4
RU		1																1
SE	8	9	4															24
SG	1	1																2
SI		1	1															2
TH		1																1
TR	21	19	2	10														76
US	29	419	70	4	44	3	1											871
VG		1																1
Total	2	109	129	9	23	101	4	3	7	1	202	1	46	1	2007	1	2007	1

Fonte: <http://200.136.214.89/~patent/DONNEES/Eadv4/Eadv4Pivot.html>

Ao avaliar a Figura 15, notam-se alguns dos resultados do cruzamento entre os indicadores “prior-Date” (data de depósito) e “country” (país no qual se solicitou a proteção da patente). Verifica-se que o primeiro depósito de uma patente sobre o tema EAD ocorreu em 13 de agosto de 1921, tendo sido solicitada pela Grã-Bretanha. No rodapé da Figura é possível verificar o total de patentes solicitadas em cada país desde 1921 até 2015, destacando-se a China (850 depósitos), Estados Unidos (704 depósitos), e Coreia do Sul (449 depósitos). Verifique que a segunda patente depositada sobre EAD ocorreu apenas em 1964, 43 anos depois. Nos últimos anos tem se notado um aumento no número de solicitações de patente sobre o tema EAD, o que indica a importância que este assunto vem tomando.

Figura 15 – Tela que demonstra a relação entre a data de depósito e o país onde a patente foi depositada.

prior-Date	country	AU	CA	CN	DE	EP	FR	GB	GR	HK	IE	IL	IN	JP	KR	LV	MX	MY	NZ	PL	RO	RU	SG	SU	TW	UA	US	WO	Totals
1921-08-13									1																				1
1964-05-25										1																			1
1976-07-23											1																		1
1976-12-16																													1
1977-03-11												1																	1
1977-11-07																													1
1977-12-15													1																1
1978-03-10																1													1
1978-05-30																	1												1
1979-03-26																		1											1
2015-06-03														1															1
2015-06-04															2														2
2015-06-06															1														1
2015-06-15																1													1
2015-06-23																													1
2015-06-30															1														1
2015-07-03															1														2
2015-07-14																													1
2015-07-21																1													1
2015-09-30																		1											1
Totals		57	67	2	12	76	9	14	2	1	6	3	9	416	489	1	11	2	2	6	2	16	10	1	82	13	109	258	3,087

Fonte: <http://200.136.214.89/~patent/DONNEES/Eadv4/Eadv4Pivot.html>

Já na Figura 16, encontram-se dispostos os resultados do cruzamento entre o indicador “year” (ano do depósito) e “Applicant-country” (país no qual se solicitou a proteção da patente). Neste cruzamento, destaque pode ser dado para as empresas Americanas, com 611 solicitações, seguidas das Sul Coreanas, com 427 solicitações, e das Chinesas, com 148 pedidos. A referida análise permite identificar, por exemplo, a evolução quantitativa dos depósitos de patentes pelas empresas de cada país nos últimos 90 anos. Ao observar mais de perto, percebe-se um forte incremento nos pedidos de patentes das empresas Americanas e Sul Coreanas nos últimos 5 anos.

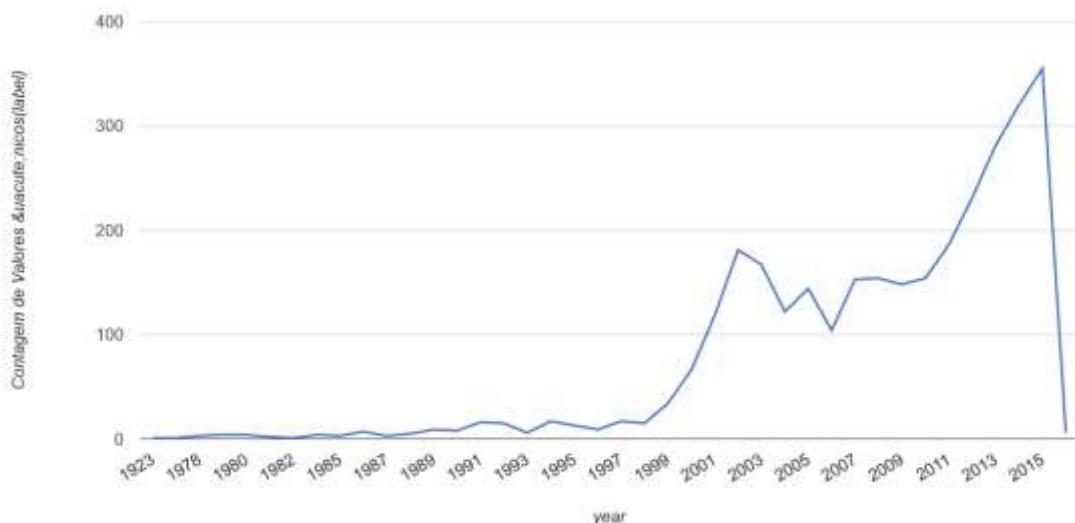
Figura 16 – Tela que demonstra a relação entre o ano de depósito e o país onde a patente foi depositada

Applicant-Country	All	AD	BR	BS	CA	CH	CN	DE	EG	FR	GB	GR	HK	IE	IL	IN	IT	JP	KR	LV	MT	MX	MY	NA	NL	NO	PL	RO	NU	SE	SG	SU	TH	TW	UA	US	VG	Total
year																																						
1929	1																																1					
1930	1																																1					
1931	1																															3						
1932	3																																6					
1933	4																																6					
1934	2																																2					
1935	1																																1					
1936	3																																6					
1937	2																																3					
1938	1																																1					
1939	1																																1					
1940	1																																1					
1941	2																																2					
1942	1																																1					
1943	3																																3					
1944	2																																3					
1945	7																																7					
Total	1	0	1	2	43	0	148	81	3	16	19	3	2	54	15	21	10	79	427	2	1	2	4	1	34	2	8	2	18	4	5	1	1	76	13	111	1	3087

Fonte: <http://200.136.214.89/~patent/DONNEES/Eadv4/Eadv4Pivot.html>

Uma forma alternativa de apresentar os dados disponíveis na Figura 16 é por meio de um gráfico de linhas, dentre outras opções disponíveis na ferramenta. Tais gráficos são montados e disponibilizados automaticamente pela ferramenta Patent2net, de maneira dinâmica e a critério do usuário, não sendo necessária a exportação de dados para outros aplicativos. Na Figura 17, os dados quantitativos apresentados na Figura 16, agora em formato de gráfico de linhas, permitem identificar o considerável aumento do número de depósitos a partir do final da década de 1990, com destaque especial para os picos de depósitos ocorridos nos anos de 2003 e 2015 com forte expansão a partir de 2011.

Figura 17 – Representação gráfica da evolução quantitativa dos depósitos das patentes em EAD.



Fonte: <http://200.136.214.89/~patent/DONNEES/Eadv4/Eadv4Pivot.html>

Os dados parciais disponíveis na Figura 18, se referem ao cruzamento dos indicadores IPCR4 com o número de citações recebidas (“citations”) por uma patente, considerando o fato de que tais citações foram recuperadas de outros documentos patentários. O número de citações, de certa forma, representa o quanto a patente citada é importante para o desenvolvimento de outras tecnologias complementares à descrita no documento que a ela se referiu. Já o IPCR4, é um código internacional que fornece algumas informações gerais sobre as tecnologias exploradas na patente depositada e classificada com a referida codificação. No exemplo, a tecnologia H04L (“Transmission of digital information, e.g. telegraphic communication”), lidera a lista com 4.680 citações, seguida da H04Q (“Selecting (switches, relays, selectors”), com 2.567 citações, e da H04M (“Telephonic communication (circuits for controlling other apparatus via a telephone cable and not involving telephone switching apparatus”)), com 1.331 citações, totalizando 48.672 citações.

Figura 18 – Parte superior (figura à esquerda) e inferior (figura à direita) da tela que avalia a relação entre o IPCR4 e o número de citações recebidas pela patente.

	Cita	NaN	Totais		
IPCR4				H04J	53
					63
	485	485		H04L	4.680
A01F	1	1			4.680
A01G	4	4		H04M	1.331
A01K	12	12			1.331
A42B	2	2		H04N	2.567
A47B	168	168			2.567
A47C	3	3		H04Q	209
A47J	2	2			209
A47L	8	8		H04R	10
A61B	138	138			10
A61F	5	5		H04S	38
					38
				H04W	725
					725
				H05B	28
					28
				H05K	46
					46
				Totais	48.672
					48.672

Fonte: <http://200.136.214.89/~patent/DONNEES/Eadv4/Eadv4Pivot.html>

Ainda levando em consideração a principal tecnologia descrita no documento de patente, a Figura 19 demonstra o resultado do cruzamento entre o indicador IPCR4 com o país do inventor (“Inventor-Country”). Neste exemplo, destacam-se as tecnologias representadas pelos códigos H04N (“Pictorial communication, e.g. television”, com 292 solicitações), H04L (246 pedidos) e H04W (“Wireless communication networks”), com 51 depósitos. Levando em consideração apenas as solicitações relacionadas à tecnologia descrita pelo código H04N, dos 292 pedidos de concessão, 34 foram feitos por inventores americanos, 33 por japoneses, e 29 por chineses. A relação entre IPCR4 e país do inventor permite avaliar em quais tipos de tecnologia os inventores de cada nacionalidade têm dedicado seu tempo de pesquisa, bem como identificar o foco dos recursos financeiros destinados para esse fim por cada país.

Figura 19 – Parte superior e inferior da tela que demonstra a relação entre o IPCR4 e o país do inventor.

Fonte: <http://200.136.214.89/~patent/DONNEES/Eadv4/Eadv4Pivot.html>

A Figura 20 disponibiliza o cruzamento do indicador IPCR4 com o país da empresa depositária (“Applicant-Country”). Neste exemplo, as mesmas tecnologias representadas pelos códigos H04N (292 solicitações), H04L (2468 pedidos) e H04W, com 51 depósitos, foram as que mais se destacaram. Detalhando os resultados relativos à tecnologia representada pelo IPCR H04N, das 292 solicitações registradas, 33 foram solicitadas por empresas japonesas, 30 por empresas americanas, e 29 por empresas chinesas. ao relação do IPCR4 com o país da empresa financiadora da patente, permite realizar uma avaliação dos interesses de cada empresa na busca pelo desenvolvimento de novas tecnologias em determinado assunto.

Figura 20 – Parte superior e inferior da tela que demonstra a relação entre o IPCR4 e o país da empresa depositante.

Applicant-Country	AE	AU	BR	BS	CA	CH	CR	DE	EG	FR	GB	GR	HK	I	IL	IN	IT	JP	KR	LV	M	MK	MY	NA	NL	NO	PL	RO	RU	SE	SG	SU	TH	TW	UA	US	VG	Total
IPC48	9	1												1	1	1	2	1	64											3	2	9	97					
A41F	1																																1					
A41G	1																																2					
A41H	1																																1					
A42B	2																																2					
A43B	6																																8					
A44C	3																																3					
A47J	1																																1					
A47L	1																																1					
A47S	17																															4	1	22				
A47F	9																																3	3				
H04J	1																																1	4				
H04L	106	2			12			17	6			4	1			3	1	8	8	1										1	48	246						
H04M	29							4	1			1						1	2												12	58						
H04N	292		2	1	28			3	1			2	4	33	27	4													2	4	30	292						
H04Q	19		2		1			1										1													5	27						
H04R	3																																2	5				
H04S	1																																4	6				
H04W	59							1	4										18													8	51					
H05B	4																		1														5					
H05K	2																		1														1	4				
Total	207	1	9	1	2	43	9	148	49	3	16	10	3	2	14	11	21	10	79	42	2	1	2	4	1	38	2	8	5	1	1	76	13	401	1	3.087		

Fonte: <http://200.136.214.89/~patent/DONNEES/Eadv4/Eadv4Pivot.html>

Seguindo a mesma linha de raciocínio apresentada nas Figura 19 e Figura 20, ou seja, com base no cruzamento do IPCR4 no eixo X, só que desta vez com o país onde se buscou proteger a patente no eixo Y (representado pelo indicador “country”), a Figura 21 também destaca as tecnologias descritas pelos códigos H04N (292 solicitações), H04L (246 pedidos) e H04W, com 51 depósitos. Neste cruzamento, com relação á tecnologia representada pelo código H04N, nota-se que a maioria das solicitações de proteção ocorreu na China (107 documentos), seguida do Japão (69 patentes), e dos Estados Unidos (43 solicitações). A relação entre o país de proteção da patente e o IPCR4, possibilita identificar quais são as principais tecnologias que as empresas ou inventores consideram importante proteger em cada país.

Figura 21 – Parte superior e inferior da tela que demonstra a relação entre o IPCR4 e o país de depósito.

IPCR4	Country	AU	CA	CN	DE	EP	FR	GB	GR	HK	IE	IL	IN	JP	KR	LV	MX	MY	NZ	PL	RO	RU	SG	SU	TW	UA	US	WO	Total	
A01F					1					1		1	2		77								7	1		7	37			
A01G														1														1		
A01K																1												2	4	
A42B																													2	
A47B																													8	
A47C																													3	
A47J																													1	
A47L																1												1		
A61B																	4											4	23	
A61F																1												1	3	
H04J																														
H04L		6	12	123			5		1	1					13	3	1						6	43	27	246				
H04M		1	5					1							20	3									14	2	50			
H04N		6	5	107			10	2			1			69	20	1	1	1					4	4	46	16	292			
H04Q		2	2												15	1									4	3	27			
H04R			2													1	1									1	1	3		
H04S																	1	1								2	2	6		
H04W				21											8	11			1						6	5	51			
H05B				4													1												5	
H05K				2																									1	4
Total	67	67	123	12	76	9	14	2	1	6	3	9	414	420	1	11	2	2	6	2	16	10	1	82	13	16	254	3.087		

Fonte: <http://200.136.214.89/~patent/DONNEES/Eadv4/Eadv4Pivot.html>

A Figura 22 associa o estatuto jurídico (“kind”, eixo Y), também com o IPCR4 (eixo X). Levando em consideração as tecnologias vinculadas aos códigos H04N (292 solicitações), H04L (246 pedidos), e H04W (51 depósitos), que mais uma vez se destacaram entre as demais, e elencando a tecnologia H04N para uma análise mais aprofundada, nota-se que 174 documentos possuem o estatuto jurídico "A", seguidos daqueles cujo estatuto jurídico é representado pelo código "A1" (72 solicitações), e daqueles cuja solicitação se enquadra no código "U" (21 documentos). Avaliar a relação entre o IPCR4 e o estatuto jurídico da patente, possibilita identificar quais são as tecnologias cujas patentes estão sob avaliação, quais foram negadas (tornando-se fonte livre de conhecimento), e quais efetivamente foram concedidas.

Figura 22 – Parte superior e inferior da tela que demonstra a relação entre o IPCR4 e o estatuto jurídico da patente.

Kind	A	A0	A1	A2	A4	B	B1	B2	C	C1	C2	E1	U	U1	Y	Y1	Totais
IPCR4																	
		B1		B	7										1	97	
A01F	1																1
A01G												1		1			2
A01K	2		1									1					4
A42B			1									1					2
A47B	3		1									2		2			8
A47C	1											2					3
A47J	1											2					1
A47L	1																1
A61B	8		2				1					11		1			23
A61F	1		1	1													3
H04J	1		2									1					4
H04L	129		72	11		6	4					22		2			240
H04M	28		15	1		1						3		2			50
H04N	174		62	8	2	8			1		31	6					392
H04Q	17		8			1						1					27
H04R	2		2	1													5
H04S	2		2	2													0
H04W	30		10			3						8					51
H05B	2											3					5
H05K	2			1								1					4
Totais	1.000	2	343	132	5	23	101	4	3	7	8	1	242	1	48	1	3.087

Fonte: <http://200.136.214.89/~patent/DONNEES/Eadv4/Eadv4Pivot.html>

Por sua vez, a Figura 23 disponibiliza a evolução temporal relacionada ao surgimento de determinadas tecnologias (ainda com base no IPCR4), levando em consideração para tal a data de depósito do documento patentário. No exemplo, é possível identificar que o triênio 2012-2014 foi aquele onde ocorreu o maior número de pedidos de patentes, com 425 solicitações no ano de 2012, 604 no ano de 2013, e 686 no ano de 2014. Relacionar o IPCR4 com a data de depósito da patente, permite traçar um perfil da evolução ao longo do tempo das tecnologias desenvolvidas em um assunto de interesse, como é o caso do EAD.

Figura 23 – Parte superior e inferior da tela que demonstra a relação entre o IPCR4 e a data de depósito da patente.

IPC4	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	Total
A01B																																				109		
A01C																																				3		
A01D																																				3		
A01F																																				4		
A01H																																				2		
A01J																																				0		
A01K																																				18		
A01L																																				0		
A01M																																				0		
A01N																																				12		
A01P																																				0		
A01R																																				0		
A01S																																				0		
A01U																																				3		
H01R																																				2		
H02J																																				6		
H02P																																				0		
H03B																																				28		
H04H																																				20		
H04U																																				4		
H04V																																				0		
H04W																																				33		
H04W8																																				33		
H04W88																																				7		
H04W96																																				0		
H05B																																				37		
H05B9																																				2		
H05B95																																				0		
H05B99																																				0		
H05B99.1																																				0		
H05B99.2																																				0		
H05B99.3																																				0		
H05B99.4																																				0		
H05B99.5																																				0		
H05B99.6																																				0		
H05B99.7																																				0		
H05B99.8																																				0		
H05B99.9																																				0		
Total	0	0	4	1	8	2	4	10	19	7	18	12	8	21	21	48	94	177	268	327	192	162	132	129	106	209	181	187	167	169	154	154	154	4 888				

Fonte: <http://200.136.214.89/~patent/DONNEES/Eadv4/Eadv4Pivot.html>

4.1.4 Mapa de geolocalização dos países (inventores, das empresas solicitantes, onde a proteção da tecnologia foi solicitada)

Outra forma de se analisar as patentes é por meio da identificação visual dos países dos inventores, das empresas solicitantes, além dos países onde a proteção da tecnologia foi efetivamente solicitada. Tais mapas de geolocalização podem ser acessados clicando no segundo *link* demonstrado na Figura 9 ("Attractivity: Geolocalisation of patent covering (without EP, WO), Applicants, Inventors when available"). A Figura 24 possibilita identificar que o maior número de inventores das tecnologias voltadas ao Ensino a Distância declarou os Estados Unidos como país de origem, seguidos de Taiwan e China.

Figura 24 – Mapa de geolocalização levando em consideração os países declarados como sendo de origem pelos inventores das tecnologias em EAD.



Fonte: <http://200.136.214.89/~patent/DONNEES/Eadv4/Eadv4CartoInventor.html>

A Figura 25 permite verificar que a maioria das empresas que investiram no desenvolvimento de tecnologias em EAD declarou como país de origem os Estados Unidos, também seguidos de Taiwan e China.

Figura 25 – Mapa de geolocalização levando em consideração os países declarados pelas empresas detentoras da tecnologia cuja patente foi solicitada



Fonte: <http://200.136.214.89/~patent/DONNEES/Eadv4/Eadv4CartoApplicant.html>

Já a Figura 26 demonstra que a maior parte dos pedidos de proteção ocorreu na China, seguida por Estados Unidos e Japão, dentre outros países cujo número de depósitos é menos expressivo. Nesse ponto, vale ressaltar que o referido mapa não utiliza como base o número de solicitações de patentes por país, seja com base no país do inventor ou da empresa solicitante, mas, sim, leva em consideração o país no qual a proteção foi solicitada.

Figura 26 – Mapa de geolocalização levando em consideração os países onde foi solicitada a proteção da tecnologia por meio do registro das patentes.



Fonte: <http://200.136.214.89/~patent/DONNEES/Eadv4/Eadv4Carto.html>

4.1.5 Mapeamento das patentes concedidas em EAD

Após a apresentação inicial, que levou em consideração todas as patentes depositadas sobre o tema EAD, independentemente de seu estatuto jurídico, serão apresentadas, a partir deste ponto, as análises que englobaram apenas as patentes efetivamente concedidas, ou seja, aquelas cujo estatuto jurídico se enquadra nas categorias B e C. Para acessar estes dados basta clicar sobre o primeiro *link* disponível na Figura 9 ("Patents datatable, Pivot table"), todavia marcando no indicador "kind" (estatuto jurídico da patente), apenas os códigos relacionados às patentes concedidas ou em processo de concessão.

4.1.6 Possibilidades de cruzamento relacionados às patentes concedidas ou em processo de concessão sobre o tema EAD

A Figura 27 é apresentada no formato de "mapa de calor", em que a intensidade do vermelho é diretamente proporcional à frequência absoluta de um evento específico, neste caso representado pelo número de patentes concedidas ou em processo de concessão, distribuídas

com base no país declarado pelo inventor. Na Figura, é possível identificar que, das 146 patentes concedidas ou em processo de concessão, os Estados Unidos, com 55 patentes concedidas ou em processo de concessão em EAD, lideram o *ranking* entre os países de origem dos inventores, seguidos da Coreia do Sul (37 patentes), e Taiwan (14 patentes).

Figura 27 – Tela que demonstra a relação entre as patentes concedidas e o país do inventor.

Inventor-Country	kind	B	B1	B2	C	C1	C2	Totals
		7	4		2			13
AU		1						1
CA		3						3
CN		4						4
DE		1	2					3
FR		1						1
IE		1						1
IN		1	1					2
JP		2	3					5
KR		1	1					2
NL		1						1
RU				5	5			10
TR		1						1
TW		14						14
UA		4	1			3		8
US		4	10	3				17
Totals		23	10	4	3	7	8	146

Fonte: <http://200.136.214.89/~patent/DONNEES/Eadv4/Eadv4Pivot.html>

Por sua vez, A Figura 28, por meio de um “mapa de calor”, demonstra a possibilidade de uma análise com múltiplos indicadores alocados no eixo X, mas que poderiam também ser alocados no eixo Y, a critério de quem realiza as análises. Na figura, os indicadores “Applicant-Country” (país da empresa depositária) e “Applicant” (empresa), foram cruzados com o indicador “kind” (estatuto jurídico). Este cruzamento possibilita identificar quais são as empresas detentoras das tecnologias cujas patentes foram efetivamente concedidas ou se encontram em processo de concessão, bem como em que país tais empresas estão localizadas.

Embora na Figura 28 tenham sido apresentadas apenas as partes superior e inferior da tela, é possível destacar o termo “empty”, que na realidade representa empresas que não declararam sua razão social ou nome de fantasia no momento da solicitação do patenteamento, tão pouco seu país de origem, com 15 patentes concedidas ou em processo de concessão na área de Ensino a Distância, sendo 12 patentes classificadas como B1, e 3 como C2. Em seguida, destaca-se nessas telas a empresa “Unext.com”, com 2 patentes concedidas no código B1, e que

declarou os Estados Unidos como país de origem. Na Figura ainda é possível identificar que o código B1 é aquele com maior número de concessões (101 patentes).

Figura 28 – Parte superior e inferior da tela que demonstra a relação entre as patentes concedidas ou em processo de concessão, a empresa depositante e o seu país de origem.

Applicant/Country	kind	B	B1	B2	C	C1	C2	Total
ABRAMOVICH GREGORIJ MIKHAILOVI					1			1
ACTV INC			1					1
DIGITAL EQUIPMENT CORP				1		1		2
empty								
IBM		1				3		4
LUTAEERIKO VJACHESLAV FEDOROVICH								
MATSIEVSKIY JURII BRONISLAVOV								
MPOB STROJTECHAUTOMATIKA AGZT								
PERFIL EV SERGEJ OLEGOVICH								
SOCIAL IT PTY LTD			1					1
TSENTRAL MYU NI ISPITATEL MYU						1		1
VOENNYJ INST RADIOTELEKTRONIKI								
WALTER AG			1					1
AU								
GOUGOURIS KAMELLA				1				1
MICNAMARA MICHELLE				1				1
SABBAGH-JAFARI SEYED MOJTABA				1				1
SOCIAL IT PTY LTD				1				1
RAMACHANDRAN RAMKUMAR					1			1
RICHARDS WILL					1			1
RIVAS ALVAREZ VICTOR					1			1
SAMSUNG ELECTRONICS CO LTD					1			1
SHAPIRO ARTHUR						1		1
SIEMENS CORP RES INC					1			1
SONY ELECTRONICS INC					1			1
SORDO MARK					1			1
TYCO ELECTRONICS CORP					1			1
UNEXIT COM						2		2
UNEXIT COM L L C						1		1
UNIV PRINCETON						1		1
UNIVERSAL ELECTRONICS INC						1		1
VON SIMSON CHRISTOPHER						1		1
WALTER AG						1		1
XILINX INC						1		1
YAHAMA CORP						1		1
Total: 20 101 4 3 7 8 146								

Fonte: <http://200.136.214.89/~patent/DONNEES/Eadv4/Eadv4Pivot.html>

Mais uma possibilidade de análise é apresentada na Figura 29, onde foram cruzados os dados referentes ao indicador “country” (país de depósito da patente) no eixo X, e “kind” (estatuto jurídico da patente) no eixo Y, também abordando apenas as classificações B e C, que se referem às patentes efetivamente concedidas ou em processo de concessão. Na Figura é possível observar claramente que o país em que existe o maior número de documentos registrados são os Estados Unidos (59 patentes no código B1), seguidos de Coreia do Sul (35 patentes também no código B1), e Taiwan (23 patentes no código B). Ainda, verifica-se que, das 146 patentes concedidas ou em processo de concessão, 101 se referem à classificação B1.

O referido cruzamento permite, por exemplo, identificar quais são os países onde existe o maior número de patentes protegidas, além de identificar qual tipo de tecnologia representa o maior interesse no tocante ao processo de patenteamento.

Figura 29 – Tela que demonstra a relação entre as patentes concedidas ou em processo de concessão e o país de depósito.

country	kind	B	B1	B2	C	C1	C2	Totais
		1	4					5
CA					3			3
JP		6						6
KR		35					35	
RU					7	5	12	
TW		23						23
UA						3		3
US			59					59
Totais		23	101	4	3	7	8	146

Fonte: <http://200.136.214.89/~patent/DONNEES/Eadv4/Eadv4Pivot.html>

Os dados apresentados na Figura 30 se referem ao cruzamento do indicador “year”, que representa o ano de depósito da patente e está localizado no eixo X, com o indicador “kind” (estatuto jurídico), localizado no eixo Y, levando-se em consideração apenas as patentes concedidas ou em processo de concessão. Na Figura é possível visualizar que, no ano de 1999, houve um considerável aumento no número de patentes (6 documentos, sendo 5 classificados como B1 e 1 classificado como B), se comparado aos anos anteriores. Ainda, verifica-se claramente que, no ano de 2001, ocorreu uma notável contribuição no tocante ao depósito de patentes em EAD, totalizando 16 documentos (12 classificados como B1, como B e 1 como C2). Assim como nos exemplos anteriores, a tecnologia representada pelo código B1 foi a que obteve maior número de concessões. A referida análise permite identificar qual foi o período que mais foram concedidas patentes no assunto EAD, além de identificar qual modalidade de tecnologia estava presente nessas solicitações.

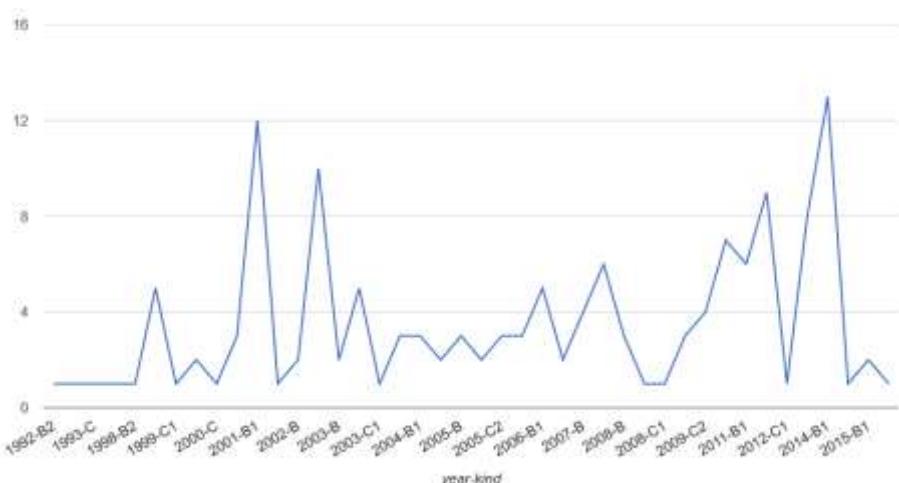
Figura 30 – Tela que demonstra a relação entre as patentes concedidas ou em processo de concessão e a data de depósito.

kind	B	B1	B2	C	C1	C2	Totals
year							
1992			1	1			2
1993				1			1
1995		1					1
1998			1				1
1999		5			1		6
2000	2		1				3
2001	3	10				1	14
2002	2	10					12
2003	2	5			1		8
2004	3	3	2				8
2005	3	2			3		8
2006	3	5		2			10
2007	4	6					10
2008	3	1		1			5
2009		3			4		7
2010		7					7
2011		6					6
2012		9		1			10
2013		8					8
2014		10		1			11
2015		2					2
2016		1					1
Totals	23	69	4	3	7	8	146

Fonte: <http://200.136.214.89/~patent/DONNEES/Eadv4/Eadv4Pivot.html>

A Figura 31 demonstra em formato de gráfico de linhas quais foram as tecnologias mais citadas nas patentes concedidas em cada ano.

Figura 31 – Gráfico de linhas que demonstra quais foram as tecnologias mais citadas nas patentes concedidas em cada ano.



Fonte: <http://200.136.214.89/~patent/DONNEES/Eadv4/Eadv4Pivot.html>

A Figura 32 apresenta os resultados provenientes do cruzamento do indicador IPCR7, que por sua vez representa a classificação da tecnologia patenteada com base em um código internacional de sete dígitos, em conjunto com o indicador “kind”, levando em consideração apenas as solicitações concedidas ou em processo de concessão, ambos alocados no eixo X, com o indicador "Applicant-country" (país da empresa detentora da concessão).

Apesar de estarem representadas apenas as partes superior e inferior da tela contendo a análise, neste exemplo, destaque pode ser dado às empresas que declararam os Estados Unidos como sendo o seu país de origem (51 concessões), seguidos da Coreia do Sul (36 concessões), e Taiwan (16 concessões). Nessas telas, verifica-se que a maior parte das patentes se refere à tecnologia representada pelo código A61N1 (“Electrotherapy; magnetotherapy; radiation therapy; ultrasound therapy”), com 4 concessões, sendo 3 delas holandesas e 1 onde o país da empresa não foi declarado no momento da solicitação.

Figura 32 – Parte superior e inferior da tela que demonstra a relação entre o IPCR7 e o país da empresa solicitante

		Applicant-Country	AU	BS	CA	DE	IE	JP	KR	NL	RU	TW	UA	US	Totais
Kind	IPCR7														
B1	A61B5												1	1	
	A61B19												1	1	
	A61L9								1					1	
	A61N1	1								3				4	
	B25J9							2						2	
	B25J13							1						1	
	B64D23							1						1	
	B64G1							1						1	
	B64G3							1						1	
	F01L13							1						1	
		G09B5								1				1	
		G09B9								1				1	
		G09B23									1			1	
		G09G5	1											1	
		H04B10									1			1	
		H04N7									1			1	
C	G06F9											1		1	
	G09B5											1		1	
	G09B7											1		1	
	H04H60											1		1	
		Totais	22	1	1	2	2	1	3	36	3	8	16	3	146

Fonte: <http://200.136.214.89/~patent/DONNEES/Eadv4/Eadv4Pivot.html>

Na Figura 33 encontram-se disponíveis as partes superior e inferior da tela que disponibiliza os resultados do cruzamento do indicador representativo da tecnologia patenteada, o IPCR7, alocado no eixo X juntamente com o indicador “kind” (estatuto jurídico da patente), e o indicador “Inventor-Country” (país do inventor), alocado no eixo Y. Nesta análise, que também levou em consideração apenas as patentes já concedidas ou em processo de concessão, observa-se como exemplo, que a tecnologia representada pelo código A61N1 (“Electrotherapy; Circuits therefor”) é uma das mais significativas, com 4 documentos (classificados como B1), todos citando os Estados Unidos como país do inventor. Os Estados Unidos ainda lideram o ranking entre os países citados pelos inventores de patentes concedidas (55 patentes), seguidos da Coreia do Sul (37 patentes) e Taiwan (14 documentos).

Figura 33 – Parte superior e inferior da tela que demonstra a relação entre o IPCR7 e o país da empresa depositante

IPCR7	kind	Inventor-Country		AU	CA	CN	DE	FR	IE	IN	JP	KR	NL	RU	TR	TW	UA	US	Total
		Count	Country																
A61B5	B1																1	1	
A61B19	B1																1	1	
A61L9	B1											1						1	
A61N1	B1															4	4		
B01D39	C2															1		1	
B25J9	B1										2						2		
B25J13	B1										1						1		
B64D23	B1										1						1		
B64G1	B1										1						1		
B64G3	B1										1						1		
Total		13	1	3	4	3	1	1	2	5	37	1	10	1	14	3	59	146	
H04W4																1	1		
H04W4	B																1		
	B1											1					1		
H04W24	B1											1					1		
H04W28	B1			1													1		
H04W36	B1			1													1		
H04W40	B1			1													1		
H04W64	B1										1						1		
H04W76	B1			1													1		
H04W88	B1										1						1		
H04W92	B1			1													1		
Total		13	1	3	4	3	1	1	2	5	37	1	10	1	14	3	59	146	

Fonte: <http://200.136.214.89/~patent/DONNEES/Eadv4/Eadv4Pivot.html>

Os resultados apresentados na Figura 34 se referem, mais uma vez, ao cruzamento de dados relacionados às tecnologias cujas patentes foram efetivamente concedidas ou em processo de concessão (representadas pelo código IPCR7), desta vez com os dados relacionados ao país (“country”) onde a proteção da referida tecnologia foi solicitada. Embora apenas as partes superior e inferior da tela com os resultados totais estejam representadas na Figura, é possível identificar, por exemplo, que a tecnologia A61N1 é aquela com o maior número de concessões (4 documentos, todos sob o código B1 e protegidos nos Estados Unidos). Os Estados Unidos representam ainda o país onde existe o maior número de patentes protegidas, totalizando 59 documentos, seguidos da Coreia com 35 concessões, e Taiwan com 23 patentes.

Figura 34 – Parte superior e inferior da tela que demonstra a relação entre o IPCR7 e o país de depósito (apenas patentes concedidas ou em processo de concessão).

IPCR7	kind	country	AU	CA	JP	KR	RU	TW	UA	US	Totais
A61B5	B1								1	1	
A61B19	B1								1	1	
A61L9	B1				1						1
A61N1	B1								4	4	
B01D39	C2							1		1	
B25J9	B1				2					2	
B25J13	B1				1					1	
B64D23	B1				1					1	
B64G1	B1				1					1	
B64G3	B1				1					1	
H04Q9	B1				1						1
H04W4	B						1				1
	B1				1						1
H04W24	B1								1	1	
H04W28	B1				1					1	
H04W36	B1				1					1	
H04W40	B1				1					1	
H04W64	B1								1	1	
H04W76	B1				1					1	
H04W88	B1				1					1	
H04W92	B1				1					1	
Totais		5	3	6	35	12	23	3	100	146	

Fonte: <http://200.136.214.89/~patent/DONNEES/Eadv4/Eadv4Pivot.html>

Na Figura 35 estão dispostos os resultados do cruzamento entre os indicadores IPCR7 e “date”, representando, respectivamente, a tecnologia patenteada e a data de depósito da patente, ambos alocados no eixo X, e levando-se em consideração, mais uma vez, apenas as patentes concedidas ou em processo de concessão, já que o indicar “kind” foi alocado no eixo Y. Nessa análise, embora apenas as partes superior e inferior da Figura estejam disponíveis, observa-se que a tecnologia representada pelo código A61N1 é aquela com o maior número de patentes (3 documentos no ano de 2002, e 1 documento no ano de 2003), sendo todas as solicitações classificadas com o código B1. Ainda é possível verificar o maior número de concessões ocorreu no ano de 2014 (13 patentes), seguido dos anos 2001 e 2002 (12 e 10 patentes, respectivamente), todas representando tecnologias classificadas como B1.

Figura 35 – Parte superior e inferior da tela que demonstra a relação entre o IPCR7 e a data de depósito

	year	1992	1993	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	Total	
	kind	B2	C	C	B1	B2	B1	C1	B1	C	B	B1	B2	B1	B1	B2	B1										
IPCR7																											
A61B5																											1
A61B10																											1
A61B15																											1
A61H1																											1
B01D39																											1
B25J9																											2
B25J13																											1
B64D3																											1
B64G1																											1
B64G3																											1
	Total	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	146	

H04Q9	1																										1
H04W8																											2
H04W24																											1
H04W26																											1
H04W28																											1
H04W36																											1
H04W40																											1
H04W44																											1
H04W54																											1
H04W66																											1
H04W76																											1
H04W88																											1
H04W92																											1
	Total	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	146	

Fonte: <http://200.136.214.89/~patent/DONNEES/Eadv4/Eadv4Pivot.html>

Outro item importante na análise das patentes é a possibilidade de identificar o número de citações que um documento patentário recebeu de outras patentes (representado pelo indicador “citations”, no eixo Y), confirmando a importância da tecnologia descrita no documento citado. No exemplo disposto na Figura 36, que levou em consideração apenas as patentes concedidas ou em processo de concessão (indicador “kind”, também no eixo Y), em conjunto com o IPCR7, observa-se que a tecnologia representada pelo código A61N1 é a mais citada (4 citações). Ainda, das 146 patentes concedidas, aquelas cujas tecnologias são classificadas como B1 são as mais citadas por outros documentos patentários, totalizando 101 citações.

Figura 36 – Parte superior e inferior da tela que demonstra a relação entre o IPCR7 e o número de citações recebidas pela patente

Cita kind	Cita	NaN	NaN	NaN	NaN	NaN	NaN	Totais
	B	B1	B2	C	C1	C2		
IPCR7								
A61B5		1						1
A61B19		1						1
A61L9		1						1
A61N1		4						4
B01D39						1		1
B25J9		2						2
B25J13		1						1
B64D23		1						1
B64G1		1						1
B64G3		1						1
H04Q9		1						1
H04W4	1	1						2
H04W24		1						1
H04W28		1						1
H04W36		1						1
H04W40		1						1
H04W64		1						1
H04W76		1						1
H04W88		1						1
H04W92		1						1
Totais	23	101	4	3	7	8		146

Fonte: <http://200.136.214.89/~patent/DONNEES/Eadv4/Eadv4Pivot.html>

4.1.7 Possibilidades de seleção de patentes com base na combinação de indicadores

Outra interessante funcionalidade do Patent2net é fornecer o acesso a uma *interface* onde os indicadores relacionados às patentes podem ser selecionados e combinados diretamente em uma única tela, permitindo que o usuário altere sua seleção de diversas maneiras com vistas a identificar as patentes de interesse de forma bastante específica e direcionada. O Acesso à referida *interface* é permitido ao clicar sobre o primeiro *link* disponível na Figura 9 ("Patents datatable, Pivot table").

A Figura 37 representa as partes superior e inferior dessa *interface*, que apresenta vinte pequenas janelas, cada uma representando um indicador específico e relacionado às patentes, que são, da esquerda para a direita, o país de depósito ("Country"), o título da patente ("Title"), o inventor ("Inventor"), o país do inventor ("From"), a empresa solicitante ("Applicant"), o país da empresa solicitante ("From"), o código de classificação de tecnologia com 11 dígitos

(“IPCR11”), o “CPC” (“*Common Patents Classification*”), que substituiu o IPC (“*International Patents Classification*”), a data exata do depósito da patente (“Prior-Date”), o ano de publicação do documento (“pub year”), o rótulo da patente (“Label”), o estatuto jurídico do documento (“Kind”), além de diversas modalidades de acompanhamento das citações recebidas pelas patentes (“Ref, CitP, CitO, Cited”), em conjunto com o número total de citações do documento (“Citations”), um sistema que busca conectar as patentes sobre a mesma invenção (“Equiv”), e o indicador de prioridade da tecnologia patenteada (“Priority”).

Na Figura 69, onde nenhum indicador específico foi solicitado, é possível visualizar as informações relacionadas aos indicadores citados das 3.090 patentes em EAD extraídas da base Espacenet pela ferramenta Patent2net. Na ausência de especificações relacionadas aos indicadores, a ferramenta organiza os documentos por ordem alfabética de acordo com o código que representa cada país de depósito. Dessa forma, ainda na Figura 37, as primeiras patentes são aquelas representadas pelo código “AU”, que representa a Austrália como país onde se pretendeu proteger a tecnologia, enquanto as últimas são aquelas com o código “WO”, patentes onde se buscou proteção mundial.

Figura 37 – Partes superior e inferior da tela com a visão geral das informações relacionadas às 3090 patentes em EAD.

Country	Title	Inventor	Applicant	IPC687	Prior-Date	pub year	Label	Kind	Ref	CitP	CitO	Cited	Equiv	Priority
AU	Online adaptive language learning teaching software online adapti... language learning software, adaptive language learning software, adaptive language learning software, adaptive language teaching software, adaptive language educational learning software, adaptive language teaching software.	No Nita, Lubatti Valery	No Nita	G09B7	2013-11-19	2014	AU2012254805	A1	0	AU2012254805				
AU	Web-based interaction	Hermann Erik	Informa Software Inc.	G09B7	1999-12-17, 2006-12-15	2001	AU20010001	A1	0	AU20010001				
AU	Web-based interaction	Hermann Erik	Dot	G09F9, G09F9	2000-12-06, 1999-12-15, 2006-12-15	2004	AU20027767	A1	0	AU20027767, EP09234340, CA2296248, US2002024138, WO0140806, EP2002024138				
AU	Systems and methods for monitoring, filtering system data and generating recommendations	Seed Nathan James, Chapman Kenneth James, Baker John Allan	Desire2Learn Inc.	G09B7, G09Q20	2012-07-21, 2013-05-05	2013	AU2012200268	A1	0	AU2012200268				
AU	Systems and methods for monitoring, filtering system data and generating recommendations	Seed Nathan James, Chapman Kenneth James, Baker John Allan	Desire2Learn Inc.	G09B7, G09Q20	2012-07-11	2013	AU2012200044	A1	0	SE18960, US2013089847, CA278452, AU2012200044				
AU	Agility - an enjoyable way to learn	Danielle Peter	Agileous Workplace Education Pty Ltd	G09B7	2010-12-23, 2011-12-23	2013	AU2011101696	A4	0	AU2011101696				
AU	Instant message for a virtual learning community	Gupta Rakesh Kumar	Hindustan Pet. Ltd.	G09B7	2006-03-31	2007	AU2006034042	A1	0	AU2006034042, US2008221685, WO2007114798				
AU	A system comprising a multi-user virtual learning environment	Jessica Anne, Testani Sandra, Wagner Edward, Burdick Jack	Franklin Institute	G06Q10, G09B7, G09K11	1998-07-17, 1999-07-19	2000	AU20000699	A1	0	AU20000699, WO20000478				
AU	Agent for posting entries in a virtual learning environment	Gitterman Abbieon	Gitterman Abbieon	G09B7	2001-04-05, 2000-07-11	2002	AU20014905	A1	0	US2002027722, EP1299914, AU4074901, CA2409085, WO200114				
AU	Method and system for online teaching using web pages	Tally Robert, Zeev El Shafay O. Odeh Michael	Discourse Technologies Inc.	G09B7	2001-08-12, 2001-10-22	2003	AU200248081	A1	0	WO200205278, GB2392001, AU200248081				

Showing 1 to 10 of 3,090 entries

First Previous 1 2 3 4 5 ... 309 Next Last

Fonte: <http://200.136.214.89/~patent/DONNEES/Eadv4/Eadv4.html>

Como exemplo prático de utilização da interface, a Figura 38 disponibiliza as patentes cuja empresa solicitante é brasileira. Para tal seleção, a sigla “br” foi inserida na janela representativa do país da empresa depositante (“Applicant-country”). É possível reparar que, como resultado, um único documento patentário foi encontrado. Trata-se de uma solicitação de cobertura mundial (“WO” na janela representativa do país), tendo o francês como língua de redação (janela do título), cujo inventor, “Eleuterio Marco-Antonio”, brasileiro, não foi financiado por nenhuma empresa (já que seu nome se repete na janela da empresa solicitante). A tecnologia descrita no referido documento é representada pelos códigos IPCR7 G09B7 (“Electrically-operated teaching apparatus or devices working with questions and answers”), CPC G09B7/00, depositada em 17/04/2002, tendo como rótulo o código WO03087970 (“Computational method for mediating asynchronous group discussions”), e estatuto jurídico representado pelo código A2, possuindo como equivalentes as tecnologias representadas pelos documentos AU2003201223 (“Computational method for mediating asynchronous group discussions”) e BR0201651 (também descrita na Espacenet como “Computational method for mediating asynchronous group discussions”).

Figura 38 – Recorte da tela com a visão geral das informações relacionadas à única patente em EAD depositada por uma empresa brasileira



Fonte: <http://patent2net.vlab4u.info/DONNEES/Ead/Ead.html>

Em uma breve análise crítica, levando-se em consideração que o Brasil solicitou apenas 1 das 3087 patentes em EAD, identifica-se a obsolescência do país no tocante ao desenvolvimento de tecnologias relacionadas ao tema. Por outro lado, ao se colocar as iniciais “br” na janela relacionada ao país onde a tecnologia foi protegida, verifica-se que nenhuma patente relacionada ao EAD se estende ao Brasil, fazendo com que os textos destes documentos, cujas patentes foram solicitadas, concedidas ou se encontram em processo de concessão, são uma enorme fonte de conhecimento para as empresas e inventores brasileiros que desejam trabalhar com esta temática.

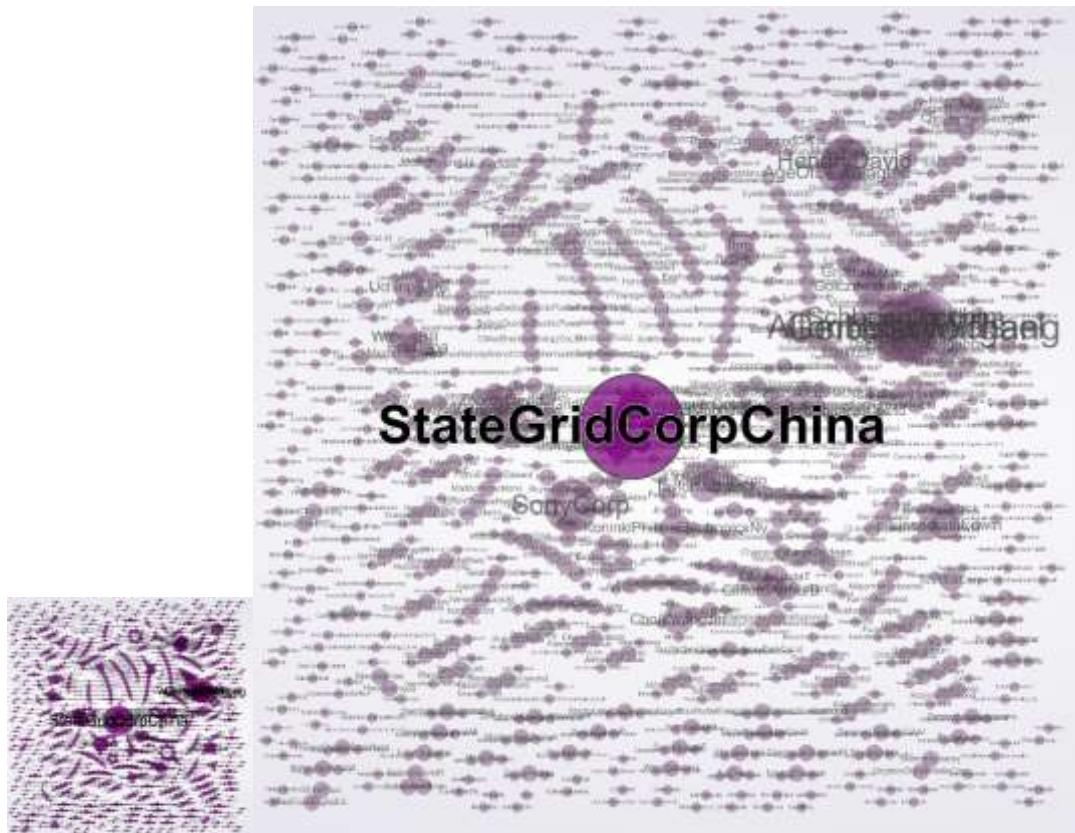
Na *interface* apresentada, é possível ainda delimitar inúmeras seleções levando em consideração cada um dos doze indicadores disponíveis, isoladamente ou em conjunto, de forma a direcionar o pesquisador a algumas poucas patentes que realmente lhe serão de interesse para, assim, realizar uma leitura completa do documento e avaliar a tecnologia descrita, bem como avaliar os seus aspectos legais e a prática de uma possível replicação.

4.1.8 Redes relacionadas às patentes em EAD

Ainda em relação às possibilidades de análise tem-se algumas redes relacionadas às patentes em EAD se encontram disponíveis a seguir. O acesso a essas redes se dá ao clicar no *link* “Networks (Inventor, Applicant, Technology)”, visível na Figura 9, para acesso às redes simples, e no *link* “Mixed Networks (Country-Technology, Inventor-Technology, Applicant-Technology, Applicant-Inventor)”, também disponível na Figura 9, para acesso às redes mistas.

Inicialmente avaliando as redes simples, que levam em consideração apenas a empresa depositante, a tecnologia desenvolvida, ou o seu inventor, a Figura 39 mostra a possibilidade de avaliar a interação entre as empresas solicitantes dessas patentes, permitindo identificar as relações entre as instituições que efetivamente têm investido nas mesmas tecnologias relacionadas ao EAD. Nesta rede, destacou-se como exemplo de elemento central a empresa “State Grid Corp China”, uma empresa Chinesa, que apresenta uma série de interações com outras empresas do ramo. Destaca-se que quanto maior o tamanho da esfera, maior a interação e importância que esta empresa tem para a rede.

Figura 39 – Rede entre as empresas depositantes das 3.090 patentes em EAD.

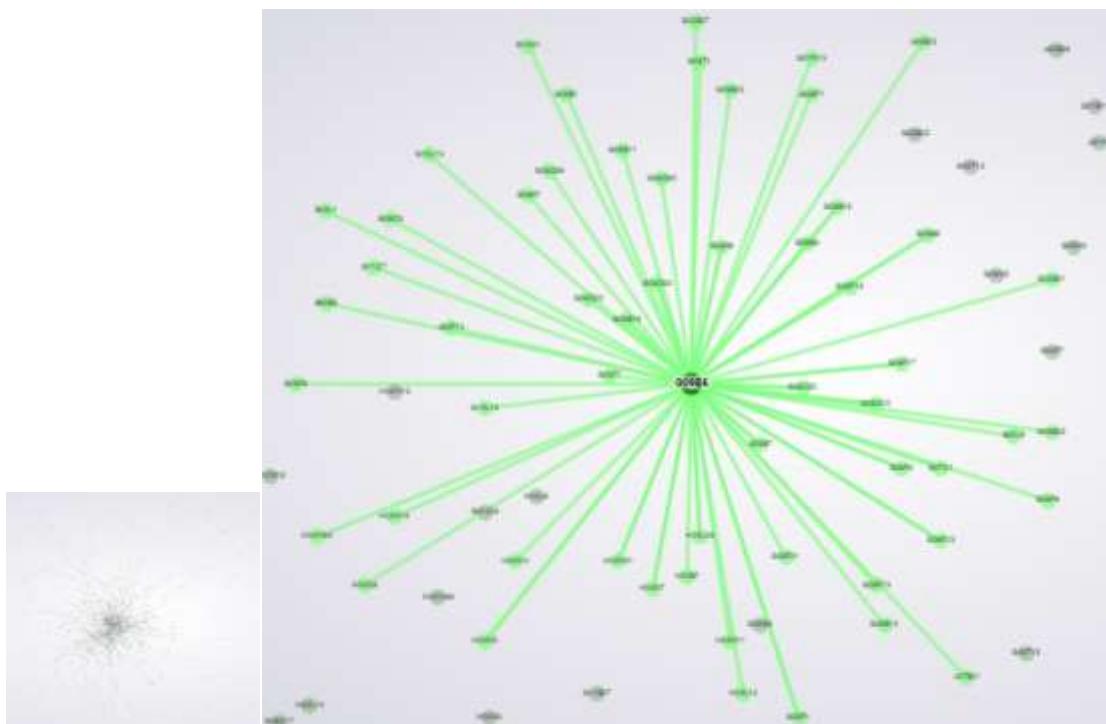


Fonte:

http://200.136.214.89/~patent/DONNEES/Eadv4/GephiFiles/Eadv4_ApplicantsJS.html

A Figura 40 demonstra a rede formada entre as diversas tecnologias cujas patentes foram solicitadas, tomando por base o IPCR7, que é o código de classificação de tecnologia com sete dígitos. A avaliação dessa rede permite identificar quais foram as tecnologias que mais geraram solicitações de patentes, assim como verificar como essas tecnologias podem ser complementares, com base na interação apresentada por elas no momento da avaliação da rede. Como elemento central na Figura, destaca-se a tecnologia representada pelo código G09B5 (“Electrically-operated educational appliances”), bem como suas ligações com outras tecnologias também voltadas ao EAD.

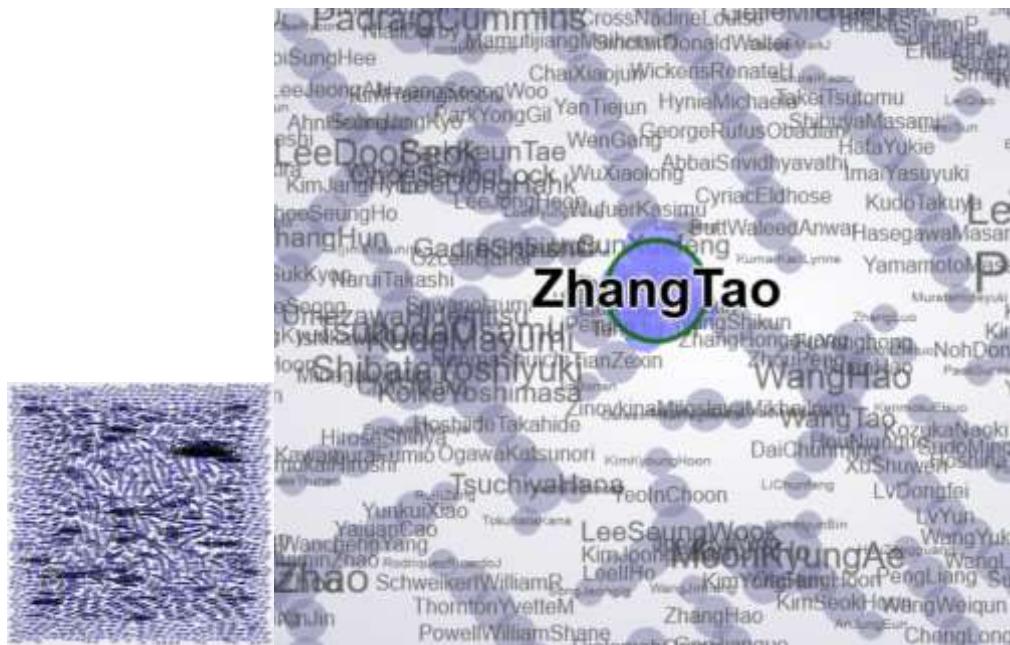
Figura 40 – Rede entre as patentes depositadas em EAD com base no IPCR7.



Fonte: http://200.136.214.89/~patent/DONNEES/Eadv4/GephiFiles/Eadv4_CrossTechJS.html

Já a Figura 41 apresenta a complexa rede de relação entre os inventores das tecnologias em EAD, de forma a permitir a avaliação e identificação de elementos centrais e seus parceiros tecnológicos. Na Figura, destacou-se como exemplo de elemento central da rede o inventor “Zhang Tao”, da China.

Figura 41 – Rede entre os países e a tecnologia da patente depositada (com base no IPCR7).

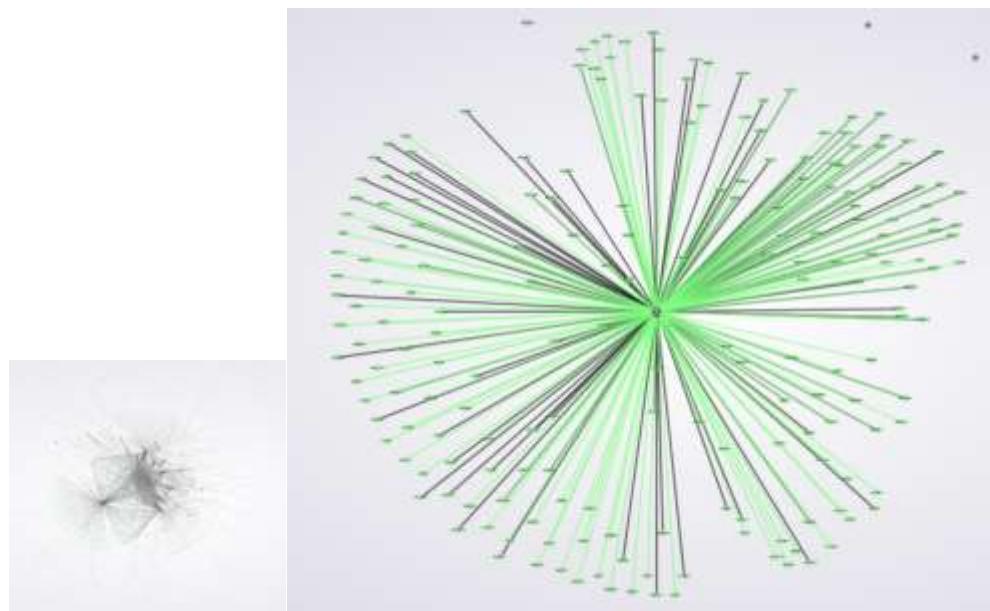


Fonte:

http://patent2net.vlab4u.info/DONNEES/Ead/GephiFiles/Ead_CountryCrossTechJS.html

Avaliando agora as redes mistas, que levam em consideração a tecnologia em relação à empresa solicitante, inventor ou país, a Figura 42 demonstra a possibilidade de identificar a relação entre a tecnologia desenvolvida e o país no qual se buscou protegê-la. Na Figura, o país eleito como elemento central foi a China, e é possível notar a grande quantidade de tecnologias voltadas ao EAD que se buscou proteger naquele país.

Figura 42 – Rede mista entre os países e a tecnologia cuja patente foi solicitada (com base no IPCR7).

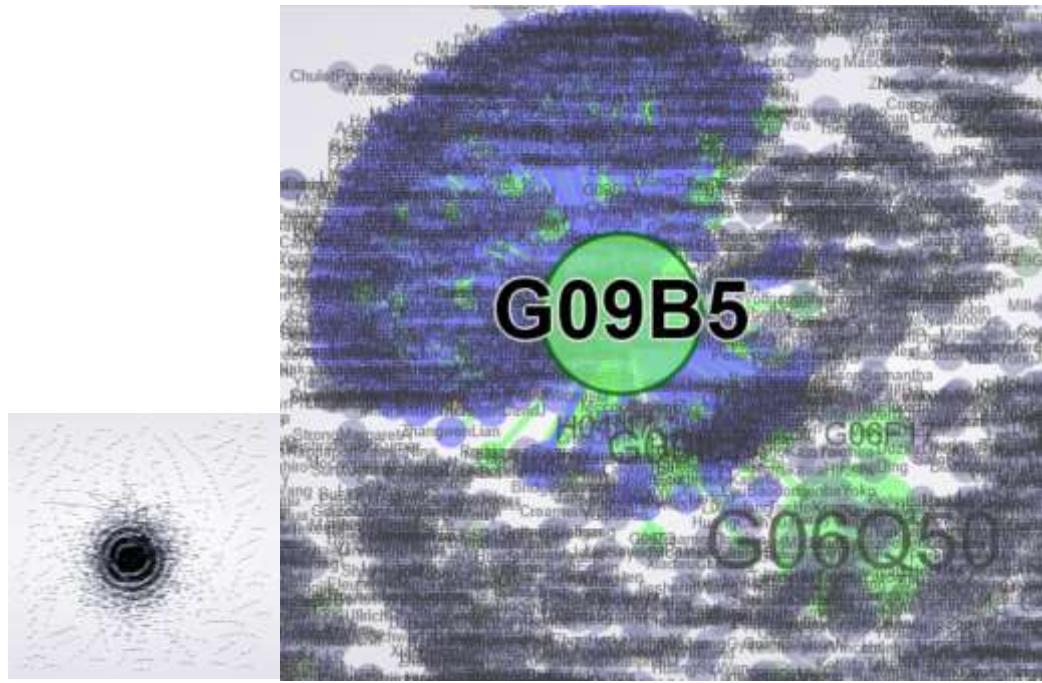


Fonte:

http://200.136.214.89/~patent/DONNEES/Eadv4/GephiFiles/Eadv4_CountryCrossTechJS.html

Na Figura 43, que utiliza os indicadores representados pela tecnologia e pelo nome do inventor, a tecnologia representada pelo código G09B5 encontra-se em destaque, sendo possível verificar a enorme quantidade de inventores que têm se dedicado ao longo dos anos ao desenvolvimento de patentes envolvendo a referida tecnologia. Neste exemplo, o número de inventores é tamanho, que para verificar seus nomes com exatidão seria necessário aumentar consideravelmente o tamanho da Figura, o que por fins didáticos não foi realizado neste exemplo.

Figura 43 – Rede mista entre os inventores e a tecnologia cuja patente foi solicitada (com base no IPCR7).

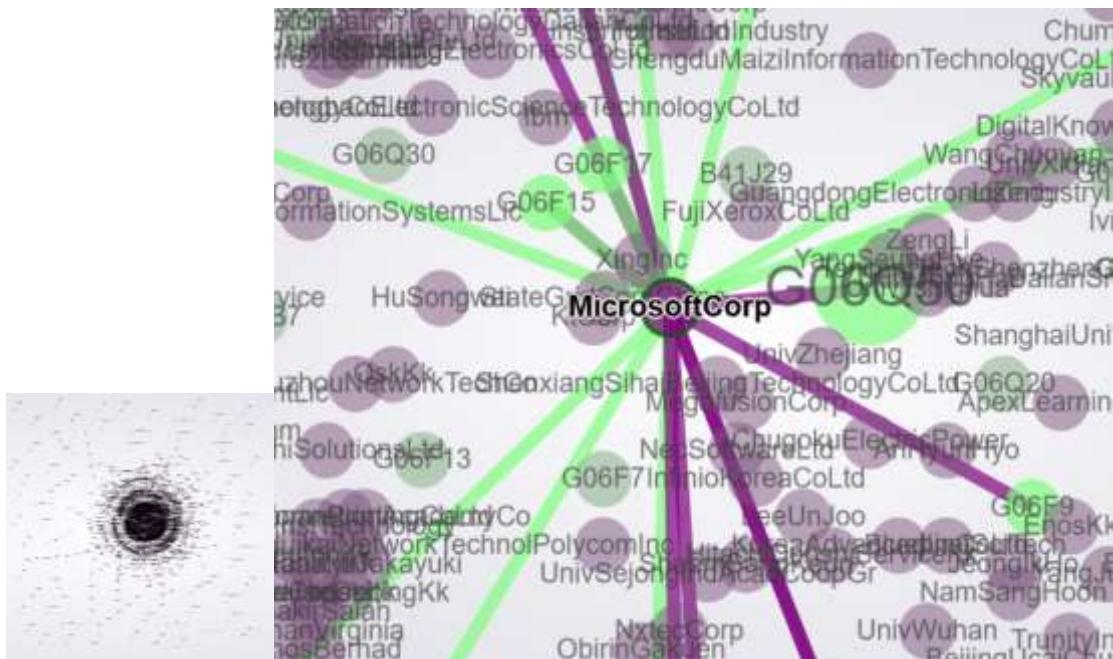


Fonte:

http://200.136.214.89/~patent/DONNEES/Eadv4/GephiFiles/Eadv4_CountryCrossTechJS.html

Já a Figura 44, que utiliza os indicadores empresa solicitante e tecnologia desenvolvida, a “MicrosoftCorp”, uma empresa americana, foi escolhida como elemento central para demonstração rede. Nesta análise, é possível verificar quais são as tecnologias de interesse desta, e de outras empresas engajadas no desenvolvimento de diversos produtos tecnológicos voltados ao Ensino a Distância.

Figura 44 – Rede mista entre as empresas e a tecnologia cuja patente foi solicitada (com base no IPCR7).



Fonte:

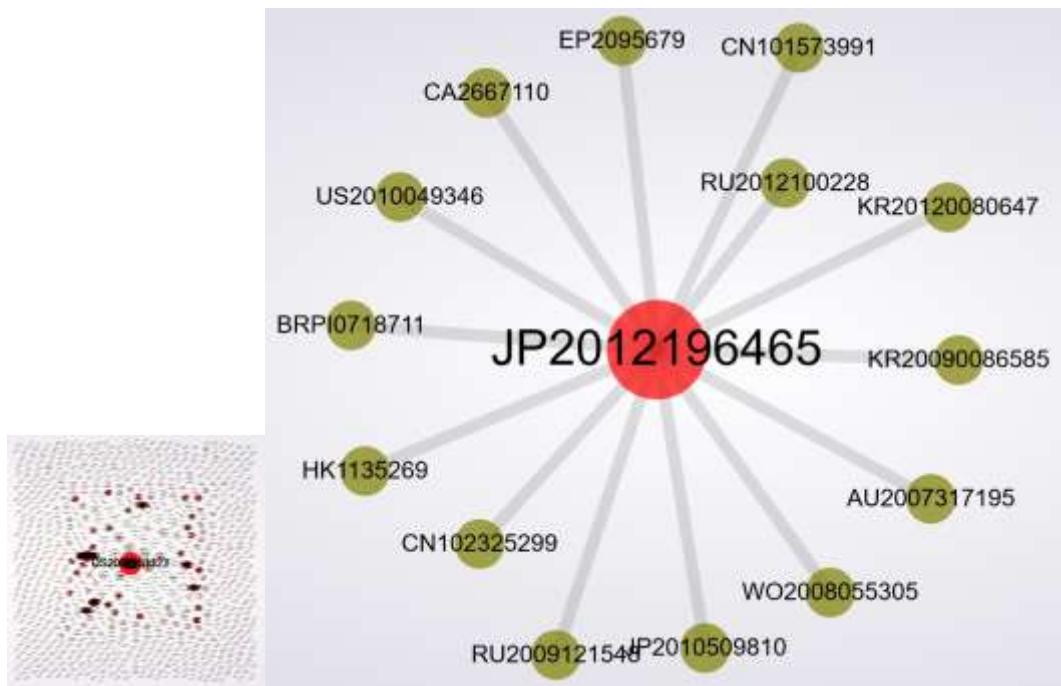
http://200.136.214.89/~patent/DONNEES/Eadv4/GephiFiles/Eadv4_Applicants_CrossTechJS.html

4.1.9 Redes de equivalência entre as tecnologias, entre os documentos citados por outras patentes ou por fontes externas, e entre as citações entre patentes

A seguir são apresentadas novas redes simples, que todavia levaram em consideração para sua construção os indicadores relacionados às possíveis tecnologias equivalentes, aos artigos científicos referenciados pelas patentes, e às citações recebidas pelas patentes. Para acessá-las é necessário clicar no quinto *link* disponível na Figura 9 (“Equivalentes, Reference (References to other patents or External references), Patents citations networks”).

Na Figura 45, que levou em consideração as tecnologias equivalentes, como exemplo, destacou-se como elemento central a tecnologia descrita no documento JP2012196465 (“Apparatus and method for use in creating audio scene”). Ainda na Figura, é possível identificar 14 outras tecnologias que, de alguma forma, apresentam certa equivalência com a tecnologia central, levando-se em consideração os documentos de patentes que as declararam.

Figura 45 – Rede simples destacando as relações entre tecnologias equivalentes voltadas ao EAD (com base no IPCR11).



Fonte:

http://200.136.214.89/~patent/DONNEES/Eadv4/GephiFiles/Eadv4_EquivalentsJS.html

Na Figura 46, a tecnologia representada pelo código US2010169677 (“Remotely Powering On-Off Network Devices via a Network Interface Device”) aparece como elemento central, ligada a 9 outras tecnologias cujas patentes citaram o mesmo artigo científico em sua lista de referências.

Figura 46 – Rede simples destacando as relações entre tecnologias que citaram os mesmos artigos científicos em suas listas de referências (com base no IPCR11).



Fonte:

http://200.136.214.89/~patent/DONNEES/Eadv4/GephiFiles/Eadv4_ReferencesJS.html

A Figura 47 traz a rede de citações entre as patentes. No exemplo, a tecnologia escolhida como central, representada pelo código WO9310615 (“Systeme for protecting and restarting computers and peripherals at remote sites which are accessible”), mostra-se ligada a 25 outras tecnologias cujos documentos patentários a citaram em sua lista de referências.

Figura 47 – Rede simples destacando as relações entre tecnologias que citaram a tecnologia central, em destaque na Figura (com base no IPCR11).



Fonte: http://200.136.214.89/~patent/DONNEES/Eadv4/GephiFiles/Eadv4_CitationsJS.html

4.1.10 Análises relacionadas às famílias de patentes

Além das análises individuais das patentes, também fez-se uma análise global dos documentos patentários, levando em consideração a família representativa da tecnologia considerada nos documentos. Os mesmos cruzamentos apresentados nas referidas Figuras podem agora englobar as patentes inseridas em uma mesma família tecnológica. Ainda, outros possíveis cruzamentos não apresentados neste momento se encontram disponíveis para realização de maneira dinâmica na *interface* de dados fornecida pela ferramenta, disponível no sexto *link* da Figura 9 ("Patent family & Pivot table") que podem ser utilizados para refinar os cruzamentos. As *interfaces* dinâmicas para realização da seleção e dos cruzamentos entre famílias se encontram exemplificadas nas Figura 48 e Figura 49. Entenda-se por interface dinâmica uma tela onde é possível digitar pequenos trechos, seja do nome do país, do inventor, da empresa, do país, dentre outros, nas janelas específicas para cada um desses indicadores, para que sejam selecionadas as patentes com base nos critérios estabelecidos na referida digitação.

Figura 48 – Interface para seleção das famílias tecnológicas das patentes.

Country	Title	Inventor	From Applicant	From EPODB1	CPC	Prior date	Year	Label	Kind	Ref	Equivalents	Priority Prior	Family length	Cat1	Cat2	Cat3
AU	Search	Search	From	From	Search	From	From	Search	From	Search	From	From	1	From	From	Search
AU	Online multiple language learning teaching software; educational computer program; learning software; multiple language; learning software; learning software; teaching software; teaching software; multiple language; instant learning; educational learning; educational learning; educational learning; software; multiple language learning; software;	Lihani, Udaya, K. N.	K. N.	G09B1/08	2012-11-18	2014	AU201221184861	0	AU2012254865	1	AU2012294863					
CN	Real time online data transmission system and remote control data transmission method.	Yan, Mingzhu	Yan, Mingzhu	H04L29/08	2014-09-19	2014	CN104262428	0	CN104262427	1	CN104262425					

Fonte: <http://200.136.214.89/~patent/DONNEES/Eadv4/FamiliesEadv4.html>

Figura 49 – Interface para cruzamento das famílias tecnológicas das patentes.



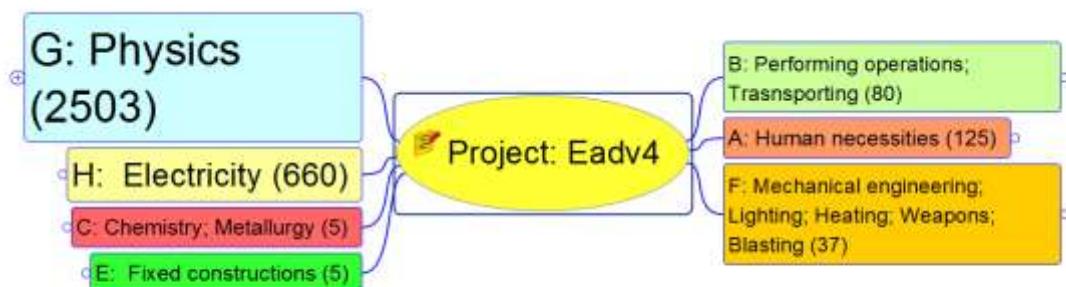
Fonte: <http://200.136.214.89/~patent/DONNEES/Eadv4/FamiliesEadv4Pivot.html>

Vale ressaltar que a tela apresentada na Figura 49 possui o intuito de demonstrar quais são os indicadores que podem ser cruzados tanto no eixo X quanto no eixo Y, todavia, levando-se em consideração as famílias das patentes, e não as patentes isoladamente. Na referida Figura, nenhum indicador foi escolhido propositalmente para que a interface pudesse ser visualizada em sua totalidade.

4.1.11 Mind map entre patentes

Para finalizar algumas das possibilidades que foram utilizadas para análise das patentes em EAD e de suas famílias, disponíveis até o presente momento, a Figura 50, Figura 51 e Figura 52 apresentam *mind maps* contendo alguns dos principais assuntos discutidos nos 3.090 documentos patentários minerados da base de patentes Espacenet pelo Patent2net. Na Figura 50 encontra-se o *mind map* com os assuntos principais, que pode ser acessado pelo link ("IPC's Mind-Map (FreePlane Plugin)"), disponível na Figura 9. As Figura 51 e Figura 52 representam as possibilidades de aprofundamento da análise, com base no *mind map*.

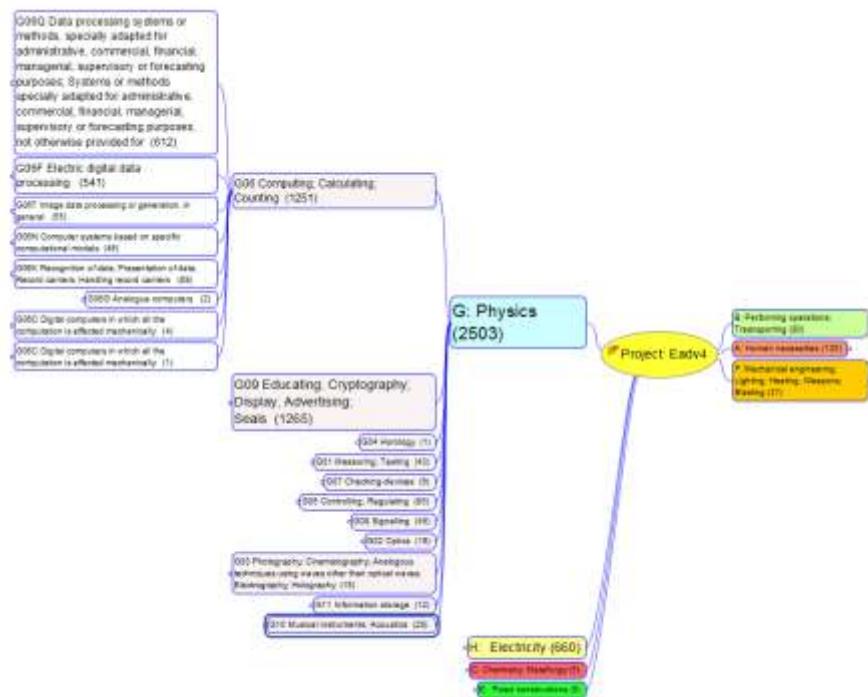
Figura 50 – Mind map com os principais assuntos extraídos pelo Patent2net das patentes em EAD depositadas na Espacenet.



Fonte: <http://200.136.214.89/~patent/DONNEES/Eadv4/Eadv4FP.html>

Na Figura 50 é possível verificar que existem na Espacenet, relacionadas ao EAD, patentes vinculadas aos temas “Physics” (em destaque), “Electricity”, Chemistry; Metallurgy”, “Fixed constructions”, “Performing operations; transporting”; “Human necessities”, e “Mechanical engineering; Lighting; Heating; Weapons; Blasting”.

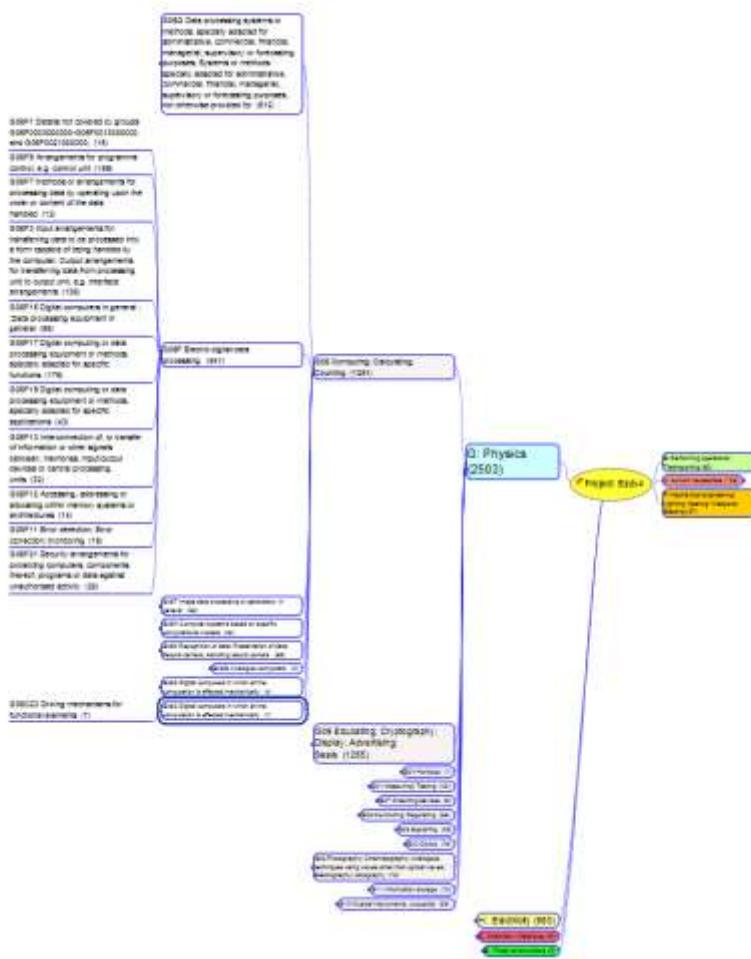
Figura 51 – Mind map com os principais assuntos extraídos pelo Patent2net das patentes em EAD depositadas na Espacenet (foco no tema principal “G: Physics (2503 documentos).



Fonte: <http://200.136.214.89/~patent/DONNEES/Eadv4/Eadv4FP.html>

A Figura 51 é um exemplo de como se podem avaliar as subcategorias de classificação das 3090 patentes em EAD. No exemplo, as patentes classificadas como “Physics” apresentam 11 subcategorias. Com relação à subcategoria “Computing; Calculating; Counting”, esta apresenta 8 categorias tecnológicas nas quais as patentes se encontram classificadas.

Figura 52 – *Mind map* com os principais assuntos extraídos pelo Patent2net das patentes em EAD depositadas na Espacenet (foco no tema central “G: Physics → G06 Computing: Calculatint: Counting”, com 1251 documentos)



Fonte: <http://200.136.214.89/~patent/DONNEES/Eadv4/Eadv4FP.html>

A Figura 52 demonstra a possibilidade de consulta das subcategorias relacionadas às tecnologias que se buscavam patentear. No exemplo, é possível identificar que a categoria tecnológica “Electric digital data processing” apresenta 11 diferentes tecnologias utilizadas para classificar de maneira mais específica cada patente depositada.

4.2 METODOLOGIA DE BUSCA DE PATENTES EM EAD POR ANÁLISE DE CONTEÚDO

Atualmente há mais de 90 milhões de pedidos de patentes depositados, com origens em mais de 70 países diferentes. São milhões de inventores e tecnologias reivindicando o

ineditismo de suas invenções, o que justificaria a concessão da patente. Para organizar o banco de dados formado pelos mais diversos depósitos de pedidos de patentes, em 1971 foi criado o IPC, por meio do acordo de Strasbourg. IPC é o acrônimo do termo em inglês *International Patent Classification* (WIPO, 2016a). O IPC é um código hierárquico que define a função e a aplicação presente no texto da patente, usado para classificar as patentes. Nesse caso, cada pedido de patente recebe pelo menos um código IPC. A literatura especializada sobre patentes, usa diferentes siglas para o mesmo significado, ou seja, IPC, CIB e CIP significam a mesma coisa. Como mencionamos, o IPC tem origem no termo em inglês, já o CIB tem origem no termo em francês (*Classification Internationale des Brevets*) e o CIP tem origem no termo em português (Classificação Internacional de Patentes). O código IPC ou CIB prove a patente, portanto, uma classificação que permite pela simples leitura do código, compreender em linhas gerais quais as principais funções e aplicações são reivindicadas pela patente, sem ter que ler todo o texto da patente. O IPC ou CIB está dividido em Seção (de A a H), Classe (de 00 a 99), Subclasse e Grupo, num total de quatro marcações, sendo a primeira com uma letra e as três restantes formadas por números e letras com dois dígitos cada (WIPO, 2016b). O Patent2Net, faz a verificação dos códigos IPC, mas usa a nomenclatura CIB, portanto, nas análises de conteúdo ou mesmo na AFC, a codificação que aparecerá, quando aplicável será a CIB. Sobre a notação do código CIB, dependendo do detalhamento pretendido, podem ser considerados os CIB1, CIB2, CIB3 ou CIB4. A sequência crescente de um a quatro, demonstra quais partes da classificação estão disponíveis. Nesse caso o CIB1, apresenta apenas a Seção do código (de A a H) e o CIB4 apresenta o código completo contendo a Seção, Classe, Subclasse e Grupo. O conceito de predominância de IPC, nas tabelas oferecidas nesse estudo, tem por finalidade apresentar os IPC de maior frequência na análise de conteúdo. Trata-se, portanto, de uma indicação qualitativa que informa a profundidade do CIB que predomina em determinado agrupamento ou *cluster* e será útil caso seja necessário escolher quais patentes deverão ser lidas integralmente de forma manual, visando o aprofundamento, dependendo dos objetivos da pesquisa.

A expressão de busca construída, recuperou 3.090 patentes relacionadas ao tema de interesse EAD. O tratamento de grandes quantidades de informações requer esforço computacional, mas também metodológico. No caso em questão, pode-se observar que a expressão de busca alcançou uma quantidade de informações superior ao que uma pessoa pode ler e compreender em tempo hábil para que a informação continue útil. Pode-se exemplificar

tal afirmação, considerando que cada patente tem um *abstract*, uma descrição tecnológica e algumas reivindicações (*claims*) não exaustivamente, que usam apenas para essas três informações cerca de uma a duas páginas formato A4. Seriam então ao menos três mil páginas de informações para serem lidas e classificadas, ou seja, não seria possível analisar essa quantidade de informações sem o uso de esforços computacionais e metodológicos.

Nessa seção, o objetivo é mostrar os procedimentos metodológicos usados para tal análise, denominada de análise de conteúdo. A análise de conteúdo pode ser realizada de várias maneiras, das quais examinou-se a classificação hierárquica descendente: Método de Reinert: a similaridade e a análise factorial de correspondência (AFC) (Silva, 2003). Tais escolhas, devem-se ao fato dessas análise estarem disponíveis no software de análise de conteúdo Iramuteq (“IRaMuTeQ”, 2016). O Patent2Net, conforme descrito, foi construído dentro dos princípios de código aberto (*open source*), assim, busca-se explorar todas as suas integrações nativas disponibilizadas para a análise das patentes, considerando que o uso de softwares livres e sem licenças consiste num diferencial positivo para a disseminação da ciência e da tecnologia.

Dessa forma, o escopo da análise em termos de qual informação analisar, foi definido pelo software Patent2Net, uma vez que suas saídas são pré-definidas. Sendo esse um software específico para recuperação de informações de patentes, as saídas disponibilizadas são: descrições, resumos, reivindicações, famílias de descrições, famílias de resumos e famílias de reivindicações. As famílias são formadas por meio de citações que determinado texto da patente faz a outra patente. Assim, famílias de resumos, por exemplo, são os resumos que estão relacionados por meio de citações ou referências que em comum. Analisou-se cada uma das saídas do Patent2Net disponibilizadas para o Iramuteq, realizando para cada uma das saídas, sequencialmente a análise da classificação hierárquica descendente, similaridade e AFC. O objetivo é oferecer uma interpretação sobre o conteúdo das 3.090 patentes, triangulando os conteúdos analisados por meio de múltiplas técnicas: Classificação Hierárquica Descendente (CHD), similaridade e Análise Fatorial de Correspondência (AFC).

4.2.1 Resumos das patentes por Classificação Hierárquica Descendente (CHD) e Análise Fatorial de Correspondência (AFC)

Foram recuperados 2.870 resumos de patentes, como resposta ao uso da expressão de busca construída. Tais resumos foram classificados em 12.365 segmentos de texto (textos com similaridades em relação a distância das formas léxicas), 8.083 formas lematizadas (ativas e suplementares) das quais analisou-se as ativas, por trazerem mais significado léxico, portanto, maior potencial de explicação sobre o conteúdo do texto. As formas ativas com frequência acima de 6 foram 2.663, sendo essas as formas mais relevantes do *corpus* analisado: 2.870 resumos de patentes. A média de formas por segmento de texto foi de 35,38, e o total de segmentos de texto tratados foi de 11.991, que representa 96,98% dos segmentos de texto classificados que foram separados em 3 *clusters*, por meio do algoritmo de classificação hierárquica descendente definido por Reinert (Reinert, 1990). A Tabela 6 resume os resultados da análise de CHD dos resumos das patentes sobre EAD.

Tabela 6 - Resultados da CHD dos resumos das patentes sobre EAD.

Número de textos	2.870
Número de segmentos de textos	12.365
Número de formas	11.027
Número de ocorrências	437.490
Número de <i>lemmas</i>	8.083
Número de formas ativas	7.273
Número de formas suplementares	810
Número de formas ativas com frequência >=6	2.663
Média de formas por segmento	35.381.318
Número de <i>clusters</i>	3
11.991 segmentos classificados em 12.365 (96.98%)	

Identificou-se a formação de três *clusters*, nesse caso, representam os três contextos: (1) Sistemas e Aprendizagem, (2) Métodos e Virtualidade, (3) Controle e Conexão; ou grupos polissêmicos, cujas formas léxicas apresentam associação em termos de significado léxico. É possível argumentar inicialmente, que nos 2.870 resumos sobre as patentes relacionadas a EAD,

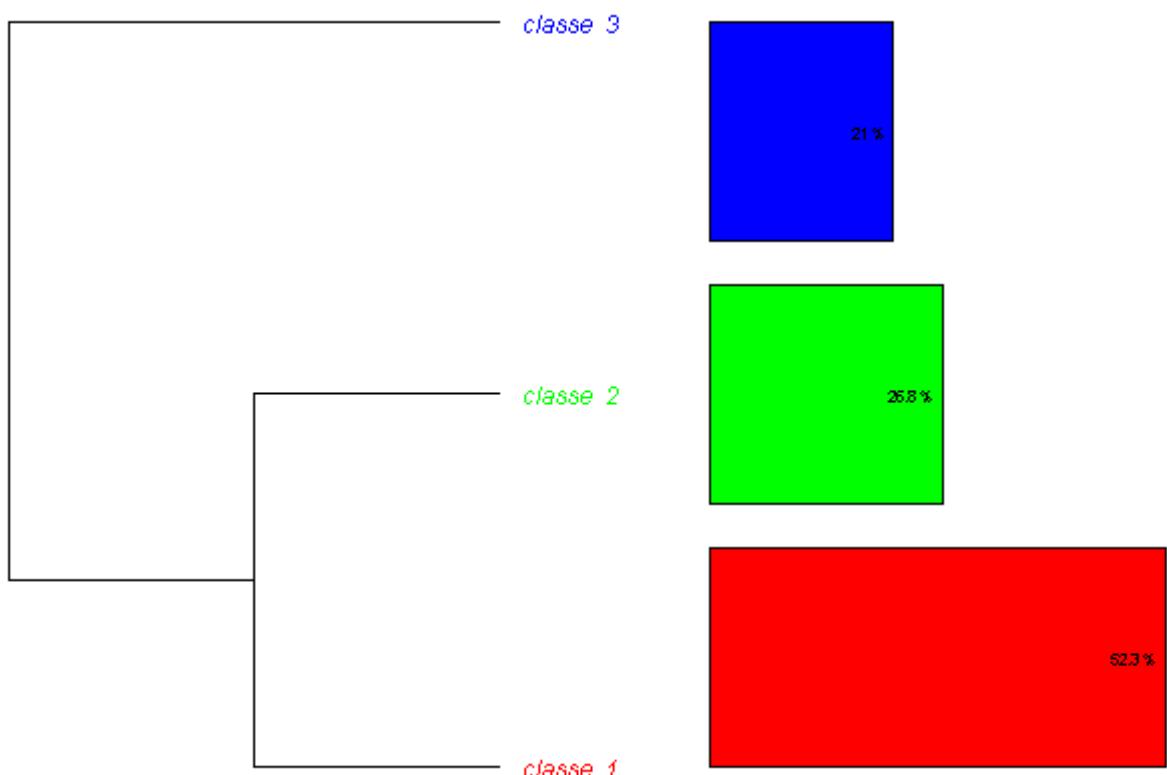
há 3 contextos que descrevem o tema, sendo o contexto 1 o representante de 62,3% do conteúdo total, o contexto 2 o representante de 26,8% do conteúdo total e o contexto 3 o representante de 21% do conteúdo total. É possível que a presença de apenas três contextos no total de patentes sobre EAD demonstre que os depositantes de patentes tem se concentrado nas mesmas explicações gerais sobre o EAD quando depositam suas patentes, uma vez que tratando dos resumos que contém explicações gerais sobre a patente. Pelas evidências verificadas na análise de conteúdo dos resumos isoladamente, é possível que as patentes sobre EAD apresentam evolução tecnológica incremental: inovações incrementais. O contexto 1, representante de 62,3% do conteúdo total dos resumos analisados, mostra que os depositantes concentraram suas explicações gerais sobre o objeto da patente de forma muito parecida.

A análise do contexto 1 será útil, portanto, para identificar o perfil tecnológico dominante nesse conjunto de patentes sobre EAD, especialmente sobre o objeto geral da patente. Por outro lado, a análise do contexto 3, que representa 21% do conteúdo total, pode revelar objetos gerais patenteados em EAD não dominantes. Analisar objetos gerais patenteados não dominantes são relevantes a medida em que podem auxiliar na compreensão de possíveis tecnologias em declínio ou nascentes, incluindo uma triangulação simples com uma tabela de patente *versus* data de depósito. Todas as tecnologias resumidas nas patentes classificadas no contexto 3, cuja frequência aumenta em função do tempo, representam tecnologias nascentes e aquelas cujas frequências diminuem em função do tempo representam tecnologias em declínio. Pode-se argumentar ainda que, se busca-se conhecer a tecnologia dominante, deve-se analisar as patentes do *cluster* 1; mas, se queremos analisar possíveis inovações radicais, a chance é maior de serem encontradas no *cluster* 3. Tal constatação é uma das questões que serão discutidas posteriormente, quando outras dimensões das patentes já estiverem analisadas permitindo, portanto, alguma triangulação.

Uma análise fatorial por correspondência (AFC) dos resumos demonstra as distâncias ou disposições de um contexto em relação ao outro. Nessa observação gráfica introdutória, observa-se que os contextos são como dimensões. Tratando de análise de conteúdo, isso significa que houve possibilidade de agrupar os temas, que chama-se de contexto em três partes, que se encontram distanciados conforme os fatores demonstrados no próximo gráfico. Tal visualização serve, numa mão para corroborar o dendograma, uma vez que a relação hierárquica se confirma entre os contextos 1 e o contexto 3, por meio do contexto 2 e, em outra mão, por

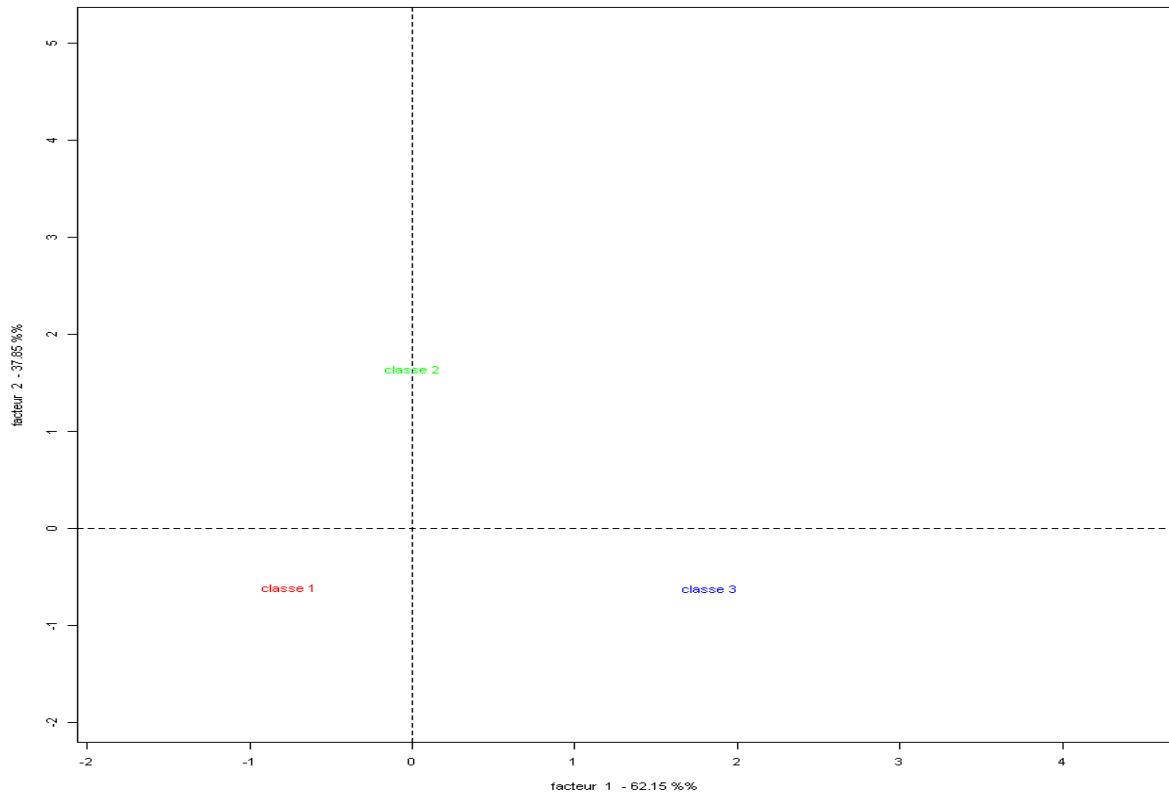
evidenciar a relevância do contexto 2, sobre o qual ainda não foi elaborado nenhum comentário. Observar-se que o contexto 2, tem carga fatorial no fator 2 (eixo Y), próximo de 2 e que os contextos 1 e 3 apresentam carga fatorial de -1 e +2 no fator 1 (eixo X), respectivamente. A verificação gráfica de que há grandes diferenças no contexto 2 quando comparado ao contexto 1 e ao contexto 3.

Figura 53 - Dendograma da CHD dos resumos das patentes sobre EAD



As análises da CHD e da AFC são, portanto, complementares, uma vez que a CHD se preocupa com a delimitação e a relação hierárquica dos *clusters* e AFC com o posicionamento espacial dos *clusters*. Nesse caso, usadas em conjunto para melhor fundamentar as análises dos conteúdos das patentes sobre EAD.

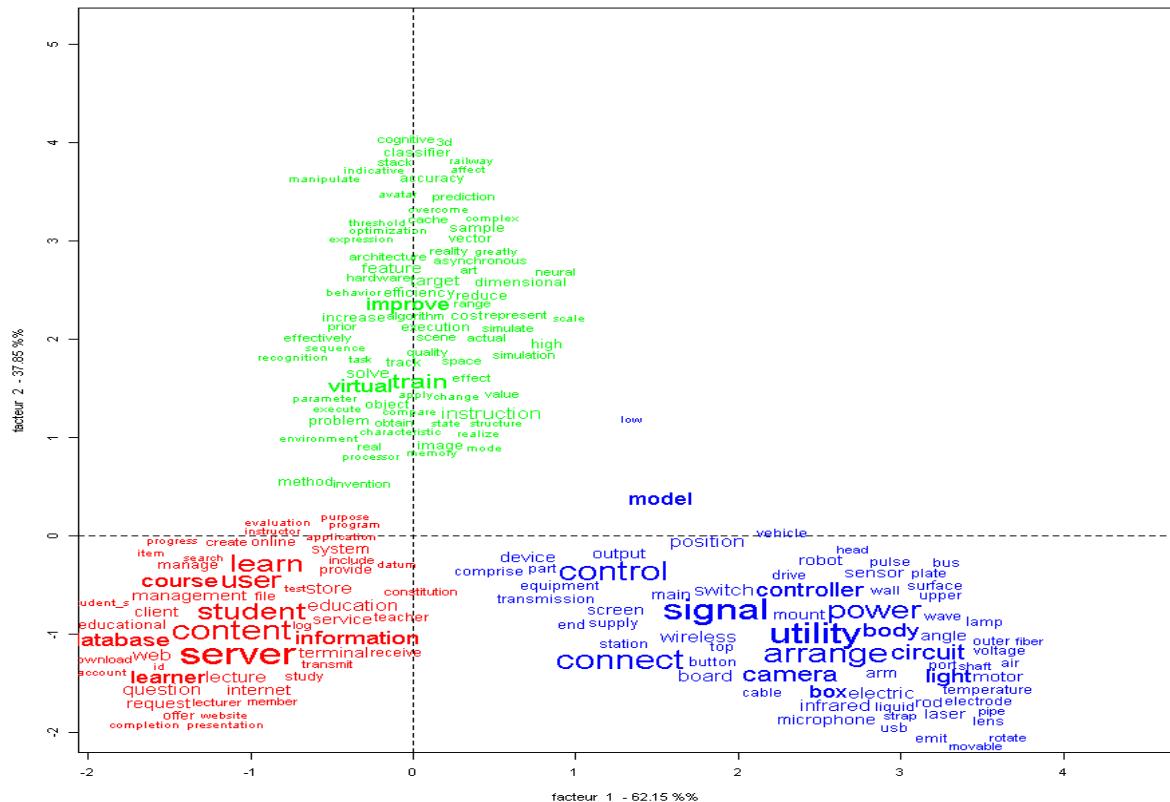
Figura 54 – AFC dos resumos das patentes sobre EAD



As verificações anteriores foram principalmente exploratórias, no entanto, para compreender mais sobre o objeto da patente. Precisa-se ir além da análise exploratória, gráfica, acessando mais detalhes, trazendo as formas que compõe os contextos. Observando as formas que compõe o contexto léxico, é possível, por meio da observação da natureza das palavras, identificar algum padrão. Entender um padrão dentro de um contexto léxico possibilita não só dar nome ao contexto, como também alcançar o dicionário para aquele contexto. O dicionário consiste nas principais palavras usadas para compor o contexto, no caso da base de dados em análise e do tema de interesse: base de patentes sobre EAD. Por conveniência, escolheu-se observar as 20 principais palavras de cada contexto. A técnica é simples e consiste na criação de um *sub-cluster*, envolvendo apenas as 20 principais palavras do *cluster* inteiro, cuja sensibilidade foi calculada pelo Iramuteq com base no qui-quadrado. Essa técnica permite observar o contexto léxico (*cluster*) de forma mais próxima e iniciar a compreensão das formas (palavras) cuja proximidade léxica, formam determinado contexto. A primeira visão das formas

que compõe o contexto está apresentada na figura 10, ainda considerando os fatores calculados na AFC.

Figura 55 – AFC dos resumos das patentes sobre EAD com as formas ativas que os compõe.



Todos os textos são compostos por formas ativas e suplementares. O interesse é nas formas ativas, porque são as formas que carregam os significados léxicos e que, portanto, em termos de linguagem natural, são as palavras que dão sentido a comunicação. Essas formas ativas passaram por um processo de lematização. A lematização consiste na flexão das formas para o masculino, singular e infinitivo, removendo as flexões verbais e temporais, uma vez que precisa-se de uma base única de comparação. Na Figura 55 observar-se as palavras que compõe cada contexto. Quanto as formas suplementares, pode-se destacar que são claramente fundamentais para a construção do texto; no entanto, nesse trabalho de decomposição do conteúdo, tais formas não conseguem explicar os sentidos ou proporcionar a separação em contextos. Mesmo assim, separou-se as formas suplementares que são apresentadas na Figura 56. Em relação aos cabeçalhos dos resumos das patentes, há informações sobre o país

depositante e sobre o código IPC da patente. O IPC (*International Patent Classification*) é o código que todas as patentes depositadas recebem, identificando sua função e aplicação. Tais informações foram oferecidas pelo Patent2Net v2 e fez-se o seu uso e separação, também com base nos contextos analisados e mostrados graficamente na

Figura 57. Pode-se observar que no *cluster* 1 predominam as patentes Koreanas (KR) e no *cluster* 3, predomina a China (CN).

Figura 56 – AFC dos resumos das patentes sobre EAD com as formas suplementares que os compõe.

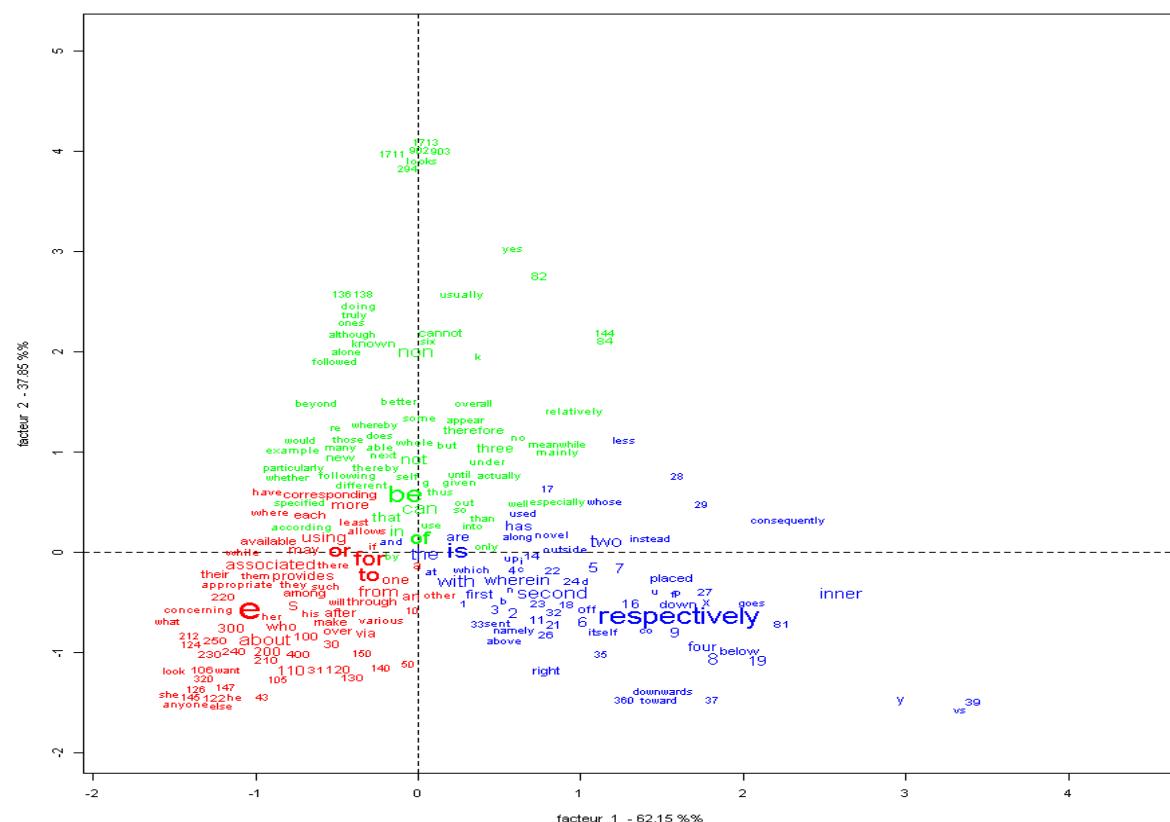
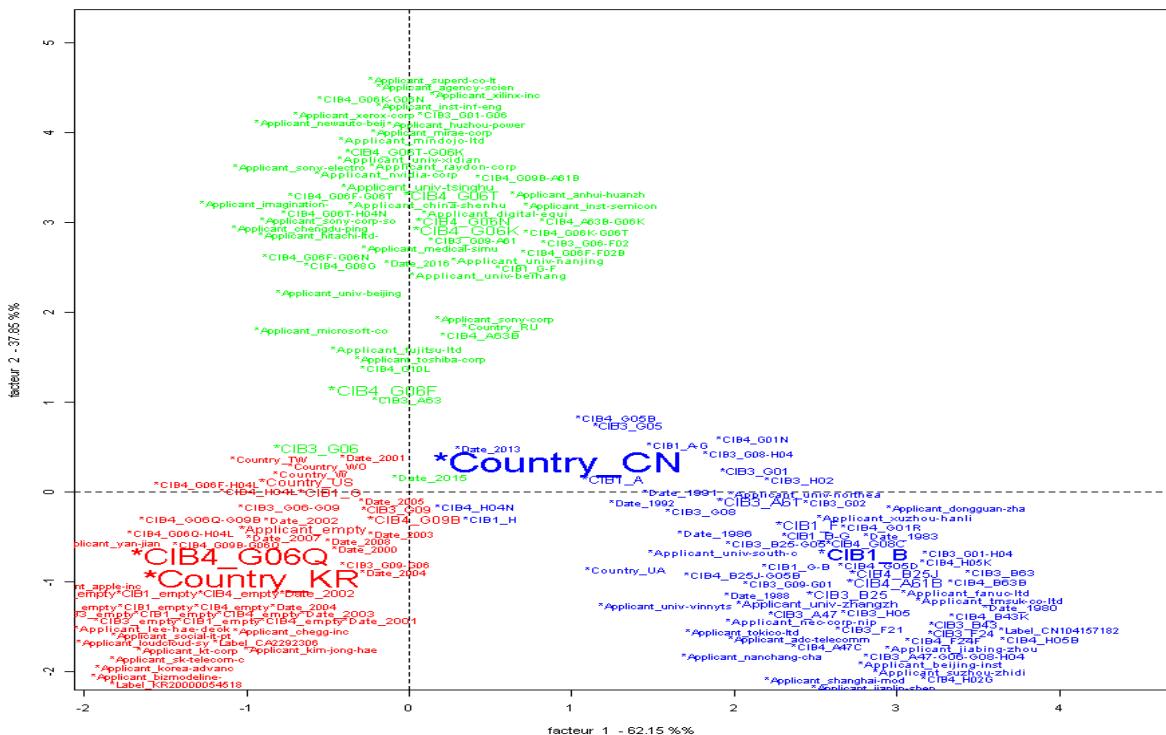


Figura 57 – AFC dos resumos das patentes sobre EAD com os cabeçalhos, indicando países depositantes e o código IPC



A seguir, separa-se cada contexto em sub-contextos, para proceder com a nomeação de cada um dos contextos com base nas formas predominantes, bem como elaborar comentários sobre a natureza das formas que compõem o contexto. Conforme descrito, serão utilizadas as primeiras vinte palavras classificadas por frequência, por conveniência de apresentação.

Figura 58 – Dendograma e sub-contexto do contexto 1 com as formas componentes.

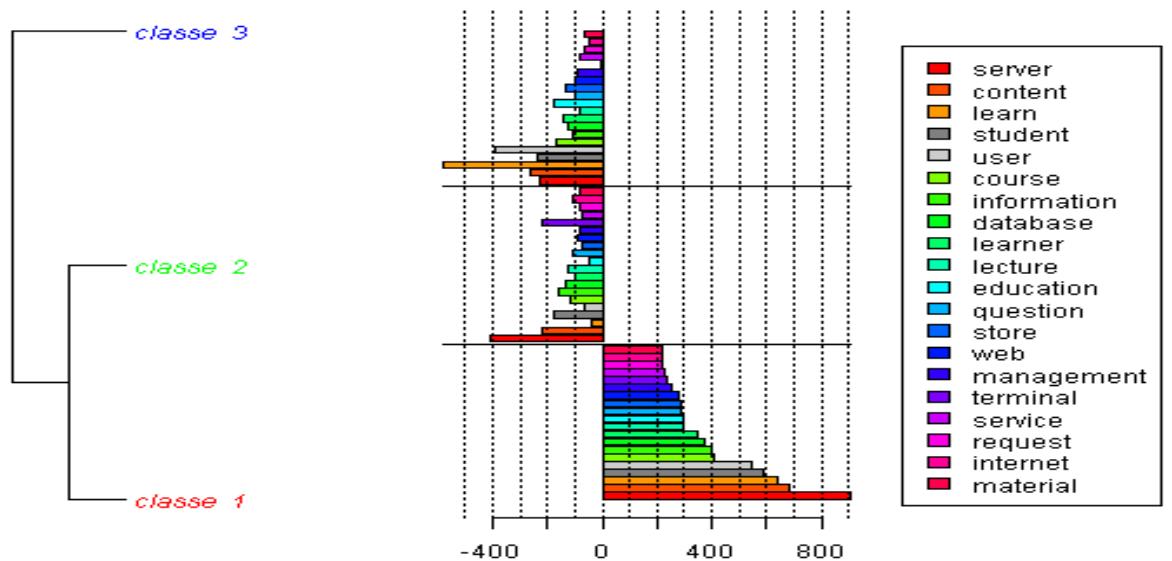


Figura 59 – Dendograma e sub-contexto do contexto 2 com as formas componentes.

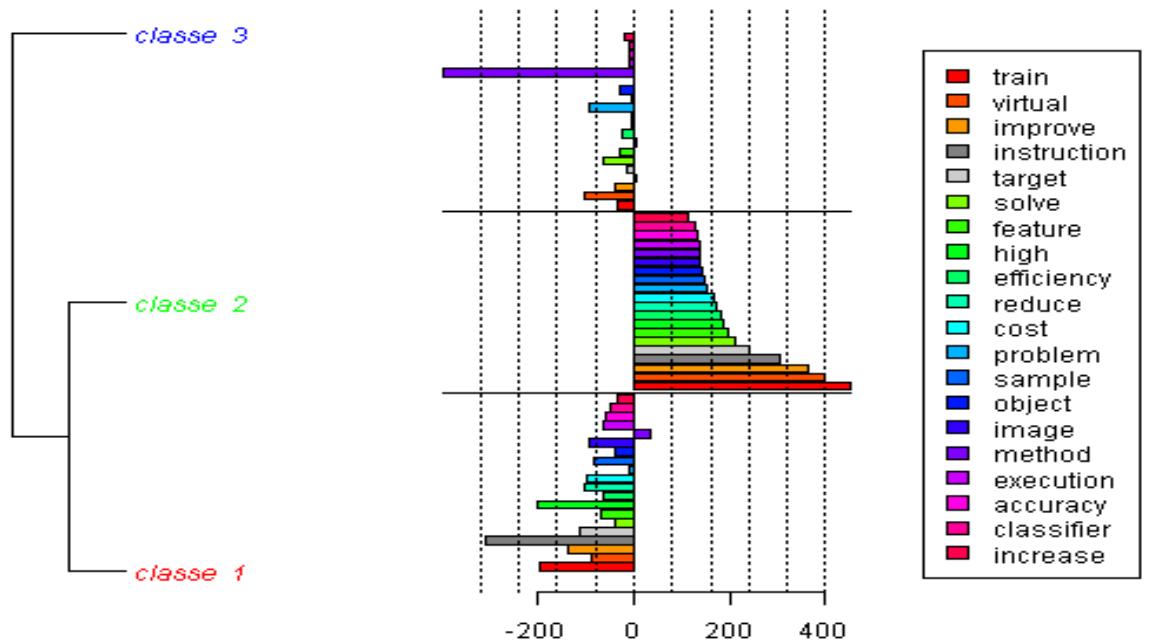
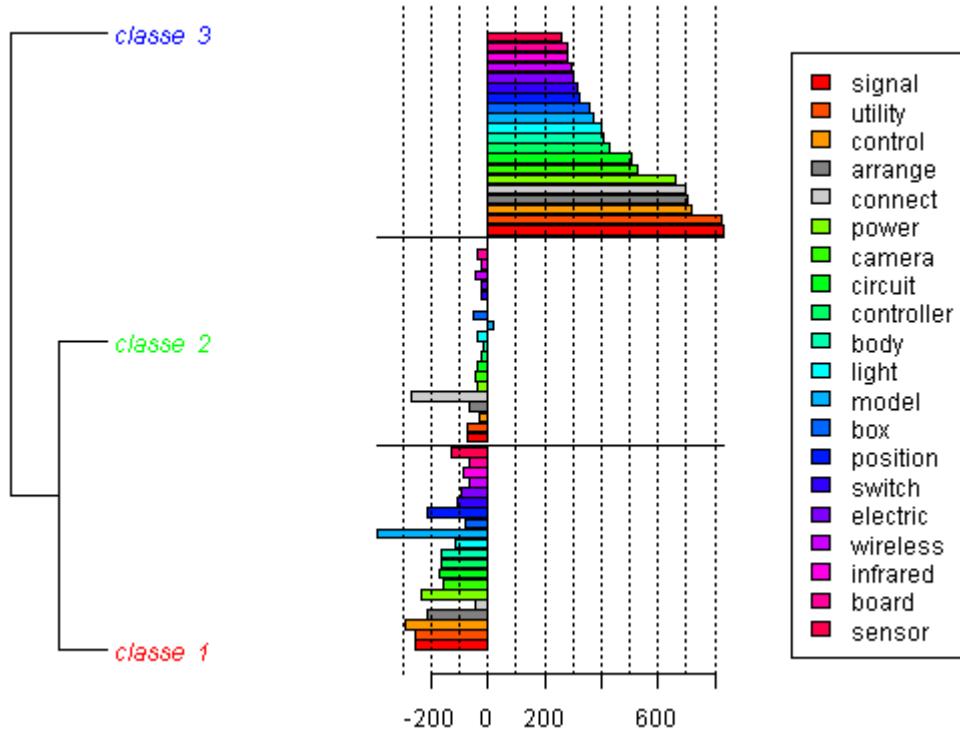
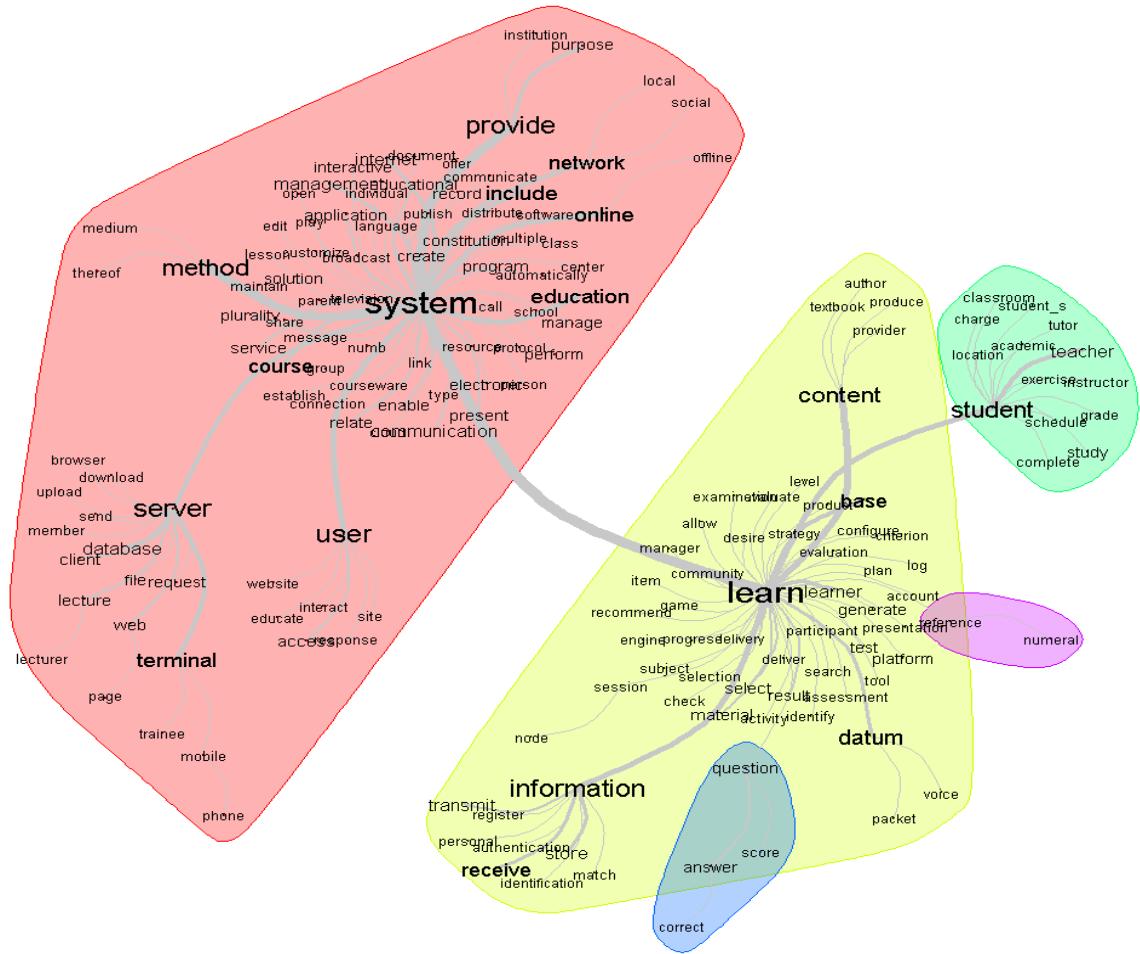


Figura 60 – Dendograma e sub-contesto do contexto 3 com as formas componentes.



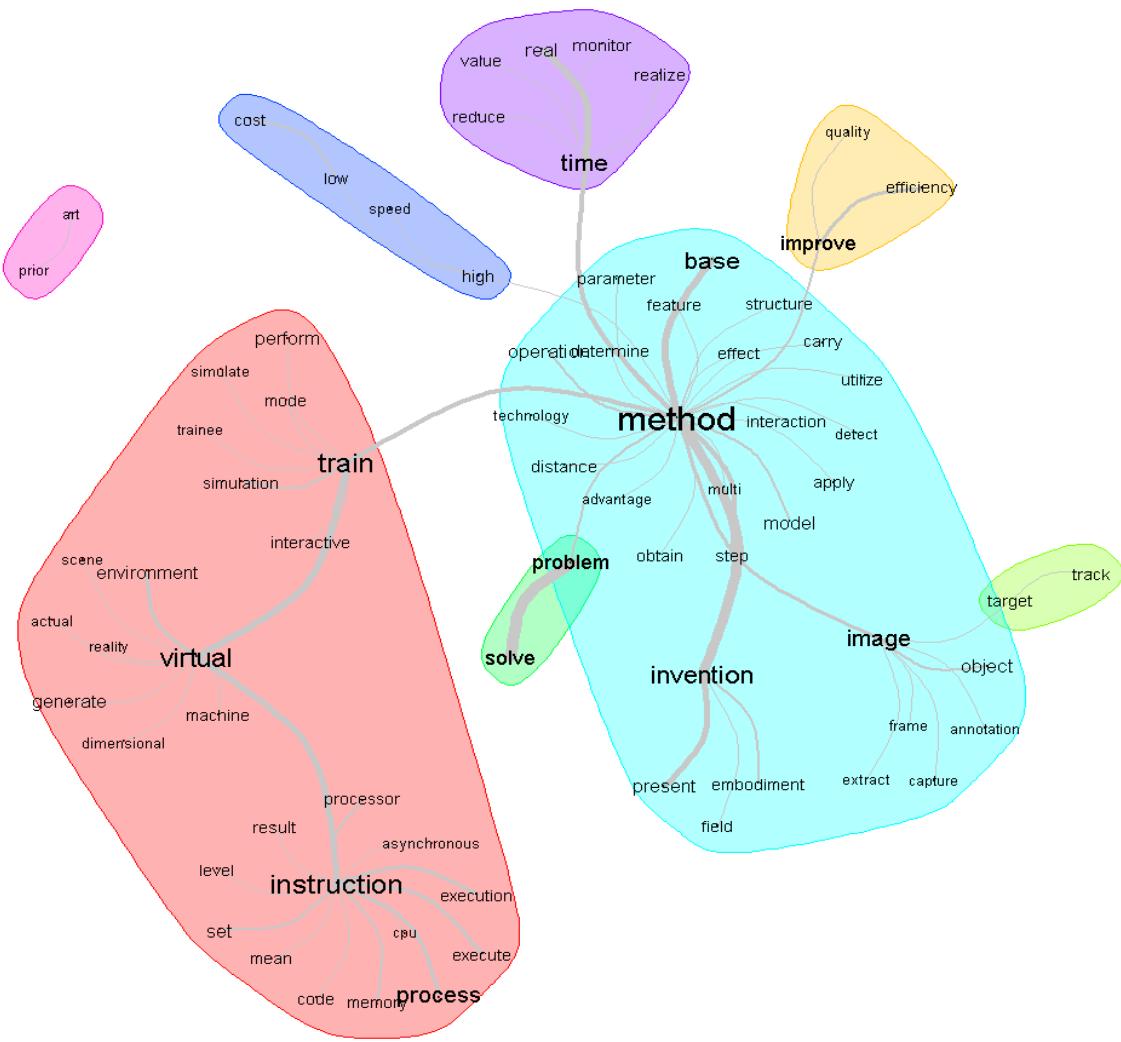
Apresentadas as palavras que compõe os contextos, preparou-se uma análise de similaridade, configurando por conveniência os limites de rede para a frequência mínima de 25 ocorrências, buscou-se destacar quais são as formas predominantes dentro do *sub-cluster*. Em termos práticos, significa detalhar o *sub-cluster*, por meio de uma técnica de construção de uma rede de formas, usando a técnica de visualização por comunidades e propagação do *label*. Essa técnica permite observar a extensão de uma determinada forma dentro do sub-contesto, destacando assim o nome que será sugerido ao contexto. Tais visões de cada sub-contesto são apresentadas a seguir e permitem a nomeação de cada contexto.

Figura 61 – Análise do Sub-contexto 1, frequência de ocorrência mínima 25 e visualização da propagação do *label*



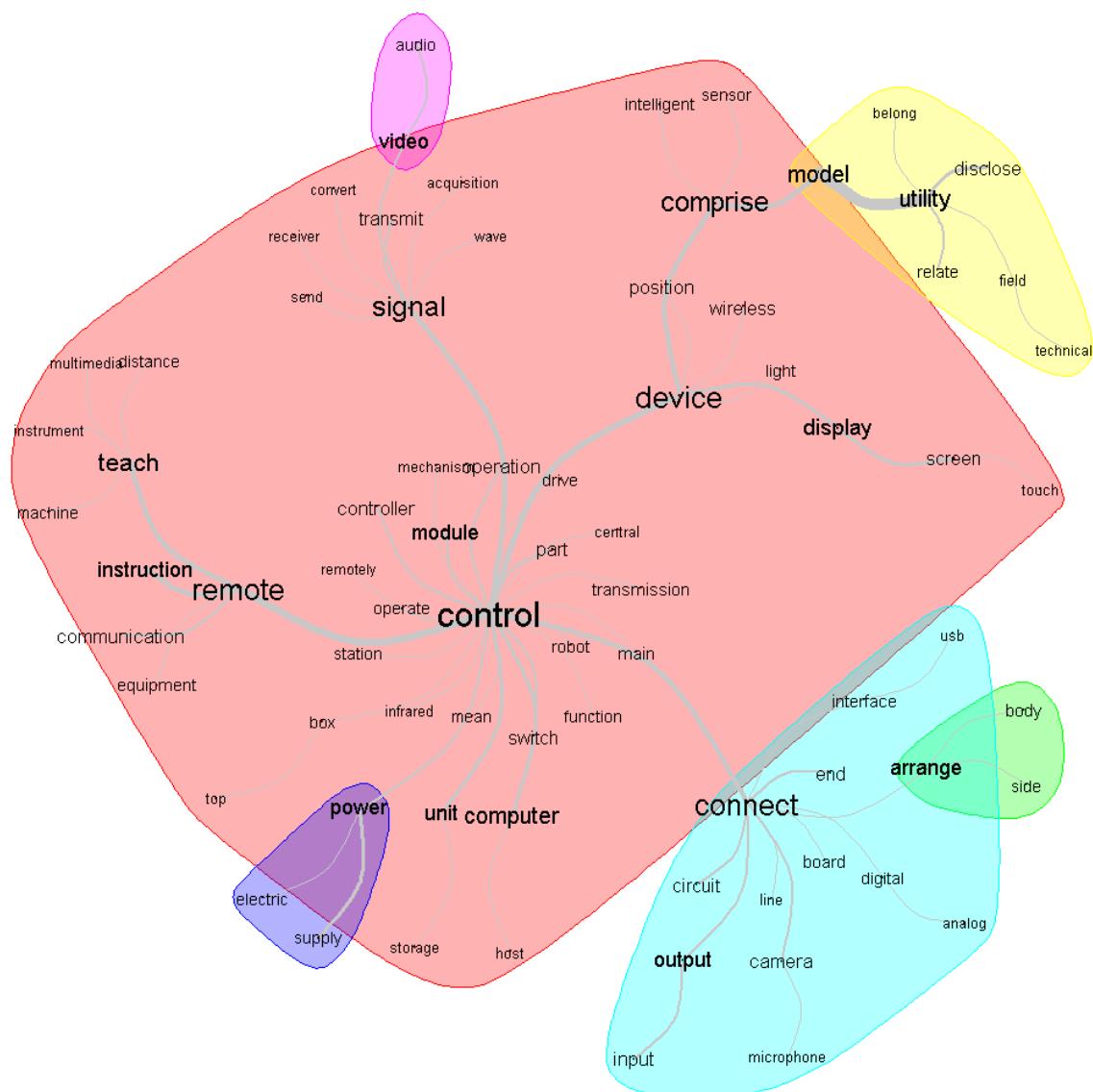
As nuvens formadas ao redor de cada forma na Figura 61, mostra o nível da propagação da forma de maior frequência, considerando que selecionou-se para apresentar apenas as formas com no mínimo 25 ocorrências. Identificou-se, portanto, no sub-contexto 1, a predominância das patentes que trama dos sistemas de aprendizagem. A mais predominante é a forma “sistema”. Nesse caso, como o sub-contexto 1 é a parte mais significativa (20 palavras) do contexto 1, nomeou-se o contexto 1: SISTEMAS e APRENDIZAGEM.

Figura 62 – Análise do Sub-contexto 2, frequência de ocorrência mínima 25 e visualização da propagação do label



As nuvens formadas ao redor de cada forma na Figura 62 mostram o nível da propagação da forma de maior frequência, considerando que foram selecionadas para apresentar apenas as formas com no mínimo 25 ocorrências. Identificou-se, portanto, no sub-contexto 2 a predominância das patentes que tratam dos métodos e de virtualidade. A mais predominante é a forma “virtualidade”. Nesse caso, como o sub-contexto 2 é a parte mais significativa (20 palavras) do contexto 2, nomeou-se o contexto 2: MÉTODO E VIRTUALIDADE.

Figura 63 – Análise do Sub-contexto 3, frequência de ocorrência mínima 25 e visualização da propagação do label

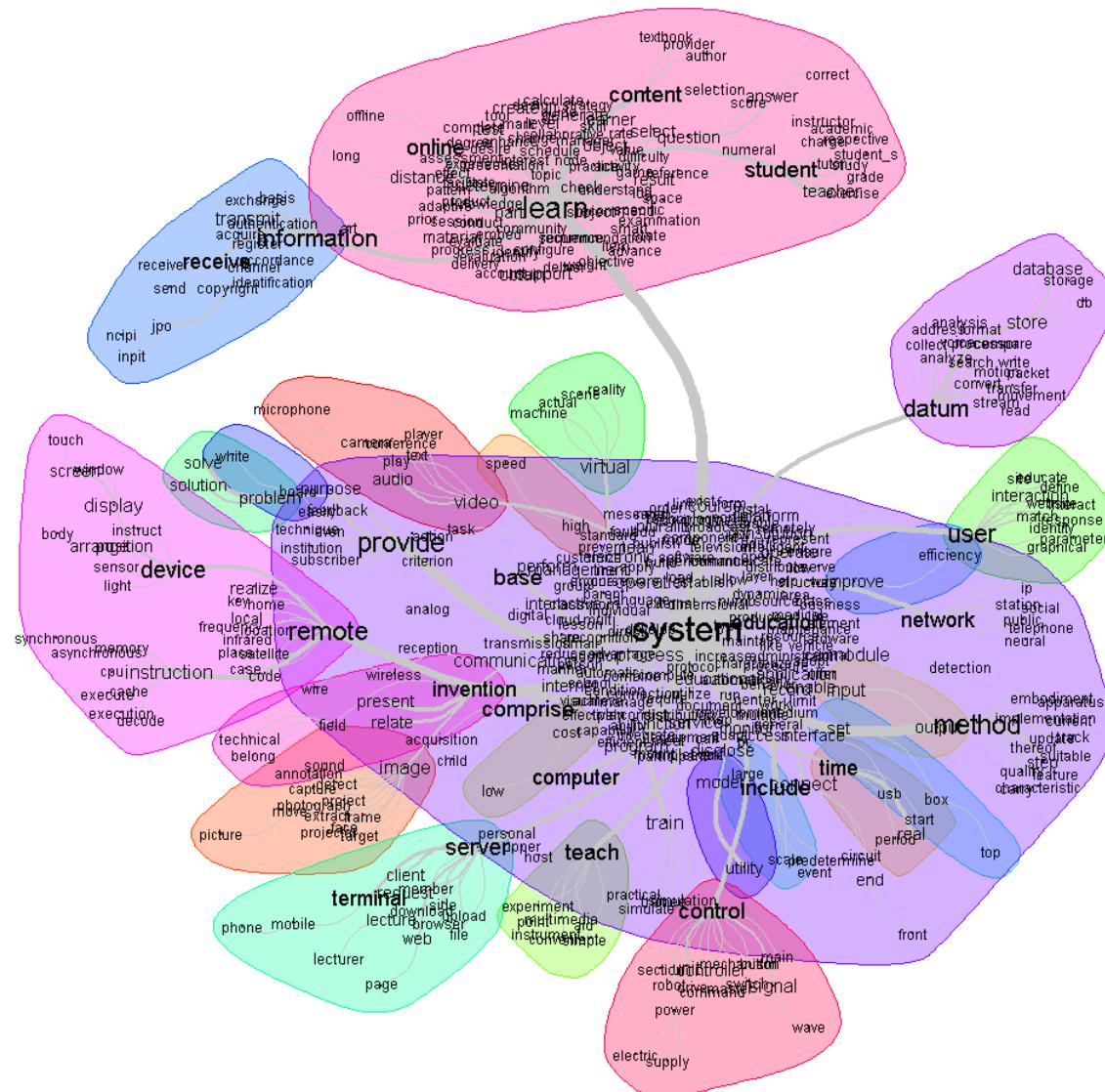


As nuvens formadas ao redor de cada forma na Figura 63 mostram o nível da propagação da forma de maior frequência, considerando que foram selecionadas para apresentar apenas as formas com no mínimo 25 ocorrências. Identificou-se, portanto, no sub-contexto 3 a predominância das patentes que tratam do controle e de conexão. A mais predominante é a forma “controle”. Nesse caso, como o sub-contexto 3 é a parte mais significativa (20 palavras) do contexto 3, nomeou-se o contexto 3: CONTROLE E CONEXÃO.

4.2.2 Resumos das patentes por similaridade

A análise de similaridade leva em consideração a co-ocorrência. Co-ocorrência, nesse estudo, caracteriza-se pela frequência maior que 1 para determinada forma. Sempre que uma forma tem frequências maiores que 1, a forma é agrupada e por conveniência de apresentação foi escolhida a técnica da comunidade multi-nível. As nuvens, nesse caso, representam o agrupamento das formas com maiores frequências. Manteve-se o mesmo parâmetro de frequência mínima, escolhido por conveniência (25 ocorrências) como limite dos nós. Pode-se observar que os principais *clusters* são o de sistema, aprendizagem, remoto, dispositivo, controle e servidor-terminal. A predominância dos depósitos de patentes sobre EAD, tratam, portanto, a respeito das invenções que descrevem sistemas, muito mais dos que ensino ou método (*teach* ou *method*) por exemplo, como se poderia intuir. A Figura 64 apresenta uma rede de formas obtidas por meio da análise de conteúdo dos resumos das patentes sobre EAD usando a técnica da similaridade.

Figura 64 – Análise de similaridade de formas ativas de 2.870 resumos de patentes sobre EAD



4.2.3 Resumos das patentes por AFC (análise de 5 modalidades)

Após realizar as análises dos conteúdos pela CHD, AFC e Similaridade, pode-se apresentar os resultados sobre o conteúdo dos resumos das patentes sobre EAD. As separações dos contextos léxicos trazem a visão geral da dinâmica dos depósitos de patentes, revelando que sobre o tema EAD há três contextos: *Cluster 1* = Contexto 1, *Cluster 2* = Contexto 2 e *Cluster 3* = Contexto 3. Conforme definido, contexto são os agrupamentos que se foram pelas

formas que apresentam proximidade léxica, que, em termos práticos, significa que, pelo tratamento dos algoritmos de Reinert (1990), ficam separados em contextos diferentes os conteúdos dos resumos das patentes que estão tratando de coisas diferentes, resolvendo, assim, os efeitos polissêmicos. As decomposições em sub-contextos permitiram nomear cada contexto, bem como revelar as 20 principais formas que compõe cada contexto. Tal detalhamento é especialmente importante numa mão, para nomear o contexto e, em outra mão, para expandir o *thesaurus* sobre o tema EAD. As análises AFC representam um primeiro ponto de triangulação e mostraram as distâncias entre os contextos, em dois eixos fatoriais. A visão final, considerando todos os resumos recuperados, foi apresentada pela rede de formas que usa a técnica de similaridade ou co-ocorrência definida anteriormente. No sentido de consolidar as análises realizadas de forma resumida, apresenta-se a matriz de consolidação na Tabela 7.

Tabela 7 – Síntese de análise de conteúdo dos resumos de patentes sobre EAD.

Síntese de análise de conteúdo

Número de resumos de patentes sobre EAD tratados: 2870

	Distribuição do Conteúdo*	Nome do Contexto	Conteúdo Predominante**	País Predominante	IPC Predominante
Contexto1	62.30%	Sistemas e Aprendizagem	servidor, conteúdo, aprender, estudante, usuário, curso, informação, bando de dados e aprendiz	Korea	CIB4
Contexto2	26.80%	Método e Virtualidade	treino, virtual, melhorar, instrução, alvo, resolver, característica, alto, eficiêcia e reduz	Russia	CIB4
Contexto 3	21%	Controle e Conexão	sinal, utilidade, controle, arranjo, conectar, força, camera, circuito, controlador e corpo	China	CIB1

* Há sobreposição, portanto, a soma da distribuição pode ser diferente de 100%

** Aqui as 10 principais formas, das 20 analisadas nos sub contextos

4.2.4 Descrições das invenções nas patentes sobre EAD por CHD e AFC

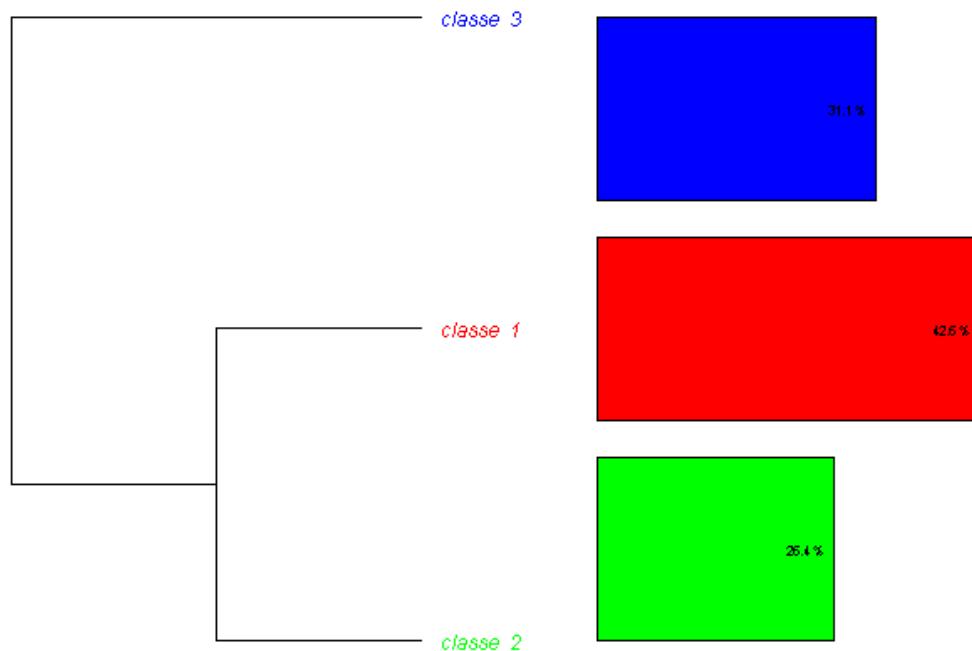
Foram recuperadas 340 descrições de invenções das patentes sobre EAD, como resposta ao uso da expressão de busca construída. Tais descrições de invenções foram classificadas em 83.458 segmentos de texto (textos com similaridades em relação a distância das formas léxicas), 8.083 formas lematizadas (ativas e suplementares) das quais analisou-se as

formas ativas, por trazerem mais significado léxico; portanto, maior potencial de explicação sobre o conteúdo dos textos. As formas ativas com frequência acima de 40 foram 2.954, sendo essas as formas mais relevantes do *corpus* analisado: 340 descrições de invenções nas patentes sobre EAD. A média de formas por segmento de texto foi de 36,03, e o total de segmentos de texto tratados foi de 74.511, que representa 89,28% dos segmentos de texto classificados que foram separados em 3 *clusters*, por meio do algoritmo de classificação hierárquica descendente (CHD) definido por Reinert (Reinert, 1990). A Tabela 8, resume os resultados da análise de CHD das descrições das invenções nas patentes sobre EAD.

Tabela 8 – Resultados da análise de CHD das descrições das invenções nas patentes sobre EAD.

Número de textos	340
Número de segmentos de textos	83.458
Número de formas	34.212
Número de ocorrências	3.007.482
Número de <i>lemmas</i>	28.491
Número de formas ativas	24.074
Número de formas suplementares	4.417
Número de formas ativas com frequência ≥ 40	2.954
Média de formas por segmento	36.035.874
Número de <i>clusters</i>	3
74.511 segmentos classificados em 83.458 (89.28%)	

Figura 65 – Dendograma da CHD das descrições das invenções nas patentes sobre EAD



Identificou-se a formação de três *clusters*, nesse caso, originados das descrições das invenções encontradas nas patentes sobre EAD, que representam os três contextos ou grupos polissêmicos, cujas formas léxicas apresentam associação em termos de significado léxico. Pode-se argumentar inicialmente que, nas 340 descrições de invenções das patentes sobre EAD, há três contextos que descrevem o tema, sendo o contexto 1 o representante de 42,6% do conteúdo total, o contexto 2 o representante de 26,4% do conteúdo total e o contexto 3 o representante de 31,1% do conteúdo total. Parece que a presença de apenas três contextos no total de patentes sobre EAD demonstra que os depositantes de patentes têm se concentrado nas mesmas descrições de invenções sobre a EAD quando depositam suas patentes; e que as descrições das invenções das patentes contêm mais detalhes sobre as patentes do que os resumos analisados inicialmente. Pelas evidências verificadas na análise de conteúdo das descrições das invenções, isoladamente, aparenta que as patentes sobre EAD apresentam evolução tecnológica incremental: inovações incrementais.

O contexto 1, representante de 42,6% do conteúdo total das descrições de invenção analisadas, mostra que os depositantes concentraram suas explicações sobre o objeto da patente de forma muito parecida. A análise do contexto 1 será útil, portanto, para identificar o perfil tecnológico dominante nesse conjunto de patentes sobre EAD, especialmente sobre o os

princípios que fundamentam a patente. Por outro lado, a análise do contexto 2, que representa 26,4% do conteúdo total, pode revelar a presença de princípios tecnológicos patenteados em EAD não dominantes. Analisar descrições de invenções não dominantes são relevantes a medida em que podem auxiliar na compreensão de possíveis tecnologias em declínio ou nascentes, bastando, para isso, incluir uma triangulação simples com uma tabela de patente x data de depósito.

Todas as tecnologias descritas nas patentes classificadas no contexto 2, cuja frequência aumenta em função do tempo, representam tecnologias nascentes e aquelas cujas frequências diminuem em função do tempo representam tecnologias em declínio. Pode-se argumentar ainda que, se busca-se conhecer a tecnologia dominante, deve-se analisar as patentes do *cluster* 1; mas, se opta-se por analisar possíveis inovações radicais, há maior chance de serem encontradas no *cluster* 2. Tal constatação é uma das questões que são discutidas posteriormente, quando outras dimensões das patentes já estiverem analisadas, permitindo, portanto, alguma triangulação.

Uma análise factorial por correspondência (AFC) das descrições das invenções demonstra as distâncias ou disposições de um contexto em relação ao outro. Nessa observação gráfica introdutória, observa-se que os contextos são como dimensões. Como trata-se de análise de conteúdo, isso significa que houve possibilidade de agrupar os temas, aqui chama-se de contextos, em três partes, que se encontram distanciados conforme os fatores demonstrados na Figura 66.

Conforme mencionado anteriormente, todos os textos são compostos por formas ativas e suplementares. O interesse é nas formas ativas, porque são as formas que carregam os significados léxicos e que, portanto, em termos de linguagem natural, são as palavras que dão sentido à comunicação. Essas formas ativas passaram por um processo de lematização.

A lematização consiste na flexão das formas para o masculino, singular e infinitivo, removendo as flexões verbais e temporais, uma vez que precisa-se de uma base única de comparação. Na Figura 67, Figura 68 e Figura 69, pode-se observar as palavras que compõe cada contexto. Quanto as formas suplementares, pode-se destacar que são claramente fundamentais para a construção do texto, no entanto, nesse trabalho de decomposição do conteúdo, tais formas não conseguem explicar os sentidos ou proporcionar a separação em

contextos. Mesmo assim, separou-se as formas suplementares que são apresentadas na Figura 68. Em relação aos cabeçalhos dos resumos das patentes, há algumas informações sobre o país depositante e sobre o código IPC da patente. O IPC (*International Patent Classification*) é o código que todas as patentes depositadas recebem, identificando sua função e aplicação. Tais informações foram oferecidas pelo Patent2Net v2 e fez-se o seu uso e separação, também com base nos contextos analisados e mostrados graficamente na Figura 69. Pode-se observar que no *cluster* 1 predominam as patentes IPC CIB1, no *cluster* 2 a predominância do depositante “AC Propertie” e no *cluster* 3 predominam as patentes com o IPC CIB3 e CIB1.

Figura 66 – AFC das descrições das invenções das patentes sobre EAD

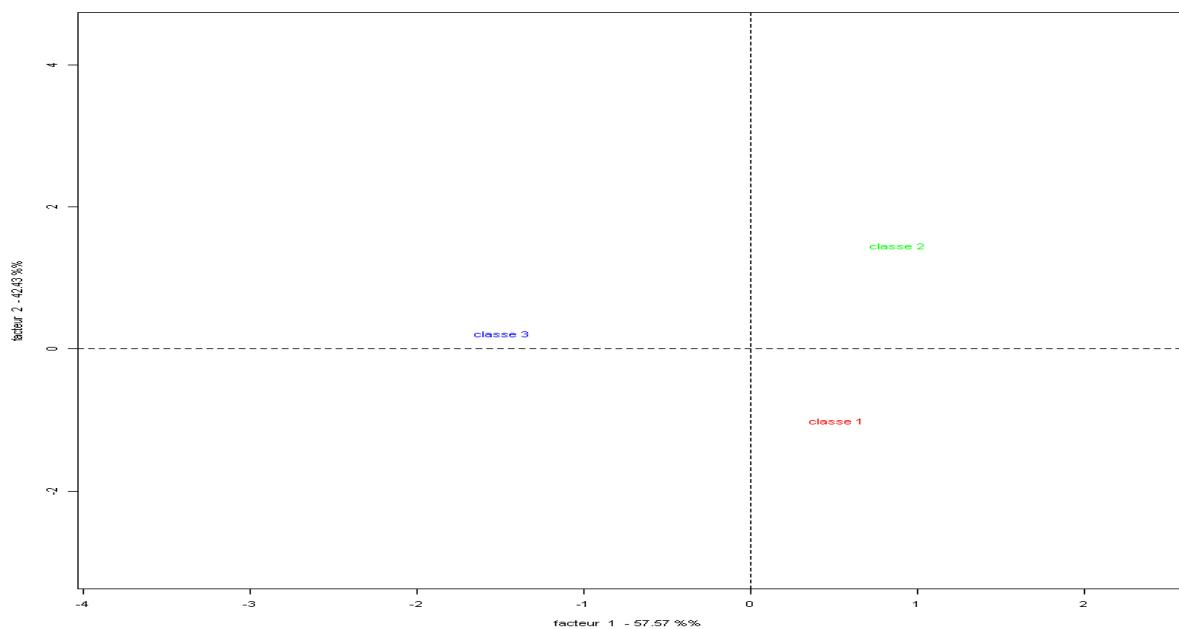


Figura 67 – AFC das descrições das invenções das patentes sobre EAD com as formas ativas que as compõe

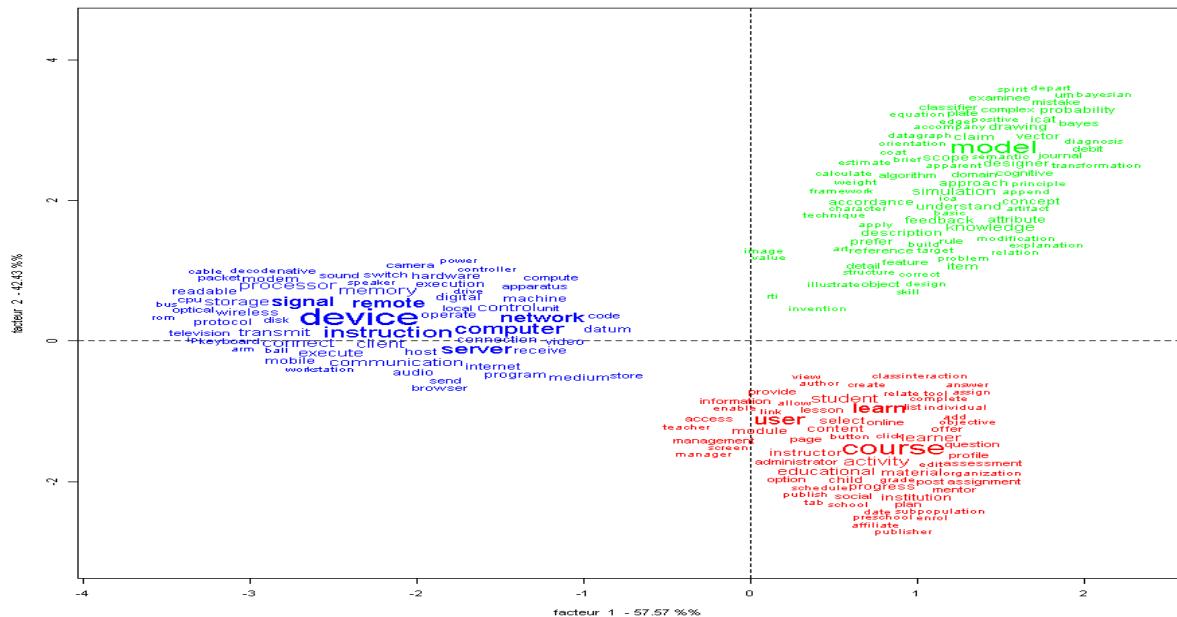


Figura 68 – AFC das descrições das invenções das patentes sobre EAD com as formas suplementares que as compõe

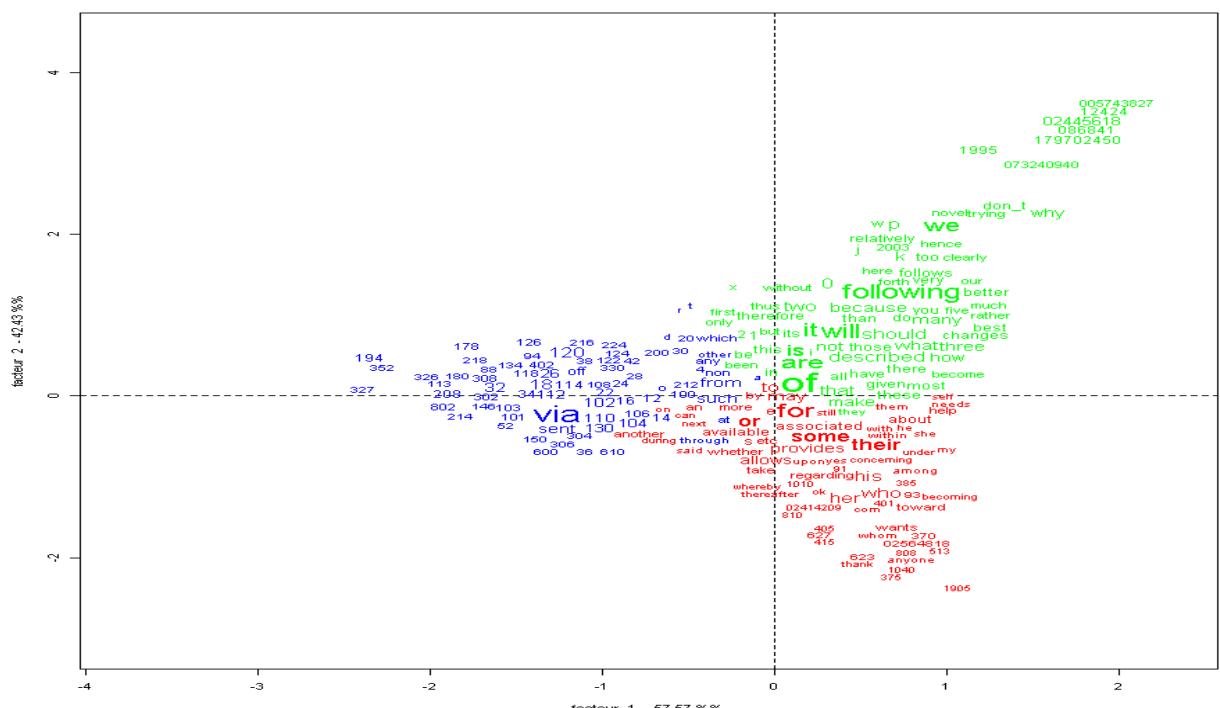
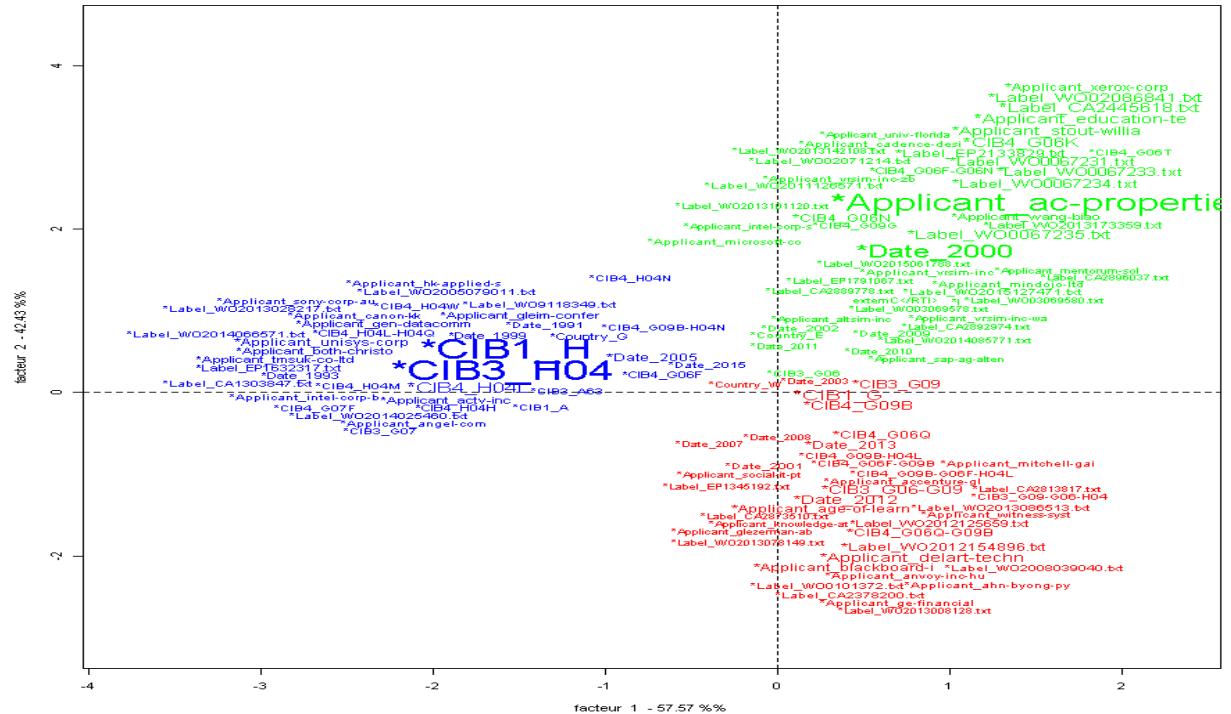


Figura 69 – AFC das descrições das invenções das patentes sobre EAD com os cabeçalhos, indicando países depositantes, depositantes e o código IPC



A seguir, apresenta-se a separação de cada contexto em sub-contextos, para proceder com a nomeação de cada um dos contextos com base nas formas predominantes, bem como elaborar comentários sobre a natureza das formas que compõe o contexto. Conforme descrito, usou-se as primeiras vinte palavras classificadas por frequência, por conveniência de apresentação. A essa subdivisão, da mesma forma como feito na análise do conteúdo dos resumos, chama-se de sub-contextos. As Figura 70, Figura 71 e Figura 72 apresentam o resultado gráfico da separação das vinte formas de maior frequência.

Figura 70 – Dendograma e sub-contesto 1 do contexto 1 com as formas componentes

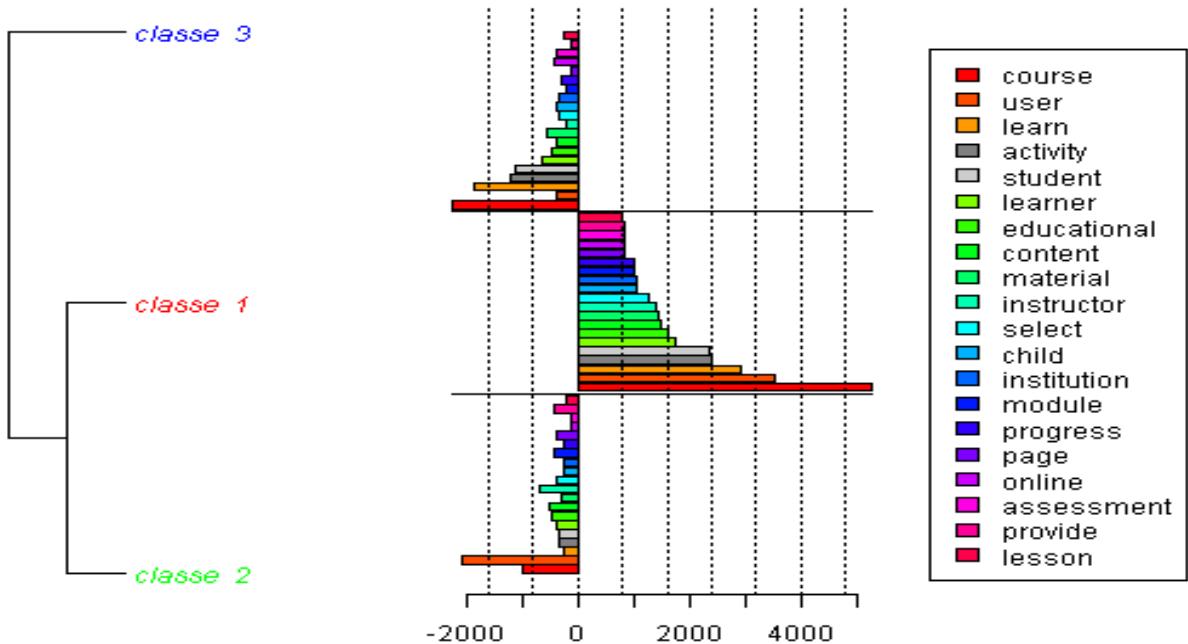


Figura 71 – Dendograma e sub-contesto 2 do contexto 2 com as formas componentes

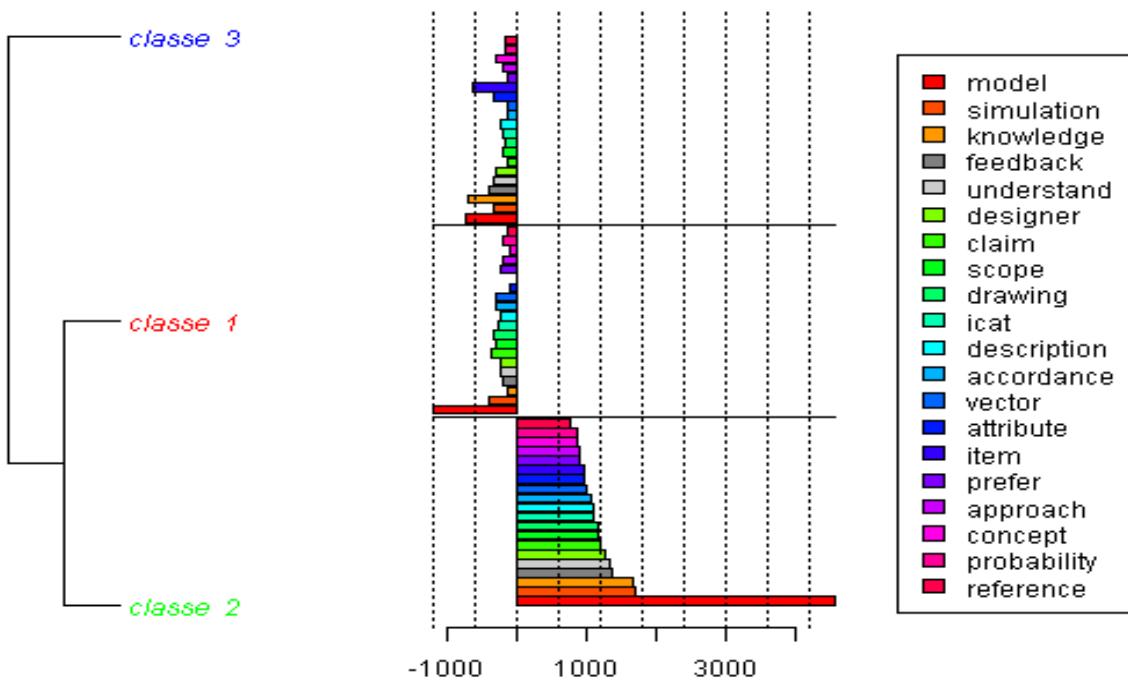
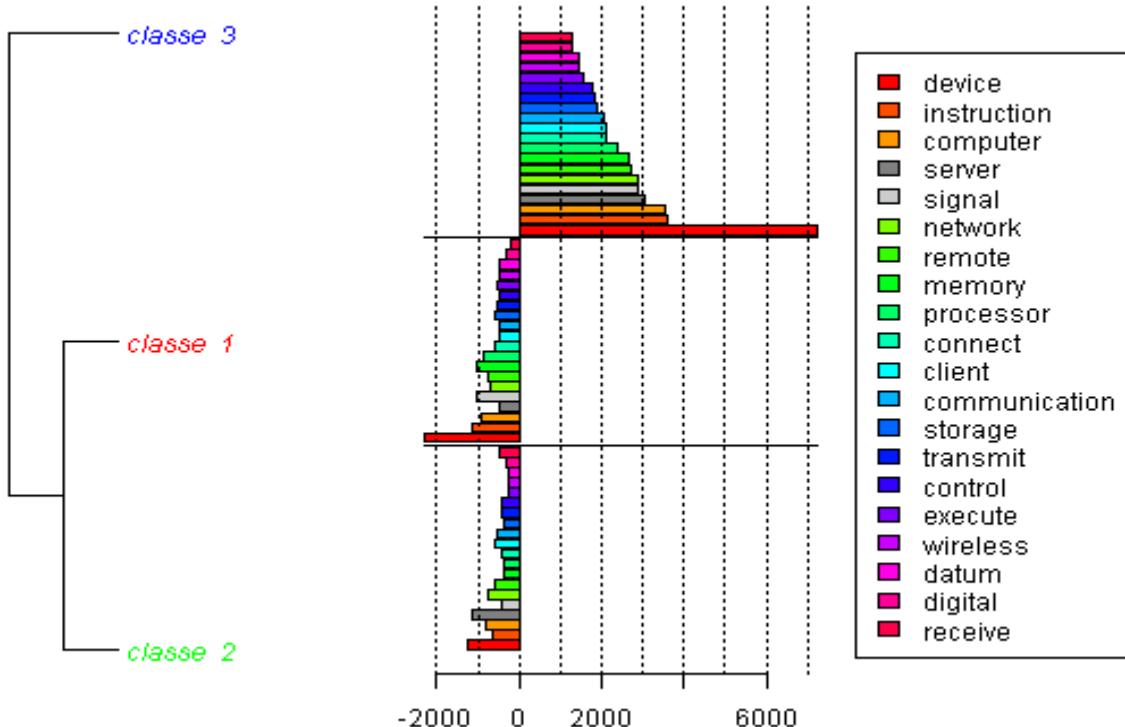
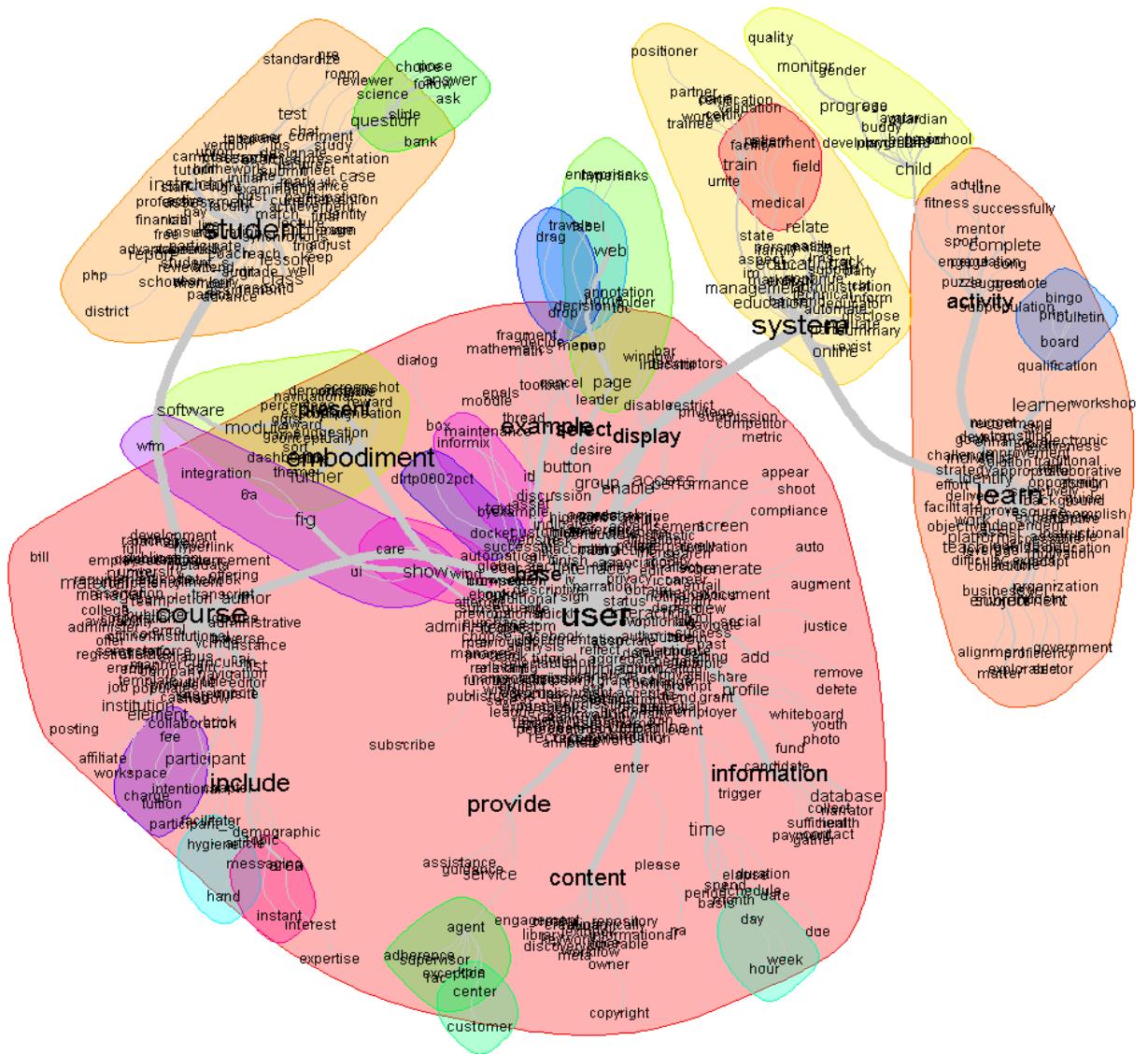


Figura 72 – Dendograma e sub-contesto 3 do contexto 3 com as formas componentes



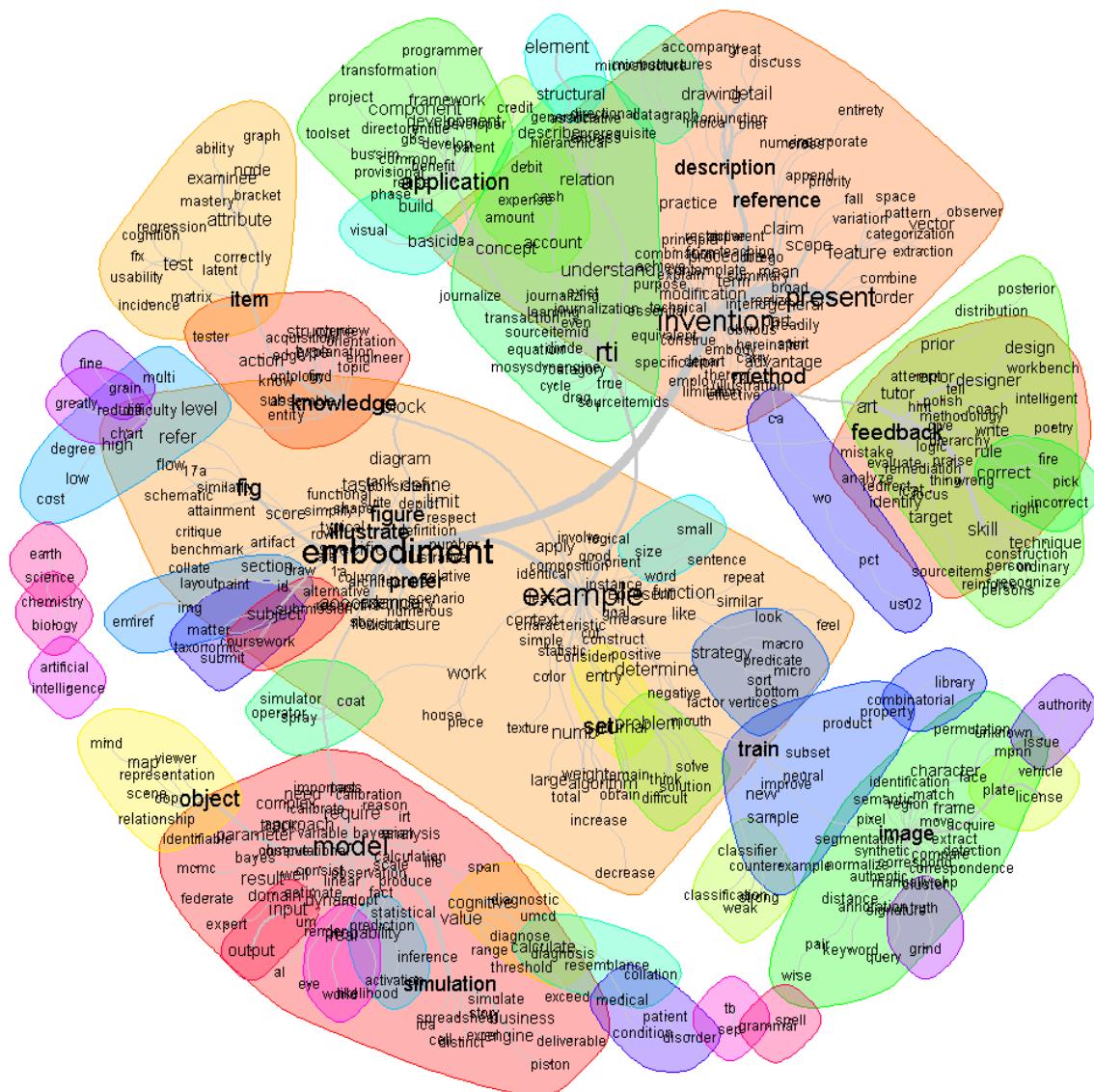
Apresentadas as palavras que compõe os contextos, preparou-se uma análise de similaridade, configurando, por conveniência, os limites de rede para a frequência mínima de 25 ocorrências, buscando destacar quais são as formas predominantes dentro do *sub-cluster*. Em termos práticos, significa detalhar o *sub-cluster*, por meio de uma técnica de construção de uma rede de formas, usando a técnica de visualização por “comunidades e propagação do *label*”. Essa técnica permite observar a extensão de uma determinada forma dentro do sub-contesto, destacando assim o nome que sugeriu-se ao contexto. Tais visões de cada sub-contesto serão apresentadas a seguir e permitirão a nomeação de cada contexto.

Figura 73 – Análise do Sub-contexto 1, frequência de ocorrência mínima 25 e visualização do tipo propagação do *label*



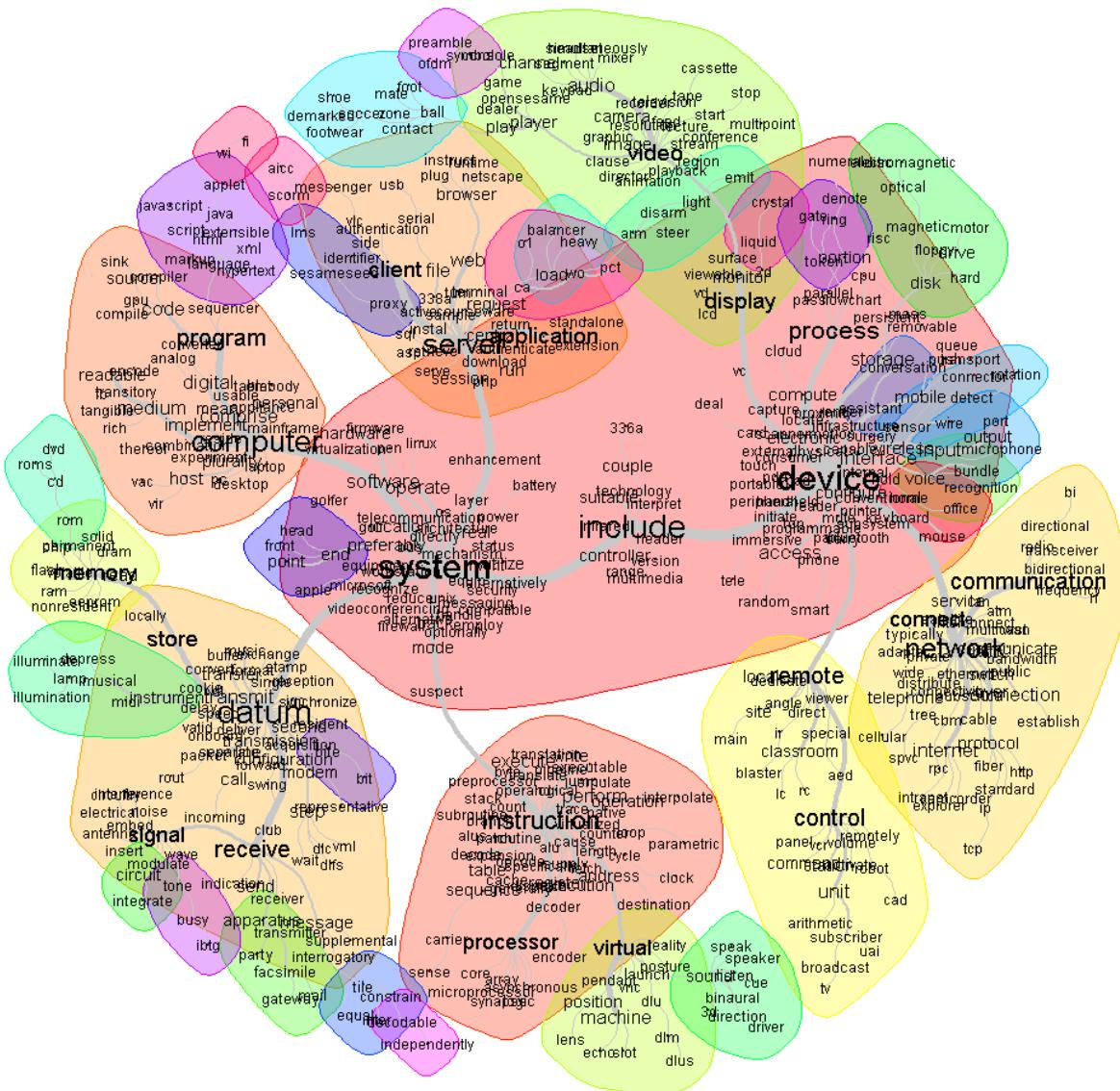
As nuvens formadas ao redor de cada forma na Figura 73 mostram o nível da propagação da forma de maior frequência, considerando que selecionou-se para apresentar apenas as formas com no mínimo 25 ocorrências. Identificou-se, portanto, no sub-contexto 1 a predominância das patentes que tratam do curso, do usuário, do estudante e de técnicas para aprender. A forma predominante é “curso e usuário”. Nesse caso, como o sub-contexto 1 é a parte mais significativa (20 palavras) do contexto 1, nomeou-se o contexto 1: USUÁRIO.

Figura 74 – Análise do Sub-contexo 2, frequência de ocorrência mínima 25 e visualização do tipo propagação do label



As nuvens formadas ao redor de cada forma na Figura 74, mostram o nível da propagação da forma de maior frequência, considerando que selecionou-se para apresentar apenas as formas com no mínimo 25 ocorrências. Identificou-se, portanto, no sub-contexro 2 a predominância das patentes que tratam da personalização, dos modelos e da simulação. A forma predominante é a “personalização”. Nesse caso, como o sub-contexro 2 é a parte mais significativa (20 palavras) do contexto 2, nomeou-se o contexto 2: PERSONALIZAÇÃO e MODELOS.

Figura 75 – Análise do Sub-contexto 3, frequência de ocorrência mínima 25 e visualização do tipo propagação do *label*

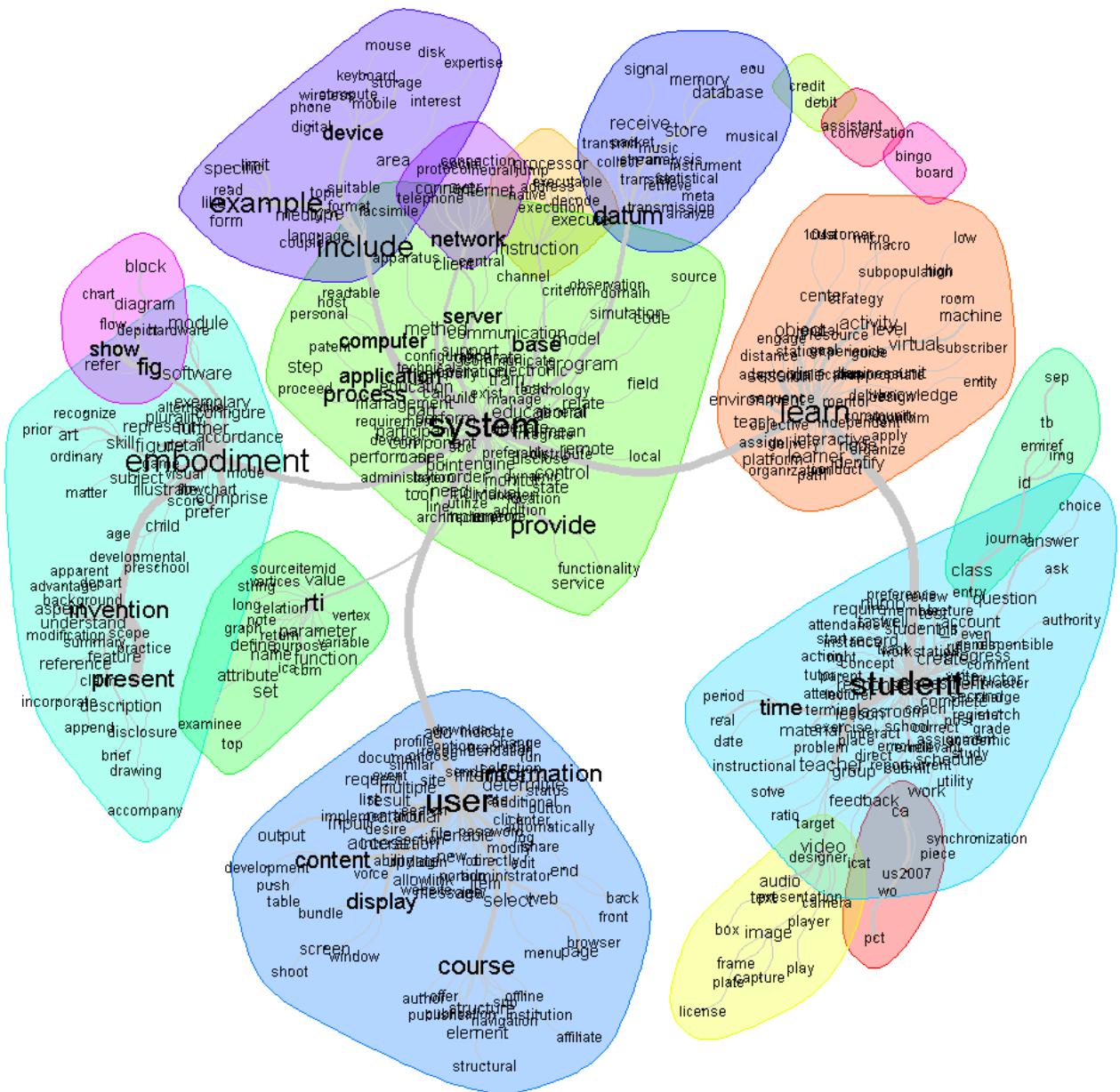


As nuvens formadas ao redor de cada forma na Figura 75 mostram o nível da propagação da forma de maior frequência, considerando que selecionou-se para apresentar apenas as formas com no mínimo 25 ocorrências. Identificou-se, portanto, no sub-contexto 3 a predominância das patentes que tratam dos sistemas, dispositivos, computadores e da instrução. A forma predominante é a “dispositivo”. Nesse caso, como o sub-contexto 3 é a parte mais significativa (20 palavras) do contexto 3, nomeou-se o contexto 3: SISTEMAS E DISPOSITIVOS.

4.2.5 Descrições das invenções das patentes por similaridade

A análise de similaridade leva em consideração a co-ocorrência. Co-ocorrência, nesse estudo, caracteriza-se pela frequência maior que 1 para determinada forma. Sempre que uma forma tem frequências maiores que 1, a forma é agrupada e, por conveniência de apresentação, foi escolhida a técnica da comunidade multi-nível. As nuvens, nesse caso, representam o agrupamento das formas com maiores frequências. Manteve-se o mesmo parâmetro de frequência mínima, escolhido por conveniência (25 ocorrências) como limite dos nós da rede de formas. Pode-se observar que os principais clusters são, o de sistemas, de aprendizagem, de aspectos sobre o estudante e o de personalização. A predominância dos depósitos de patentes sobre EAD, tratam, portanto, a respeito das invenções que descrevem sistemas e suas interfaces com usuários e estudantes, com especial destaque a personalização. Também aqui, nas descrições da invenção, assim como nos resumos das patentes, são destaques os dispositivos eletrônicos usados para viabilizar o EAD, muito mais do que o ensino per se ou método (teach ou method) por exemplo, como se poderia intuir. A Figura 76 apresenta a rede de formas obtidas por meio da análise de conteúdo das descrições das invenções, presentes nas patentes sobre EAD usando a técnica da similaridade.

Figura 76 – Análise de similaridade de formas ativas de 344 descrições das invenções das patentes sobre EAD



Após realizar as análises dos conteúdos pela CHD, AFC e Similaridade, pode-se apresentar os resultados sobre o conteúdo das descrições das invenções presentes nas patentes sobre EAD. As separações dos contextos léxicos nos trazem a visão geral da dinâmica dos depósitos de patentes, revelando que, sobre o tema EAD, há três contextos: *Cluster 1 = Contexto 1*, *Cluster 2 = Contexto 2* e *Cluster 3 = Contexto 3*. Conforme definido, contextos são os

agrupamentos que se foram pelas formas que apresentam proximidade léxica, que em termos práticos, significam que, pelo tratamento dos algoritmos de Reinert (1990), ficam separados em contextos diferentes, os conteúdos das descrições das invenções das patentes que estão tratando de coisas diferentes, resolvendo assim os efeitos polissêmicos. A decomposição em sub-contextos permite nomear cada contexto, bem como revelar as 20 principais formas que compõe cada contexto. Tal detalhamento é especialmente importante, numa mão, para nomear o contexto e, em outra mão, para expandir o *thesaurus* sobre o tema EAD. As análises AFC representam um primeiro ponto de triangulação e mostram as distâncias entre os contextos, em dois eixos fatoriais. A visão final, considerando todos as descrições recuperadas: 340 descrições de invenções de uma amostra de 3.090 patentes recuperadas, foi apresentada pela rede de formas que usa a técnicas de similaridade ou co-ocorrência definida anteriormente. No sentido de apresentar as análises realizadas de forma sintética, sugeriu-se a matriz de consolidação a seguir, na Tabela 9.

Tabela 9 – Síntese de análise de conteúdo das descrições das invenções de patentes sobre EAD

Síntese de análise de conteúdo

Número de descrições de invenções de patentes sobre EAD tratados: 340

	Distribuição do Conteúdo*	Nome do Contexto	Conteúdo Predominante**	País Predominante***	IPC Predominante
Contexto1	42.60%	Usuário	curso, usuário, aprender, atividade, estudante, aprendiz, educacional, conteúdo, material e instrutor	Não se Aplica	CIB1
Contexto2	26.40%	Personalização e Modelos	modelo, simulação, conhecimento, feedback, entendimento, projetista, reivindicação, escopo e desenhando	Não se Aplica	CIB4
Contexto 3	31.10%	Sistemas e Dispositivos	dispositivo, instrução, computador, servidor, sinal, rede, remoto, memória, processador e conectar	Não se Aplica	CIB1 e CIB3

* Há sobreposição, portanto, a soma da distribuição pode ser diferente de 100%

** Aqui as 10 principais formas, das 20 analisadas nos sub contextos

*** As descrições de invenções não apresentaram referência a países nessa amostra

4.2.6 Reivindicações das patentes de EAD por CHD e AFC

Foram recuperados 322 textos sobre reivindicações das patentes sobre EAD, como resposta ao uso da expressão de busca construída. Destaca-se aqui, que as reivindicações (*claims*) consistem nos apontamentos específicos da propriedade industrial que se pretende obter. Tratam-se, portanto, dos elementos entendidos como de potencial valor, ou fundamentais, sobre os quais, os depositantes pretendem obter a patente. Tais reivindicações foram classificadas em 11.548 segmentos de texto: textos com similaridades em relação a distância das formas léxicas, 7.095 formas lematizadas (ativas e suplementares) das quais analisou-se as formas ativas, por trazerem mais significado léxico, portanto, maior potencial de explicação sobre o conteúdo dos textos. As formas ativas com frequência acima de 3 foram 2.868, sendo essas as formas mais relevantes do *corpus* analisado: 322 reivindicações presentes nas patentes sobre EAD. A média de formas por segmento de texto foi de 35,5 e o total de segmentos de textos tratados foi de 10.733, que representa 92,94% dos segmentos de texto classificados que foram separados em 2 *clusters*, por meio do algoritmo de classificação hierárquica descendente (CHD) definido por Reinert (Reinert, 1990). A Tabela 10, resume os resultados da análise de CHD das descrições das invenções nas patentes sobre EAD.

Tabela 10 – Resultados da CHD das reivindicações das patentes sobre EAD

Número de textos	322
Número de segmentos de textos	11548
Número de formas	7095
Número de ocorrências	410576
Número de <i>lemmas</i>	5205
Número de formas ativas	4577
Número de formas suplementares	628
Número de formas ativas com frequência ≥ 3	2868
Média de formas por segmento	35.553862
Número de <i>clusters</i>	2
10.733 segmentos classificados em 11.548 (92.94%)	

Identificou-se a formação de dois *clusters*, que, nesse caso, representam os dois contextos ou grupos polissêmicos, cujas formas léxicas apresentam associação em termos de significado léxico. Pode-se argumentar inicialmente que, nos 322 textos sobre reivindicações

presentes nas patentes sobre EAD, há dois contextos que descrevem o tema, sendo o contexto 1 o representante de 82,3% do conteúdo total, o contexto 2 o representante de 17,7% do conteúdo total. Parece que a presença de apenas dois contextos no total de patentes sobre EAD demonstra que os depositantes de patentes, tem se concentrado em reivindicações muito parecidas sobre o EAD, quando depositam suas patentes, uma vez que trata-se das reivindicações, que contém elementos específicos que se pretende patentear. Por meio das evidências verificadas na análise de conteúdo das reivindicações isoladamente, observou-se que as patentes sobre EAD, apresentam evolução tecnológica incremental: inovações incrementais, o que explica a similaridade das reivindicações.

O contexto 1, representante de 82,3% do conteúdo total das reivindicações analisadas, mostra que os depositantes concentraram seus objetivos de patenteamento em conteúdos muito parecidos. A análise do contexto 1, será útil, portanto, para identificar o perfil das especificidades tecnológicas entendidas como de maior valor potencial, representando, aquelas reivindicações que mais geram interesse do depositante, o que pode ser chamado de reivindicações dominantes nesse conjunto de patentes sobre EAD. Por outro lado, a análise do contexto 2, que representa 17,7% do conteúdo total, pode revelar reivindicações sobre patentes EAD, não dominantes. Analisar reivindicações não dominantes, é relevante a medida em que podem auxiliar na compreensão de possíveis interesses tecnológicos em declínio ou nascentes, da mesma forma como descrito nas análises anteriores. Pode-se argumentar ainda que, se busca-se conhecer a tecnologia dominante, deve-se analisar as patentes do *cluster* 1, mas se o objetivo é a analisar possíveis inovações radicais, há maior chance de encontrá-las no *cluster* 2. Tal constatação é uma das questões que serão discutidas posteriormente, quando outras dimensões das patentes já estiverem analisadas permitindo, portanto, maior fundamentação.

Passa-se agora a separar cada contexto em sub-contextos, para proceder a nomeação de cada um dos contextos com base nas formas predominantes, bem como elaborar comentários sobre a natureza das formas que compõe o contexto. Conforme descrito, usou-se as primeiras 20 palavras classificadas por frequência, por conveniência de apresentação. A essa subdivisão, da mesma forma como feito na análise do conteúdo dos resumos, chama-se de sub-contextos. As Figura 77 e Figura 78 apresentam o resultado gráfico da separação das 20 formas de maior frequência.

Figura 77 – Dendograma e sub-contesto 1 do contexto 1 com as formas componentes

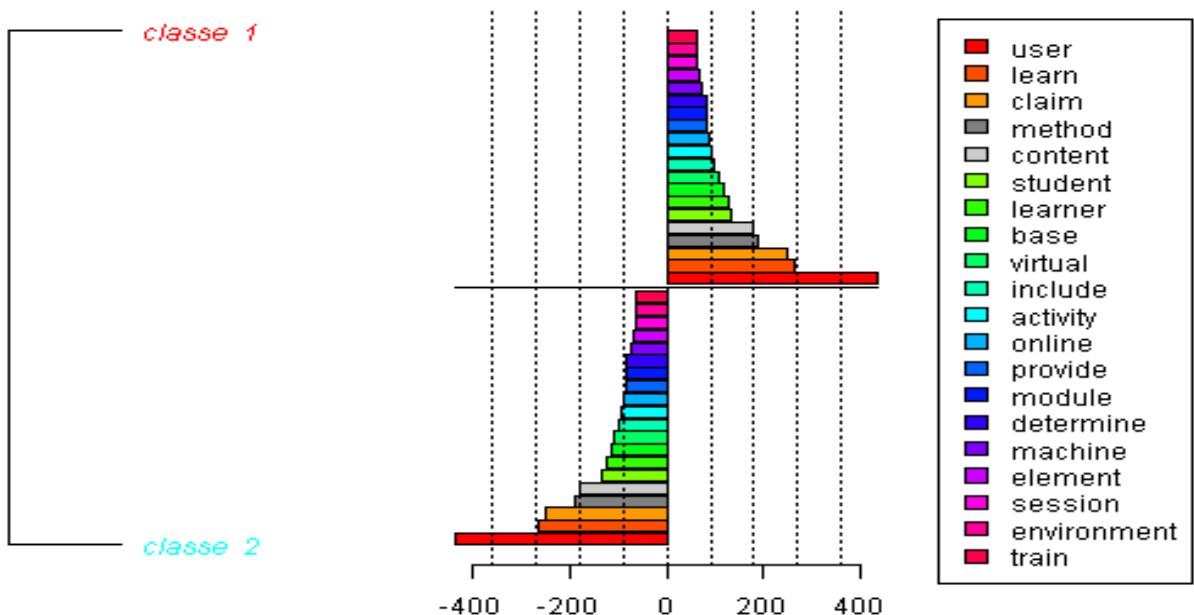
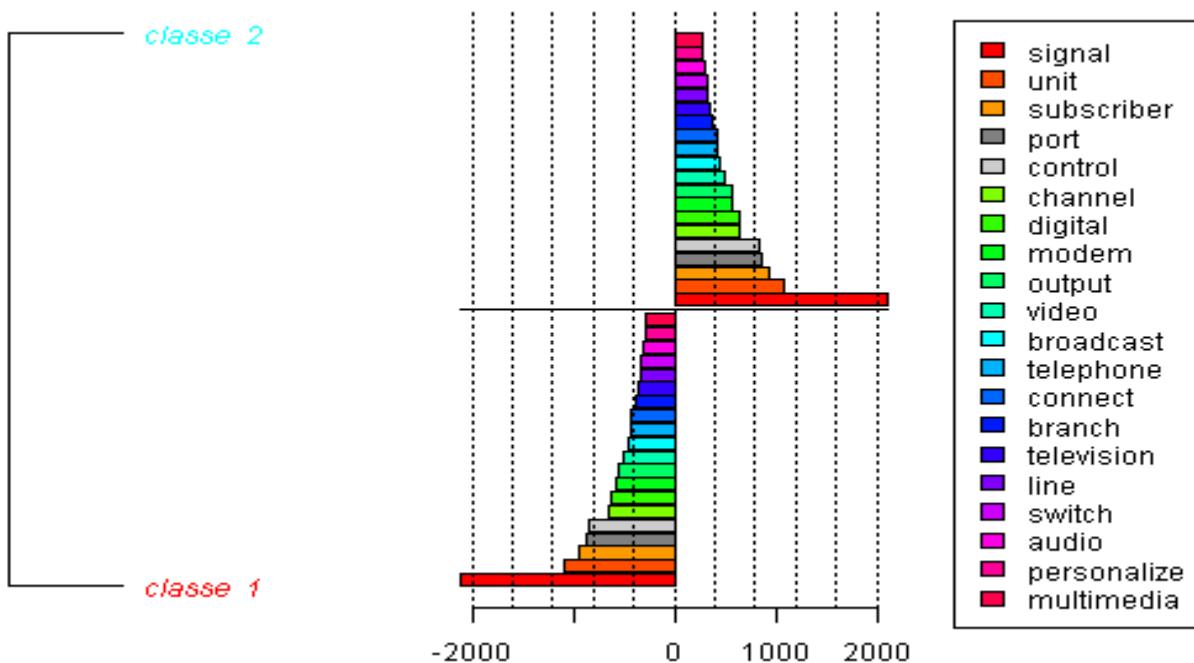
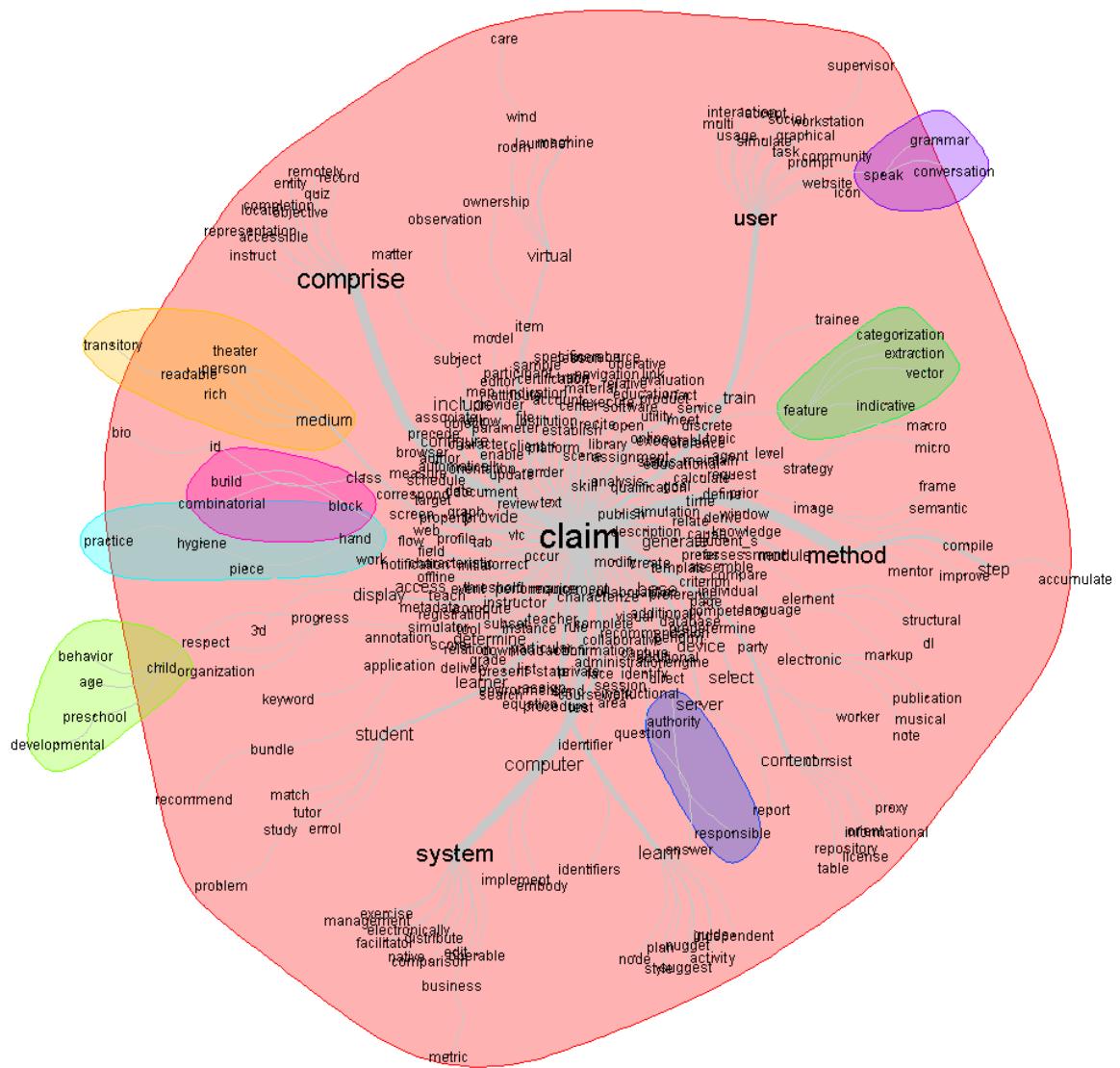


Figura 78 – Dendograma e sub-contesto 2 do contexto 2 com as formas componentes



Apresentadas as palavras que compõe os contextos, preparou-se uma análise de similaridade, configurando por conveniência, os limites de rede para a frequência mínima de 25 ocorrências, buscando destacar quais são as formas predominantes dentro do *sub-cluster*. Em termos práticos, significa detalhar o *sub-cluster*, por meio de uma técnica de construção de uma rede de formas, usando a técnica de visualização por comunidades e propagação do *label*. Essa técnica permite observar a extensão de uma determinada forma dentro do sub-contesto, destacando assim o nome que será sugerido ao contexto. Tais visões de cada sub-contesto serão apresentadas a seguir e permitirão a nomeação de cada contexto.

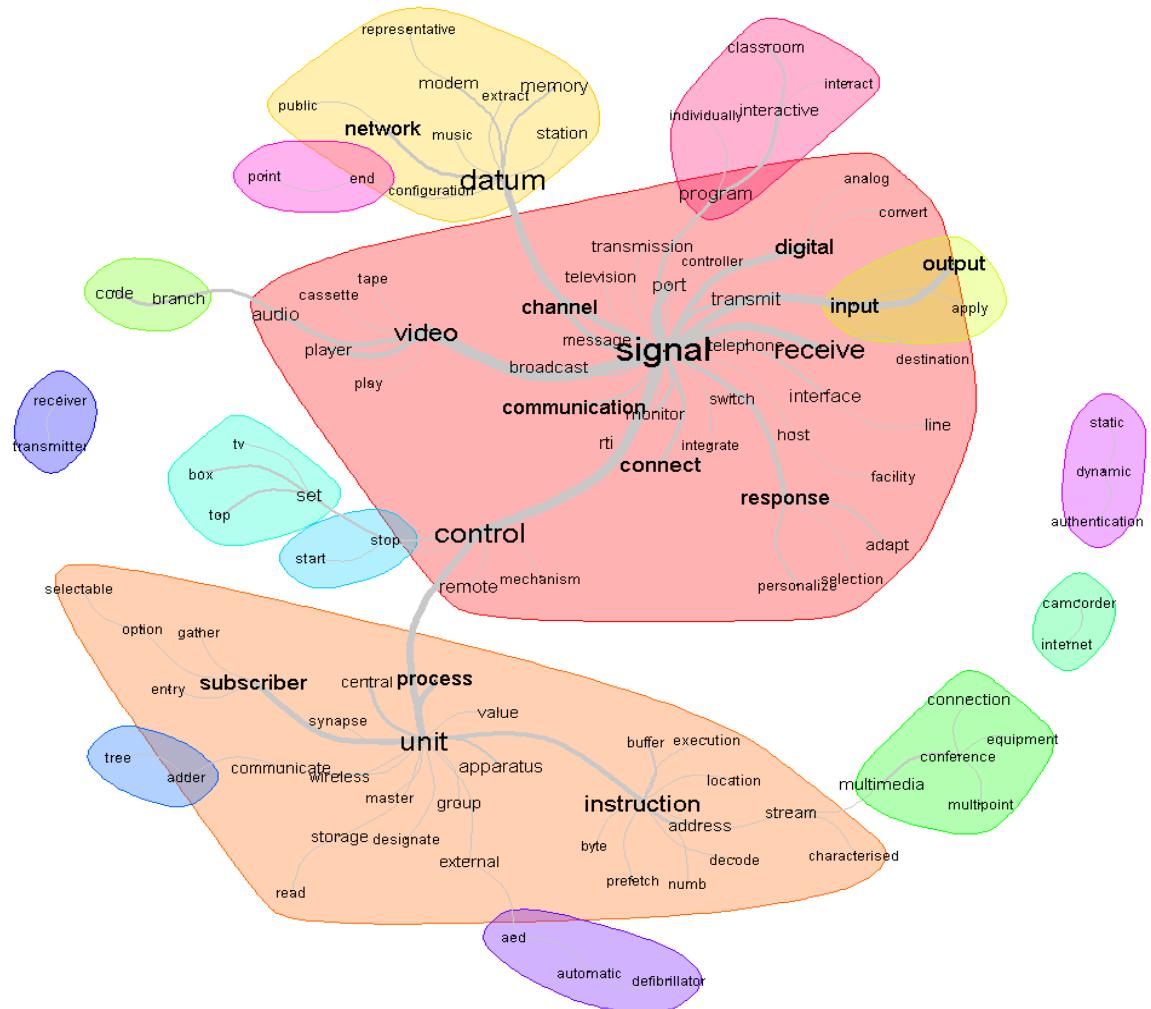
Figura 79 – Análise do Sub-contexto 1, frequência de ocorrência mínima 25 e visualização do tipo propagação do *label*



As nuvens formadas ao redor de cada forma na Figura 79 mostram o nível da propagação da forma de maior frequência, considerando que selecionou-se para apresentar apenas as formas com no mínimo 25 ocorrências. Identificou-se, portanto, no sub-contesto 1 a predominância das patentes que tratam do usuário, método, sistema e abrangência. A forma predominante é “abrangência”. Como a reivindicação tem por objetivo delimitar e especificar os elementos sob os quais a patente será concedida, a forma “abrange ou abrangência” precisa ser observada de forma composta para alcançar sentido. Serão, portanto, considerados quatro

termos compostos envolvendo o termo abrange; “abrange instrução, abrange quiz, abrange remotamente e abrange acessibilidade”. Outro resultado observado é de que no universo das patentes sobre EAD, os estudantes são considerados usuários dos sistemas, o que demonstra o viés técnico do vocabulário usado, em detrimento da natureza educacional, que normalmente usaria em seu vocabulário a palavra estudante. Nesse caso, como o sub-contexto 1 é a parte mais significativa (20 palavras) do contexto 1, nomeou-se o contexto 1: PROCESSOS DE ENSINO E ACESSIBILIDADE AO USUÁRIO.

Figura 80 – Análise do Sub-contexto 2, frequência de ocorrência mínima 25 e visualização do tipo propagação do label

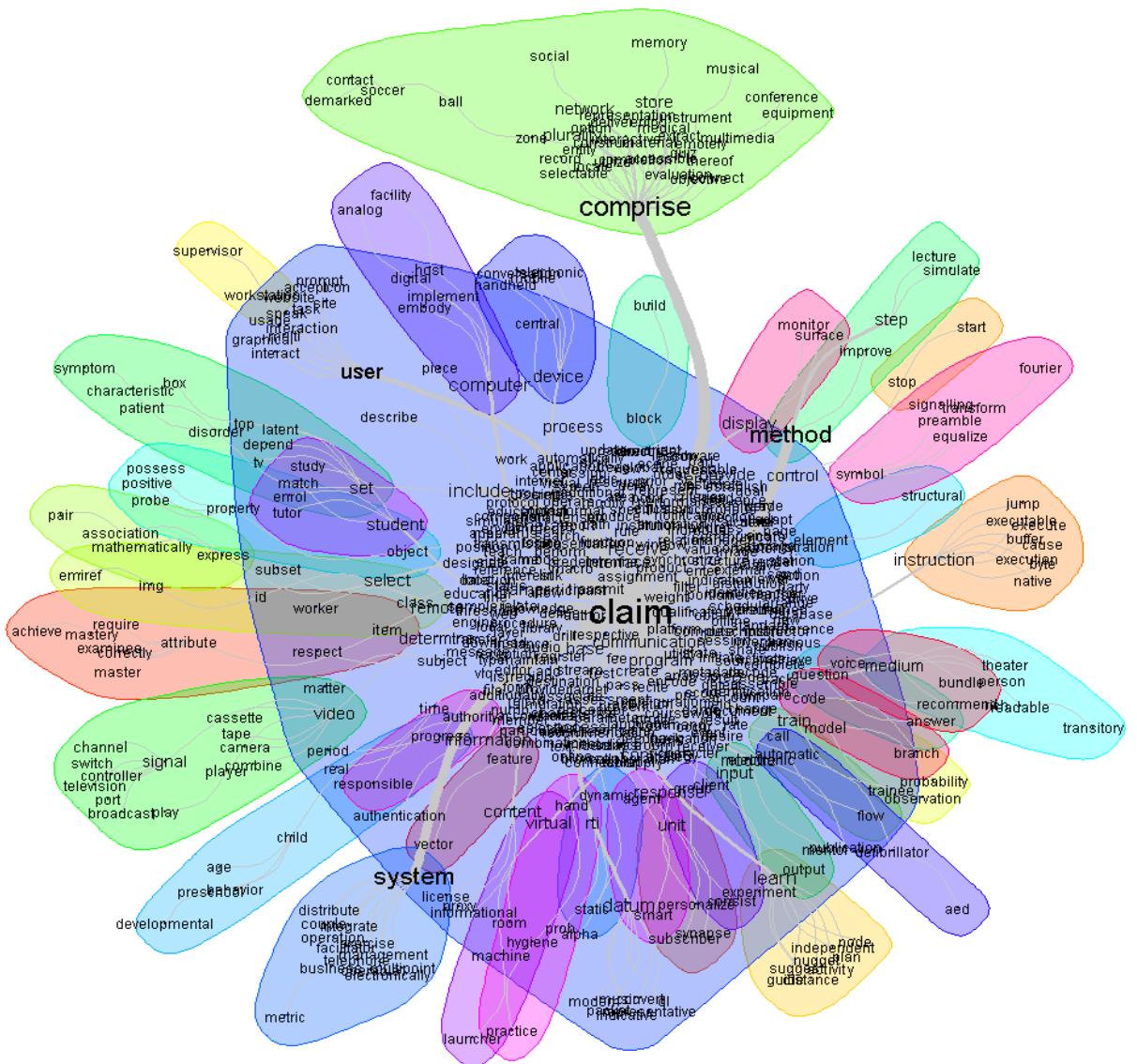


As nuvens formadas ao redor de cada forma na Figura 80, mostram o nível da propagação da forma de maior frequência, considerando que selecionou-se para apresentar apenas as formas com no mínimo 25 ocorrências. Identificou-se, portanto, no sub-contexto 2 a predominância das patentes que tratam dos detalhes sobre a transmissão dos conteúdos aos usuários, incluindo o tipo de sinal (digital, analógico e seus conversores), canais de transmissão, meios de apresentação de conteúdo (vídeo e áudio), meios de armazenamento e disponibilização de conteúdo (cassete e *tape*), além das reivindicações a respeito dos dispositivos eletrônicos. Esse contexto parece alocar as inovações em produto, mais do que em serviços por exemplo, dada a alta frequência das reivindicações focadas em elementos eletroeletrônicos e de telecomunicações. A forma predominante é a “sinal e unidade”. Nesse caso, como o sub-contexto 2 é a parte mais significativa (20 palavras) do contexto 2, nomeou-se o contexto 2: UNIDADES ELETRÔNICAS E DE TELECOMUNICAÇÕES.

4.2.7 Reivindicações das patentes por similaridade

A análise de similaridade leva em consideração a co-ocorrência. Co-ocorrência, nesse estudo, caracteriza-se pela frequência maior que um para determinada forma. Sempre que uma forma, tem frequências maiores que um, a forma é agrupada e por conveniência de apresentação foi escolhida a técnica da comunidade multi-nível. As nuvens nesse caso, representam o agrupamento das formas com maiores frequências. Manteve-se o mesmo parâmetro de frequência mínima, escolhido por conveniência (25 ocorrências) como limite dos nós. Pode-se observar que os principais *clusters* são os de sistemas, métodos, usuários e de abrangência. A predominância dos depósitos de patentes sobre EAD, tratam, portanto, a respeito das reivindicações sobre a abrangência dos sistemas de ensino envolvendo seus componentes e os elementos eletrônicos de telecomunicações. Diferentemente das descrições da invenção e dos resumos das patentes, percebe-se nas reivindicações o interesse em proteção patentária relacionada a equipamentos para operacionalização envolvendo o processamento, armazenamento, transmissão e recepção do conteúdo, bem como nos métodos e formas de provocar a aprendizagem, evidenciado principalmente pela forma predominante “*abranger-comprise*”. A Figura 81 apresenta uma rede de formas obtidas por meio da análise de conteúdo das reivindicações presentes nas patentes sobre EAD, usando a técnica da similaridade.

Figura 81 – Análise de similaridade de formas ativas de 322 reivindicações presentes nas patentes sobre EAD



Após realizar as análises de conteúdo das reivindicações das patentes sobre EAD, pode-se apresentar os resultados sobre o conteúdo das mesmas. As separações dos contextos léxicos, nos trazem a visão geral da dinâmica dos depósitos de patentes, revelando que sobre o tema EAD, do ponto de vistas das reivindicações, há dois contextos: *Cluster 1 = Contexto 1* e *Cluster 2 = Contexto 2*. Conforme definido, contextos são os agrupamentos que se foram pelas formas que apresentam proximidade léxica, que em termos práticos, significam que pelo tratamento dos algoritmos de Reinert (1990), ficam separados em contextos diferentes os

conteúdos das reivindicações das patentes que estão tratando de coisas diferentes, resolvendo assim os efeitos polissêmicos. As decomposições em sub-contextos, permite nomear cada contexto, bem como revelar as 20 principais formas que compõe cada contexto. Tal detalhamento é especialmente importante numa mão, para nomear o contexto e em outra mão para expandir o *thesaurus* sobre o tema EAD. As análises AFC, não foram geradas pelo software Iramuteq para as reivindicações devido a existência de apenas duas dimensões de análise, ou dois contextos como detalhados anteriormente. A visão final, considerando todas as reivindicações recuperadas, 322 reivindicações de uma amostra de 3090 patentes recuperadas, foi apresentada pela rede de similaridade das formas, por meio das técnicas de agrupamento por similaridade ou co-ocorrência, definida anteriormente. No sentido de apresentar as análises realizadas de forma sintética, escolheu-se a matriz de consolidação, a seguir, na Tabela 11.

Tabela 11 – Síntese de análise de conteúdo dos resumos de patentes sobre EAD

Síntese de análise de conteúdo

Número de conjuntos de reivindicações de patentes sobre EAD tratados: 322

		Distribuição do Conteúdo*	Nome do Contexto	Conteúdo Predominante**	País Predominante***	IPC Predominante
Contexto1	82.30%		Processos de Ensino e Acesso do Usuário	usuário, aprender, reivindicar, método, conteúdo, estudante, aprendiz, base, virtual e inclui	Não se Aplica	Não se Aplica
Contexto2	17.70%		Unidades Eletrônicas e de Telecomunicações	sinal, unidade, assinante, porta, controle, canal, digital, modem, saída e vídeo	Não se Aplica	Não se Aplica

* Há sobreposição, portanto, a soma da distribuição pode ser diferente de 100%

** Aqui as 10 principais formas, das 20 analisadas nos sub contextos

*** As reivindicações não apresentaram referência a países nessa amostra

4.2.8 Consolidação das análises de conteúdo das patentes sobre EAD

Conforme descrito na seção sobre a expressão de busca das patentes sobre EAD, recuperou-se 3.090 patentes para análise. Dessa amostra, foram recuperados 2.870 resumos, 340 descrições de invenções e 322 reivindicações. Detalhou-se anteriormente a natureza distinta entre cada um desses *corpus* de texto, sendo os resumos, as apresentações gerais sobre a invenção, as descrições as apresentações detalhadas sobre a invenção e as reivindicações as

solicitações específicas sobre o que se pretende proteger por meio da patente, incluindo vários itens, objetivamente circunscritos em termos de abrangência e características. Quando consolidou-se os resultados de cada análise, verifica-se que há distinção entre elas, a respeito do conteúdo predominante. Considerando a natureza dos textos, pode-se interpretar que a natureza das descrições das invenções apresentam, na amostra, mais detalhes sobre a invenção do que o resumo e do que a reivindicação. É na descrição da invenção que o depositante explica como o objeto da invenção irá funcionar e quais suas características principais e diferenciadoras. Percebe-se, considerando esse pressuposto, que o “usuário”, a “personalização e os modelos de interface” e os “Sistemas de Disponibilização” do EAD, representam as principais funções e aplicações tecnológicas pretendidas pelos depositantes. Percebe-se ainda que a estratégia usada para alcançar a patente e, portanto, o título de propriedade sobre tais tecnologias são os métodos de ensino.

Os depositantes focam no método de ensino e nos equipamentos envolvidos na disponibilização do EAD para solicitar suas patentes. Pode-se fundamentar o argumento, observando as reivindicações presentes nessas patentes analisadas. Observa-se que a abrangência de proteção patentária pretendida, especifica “processos de ensino, processos de acesso para o usuário” e as unidades eletrônicas de processamento, armazenagem, transmissão e recepção dos conteúdos EAD. Deve-se destacar, que a concentração das reivindicações, encontram-se em “processos de ensino e acesso do usuário” (82,30%). Isso corrobora com o contexto¹ das descrições da invenção (Usuário). Um dos resultados não esperados é a inversão dos pesos percentuais dos conteúdos verificados entre os resumos e as reivindicações. O peso dos “processos de ensino” e das “unidades eletrônicas” nas reivindicações das patentes é de 82,3% e 17,70%, respectivamente. Quando observa-se os resumos das patentes, pode-se observar a ênfase aos sistemas (62,30%) e menor ênfase ao “método e virtualidade” (26,8%). Tal efeito pode ser reflexo da diferença no tamanho das amostras conseguidas (870 Resumos x 322 Reivindicações), mas entende-se que se trata de algo a ser aprofundado em outros estudos. A matriz consolidada pode ser vista na Tabela 12.

Tabela 12 – Consolidação das análises de conteúdo das patentes de EAD: Resumo, Descrições e Reivindicações

Consolidação da análise de conteúdo das Patentes de EAD

	Distribuição do Conteúdo*	Nome do Contexto	Conteúdo Predominante**	País Predominante	IPC Predominante
Análise de Conteúdo dos Resumos das Patentes EAD	Contexto1 62.30%	Sistemas e Aprendizagem	servidor, conteúdo, aprender, estudante, usuário, curso, informação, bando de dados e aprendiz	Korea	CIB4
	Contexto2 26.80%	Método e Virtualidade	treino, virtual, melhorar, instrução, alvo, resolver, característica, alto, eficiência e reduz	Russia	CIB4
	Contexto 3 21%	Controle e Conexão	sinal, utilidade, controle, arranjo, conectar, força, camera, circuito, controlador e corpo	China	CIB1
Análise de Conteúdo das Descrições das Invenções nas Patentes EAD	Contexto1 42.60%	Usuário	curso, usuário, aprender, atividade, estudante, aprendiz, educacional, conteúdo, material e instrutor	Não se Aplica	CIB1
	Contexto2 26.40%	Personalização e Modelos	modelo, simulação, conhecimento, feedback, entendimento, projetista, reivindicação, escopo e desenhando	Não se Aplica	CIB4
	Contexto 3 31.10%	Sistemas e Dispositivos	dispositivo, instrução, computador, servidor, sinal, rede, remoto, memória, processador e conectar	Não se Aplica	CIB1 e CIB3
Análise de Conteúdo das Reivindicações presentes das Patentes EAD	Contexto1 82.30%	Processos de Ensino e Acesso do Usuário	usuário, aprender, reivindicar, método, conteúdo, estudante, aprendiz, base, virtual e inclui	Não se Aplica	Não se Aplica
	Contexto2 17.70%	Unidades Eletrônicas e de Telecomunicações	sinal, unidade, assinante, porta, controle, canal, digital, modem, saída e vídeo	Não se Aplica	Não se Aplica

* Há sobreposição, portanto, a soma da distribuição pode ser diferente de 100%

** Aqui as 10 principais formas, das 20 analisadas nos sub contextos

*** As reivindicações não apresentaram referência a países nessa amostra

4.2.9 Patentes EAD por análise de agrupamento: *clustering*

A análise de agrupamento leva em consideração a contagem de palavras em cada documento e seu peso relativo dentro de tal documento. A comparação dos pesos relativos entre vários documentos coloca-os em grupos, cujos pesos são parecidos, nesse caso, formando um grupo ou *cluster*. A relevância dessa análise pode ser explicada por dois motivos principais,

sendo o primeiro o fato de analisar documentos distintos quanto a origem geográfica, depositante, tecnologia e data.

Considerando que se possui uma amostra de 3.090 patentes sobre EAD que foram depositadas por interessados distintos, em diversos países, em datas variadas e com descrições tecnológicas e reivindicações distintas, parece relevante identificar se há algum padrão nessa amostra. Identificar padrões ajuda na compreensão das informações das patentes recuperadas, oferecendo melhor entendimento do desdobramento tecnológico e da evolução de determinada técnica, permitindo em alguns casos desenvolver possíveis tendências. O segundo motivo que torna esse tipo de análise relevante é sua natureza de simplificação e organização das informações. Uma vez mais, deve-se lembrar que se tratam de 3.090 patentes. Portanto, ler todas, manualmente, não seria possível em tempo hábil. Ainda assim, alguma leitura será necessária, uma vez que se trata de estudo exploratório e de análise. Para escolher o que ler primeiro e até mesmo o que ler e o que descartar, a análise de *cluster* torna-se especialmente útil.

Considerando que as patentes recuperadas sobre o tema EAD podem ser agrupadas por afinidade de termos, peso relativo das palavras similares entre documentos, pode-se ter uma classificação que separa as patentes por similaridades e nomeia automaticamente cada *cluster*, considerando o referido peso relativo descrito (CEINE, 2015). Evidentemente, os diferentes algoritmos usados nos cálculos dos pesos relativos oferecem diferentes resultados e desempenhos que são explicados a seguir, no entanto, detalhando com algum detalhe a relação entre o Patent2Net v2 e sua saída para Carrot2.

O Carrot2 é um software *open source* que aplica algoritmos a conteúdos Web ou XML com objetivo de separar esses conteúdos em grupos ou *clusters*. O Carrot2 permite a “clusterização” de acordo com o algoritmo escolhido, dos quais destaca-se o *lingo* e o *k-means*. Em termos gerais, cada algoritmo apresenta pontos fortes e fracos, sendo o Lingo mais granular, portanto, com maior chance e visualizar *outliers* (importante para identificar possíveis inovações ou inovações em estágios iniciais). Por outro lado, gera uma quantidade maior de grupos que o *K-means*. O *K-means* por sua vez condensa mais os grupos, oferecendo menos grupos, com nomenclatura mais precisa do que com o *Lingo*.

Na Figura 82, apresenta-se os resultados das análises de *cluster*, sobre os resumos das patentes. Tais análises foram realizadas sobre os resumos das patentes, uma vez que o Patent2Net v2 oferece tal saída e o mesma conveniência se deve a escolha do Carrot2 como software de análise de *cluster*. Nesse caso, a saída do Patent2Net está no formato XML para Carrot2.

Figura 82 – Clusters dos resumos das patentes sobre EAD usando o algoritmo *Lingo* e a visualização *Foam tree* (88 clusters)

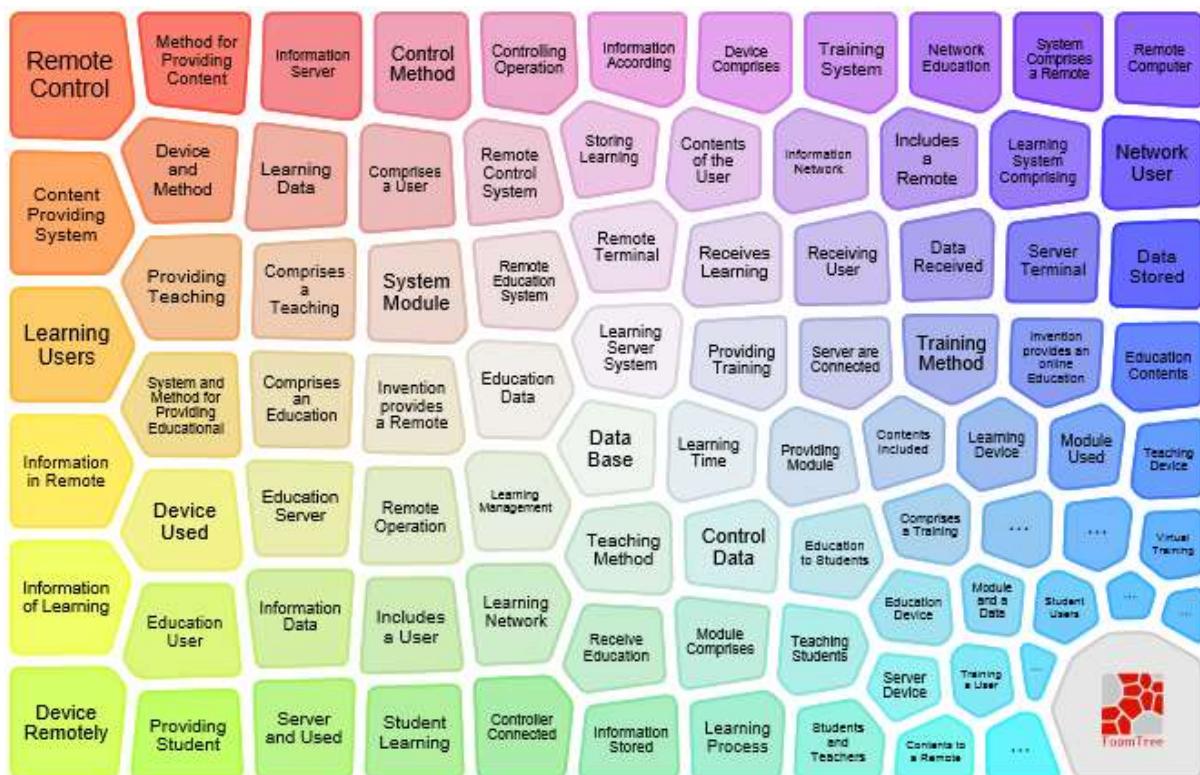


Figura 83 – Clusters dos resumos das patentes sobre EAD usando o algoritmo *K-means* e a visualização *Foam tree* (25 clusters)

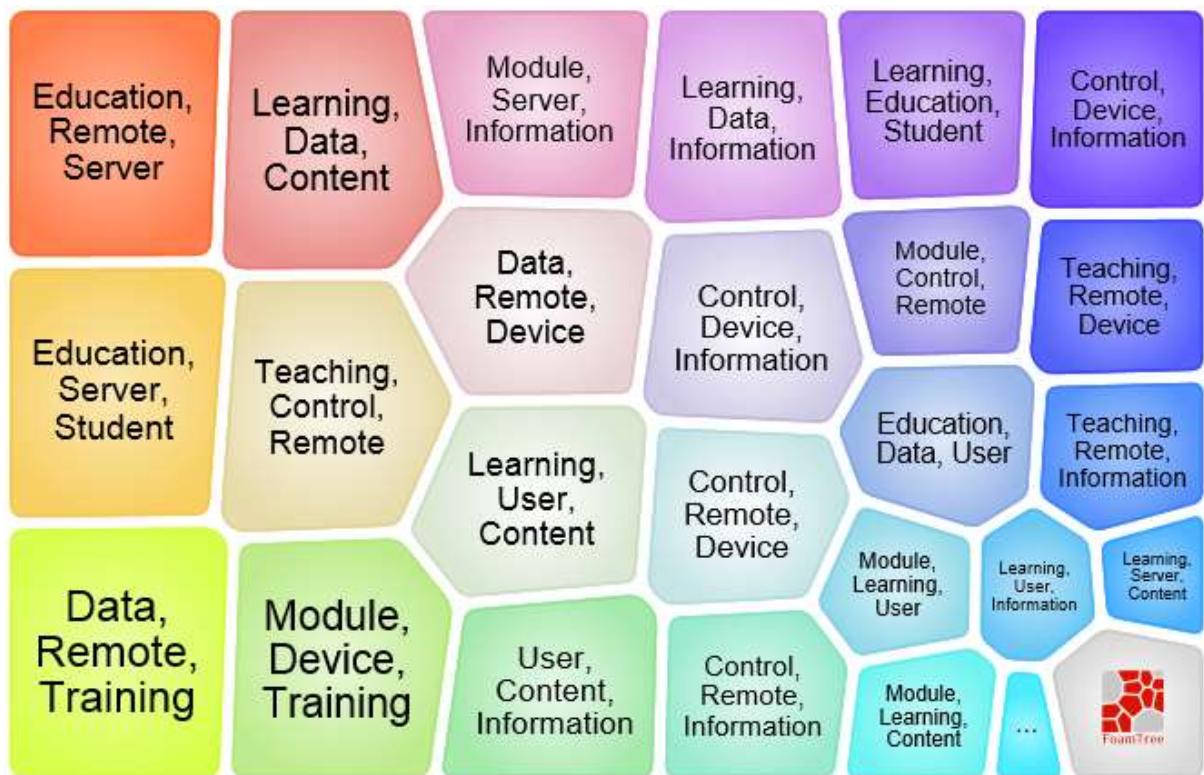
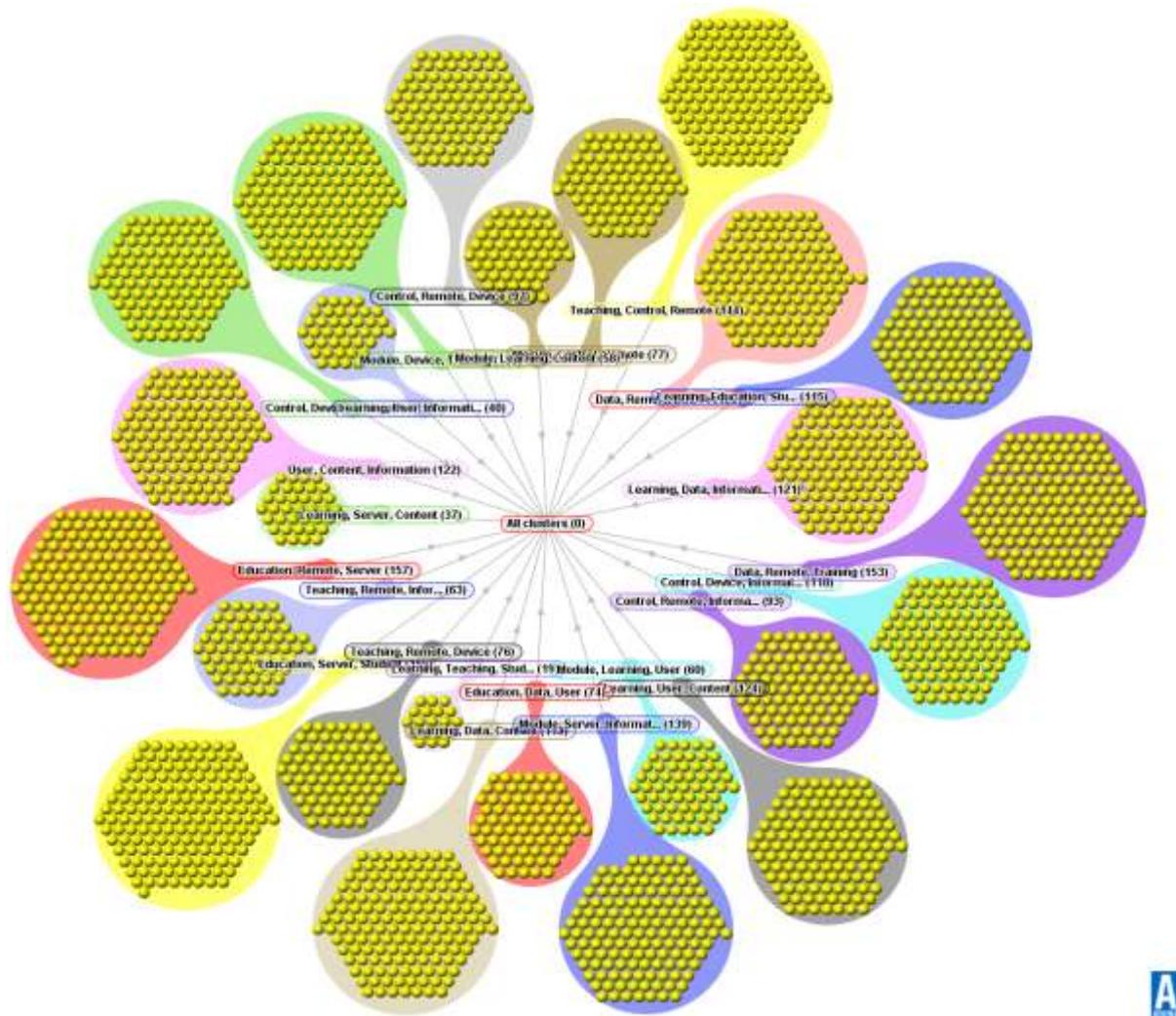


Figura 84 – Clusters dos resumos das patentes sobre EAD usando o algoritmo *K-means* e a visualização Aduna (25 Clusters)



4.3 MAPEAMENTO DAS PATENTES EM EAD EM UNIVERSIDADES

Na Figura 85 estão disponíveis as partes superior e inferior da tela que permite a avaliação dos resultados do cruzamento entre os indicadores “Applicant-Country”, que se refere ao país de origem da empresa que solicitou a patente, com o indicador “applicant”, que representa a empresa solicitante do patenteamento. Na referida Figura, levou-se em consideração apenas as patentes solicitadas por universidades, excluindo-se então aquelas registradas por pessoas físicas e por empresas que efetivamente não se dedicam ao ensino e a pesquisa. Das 3.090 patentes em EAD, 2.933 documentos foram solicitados por universidades. Destes, em 2.042 o país de origem da instituição não foi declarado. Dentre as restantes, observa-se a liderança das universidades chinesas, com 353 solicitações, seguidas das instituições coreanas, com 180 pedidos, e das americanas, com 123 solicitações. Uma análise completa do cruzamento possibilita apontar a “UNIV SHANGHAI JIAOTONG”, que depositou 113 patentes declarando a China como seu país de origem, e 52 documentos sem declarar nenhum país, totalizando 165 solicitações, seguida da “UNIV TSINGHUA”, que em 64 solicitações declarou a China como país de origem, o Japão em 48 documentos, e nenhum país em 35 solicitações, totalizando 147 pedidos, e a “UNIV HEBEI TECHNOLOGY”, que não declarou o país de origem em nenhum dos seus 126 pedidos de patenteamento.

Figura 85 – Parte superior e inferior da tela que demonstra a relação entre as universidades solicitantes de patentes em EAD e o país de origem dessas instituições

applicant	Applicant-Country	CA	CN	JP	KR	LV	TW	UA	US	Total	
ACADEMY OF APPLIED SCIENCES									4	4	
ALTSIM INC AND UNIVERSITY OF S									6	6	
HEZHOU UNIVERSITY		2								2	
INST OF COMM ENGINEERING PLA UNIVERSITY OF SCIENCE AND TECHNOLOGY		4								4	
JIANGSU OPEN UNIVERSITY		54								54	
KOREA NAT OPEN UNIVERSITY INDU						4				4	
MICHIGAN VIRTUAL UNIVERSITY		8							2	6	
NAT TAIPEI UNIVERSITY OF NURSING AND HEALTH SCIENCES							3			3	
NAT TAIPEI UNIVERSITY OFTECHNOLOGY							7			7	
NAT UNIV CHONBUK IND COOP FOUN						4				4	
UNIV ZHEJIANG		91								91	
UNIV ZHEJIANG GONGSHANG		38								38	
UNIV ZHEJIANG TECHNOLOGY		5								5	
UNIV ZHONGSHAN					3					3	
UNIVERSAL ELECTRONICS INC								1	1	1	
USTC UNIV SCIENCE TECH CHN		11								11	
WASEDA UNIVERSITY INTERNAT CO		6								6	
WESTERN MICHIGAN UNIVERSITY RES FOUNDATION								3	3	3	
XI AN JIAOTONG UNIVERSITY CITY COLLEGE		4							4	4	
ZHONGSHAN HOSPITAL FUDAN UNIV		16								16	
Total		114	36	353	60	100	21	100	18	123	2.933

Fonte: <http://200.136.214.89/~patent/DONNEES/Eadv4/Eadv4Pivot.html>

A Figura 86 demonstra parcialmente os resultados do cruzamento entre o indicador relativo à universidade solicitante do patenteamento e o país onde se buscou proteger a tecnologia descrita no documento (“country”). Das 2.933 solicitações de proteção requeridas por instituições de ensino, 2.157 foram feitas na China, seguidas daquelas solicitadas no Japão e nos Estados Unidos (143 pedidos de proteção em cada país), além daquelas solicitadas na Coreia (134 pedidos).

Figura 86 – Parte superior e inferior da tela que demonstra a relação entre as universidades solicitantes de patentes em EAD e o país onde a patente foi solicitada.

applicant	country	AU	CA	CN	EP	JP	KR	LV	TW	UA	US	WD	Total
ACADEMY OF APPLIED SCIENCES										4			4
ALTSIM INC AND UNIVERSITY OF S										6			6
HEZHOU UNIVERSITY						2							2
INST OF COMM ENGINEERING PLA UNIVERSITY OF SCIENCE AND TECHNOLOGY						4							4
JIANGSU OPEN UNIVERSITY				54								54	54
KOREA NAT OPEN UNIVERSITY INDU									4				4
MICHIGAN VIRTUAL UNIVERSITY											2		6
NAT TAIPEI UNIVERSITY OF NURSING AND HEALTH SCIENCES									3				3
NAT TAIPEI UNIVERSITY OF TECHNOLOGY									3				3
NAT UNIV CHONBUK IND COOP FOUR								4					4
UNIV ZHEJIANG				31								31	31
UNIV ZHEJIANG GONGSHANG				30								30	30
UNIV ZHEJIANG TECHNOLOGY				5									5
UNIV ZHONGSHAN				3									3
UNIVERSAL ELECTRONICS INC										1			1
USTC UNIV SCIENCE TECH CN				11									11
WASEDA UNIVERSITY INTERNAT CD								6					6
WESTERN MICHIGAN UNIVERSITY RES FOUNDATION											3		3
XI'AN JIAOTONG UNIVERSITY CITY COLLEGE				4									4
ZHONGSHAN HOSPITAL FUDAN UNIV				31								18	18
Totals		29	38	167	143	138	12	35	18	143	32	2.933	

Fonte: <http://200.136.214.89/~patent/DONNEES/Eadv4/Eadv4Pivot.html>

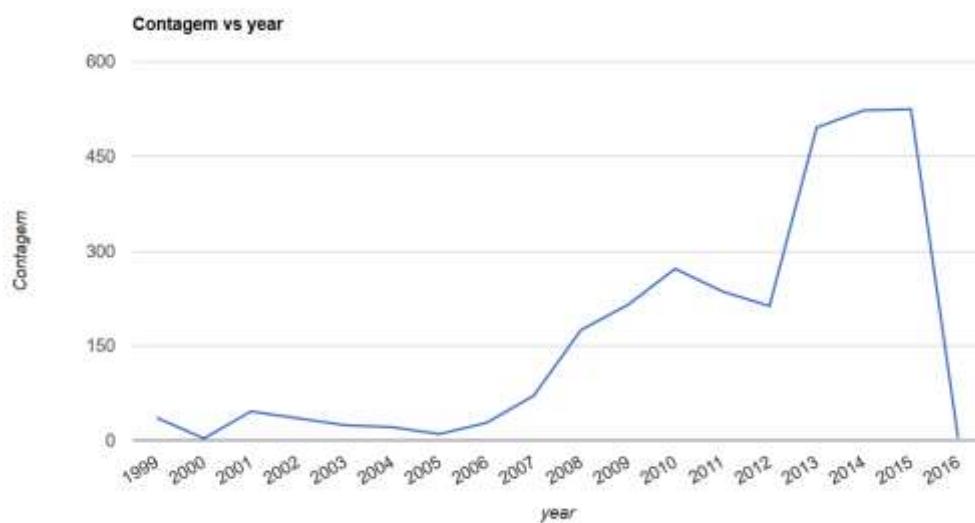
Os resultados parciais relativos ao cruzamento do indicador representativo da universidade solicitante da patente e a data de depósito do documento (“year”) são apresentados nas Figura 87 e Figura 88. Verifica-se que a solicitação de proteção de tecnologias voltadas ao EAD começou a crescer no final da década de 1990, com 36 depósitos em 1999, com grandes saltos em 2010 (272 solicitações), 2013 (493 pedidos) e 2015 (525 solicitações).

Figura 87 – Parte superior e inferior da tela que demonstra a relação entre as universidades solicitantes de patentes em EAD e o país onde a patente foi solicitada.

applicant	year	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	Total
ACADEMY OF APPLIED SCIENCES		4																	4	
ALTSIM INC AND UNIVERSITY OF S							6												6	
HEZHOU UNIVERSITY																		2		
INST OF COMM ENGINEERING PLA UNIVERSITY OF SCIENCE AND TECHNOLOGY																		4		
JIANGSU OPEN UNIVERSITY																		56		
KOREA NAT OPEN UNIVERSITY INDU																		8		
MINHWA VIRTUAL UNIVERSITY																		6		
NAT TAIPEI UNIVERSITY OF NURSING AND HEALTH SCIENCES																		3		
NAT TAIPEI UNIVERSITY OF TECHNOLOGY																		3		
NAT UNIV CHONBUK IND COOP FOUND																		6		
UNIV ZHEJIANG																	35			
UNIV ZHEJIANG GONGSHANG																	38			
UNIV ZHEJIANG TECHNOLOGY																	9			
UNIV ZHONGSHAN																	3			
UNIVERSAL ELECTRONICS INC																	1			
USTC UNIV SCIENCE TECH CHN																	11			
WALIEDA UNIVERSITY INTERNAT CO																	6			
WESTERN MICHIGAN UNIVERSITY HES FOUNDATION																	3			
XIA JIADONG UNIVERSITY CITY COLLEGE																	4			
ZHONGSHAN HOSPITAL FUDAN UNIV																	56			
Total:		36	3	48	25	24	21	19	28	21	129	219	222	236	213	4	2,333			

Fonte: <http://200.136.214.89/~patent/DONNEES/Eadv4/Eadv4Pivot.html>

Figura 88 – Gráfico de linhas representando a evolução temporal das solicitações de patenteamento de tecnologias voltadas ao EAD por universidades.

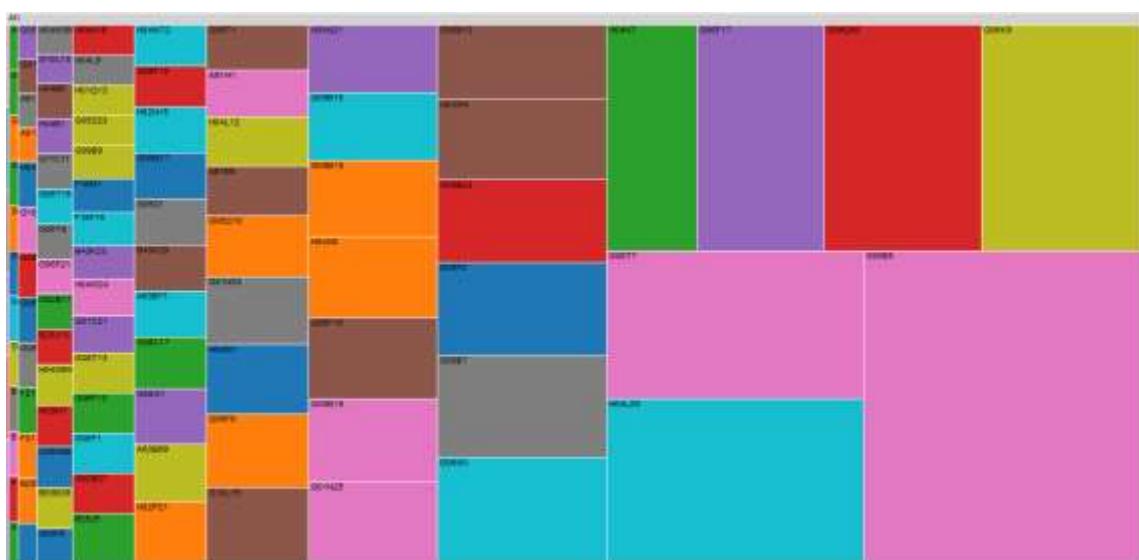


Fonte: <http://200.136.214.89/~patent/DONNEES/Eadv4/Eadv4Pivot.html>

A Figura 89 demonstra os resultados do cruzamento entre as patentes solicitadas pelas universidades e a tecnologia descrita no documento, representada pelo indicador “IPCR7”. Em

uma análise completa dos resultados é possível averiguar que, dos 2.033 pedidos, a tecnologia representada pelo código G09B5 (“Electrically-operated educational appliances”) foi a mais solicitada, com 419 pedidos, seguida de H04L29 (“Arrangements, apparatus, circuits or systems, not covered by a single one of groups”), com 204 pedidos, e G06T7 (“Image analysis, e.g. from bit-mapped to non bit-mapped”), com 182 depósitos.

Figura 89 – Gráfico de área representando a tecnologia em EAD com os maiores números de patentes solicitadas por universidades, com base no indicador IPCR7



Fonte: <http://200.136.214.89/~patent/DONNEES/Eadv4/Eadv4Pivot.html>

Já a Figura 90 ilustra parcialmente a tela com resultados do cruzamento entre as patentes solicitadas por universidades e o número de vezes que estes documentos foram citados por outros pedidos de patentes em EAD. Nela é possível verificar que as patentes mais citadas são da instituição “UNIV SHANDONG” e representam a tecnologia G09B5, com 51 citações. A instituição que mais possui solicitações de patentes com citações é a “UNIV SHANGHAI JIAOTONG”, com 165 menções. Por fim, a tecnologia com maior número de citações é aquela representada também pelo código G09B5, com 419 citações, seguida da H04L29, com 204 menções, e da G06T7, citada 182 vezes.

Figura 90 – Parte superior e inferior da tela que demonstra a relação entre as universidades solicitantes de patentes em EAD, o número de citações recebidas pelos documentos, e a tecnologia presente nas solicitações de patenteamento.

IPCRI	applicant	Cita	NaN	Totais
	UNIV UTAH RES FOUND		4	4
A01G9	UNIV CHONGQING ARTS & SCIENCES		3	3
A47B21	UNIV LONGDONG		2	2
A47B41	UNIV LONGDONG		2	2
A47C7	UNIV WEIFANG MEDICAL		1	1
A61B5	UNIV ZHANGZHOU NORMAL		24	24
A61H1	UNIV JIANGSU SCIENCE & TECH		15	15
	ZHONGSHAN HOSPITAL FUDAN UNIV		8	8
A61M25	UNIV BEIHANG		3	3
A63B69	NAT UNIV CHONBUK IND COOP FOUN		4	4
	UNIV HANGZHOU DIANZI		10	10
H04N21	UNIV IND & ACAD COLLABORATION		14	14
	UNIV INNER MONGOLIA		26	26
	UNIV XIAMEN		2	2
H04W4	SHAANXI RADIO AND TELEVISION UNIVERSITY		34	34
	UNIV EAST CHINA NORMAL		18	18
	UNIV INNER MONGOLIA		12	12
	UNIV SEJONG IND ACAD COOP GR		1	1
H04W24	UNIV DREXEL		9	9
	UNIV SOONGSIL RES CONSORTIUM		2	2
H04W36	USTC UNIV SCIENCE TECH CN		5	5
H04W64	UNIV INNER MONGOLIA SCI & TECH		5	5
	UNIV SOONGSIL RES CONSORTIUM		2	2
H04W72	UNIV DREXEL		9	9
	USTC UNIV SCIENCE TECH CN		5	5
H04W74	INST OF COMM ENGINEERING PLA UNIVERSITY OF SCIENCE AND TECHNOLOGY		4	4
		Totais	2.933	2.933

Fonte: <http://200.136.214.89/~patent/DONNEES/Eadv4/Eadv4Pivot.html>

4.4 USO DE INFORMAÇÕES PATENTÁRIAS NA DEFINIÇÃO DE ESTRATÉGIA EM IES À LUZ DO RBV

Os resultados encontrados nas diversas análises parecem mostrar a importância da inovação para a reputação das universidades e a relação com sistema de inovação do país (Fan, 2014). Independentemente de outros resultados de diferentes países e empresas, é a China e as universidades chinesas que se destacaram. As universidades chinesas se destacam e estão muito à frente das demais, principalmente nos anos recentes, o que coincide com a explosão de oferta de cursos EAD a partir de 2010. Dentre as universidades que se destacam, estão a Universidade Shanghai Jiaotong, a Universidade Tsinghua, a Universidade de Shandong e a Universidade Hebei de Ciência e Tecnologia. Com exceção desta última, que resultou da fusão entre diversas

universidades em 1996, as demais são universidades chinesas centenárias. Todas, sem exceção, investem fortemente na pesquisa e buscam estar entre as melhores do mundo. Vale ressaltar que são inúmeras áreas de pesquisa além, do EAD, que é foco deste trabalho.

A despeito da inferência que se pode fazer da importância do EAD para as próprias universidades, em função da dimensão da China, de sua da população e mesmo da dimensão geográfica do país, os resultados coincidem com outros estudos sobre inovação que mostram a pujança chinesa alcançada nos últimos anos. Preschitschek e Bresser (2010), investigando o campo de nanotecnologia a partir de dados de patentes, ressaltam a predominância chinesa em diversas áreas como tecnologia em química e medicina. Motohashi e Yuan (2009) apresenta resultados semelhantes, também num estudo sobre patentes chinesas.

Os resultados deste trabalho concordam com os de Preschitschek e Bresser (2010), que são as universidades chinesas, e não as empresas, que movem a inovação, ao contrário do que acontece eventualmente em países ocidentais. Fan (2014) argumenta, a partir de seu levantamento, que o sistema de inovação do governo chinês, após a reforma, foi fundamental para as inovações tecnológicas na China. O autor chega a afirmar a possibilidade de ameaçar o monopólio passado dos países ocidentais em tecnologia.

Um aspecto fundamental da concorrência entre empresas tem sido a disputa pelo conhecimento e pela inovação (McGrath, Tsai, Venkataraman, & MacMillan, 1996; Teece, 2000). Os rendimentos futuros das empresas está relacionado com a capacidade das empresas gerarem e gerenciarem as inovações (Ferreira, Serra, & Maccari, 2012). Os estudos em estratégia, que atestam a predominância da RBV como abordagem teórica, têm mostrado que os trabalhos estão integrados às áreas de conhecimento, aprendizagem organizacional e inovação. Uma das possíveis inferências, é que estas áreas de conhecimento possam “vias para a vantagem competitiva” (Serra et al., 2012, p. 264). A inovação é sem dúvida importante para as empresas (Serra et al., 2012), e também para as universidades. A base para que as empresas possam competir, depende da possibilidade de ter acesso a ativos que sejam próprios e difíceis de imitar (Amit & Schoemaker, 1993). Os resultados do trabalho e a discussão apresentada acima mostram que sem dúvida a capacidade de inovar tem a possibilidade de gerar vantagem competitiva para as universidades. Especificamente, foram estudadas patentes relacionadas

com EAD, que no entanto não refletem uma prioridade de pesquisa das universidades que se destacaram.

Ao considerar a necessidade de resguardar as inovações/desenvolvimentos e evitar imitações dos produtos/processos visando proteger o desempenho e rendimentos futuros, os mecanismos de proteção da inovação (McEvily, Eisenhardt, & Prescott, 2004; Teece, 1988, 1998, 2000), especificamente as patentes, têm sido consideradas importantes (Hurmelinna-Laukkanen & Puumalainen, 2007). As patentes têm como objetivo proteger o conhecimento que está embutido nas inovações (McEvily et al., 2004; McGrath et al., 1996), inclusive como mecanismo criador de barreiras que possam impedir a entrada de competidores potenciais (Cohen, Nelson, & Walsh, 2000; Zhang, 2010). As patentes permitem proteger os direitos de propriedade e prevenir a cópia pelos concorrentes (Deardorff, 1992).

As patentes também demonstram a capacidade de inovação, influenciando o valor da empresa no mercado (Blind, Edler, Frietsch, & Schmoch, 2006). Independentemente dos diversos benefícios ficou claro que as universidades que se destacaram nesse estudo, observou-se que a partir das inovações e da pesquisa – não só em EAD numa primeira observação no site dessas universidades, ganharem notoriedade internacional e se colocarem nos rankings entre as melhores universidades do mundo. A compreensão da origem e da alavancagem de recursos e capacidades é um tema importante e fundamental para a compreensão da vantagem competitiva a partir da RBV (Barney et al., 2011).

O estudo também alerta para a melhor compreensão e a continuação dos estudos de sistemas de inovação de países e das diferenças entre as filosofias de desenvolvimento de países asiáticos e ocidentais. Uma perspectiva interessante seria a colocada no livro *A Ordem Mundial*, de Kissinger (2014).

A própria capacidade da empresa gerar e colocar suas patentes é uma forma de apropriação de recursos e capacidades que podem gerar vantagem competitiva (Granstrand, 1999; Reitzig, 2004; Reitzig & Puranam, 2009). Pode-se inferir pelos dados, pela existência de um sistema chinês de inovação (Fan, 2014) e mesmo pela extensa produção de inovações medida a partir das patentes de universidades chinesas (Preschitschek & Bresser, 2010) que a própria capacidade de gerar patentes é importante para a organização do conhecimento e para o reconhecimento e a reputação das universidades chinesas.

Este trabalho contribui e estende outros trabalhos que investigaram patentes, na medida em que a partir da análise de um volume extenso de dados dos pedidos e patentes concedidas, reforça a importância da inovação como fonte de vantagem competitiva (Teece, 2000), e da sua ligação com os recursos e com as capacidades dinâmicas das organizações (Serra et al., 2012). Os resultados reforçam também o fato das patentes como medida para recursos e capacidades intangíveis, considerados uma fonte importante de vantagem competitiva (Hall, 1993).

O trabalho também traz uma contribuição prática para os elaboradores de políticas públicas ao mostrar a relevância de um sistema de inovação e de desenvolvimento a partir da educação e da pesquisa dos mais de 30 anos da reforma chinesa. Também contribui para mantenedores e gestores universitários ao mostrar o grande valor dado a pesquisa para a reputação e o reconhecimento das universidades.

Independentemente dos critérios fundamentais para a uma vantagem competitiva sustentável a partir de recursos VRIN (Barney, 1991), este estudo ressalta a importância das patentes e do seu volume como fator de inimitabilidade (Ferreira et al., 2014). Neste sentido, pela predominância de quatro universidades chinesas como detentoras de patentes e de pedidos de patente, vale um estudo mais profundo de compreender que mesmo as patentes possuindo limitações na proteção do conhecimento (Ferreira et al., 2014), uma quantidade grande e evolutiva de inovações poderia ser estuda e verificada a partir das patentes pedidas e concedidas. Isso permitirá confirmar a hipótese da existência de capacidades dinâmicas e até se este processo dinâmico de pedidos e registros de patentes poderia servir como barreira aos competidores pelo avanço daqueles que predominam.

5 CONCLUSÕES E RECOMENDAÇÕES

A educação superior na modalidade a distância busca novos modelos educacionais para que seus processos de mediações possam atender as pessoas em espaços e tempos diferentes, em contextos diversos, com a finalidade de dissociar a dicotomia espaço-tempo. Neste modelo professores e alunos podem estar em lugares e em tempos diferentes da instituição de ensino, o que exige o uso de Tecnologias de Informação e Comunicação (TICs) e recursos organizacionais especiais.

A EAD é um dos principais fatores da difusão e democratização da educação e revela-se como uma opção para a inclusão social e para a melhoria do processo educacional; face à limitação do sistema educativo convencional (presencial) de responder às demandas pleiteadas pela evolução da sociedade e dos novos processos midiáticos de comunicação. Além disso, a globalização gerou mudanças no cenário competitivo de organizações. Em especial, em instituições de ensino superior, ocorreu aumento do acesso a cursos superiores.

O acirramento da competição no ensino superior fez com que as organizações desse setor reavaliassem as bases de sustentação de suas vantagens competitivas. Tecnologia, internacionalização e inovações se tornaram as bases da competição das instituições de ensino superior. Isso faz com que se busque lentes teóricas que possam explicar o uso de certos recursos que as IES possuem como fonte de vantagem competitiva. Nesse sentido, optou-se pela *Resource-based view* (RBV) como uma das principais abordagens teóricas desta tese.

Para realizar este estudo, traçou-se os seguintes objetivos específicos: (a) apresentar o diagnóstico sobre as patentes relacionadas ao Ensino a Distância no mundo; (b) desenvolver uma metodologia, com o uso do KDD, para análise de conteúdo de informações patentárias em EAD para IES; (c) descrever o mapeamento das patentes em EAD em Universidades; e (d) avaliar como o uso de informações patentárias podem ser utilizadas na definição de estratégias de adoção de Ensino a Distância em IES, à luz do RBV. A seguir, apresenta-se as conclusões sobre estes objetivos.

5.1 CONCLUSÕES

Em relação ao primeiro objetivo - apresentar o diagnóstico sobre as patentes relacionadas ao EAD no mundo - as principais conclusões desse estudo foram:

- 1- Por meio de doze expressões de busca que utilizou um processo de lógica combinatória e agrupamento de 10 em 10 termos ocasionando em 52 palavras-chave, levantou-se um total de 3.090 patentes relacionadas a EAD no mundo;
- 2- A primeira patente sobre o tema EAD foi depositada em 1921 na Grã-Bretanha e a segunda patente depositada ocorreu somente em 1964. O número de depósitos aumentou significativamente somente nos últimos 15 anos (2000-2015);
- 3- Desde 1921 até 2015, a China se destaca como principal país solicitante com 850 depósitos, seguido dos EUA (704 depósitos) e Coréia do Sul (449 depósitos). Há forte incremento nos pedidos de patentes das empresas americanas e sul coreanas nos últimos 5 anos;
- 4- As três principais tecnologias citadas nas patentes, de acordo com a codificação *International Patent Classification (IPC)*, são: transmissão de informação digital (4.680 citações); mecanismos de seleção e *relays* (2.567 citações) e comunicação telefônica (1.331 citações) e
- 5- A maioria das organizações que investiram no desenvolvimento de tecnologias em EAD possuem país de origem como EUA, Taiwan e China; e os dois principais assuntos das patentes são: física e eletricidade.

Em relação ao segundo objetivo - desenvolver uma metodologia, com o uso do KDD, para análise de conteúdo de informações patentárias em EAD para IES - aponta-se as seguintes principais conclusões:

- 1- Foram recuperados 2.870 resumos de patentes, como resposta ao uso da expressão de busca construída. Ao filtrar-se os resultados, por meio de algoritmo de Classificação Hierárquica Descendente e atentando-se para os fenômenos de ruído e silêncio inerentes da busca, produzindo 3 *clusters* distintos que apresentam associação em termos de significado léxico:

- a. O *cluster* Sistemas e Aprendizagem contém 62% do conteúdo dos resumos das patentes e o país predominante é a Coréia. O *cluster* teve uma predominância da ocorrência da palavra *system* em um sub-contesto, *learn* em outro sub-contesto e *student* em um terceiro sub-contesto;
 - b. O *cluster* Método e Virtualidade contém 27% do conteúdo dos resumos das patentes e o país predominante é a Rússia. O *cluster* teve uma predominância da ocorrência da palavra *method* em um sub-contesto, *virtual* em outro sub-contesto e *time* em um terceiro sub-contesto;
 - c. O *cluster* Controle e Conexão contém 21% do conteúdo dos resumos das patentes e o país predominante é a China. O *cluster* teve uma predominância da ocorrência da palavra *control* em um sub-contesto, *connect* em outro sub-contesto e *utility* em um terceiro sub-contesto e
 - d. Conclui-se, por meio da análise dos resumos de patentes, que as principais tecnologias dominantes são as relacionadas com sistemas de aprendizagem e que as inovações menos predominantes são relacionadas com método e virtualidade, e conexão e controle; sendo esta última a melhor candidata para inovações radicais.
- 2- Foram recuperadas 340 descrições das invenções de patentes, como resposta ao uso da expressão de busca construída. Ao filtrar-se os resultados, por meio de algoritmo de Classificação Hierárquica Descendente e atentando-se para os fenômenos de ruído e silêncio inerentes da busca, produzindo 3 *clusters* distintos que apresentam associação em termos de significado léxico:
- a. O *cluster* Usuário contém 43% do conteúdo das invenções das patentes. O *cluster* teve uma predominância da ocorrência da palavra *user* em um sub-contesto, *student* em outro sub-contesto e *learn* em um terceiro sub-contesto;
 - b. O *cluster* Personalização e Modelos contém 27% do conteúdo das invenções das patentes. O *cluster* teve uma predominância da ocorrência da palavra *embodiment* em um sub-contesto, *model* em outro sub-contesto e *invention* em um terceiro sub-contesto;
 - c. O *cluster* Sistemas e Dispositivos contém 31% do conteúdo das invenções das patentes. O *cluster* teve uma predominância da ocorrência da palavra *system* em

- um sub-contesto, *device* em outro sub-contesto e *computer* em um terceiro sub-contesto e
- d. Conclui-se, por meio da análise das descrições das invenções de patentes, que as principais tecnologias dominantes são as relacionadas com usuários e que as inovações menos predominantes são relacionadas com personalização e modelos, e sistemas e dispositivos; sendo esta última a melhor candidata para inovações radicais.
- 3- Foram recuperadas 322 textos sobre reivindicações das patentes sobre EAD, como resposta ao uso da expressão de busca construída. Ao filtrar-se os resultados, por meio de algoritmo de Classificação Hierárquica Descendente e atentando-se para os fenômenos de ruído e silêncio inerentes da busca, produzindo 2 *clusters* distintos que apresentam associação em termos de significado léxico, produzindo 2 *clusters* distintos:
- a. O *cluster* Processos de Ensino e Acessibilidade ao Usuário contém 82% do conteúdo total das reivindicações analisados. O *cluster* teve uma predominância da ocorrência da palavra *comprise* em um sub-contesto, *method* em outro sub-contesto e *user* em um terceiro sub-contesto;
 - b. O *cluster* Unidades Eletrônicas e de Telecomunicações contém 18% do conteúdo total das reivindicações analisados. O *cluster* teve uma predominância da ocorrência da palavra *signal* em um sub-contesto, *network* em outro sub-contesto e *unit* em um terceiro sub-contesto; e
 - c. Conclui-se, por meio da análise das reivindicações das patentes, que as principais tecnologias dominantes são as relacionadas com processos de ensino e acessibilidade ao usuário e que as inovações menos predominantes são relacionadas com unidades eletrônicas e de telecomunicações; sendo esta última a melhor candidata para inovações radicais.
- 4- Ressalta-se que, geralmente ao analisar-se o *cluster* 1 (mais predominante), busca-se conhecer a tecnologia dominante e, ao analisar o *cluster* 2 e/ou *cluster* 3 com menor predominância, busca-se conhecer as inovações radicais.

Em relação ao terceiro objetivo - descrever o mapeamento das patentes em EAD em Universidades – arrola-se as seguintes principais conclusões

- 1- Das 3.090 patentes em EAD, 2.933 documentos foram solicitados por universidades. Apenas em 891 documentos, dentre os 2.933, o país de origem da instituição foi declarado;
- 2- Em primeiro lugar se encontram as instituições de ensino chinesas (353 solicitações), seguido das instituições coreanas (180 solicitações) e das americanas (123 solicitações);
- 3- A instituição que mais depositou patentes foi a “UNIV SHANGHAI JIAOTONG” com 165 patentes, seguido da “UNIV TSINGHUA” com 147 patentes, e, em terceiro lugar, a “UNIV HEBEI TECHNOLOGY” com 126 patentes;
- 4- A solicitação de patentes por instituições de ensino começou a crescer no final da década de 1990, com 36 depósitos em 1999, com grandes saltos em 2010 (272 solicitações), 2013 (493 pedidos), e 2015 (525 solicitações)

Em relação ao quarto objetivo - avaliar como o uso de informações patentárias podem ser utilizadas na definição de estratégias de adoção de Ensino a Distância em IES, à luz do RBV – elenca-se as seguintes principais conclusões:

- 1- As universidades chinesas se destacam e estão muito a frente das demais, principalmente nos anos recentes, o que coincide com a explosão de oferta de cursos EAD a partir de 2010. Todas, sem exceção, investem fortemente na pesquisa e buscam estar entre as melhores do mundo;
- 2- A importância do EAD para as próprias universidades chinesas, em função da dimensão da China, de sua população e mesmo da dimensão geográfica do país, coincide com outros estudos sobre inovação que mostram a pujança chinesa alcançada nos últimos anos;
- 3- São as universidades chinesas e não as empresas que movem a inovação no país ao contrário do que acontece eventualmente em países ocidentais;
- 4- A capacidade de inovar, por meio de patentes, tem a possibilidade de gerar vantagem competitiva e é importante para a organização do conhecimento e para o reconhecimento e para a reputação das universidades chinesas

Este estudo como todos possui limitações. A principal limitação, apesar da quantidade e qualidade dos dados, e da tentativa de algum tratamento estatístico, é o fato de ainda ser um tanto descritivo. Neste caso, esta limitação pode ser atenuada pela possibilidade de exploração

de uma grande quantidade de dados e de forma mais intensa, e com diversas instâncias de análise, aparentemente não encontrada nos estudos que serviram de base para discussão apresentados neste trabalho, como por exemplo os de patentes chinesas (Fan, 2014; Motohashi & Yuan, 2009; Preschitschek & Bresser, 2010).

5.2 RECOMENDAÇÕES

A partir das observações contidas na apresentação dos resultados, algumas recomendações tornam-se pertinentes, entre elas:

- (a) como pesquisa futura vale verificar, em se tratando de país com dimensão territorial e população extensa, porque estas universidades se destacaram em relação às demais chinesas.
- (b) vale como estudos futuros buscar compreender junto às universidades que se destacaram e também pelo exame das patentes, quais são estes desenvolvimentos e como estão sendo utilizados, se pela própria universidade ou para a comercialização de tecnologia.

REFERÊNCIAS

- ABED. (2013). *Censo EAD - Relatório Analítico da Aprendizagem a Distância no Brasil*. Brasil.
- Altbach, P. G., & Knight, J. (2007). The internationalization of higher education: Motivations and realities. *Journal of studies in international education*, 11(3-4), 290–305.
- Amit, R., & Schoemaker, P. J. (1993). Strategic assets and organizational rent. *Strategic management journal*, 14(1), 33–46.
- Barney, J. (1991). Firm Resources and Sustained Competitive Advantage. *Journal of Management*, 17(1), 99.
- Barney, J. B., & Hesterly, W. S. (2008). *Administração estratégica e vantagem competitiva*. Pearson Prentice Hall.
- Barney, J. B., Ketchen, D. J., & Wright, M. (2011). The Future of Resource-Based Theory Revitalization or Decline? *Journal of Management*, 37(5), 1299–1315. <http://doi.org/10.1177/0149206310391805>
- Barney, J., Wright, M., & Ketchen, D. J. (2001). The resource-based view of the firm: Ten years after 1991. *Journal of management*, 27(6), 625–641.
- Belloni, M. L. (2002). Dos audiovisuais à multimídia: análise histórica das diferentes dimensões de uso dos audiovisuais na escola. *A formação na sociedade do espetáculo*. São Paulo: Loyola, 47–72.
- Belloni, M. L. (2005). Educação a distância e inovação tecnológica. *Trabalho, Educação e Saúde, Rio de Janeiro*, 3(1), 187–198.
- Blind, K., Edler, J., Frietsch, R., & Schmoch, U. (2006). Motives to patent: Empirical evidence from Germany. *Research Policy*, 35(5), 655–672.
- Boyd, D., & Crawford, K. (2012). Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, communication & society*, 15(5),

- 662–679.
- Brasil. Lei de diretrizes e bases da educação nacional, Pub. L. No. 9.394/96 (1996).
- Buhl, H. U., Röglinger, M., Moser, F., Heidemann, J., & others. (2013). Big data. *Business & Information Systems Engineering*, 5(2), 65–69.
- Capper, J. (2001). E-learning growth and promise for the developing world. *TechKnowLogia*, 2(2), 7–10.
- CEINE. (2015). The KDD Process for Extracting Knowledge from Big-Data | CEINE - Centro de Investigación en Inteligencia de Negocios. Recuperado 17 de fevereiro de 2016, de <http://www.chine.cl/the-kdd-process-for-extracting-useful-knowledge-from-volumes-of-data/>
- Chen, H., Chiang, R. H., & Storey, V. C. (2012). Business Intelligence and Analytics: From Big Data to Big Impact. *MIS quarterly*, 36(4), 1165–1188.
- Clarke, G. (1997). Reassessing resource allocation strategies in higher education: methods for analysis. *International Journal of Educational Management*, 11(6), 286–292.
- Cohen, W. M., Nelson, R. R., & Walsh, J. P. (2000). *Protecting their intellectual assets: Appropriability conditions and why US manufacturing firms patent (or not)*. National Bureau of Economic Research. Recuperado de <http://www.nber.org/papers/w7552>
- Collis, D. J. (1994). Research Note: How Valuable are Organizational Capabilities? *Strategic Management Journal*, 15(S1), 143–152. <http://doi.org/10.1002/smj.4250150910>
- Crubellate, J. M., Pascucci, L., & Grave, P. S. (2008). Contribuições para uma visão baseada em recursos legítimos. *RAE-Revista de Administração de Empresas*, 48(4), 8–19.
- Cukier, K., & Mayer-Schoenberger, V. (2013). Rise of Big Data: How it's Changing the Way We Think about the World, The. *Foreign Aff.*, 92, 28.
- Davies, R. S., Howell, S. L., & Petrie, J. A. (2010). A review of trends in distance education scholarship at research universities in North America, 1998-2007. *The International*

- Review of Research in Open and Distributed Learning, 11(3), 42–56.*
- Dean, J., & Ghemawat, S. (2008). MapReduce: simplified data processing on large clusters. *Communications of the ACM, 51*(1), 107–113.
- Deardorff, A. V. (1992). Welfare effects of global patent protection. *Economica, 35*–51.
- Dierickx, I., & Cool, K. (1989). Asset stock accumulation and sustainability of competitive advantage. *Management science, 35*(12), 1504–1511.
- DiMaggio, P. J., & Powell, W. W. (1983). The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. *American sociological review, 147*–160.
- Docebo. (2014). *E-Learning Market Trends & Forecast 2014 - 2016 Report*.
- Dou, H. (2010). Information brevet 2.0, transfert de technologies et valorisation des ressources. *chapitre), Hermès Lavoisier publisher, Luc Quoniam editor, dans Competitive intelligence, 2.*
- Duarte, Z. M. C. (2011). Educação a Distância (EAD): Eestudo dos fatores críticos de sucesso na gestão de cursos da região metropolitana de Belo horizonte na visão dos Tutores. *Belo Horizonte: Fumec.* Recuperado de http://www.fumec.br/anexos/cursos/mestrado/dissertacoes/completa/zalina_maria.pdf
- Durham, E. R. (1990). A institucionalização da avaliação. *Núcleo, 8*, 90.
- Dyson, R. G. (2004). Strategic development and SWOT analysis at the University of Warwick. *European journal of operational research, 152*(3), 631–640.
- Eisenhardt, K. M., & Martin, J. A. (2000). Dynamic capabilities: what are they? *Strategic Management Journal, 21*(10-11), 1105–1121. [http://doi.org/10.1002/1097-0266\(200010/11\)21:10/11<1105::AID-SMJ133>3.0.CO;2-E](http://doi.org/10.1002/1097-0266(200010/11)21:10/11<1105::AID-SMJ133>3.0.CO;2-E)
- English Grammar - English Suffixes - Learn English. (2016). Recuperado 13 de fevereiro de 2016, de <http://www.learnenglish.de/grammar/suffixtext.html>

- Ésther, A. B., & Melo, M. C. de O. L. (2008). A Construção da Identidade Gerencial dos Gestores da Alta Administração de Universidades Federais em Minas Gerais. *Cadernos EBAPE.BR*, 0(1), 1–17.
- Estrutura das palavras. (2016). Recuperado 13 de fevereiro de 2016, de <http://portugues.uol.com.br/gramatica/estrutura-das-palavras-.html>
- Eyal, O., Berkovich, I., & Schwartz, T. (2011). Making the right choices: ethical judgments among educational leaders. *Journal of Educational Administration*, 49(4), 396–413.
- Fallows, S., & Steven, C. (2000). Building employability skills into the higher education curriculum: a university-wide initiative. *Education+ training*, 42(2), 75–83.
- Fan, P. (2014). Innovation in China. *Journal of Economic Surveys*, 28(4), 725–745.
- Fatimah, H., Gazi, M. A., & Saedah, S. (2010). Information and communication technology for participatory based decision-making-E-management for administrative efficiency in Higher Education. *International Journal of Physical Sciences*, 5(4), 383–392.
- Ferreira, M. P., Serra, F. R., & Maccari, E. A. (2012). When the innovator fails to capture rents from innovation. *Latin American Business Review*, 13(3), 199–217.
- Ferreira, M., Reis, N., Serra, F., & Costa, B. (2014). Understanding the Footprint of the RBV in International Business Studies: the Last Twenty Years of Research. *Brazilian Business Review*, 11(4), 53–83. <http://doi.org/10.15728/bbr.2014.11.4.3>
- Fiorillo, A., Maccari, E. A., & Martins, C. B. (2015). A EAD NO BRASIL E A IMPORTÂNCIA DE COMPETÊNCIAS SUBJACENTES PARA O RECONHECIMENTO DE CURSOS DE ACORDO COM A PERCEPÇÃO DE COORDENADORES DE CURSO/DL IN BRAZIL AND THE IMPORTANCE OF THE UNDERLYING FOR COURSE RECOGNITION ACCORDING TO COURSE COORDINATORS. *Administração: Ensino e Pesquisa*, 16(1), 141.
- Fujino, A., Stal, E., & Plonski, G. A. (1999). A proteção do conhecimento na universidade. *Revista de Administração da Universidade de São Paulo*, 34(4). Recuperado de

http://www.rausp.usp.br/busca/artigo.asp?num_artigo=87

Furrer, O., Thomas, H., & Goussevskaia, A. (2008). The structure and evolution of the strategic management field: A content analysis of 26 years of strategic management research. *International Journal of Management Reviews*, 10(1), 1–23. <http://doi.org/10.1111/j.1468-2370.2007.00217.x>

George, G., Haas, M. R., & Pentland, A. (2014). Big data and management. *Academy of Management Journal*, 57(2), 321–326.

Granstrand, O. (1999). *The economics and management of intellectual property*. Cheltenham, UK: Edward Elgar. Recuperado de <https://ideas.repec.org/b/elg/eebook/1651.html>

Greenhow, C., Robelia, B., & Hughes, J. E. (2009). Learning, teaching, and scholarship in a digital age Web 2.0 and classroom research: What path should we take now? *Educational researcher*, 38(4), 246–259.

Guri-Rosenblit, S. (2005). “Distance education”and “e-learning”: Not the same thing. *Higher education*, 49(4), 467–493.

Haase, H., de Araújo, E. C., & Dias, J. (2009). Inovações Vistas pelas Patentes: Exigências Frente às Novas Funções das Universidade. *Revista Brasileira de Inovação*, 4(2 jul/dez), 329–362.

Hall, R. (1993). A framework linking intangible resources and capabilities to sustainable competitive advantage. *Strategic management journal*, 14(8), 607–618.

Hara, N. (2000). Student distress in a web-based distance education course. *Information, Communication & Society*, 3(4), 557–579.

Helpat, C. E., & Peteraf, M. A. (2003). The dynamic resource-based view: Capability lifecycles. *Strategic management journal*, 24(10), 997–1010.

Hemsley-Brown, J., & Oplatka, I. (2006). Universities in a competitive global marketplace: A systematic review of the literature on higher education marketing. *International Journal of Public Sector Management*, 19(4), 316–338.

- Hirata, D., Kniess, C. T., Cortese, T. T. P., & Quoniam, L. (2015). O USO DE INFORMAÇÕES PATENTÁRIAS PARA A VALORIZAÇÃO DE RESÍDUOS INDUSTRIAIS: O CASO DO LODO DE TRATAMENTO DE ESGOTO DOMÉSTICO. *Revista de Ciências da Administração*, 1(1), 55–71.
- HOPER Educação. (2015). *Análise Setorial da Educação Superior Privada*. Brasil.
- Hoskisson, R. E., Hitt, M. A., Wan, W. P., & Yiu, D. (1999). Theory and research in strategic management: Swings of a pendulum. *Journal of management*, 25(3), 417–456.
- Hurmelinna-Laukkonen, P., & Puimalainen, K. (2007). Nature and dynamics of appropriability: strategies for appropriating returns on innovation. *R&d Management*, 37(2), 95–112.
- INPI. (2013). Portal INPI. Recuperado de <http://www.inpi.gov.br/portal/>
- INPI. (2015). Recuperado 13 de fevereiro de 2016, de <http://www.inpi.gov.br/home>
- IRaMuTeQ. (2016). Recuperado 13 de fevereiro de 2016, de <http://www.iramuteq.org/>
- Itacaramby Pardim, V., & Maccari, E. A. (2014). A educação on-line na pós-graduação stricto sensu: a experiência de um mestrado profissional em Gestão e Avaliação da Educação Pública. *RBPG. Revista Brasileira de Pós-Graduação*, 11(24). Recuperado de <http://search.ebscohost.com/login.aspx?direct=true&profile=ehost&scope=site&authtyp=e=crawler&jrnl=18068405&AN=102491269&h=7gqUTFjDFKoOAYxYW1lF1Qel6PHPhKQbj2YOIQFcKqPKbh1Gx9TpEsg4nMNx5Ej6qjfHG9EAe2YItSjU8z4ZrQ%3D%3D&crl=c>
- Ivy, J. (2001). Higher education institution image: a correspondence analysis approach. *International Journal of Educational Management*, 15(6), 276–282.
- Kerr, M. S., Ryneanson, K., & Kerr, M. C. (2006). Student characteristics for online learning success. *The Internet and Higher Education*, 9(2), 91–105.
- Kissinger, H. (2014). *A Ordem Mundial*. Leya.

- Laurillard, D. (2013). *Rethinking university teaching: A conversational framework for the effective use of learning technologies.* Routledge. Recuperado de <https://books.google.com.br/books?hl=pt-BR&lr=&id=KBc3T0aTM3UC&oi=fnd&pg=PP2&dq=Rethinking+university+teaching+A+conversational+framework+for+the+effective+use+of+learning+technologies&ots=2bTrLWnk2z&sig=Vtf5zTm8BCzqFNW2cCtJprJ2HF4>
- LaValle, S., Lesser, E., Shockley, R., Hopkins, M. S., & Kruschwitz, N. (2011). Big data, analytics and the path from insights to value. *MIT Sloan management review*, 52(2), 21.
- Lawrence, T. B., Winn, M. I., & Jennings, P. D. (2001). The Temporal Dynamics of Institutionalization. *Academy of Management Review*, 624–644.
- LearnThat Foundation. (2013, novembro 28). In *Wikipedia, the free encyclopedia*. Recuperado de https://en.wikipedia.org/w/index.php?title=LearnThat_Foundation&oldid=583683523
- Levy, Y. (2007). Comparing dropouts and persistence in e-learning courses. *Computers & education*, 48(2), 185–204.
- Liaw, S.-S. (2008). Investigating students' perceived satisfaction, behavioral intention, and effectiveness of e-learning: A case study of the Blackboard system. *Computers & Education*, 51(2), 864–873.
- Liaw, S.-S., Huang, H.-M., & Chen, G.-D. (2007). Surveying instructor and learner attitudes toward e-learning. *Computers & Education*, 49(4), 1066–1080.
- Maassen, P. A., & Potman, H. P. (1990). Strategic decision making in Higher Education. *Higher Education*, 20(4), 393–410.
- Maccari, E. A., & Quoniam, L. (2008). Inovação no Modelo de Ensino Na União Européia. Apresentado em XI SemeAd, São Paulo: FEA-USP.
- Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Byers, A. H. (2011). Big data: The next frontier for innovation, competition, and productivity. Recuperado de

<http://www.citeulike.org/group/18242/article/9341321>

March, J. G. (1991). Exploration and exploitation in organizational learning. *Organization science*, 2(1), 71–87.

Marginson, S. (2006). Dynamics of national and global competition in higher education. *Higher education*, 52(1), 1–39.

Mazzarol, T., & Soutar, G. N. (1999). Sustainable competitive advantage for educational institutions: a suggested model. *International Journal of Educational Management*, 13(6), 287–300.

McEvily, S. K., Eisenhardt, K. M., & Prescott, J. E. (2004). The global acquisition, leverage, and protection of technological competencies. *Strategic Management Journal*, 25(8-9), 713–722.

McGrath, R. G., Tsai, M.-H., Venkataraman, S., & MacMillan, I. C. (1996). Innovation, competitive advantage and rent: a model and test. *Management Science*, 42(3), 389–403.

Meyer, J. W., & Rowan, B. (1977). Institutionalized organizations: Formal structure as myth and ceremony. *American journal of sociology*, 340–363.

Meyer, J. W., & Rowan, B. (1983). *The structure of educational organizations*. London, New Delhi: Sage publications.

Meyer, J. W., & Scott, W. R. (1983). Centralization and the legitimacy problems of local government. *Organizational environments: Ritual and rationality*, 199–215.

Ministério da Educação. (2013). Censo do Ensino Superior - 2012.

Ministério da Educação. (2014). Plano Nacional de Educação.

Moore, M. G. (2013). *Handbook of distance education*. Routledge. Recuperado de https://books.google.com.br/books?hl=pt-BR&lr=&id=Up_fAmUHaXsC&oi=fnd&pg=PR3&dq=Handbook+of+Distance+Learning&ots=7dHKSpH3FP&sig=mewuzJ21CIGqf3Jao87z3yAekMo

- Moore, M. G., & Kearsley, G. (2011). *Distance education: A systems view of online learning*. Cengage Learning. Recuperado de <https://books.google.com.br/books?hl=pt-BR&lr=&id=dU8KAAAAQBAJ&oi=fnd&pg=PR4&dq=Distance+learning+moore&ots=D10fY1Gwky&sig=xGJ7A1RaB73veARSPXQPZAgp-c>
- Moore, M. G., Kearsley, G., Galman, R., & Mello, A. (2007). *Educação a distância: uma visão integrada*. Cengage Learning.
- Moran, J. M. (2007). *A educação que desejamos: novos desafios e como chegar lá*. Papirus Editora. Recuperado de <https://books.google.com.br/books?hl=pt-BR&lr=&id=PiZe8ahPcD8C&oi=fnd&pg=PA7&dq=A+educa%C3%A7%C3%A3o+que+desejamos+novos+desafio+e+como+chegar+l%C3%A1&ots=Bp22n62EIF&sig=mtzV9AIItTZNSc36SHaSS1PSr2Tg>
- Motohashi, K., & Yuan, Y. (2009). IT, R&D and productivity of Chinese manufacturing firms. In *RIETI Working Papers, 09-E-007*. Research Institute of Economy, Trade and Industry Tokyo. Recuperado de <http://www.rieti.go.jp/jp/publications/dp/09e007.pdf>
- OECD. (2002). *Frascati Manual: Proposed standard practice for surveys on research and experimental development*. Paris: Organisation for Economic Co-operation and Development (OECD).
- Oliver, C. (1991). Strategic responses to institutional processes. *Academy of management review, 16*(1), 145–179.
- Penrose, E. T. (1959). *The theory of the growth of the firm*.
- Preschitschek, N., & Bresser, D. (2010). Research Section. *Journal of Business Chemistry*, 7(1). Recuperado de https://www.researchgate.net/profile/Dominic_Bresser/publication/41204721_Research_Section_Nanotechnology_patenting_in_China_and_Germany_-a_comparison_of_patent_landscapes_by_bibliographic_analyses/links/54b2a0740cf28ebe92e2d823.pdf
- Puhlmann, A. C. A., & Moreira, C. F. (2004). Noções gerais sobre proteção de tecnologia e

- produtos: versão inventor. *São Paulo: Instituto de Pesquisas Tecnológicas*.
- Quoniam, L., Kniess, C. T., & Mazzieri, M. R. (2014). A patente como objeto de pesquisa em Ciências da Informação e Comunicação. *Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação*, 19(39), 243–268.
- Ramos-Rodríguez, A.-R., & Ruíz-Navarro, J. (2004). Changes in the intellectual structure of strategic management research: A bibliometric study of the Strategic Management Journal, 1980–2000. *Strategic Management Journal*, 25(10), 981–1004.
- Reinert, M. (1990). Une méthode de classification des énoncés d'un corpus présentée à l'aide d'une application. *Cahiers de l'analyse des données*, 15(1), 21–36.
- Reitzig, M. (2004). Strategic management of intellectual property. *MIT Sloan Management Review*, 45(3), 35.
- Reitzig, M., & Puranam, P. (2009). Value appropriation as an organizational capability: the case of IP protection through patents. *Strategic Management Journal*, 30(7), 765–789.
- Retamal, D. R. C., Behar, P. A., & Maçada, A. C. G. (2009). Elementos de Gestão para Educação a Distância: um estudo a partir dos Fatores Críticos de Sucesso e da Visão Baseada em Recursos. *RENOTE*, 7(1). Recuperado de <http://www.seer.ufrgs.br/renote/article/download/13974/7867>
- Roesler, J. (2008). Gestão da Educação Superior Online. In *Congresso internacional ABED de educação a distância* (Vol. 14).
- Salmon, G. (2005). Flying not flapping: a strategic framework for e-learning and pedagogical innovation in higher education institutions. *Research in Learning Technology*, 13(3). Recuperado de <http://www.researchinlearningtechnology.net/index.php/rlt/article/view/11218>
- Sampaio, R. M., & Laniado, R. N. (2009). Uma experiência de mudança da gestão universitária: o percurso ambivalente entre proposições e realizações. *Rev Adm Pública*, 4(1), 151–74.
- Santos, B. S. (1989). Da ideia de universidade à universidade de ideias. *Revista Crítica de*

Ciências Sociais, 27(28), 11–62.

- Sapienza, H. J., Autio, E., George, G., & Zahra, S. A. (2006). A capabilities perspective on the effects of early internationalization on firm survival and growth. *Academy of Management Review*, 31(4), 914–933.
- Sartori, A., & Roesler, J. (2005). *Educação Superior a Distância: gestão da aprendizagem e da produção de materiais didáticos impressos e on-line*. Ed. UNISUL.
- Schofer, E., & Meyer, J. W. (2005). The worldwide expansion of higher education in the twentieth century. *American sociological review*, 70(6), 898–920.
- Schwartzman, S. (1991). Educação básica no Brasil: a agenda da modernidade. *Estudos Avançados*, 5(13), 49–60.
- Scott, W. R. (1995). Institutions and organizations. Recuperado de <http://library.wur.nl/WebQuery/clc/924652>
- Secundo, G., Margherita, A., Elia, G., & Passiante, G. (2010). Intangible assets in higher education and research: mission, performance or both? *Journal of Intellectual Capital*, 11(2), 140–157.
- Selim, H. M. (2007). Critical success factors for e-learning acceptance: Confirmatory factor models. *Computers & Education*, 49(2), 396–413.
- Serra, F. R., Ferreira, M. P., de Almeida, M. I. R., & de Souza Vanz, S. A. (2012). A pesquisa em administração estratégica nos primeiros anos do século XXI: um estudo bibliométrico de citação e cocitação no Strategic Management Journal entre 2001 e 2007. *Revista Eletrônica de Estratégia & Negócios*, 5(2), p–257.
- Sharpe, R., Benfield, G., & Francis, R. (2006). Implementing a university e-learning strategy: levers for change within academic schools. *Research in Learning Technology*, 14(2). Recuperado de <http://researchinlearningtechnology.net/coaction/index.php/rlt/article/view/10952>
- Silva, C. A. da. (2003). Sociologia (apoio à docência): “Análise das Correspondências e

- Investiga&ccedaacute;o Social Aplicada". Recuperado 13 de fevereiro de 2016, de http://evunix.uevora.pt/~casilva/txt/cs_cor.html
- Soong, M. B., Chan, H. C., Chua, B. C., & Loh, K. F. (2001). Critical success factors for on-line course resources. *Computers & Education*, 36(2), 101–120.
- Statistics MRC. (2015). *Global E-Learning Market Outlook (2014-2022)*.
- Straumsheim, C. (2014, junho 3). U.S. releases data on distance education enrollments. Recuperado 13 de fevereiro de 2016, de <https://www.insidehighered.com/news/2014/06/03/us-releases-data-distance-education-enrollments>
- Suffix. (2016, fevereiro 7). In *Wikipedia, the free encyclopedia*. Recuperado de <https://en.wikipedia.org/w/index.php?title=Suffix&oldid=703750977>
- Suffix mania! Long list of suffixes, incl. meaning, word lists, examples,. (2016). Recuperado 13 de fevereiro de 2016, de <https://www.learnthat.org/pages/view/suffix.html#d>
- Sun, P.-C., Tsai, R. J., Finger, G., Chen, Y.-Y., & Yeh, D. (2008). What drives a successful e-Learning? An empirical investigation of the critical factors influencing learner satisfaction. *Computers & education*, 50(4), 1183–1202.
- Takahashi, A. R. W., & Fischer, A. L. (2009). Aprendizagem e competências organizacionais em instituições de educação tecnológica: estudos de casos. *Revista de Administra&ccdeil;ão da Universidade de São Paulo*, 44(4). Recuperado de <http://www.rausp.usp.br/principal.asp?artigo=1378>
- Teece, D. J. (1988). Capturing value from technological innovation: Integration, strategic partnering, and licensing decisions. *Interfaces*, 18(3), 46–61.
- Teece, D. J. (1998). Capturing value from knowledge assets: The new economy, markets for know-how, and intangible assets. *California management review*, 40(3), 55–79.
- Teece, D. J. (2000). Strategies for managing knowledge assets: the role of firm structure and industrial context. *Long range planning*, 33(1), 35–54.

- Teece, D. J., Pisano, G., & Shuen, A. (1997). Dynamic capabilities and strategic management. *Strategic Management Journal*, 18(7), 509–533. [http://doi.org/10.1002/\(SICI\)1097-0266\(199708\)18:7<509::AID-SMJ882>3.0.CO;2-Z](http://doi.org/10.1002/(SICI)1097-0266(199708)18:7<509::AID-SMJ882>3.0.CO;2-Z)
- Tolbert, P. S. (1985). Institutional environments and resource dependence: Sources of administrative structure in institutions of higher education. *Administrative Science Quarterly*, 1–13.
- Vaira, M. (2004). Globalization and higher education organizational change: A framework for analysis. *Higher education*, 48(4), 483–510.
- Van der Wende, M. C. (2003). Globalisation and access to higher education. *Journal of Studies in International Education*, 7(2), 193–206.
- Wernerfelt, B. (1984). A resource-based view of the firm. *Strategic Management Journal*, 5(2), 171–180. <http://doi.org/10.1002/smj.4250050207>
- What Are the Most Common Suffixes in English? (2016). Recuperado 13 de fevereiro de 2016, de <http://grammar.about.com/od/words/a/comsuffixes.htm>
- Wiederhold, G., Tessler, S., Gupta, A., & Smith, D. B. (2009). The valuation of technology-based intellectual property in offshoring decisions. *Communications of the Association for Information Systems*, 24(1), 31.
- Wilmore, A. (2014). IT strategy and decision-making: a comparison of four universities. *Journal of Higher Education Policy and Management*, 36(3), 279–292.
- Winter, S. G. (2003). Understanding dynamic capabilities. *Strategic Management Journal*, 24(10), 991–995. <http://doi.org/10.1002/smj.318>
- WIPO. (2013). WIPO - World Intellectual Property Organization. Recuperado de <http://www.wipo.int/portal/index.html.en>
- WIPO. (2016a). International Patent Classification. Recuperado 8 de março de 2016, de <http://www.wipo.int/classifications/ipc/en/>

- WIPO. (2016b). International Patent Classification Official Publication. Recuperado 8 de março de 2016, de <http://web2.wipo.int/classifications/ipc/ipcpub/#refresh=page>
- Zajac, E. J., & Kraatz, M. S. (1993). A diametric forces model of strategic change: Assessing the antecedents and consequences of restructuring in the higher education industry. *Strategic Management Journal*, 14(S1), 83–102.
- Zhang, Y. P. (2010). Around the block. *Nature Biotechnology*, 28(12), 1239–1241.