

# Online Analytical Processing em Ambientes Virtuais de Aprendizagem da Educação a Distância

Charles N. C. Freitas<sup>1</sup>, Roberta M. M. Gouveia<sup>2</sup>, Adenilton J. Silva<sup>3</sup>

<sup>123</sup>Departamento de Estatística e Informática – Universidade Federal Rural de Pernambuco (UFRPE)

Caixa Postal CEP: 52171-900 – Recife – PE – Brasil  
cnicollas21@hotmail.com, robertammg@gmail.com,  
adenilton.silva@gmail.com

**Abstract.** *This paper presents the Knowledge Discovery in Databases process (KDD) applied in the context of Distance Education (DE) and virtual learning environments. The goal is to apply the KDD process to profile of distance education in Brazil and make an accurate analysis of the courses conducted in e-learning environments as well as reveal the main challenges of the area. The methodology consists of collecting government open data-grained systems Universidade Aberta do Brasil (UAB) and e-MEC then apply the ETL process, organize and integrate the data into a data warehouse and OLAP apply the technologies in order to obtain one contextualised and realistic view of the distance mode nationally.*

**Resumo.** *Este artigo apresenta o processo Knowledge Discovery in Databases (KDD) aplicado no contexto da Educação a Distância (EaD) e dos ambientes virtuais de aprendizagem. O objetivo consiste em aplicar o processo KDD para traçar o perfil da EaD no Brasil e efetuar uma análise acurada dos cursos realizados em ambientes e-learning, bem como revelar os principais desafios da área. A metodologia consiste em coletar dados abertos governamentais de alta granularidade dos sistemas Universidade Aberta do Brasil (UAB) e e-MEC, em seguida aplicar o processo ETL, organizar e integrar os dados em um Data Warehouse e aplicar as tecnologias OLAP, visando obter uma visão contextualizada e realista da modalidade a distância no âmbito nacional.*

## 1. Introdução

Este artigo apresenta o processo de descoberta de conhecimento em banco de dados, conhecido pelo acrônimo KDD – *Knowledge Discovery in Databases*, aplicado no contexto da Educação a Distância (EaD) e dos ambientes virtuais de aprendizagem. O processo KDD é uma área de pesquisa que tem como principal objetivo o desenvolvimento de métodos do Aprendizado de Máquina, visando explorar conjuntos de dados coletados de diversos ambientes operacionais.

Conforme Lei nº 9.394 – Lei de Diretrizes e Bases da Educação (LDB), as universidades têm autonomia para definir o processo seletivo de ingresso de seus estudantes. Já a Lei nº 12.711/2012 garante reserva de 50% das matrículas por curso e turno nas IFES a estudantes oriundos do ensino médio público. Diante desse cenário, surgem novas vagas e formas de ingresso na educação superior, seja por meio do sistema de cotas, Exame Nacional do Ensino Médio (ENEM), Sistema de Seleção

Unificada (SISU), bolsas e financiamentos, tais como Fundo de Financiamento Estudantil (Fies), entre outras. Assim, estudantes que antes não eram admitidos pelos processos seletivos tradicionais, passaram a ter oportunidade de ingressar em uma universidade/faculdade. Em geral, ao matricular-se em uma IES o estudante é motivado pelo fato de vislumbrar melhores condições de vida após conclusão do curso, bem como obter sua independência e realização profissional. Porém, a aprovação e matrícula em uma IES não garantem que a motivação continue e que o estudante permaneça e conclua o curso. Atualmente com a forma de ingresso ampliada e mais simplificada, o estudante tende a mudar de curso com mais frequência, aumentando os índices de evasão, principalmente nos cursos da área de exatas, e mais especificamente os cursos de ciência da computação, sistemas de informação e afins. Portanto, um dos principais desafios da área de computação para os próximos 10 (dez) anos é compreender o perfil desse novo estudante, a fim de evitar que ele desista do curso, o que ocorre geralmente nos primeiros semestres.

O objetivo do trabalho consiste em aplicar o processo KDD para traçar o perfil da EaD, especificamente dos cursos realizados em ambientes *e-learning*, bem como realizar uma análise minuciosa da modalidade a distância, com vista à obtenção de melhorias na qualidade dos cursos ofertados pelas Instituições Federais de Ensino Superior (IFES), e de um maior estímulo e engajamento por parte dos estudantes.

A busca por uma educação além do limite espaço-tempo, que visa transformar e evoluir o processo tradicional de aprendizagem, é uma das propostas da EaD, afirma Moran (2014). Dessa forma, a motivação do trabalho surge do interesse em investigar e diagnosticar as deficiências e obstáculos enfrentados pelos gestores da EaD, docentes e discentes nos ambientes virtuais de aprendizagem. Como resultados dos estudos de caso realizados, obtém-se o respaldo científico necessário para detectar padrões e descobrir regras significativas na tentativa de minimizar os índices de evasão e retenção em cursos a distância, e como consequência adquirir um melhor entendimento acerca da EaD.

A EaD está em constante expansão, haja vista o aumento considerável do número de instituições adotando a modalidade em seus cursos de graduação e especializações, além do incentivo do governo e uma crescente adesão dos estudantes. A maioria das IFES do Brasil que ofertam cursos presenciais, também ofertam, ou está começando a ofertar, cursos na modalidade a distância. Embora tenha aumentado o número de instituições que aderiram à EaD, o Brasil ainda está em fase de transição nessa modalidade, visto que muitos docentes estão limitando-se a reproduzir para o ambiente virtual pequenas adaptações do ensino presencial. Essas práticas contribuem para aumentar os índices de evasão e retenção dos estudantes, visto que muitas vezes as aulas são disponibilizadas do ensino presencial para o virtual sem qualquer alteração didático-pedagógica nos processos de ensino-aprendizagem.

Para dar suporte ao processo KDD no cenário da EaD este trabalho emprega o processamento analítico online, mais conhecido como pelo acrônimo OLAP – *OnLine Analytical Processing*. A aplicação de tecnologias OLAP em bases de dados educacionais possui grande potencial de diagnóstico, e como consequência, deve refletir na melhoria dos processos de ensino-aprendizagem. Com intuito de identificar contribuições relevantes publicadas na área de KDD e OLAP aplicados em Banco de Dados Educacionais, foi realizado o levantamento bibliográfico atualizado, obtendo-se o estado da arte com ênfase nos métodos e aplicações que influenciam às novas metodologias de ensino.

O presente trabalho tem um papel relevante na análise de dados de alta granularidade da EaD. A ideia fundamental consiste em coletar dados abertos governamentais dos sistemas Universidade Aberta do Brasil (UAB) e e-MEC. Organizá-los em um *Data Warehouse* (DW) e aplicar as tecnologias OLAP, visando traçar o perfil da EaD no Brasil e obter uma visão contextualizada e realista da modalidade a distância no âmbito nacional. Este estudo revela quais são os principais desafios encontrados na EaD e os fatores relacionados à melhoria da qualidade dos cursos da EaD.

## 2. DW e OLAP

Um *Data Warehouse* é um tipo especial de banco de dados que se tornou conhecido e bastante utilizado a partir da década de 90. De acordo com Inmon (2005), o termo é definido como “um depósito de dados orientado por assunto, integrado, não volátil, variável com o tempo, para apoiar as decisões da gerência”. O DW é utilizado como parte do processo de descoberta de conhecimento na base de dados, facilitando os comandos e execuções OLAP, além de todo processo de Data Mining.

Nesse trabalho o DW foi implementado visando a integração e disponibilização dos dados da EaD. A partir do DW, aplicou-se as tecnologias OLAP para análise dos dados, por meio do Modelo Dimensional (MD). Essas tecnologias proporcionaram as agregações e sumarizações dos dados contidos no DW, gerando informações úteis ao processo decisório e oferecendo uma investigação detalhada sobre EaD.

O propósito do MD, também conhecido como Modelo Multidimensional, é mudar a forma como os dados são organizados em um banco de dados relacional, para um formato que facilite a exploração. O formato do MD, com tabelas fato e tabelas dimensão, com relacionamentos simples entre elas, requer um processo de ETL relativamente simples para armazenar histórico. Praticamente toda complexidade está concentrada na leitura dos dados, já que a saída, no modelo dimensional, obedece a regras claras e pré-estabelecidas.

Um MD é formado por um conjunto de tabelas especializadas: Tabelas Fato e Tabelas Dimensão. As Tabelas Fato guardam dados de eventos previamente desconhecidos. Chama-se fato, ou observação, o resultado da repetição de um evento. Métricas são os atributos previamente desconhecidos de uma observação, e que são medidos apenas no momento da observação do fato. As Tabelas Dimensão guardam dados previamente conhecidos que descrevem as observações ou fatos. Dimensão é um conjunto de atributos dimensionais correlacionados. Atributo dimensional é algo que caracteriza uma observação, e é previamente conhecido, como por exemplo, nome, região e organização acadêmica de uma instituição.

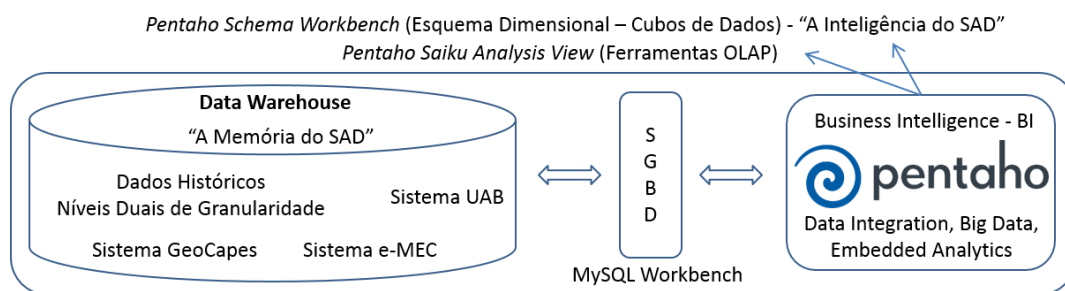
O processo KDD aplicado no contexto da EaD, especificamente a temática da Mineração de Dados Educacionais, é uma área de pesquisa científica em crescimento. Encontra-se trabalhos na literatura, tais como GOTTARDO et al. (2014); KAMPFF (2009); GUERCIO et al. (2014); RODRIGUES et al (2014); SANTANA et al. (2014); SILVA et al. (2014), publicados por pesquisadores interessados com a análise dos impacto e desafios da informática na educação. Tais trabalhos trazem reflexões fundamentadas abordando os desafios da informática na educação com perspectiva de futuro, com objetivo de potencializar os usos das tecnologias em prol do fortalecimento do processo de ensino e aprendizagem.

A modalidade de ensino a distância tem crescido bastante no âmbito nacional. Segundo dados do censo 2013 da EaD no Brasil existem mais de 300 instituições oferecendo curso a distância em todo país. Essas instituições correspondem a mais de

15 mil cursos e quase 4 milhões de alunos (ABED, 2014). Isso mostra que esta modalidade de ensino está cada vez mais acessível e crescente no país. Contudo, é preciso avaliar o desempenho desses cursos, para descobrir quais são os principais benefícios e quais são as principais dificuldades enfrentadas pelas instituições, cursos e estudantes. E a partir dessa investigação, estabelecer um melhoramento na qualidade de ensino da EaD no Brasil.

### 3. Metodologia

A metodologia proposta, proveniente da implementação de um Sistema de Apoio à Decisão (SAD) para EaD, consiste em um ambiente integrado projetado para facilitar as análises dos dados e contribuir com o processo de tomada de decisão. Conforme ilustra a Figura 1, o SAD implementado nesse trabalho é formado por três componentes básicos: os dados integrados no DW, o SGBD (*MySQL Workbench*), e as ferramentas de *Business Intelligence* (BI) – suíte do *software Pentaho*.



**Figura 1 - Componentes do ambiente de apoio à decisão implementado.**

No processo inicial de análise dos dados, utilizou-se a modelagem dimensional dos dados com a construção de um DW, para melhor manuseio e criação dos cubos de dados. Os dados foram obtidos a partir de dados abertos disponibilizados por repositórios governamentais relacionados ao Ministério da Educação (MEC).

O método utilizado para análise e descoberta de conhecimento iniciou-se com o processo ETL (*Extract, Transform and Load*) e prosseguiu com a implementação do DW, conforme etapas descritas a seguir:

- Criação do *Fact Constellation Schema*: consiste na modelagem multidimensional, com a construção das Tabelas de Fato e suas dimensões. O esquema constelação de fato favorece a otimização de consultas, fazendo com que os comandos da linguagem MDX tenham o mínimo de junções. Esse esquema auxilia o desempenho do DW devido à desnormalização das tabelas de dimensão. Nesse trabalho, fez-se necessário a modelagem multidimensional de dois esquemas do DW, um representando o sistema UAB e outro o e-MEC.
- Pré-processamento, Transformação, Seleção e Integração de Dados: consiste da limpeza e enriquecimento da base de dados. Serve para detectar os erros de cadastros e inconsistência dos dados extraídos dos ambientes operacionais do e-MEC e UAB. É realizada a limpeza e integração de dados heterogêneos, remoção de dados incompletos, ajustes de problemas de tipagem e eliminação de dados duplicados.
- Implementação do *Data Warehouse*: geração das tabelas de fato e dimensões no SGBD *MySQL Workbench*, e povoamento do DW com os dados dos sistemas UAB e e-MEC.

O estudo prossegue com a utilização de *Business Intelligence* para apoio à decisão, especificamente a aplicação de tecnologias OLAP, a fim de realizar uma

análise criteriosa e acurada dos dados armazenados no DW, visando traçar o perfil da EaD no Brasil, revelando quais os principais desafios da área, em especial de Computação, além de obter uma visão contextualizada e realista da modalidade a distância no âmbito nacional, com vistas à melhoria na qualidade dos cursos de EaD.

As tecnologias OLAP tornam-se viáveis em virtude das operações realizadas sobre os cubos de dados, bem como a utilização de funções de agregações, estatísticas e gráficas. Dentre as operações sobre os cubos de dados tem-se: *slice and dice*, *roll-up* (ou *drill-up*), *drill-down*, *drill-across*, *drill-through*, *pivot* etc. Essas operações permitem materializar diferentes perspectivas, também conhecidas como visões, além de consultas, simulações e análises interativas sobre os dados armazenados no DW.

Para construção, manipulação e análise dos cubos de dados OLAP, o presente trabalho fez uso de duas ferramentas da *Suíte Pentaho*: o *Pentaho Schema Workbench* e o *Pentaho BI Server*, mais especificamente o *plugin Saiku Analytics View*.

A ferramenta *Pentaho Schema Workbench* (PSW) é utilizada para criação dos cubos de dados (tabelas de fato, dimensões e métricas) do esquema multidimensional constelação de fato, sendo criados dois cubos de dados, um representando os dados da UAB e outro do e-MEC. Também se utiliza a ferramenta PSW para facilitar a criação do Esquema Mondrian (meta-modelo OLAP), usando interface gráfica.

O Esquema Mondrian é um arquivo XML, com estruturas que mapeiam tabelas de um banco de dados relacional em elementos de um modelo multidimensional, tais como métricas, níveis e membros. O PSW possui integração com *BI Server* e permite publicar o esquema desenvolvido diretamente nele.

Para execução das operações OLAP sobre os dois cubos de dados implementados, utiliza-se a ferramenta *Saiku Analytics View*, que por sua vez permite a visualização multidimensional e análise dos dados educacionais de cada um desses órgãos do governo federal, fornecendo ao analista informações detalhadas, sumarizadas e agregadas da área de estudo.

O modelo multidimensional desenvolvido para representar o sistema UAB possui 2 (duas) tabelas de fato e 8 (oito) tabelas de dimensão. As tabelas de fato são:

- *fato\_curso\_graduação*, contendo 2 (duas) métricas: “quantidade de instituições”, oriunda de “*dimensao\_dados\_ies*” e “Quantidade de cursos de graduação”, oriunda de “*dimensao\_dados\_curso\_graduação*”.
- *fato\_instituição\_ies*, contendo 1 (uma) métrica: “quantidade de instituições”, oriunda de “*dimensao\_dados\_ies*”.

A Tabela 1 apresenta o esquema matricial, mais conhecido como *Bus Matrix*, correspondente as 8 (oito) tabelas de dimensão e sua relação com a tabela de fato. Por exemplo, a dimensão “*dados\_curso\_graduação*” está associada apenas a uma tabela de fato (*curso\_graduação*), enquanto que as dimensões “*dados\_ies*” e “*localização*” estão associadas à ambas tabelas de fato (“*curso\_graduação*” e “*instituição\_ies*”).

O modelo multidimensional desenvolvido para representar o sistema e-MEC possui 3 (três) tabelas de fato e 14 (quatorze) tabelas de dimensão. As tabelas de fato são:

- *fato\_curso\_graduação*, contendo as métricas: “quantidade de cursos de graduação”, oriunda de “*dimensao\_dados\_curso\_graduação*”, e as métricas “Min, Max, AVG” do Conceito do Curso (CC), Conceito Preliminar de Curso (CPC) e Exame Nacional de Desempenho dos Estudantes (ENADE).
- *fato\_instituição\_ies*, contendo as métricas “quantidade de instituições”, oriunda de “*dimensao\_dados\_ies*”, e as métricas “Min, Max e AVG” do Conceito Institucional (CI) e Índice Geral de Cursos (IGC) da instituição.

- fato\_curso\_especialização, contendo a métrica “quantidade de instituições”, oriunda de dimensão\_dados\_ies.

**Tabela 1 – Bus Matrix do modelo multidimensional da UAB.**

		Tabelas Fato	
		curso_graduação	instituição_ies
Tabelas de Dimensão	dados_curso_graduação	x	
	modalidade_curso	x	
	area_cientifica	x	
	formação_professor	x	
	dados_ies	x	x
	localização	x	x
	organização_academica		x
	categoria_adm		x

A Tabela 2 apresenta o *Bus Matrix* correspondente as 14 (quatorze) tabelas de dimensão e sua relação com a tabela de fato.

**Tabela 2 – Bus Matrix do modelo multidimensional do e-MEC.**

		Tabelas Fato		
		curso_graduação	curso_especialização	instituição_ies
Tabelas de Dimensão	dados_curso_graduação	x		
	modalidade_curso	x		
	avaliação_curso_cc	x		
	avaliação_curso_cpc	x		
	avaliação_curso_enade	x		
	area_curso	x	x	
	dados_ies	x	x	x
	localização	x	x	x
	organização_academica			x
	categoria_adm			x
	dados_curso_especialização		x	
	carga_horaria		x	
	indice_ci			x
	indice_igc			x

Observa-se na Tabela 2, por exemplo, que a dimensão “organização\_academica” está associada apenas a uma tabela de fato (instituição\_ies), enquanto que as dimensões “dados\_ies” e “localização” estão associadas a três tabelas de fato (“curso\_graduação”, “curso\_especialização” e “instituição\_ies”).

#### 4. Resultados e Discussão

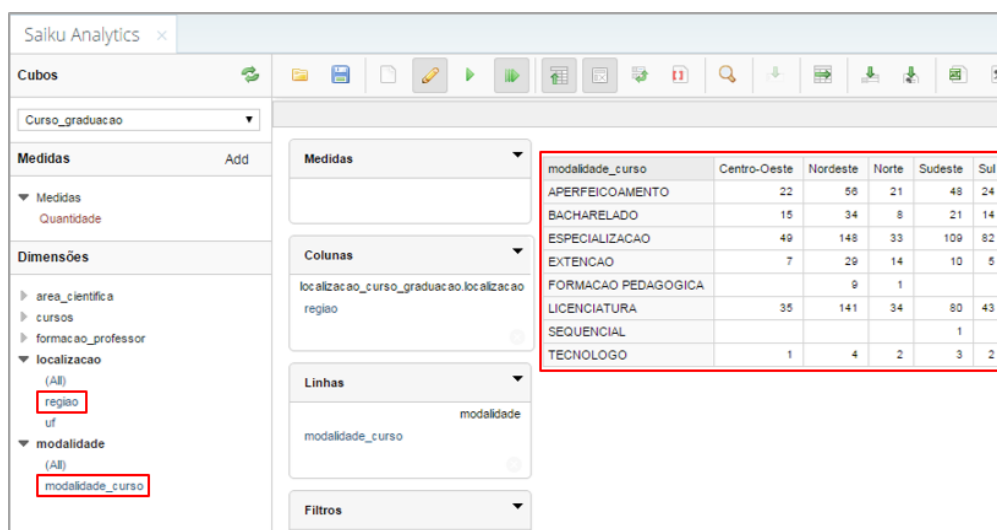
Há no cenário nacional uma crescente necessidade das instituições educacionais em implementar, bem como melhorar, suas práticas pedagógicas, em especial na EaD. Atualmente, busca-se um processo de ensino e aprendizagem não apenas presencial,

mas sim uma educação que ultrapasse o limite espaço-temporal. Isso deve-se ao fato da educação *online* flexibilizar os processos de ensino e aprendizagem, além de expandir as universidades. Por meio de plataformas *online*, que possibilitam um ensino mais interativo e autônomo, a EaD evolui e se destaca por seus formatos variados, sem afetar o aprendizado do estudante ou interferir na identidade e valores da instituição responsável pela oferta de cursos na modalidade a distância.

Uma avaliação inicial dos resultados obtidos e descritos no processo de KDD demonstra que a atuação da EaD no Brasil está se disseminando, impulsionada por programas governamentais que incentivam o acesso de alunos ao ensino superior na modalidade a distância. O crescimento da EaD nos últimos quatro anos foi de cerca de 16,43%, um aumento significativo em relação ao ensino presencial. O resultado desta análise tem o objeto norteador na avaliação da qualidade do ensino a distância no Brasil, bem como análise dos dados estatísticos referentes à expansão da EaD no âmbito nacional.

No ano de 2014 tinha-se uma média de 300 instituições ofertando cursos a distância no Brasil. Desse conjunto, 37,3% são universidades, 47,7% são faculdades, 14% são instituto federal de educação, ciência e tecnologia (IFET) e 1% são centros universitários. No cenário atual podemos observar que as faculdades representam a maioria das instituições, seguido das universidades. No contexto categoria administrativa, as faculdades privadas ainda são as instituições predominantes, contudo há um acréscimo considerável no número de instituições públicas (federais e estaduais) no âmbito da EaD.

Observa-se na Figura 3 a distribuição numérica dos cursos da EaD por região geográfica e grau de ensino (modalidade).



**Figura 3 – Quantidade de cursos, por grau de ensino, nas regiões geográficas do Brasil em 2014.** Fonte: Cubo de dados da UAB implementado no *Saiku Analytics View*.

A grande concentração está na região Nordeste (38,1%), enquanto que a menor concentração de cursos está no Norte (10,2%). Os graus de ensino mais ofertados são Especialização (421 cursos – 38%) e Licenciatura (333 cursos – 30%), conforme ilustra a Figura 3.

Na Figura 4 verifica-se a distribuição numérica de cursos de graduação por região geográfica e área científica. A área científica que possui mais cursos é Educação, com 40,1%, enquanto que a área “Ciência, Matemática e Computação” representa 23,7%.

Dos cursos pertencentes à área “Ciência, Matemática e Computação”, 9 (nove) são cursos de graduação específicos de Computação, são eles: Ciências da Computação, Computação, Informática, Sistemas de Informação, Tecnologia em Análise e Desenvolvimento de Sistemas, Tecnologia em Desenvolvimento de Sistemas de Informação, Tecnologia em Sistema para Internet, Tecnologia em Sistemas de Computação, e Tecnologia em Sistemas para Internet a Distância.

area_cientifica	Centro-Oeste	Nordeste	Norte	Sudeste	Sul	Grand Total
Agricultura e veterinaria		1	1			2
Ciencias sociais, negocios e direito	50	125	32	84	59	350
Ciencias, matematica e computacao	23	104	32	68	35	262
Educacao	45	178	42	108	71	444
Engenharia, producao e construcão				1		1
Humanidades e artes	6	9	3	6	2	26
Saude e bem estar social	5	4	3	4	3	19
Servicos				1		1
Grand Total	129	421	113	272	170	

**Figura 4 - Quantidade de cursos, por área científica, nas regiões geográficas do Brasil em 2014. Fonte: Cubo de dados da UAB implementado no Saiku Analytics View.**

Conforme mostra a Tabela 3, dos 26 cursos de graduação em Computação ofertados na modalidade a distância no Brasil, 50% estão na região Nordeste (13 cursos), 23% no Sudeste, 11,6% no Centro-oeste e Norte e 3,8 no Sul. O grau de ensino “Licenciatura” é o que prevalece na área de Computação, com 65%, seguido por “Tecnólogo” com 19%, e “Bacharelado” com 16%.

**Tabela 3 – Quantidade de cursos de graduação em Computação na EaD, grau de ensino, nas regiões geográficas do Brasil em 2014.**

	Cursos de Computação									
	Centro-Oeste		Nordeste			Norte		Sudeste		Sul
	Licenciatura	Tecnólogo	Bacharelado	Licenciatura	Licenciatura	Tecnólogo	Bacharelado	Licenciatura	Tecnólogo	Tecnólogo
Ciências da Computação				1						
Computação	2			8	1			2		
Informática				1	1			1		
Sistemas de Informação			3				1			
Tecnologia em Análise e Desenvolvimento de Sistemas									1	
Tecnologia em Desenvolvimento de Sistemas de Informação						1				
Tecnologia em Sistema para Internet		1								
Tecnologia em Sistemas de Computação									1	
Tecnologia em Sistemas para Internet a Distancia										1

Fonte: Cubo de dados da UAB implementado no Saiku Analytics View.

Das 96 instituições de ensino superior cadastradas no sistema UAB que ofertam cursos à distância, apenas duas não contemplam vagas específicas para professores leigo, ou seja, professores que ainda não são habilitados. Dos 262 cursos da área “Ciências, Matemática e Computação”, 11% não contemplam vagas específicas para professores leigos e 89% contemplam. Dentre os 26 cursos que são específicos de Computação, 10 (38%) não contemplam vagas específicas para professores leigos, e 16 (62%) contemplam.



A Tabela 4 apresenta o total de instituições com oferta de cursos em EaD por região, além da quantidade de cursos por instituição e quantidade de cursos de graduação em Computação. Das 27 unidades federativas do país, apenas o estado do Acre não possui instituição de ensino com oferta de curso na modalidade a distância.

**Tabela 4 – Visão geral quantitativa de instituições e cursos**

Região	Nº Instituições	Nº Cursos	Cursos de Graduação em Computação
Centro-Oeste	9	129	3
Nordeste	33	421	13
Norte	12	113	3
Sul	14	170	1
Sudeste	28	272	6

Fonte: Cubo de dados da UAB implementado no *Saiku Analytics View*.

As instituições que ofertam cursos da área “Ciências, Matemática e Computação” que possuem os maiores índices CC são: Universidade de São Paulo e Universidade Federal de Juiz de Fora, com índice CC igual a 5 (cinco). E as instituições com os maiores índices ENADE são: Universidade Federal de Pelotas e Universidade Federal Fluminense, com índice ENADE igual a 5 (cinco). Já a instituição que oferta curso da área “Ciências, Matemática e Computação” com o menor índice CC é a Universidade Federal do Pará, com índice CC igual a 2 (dois). E as instituições com os menores índices ENADE são: Universidade Federal de Alagoas, Universidade Federal do Amazonas, Universidade Federal do Rio de Janeiro e Universidade Federal Rural de Pernambuco.

Quanto aos cursos de especialização, a área científica “Ciências sociais, negócios e direito” é a que dispõe de mais cursos (347 cursos – 46%). Em segundo lugar tem-se a área científica “Educação”, com 33%, seguida pela área “Ciências, Matemática e Computação”, com 6%, ou seja, 48 cursos de especialização. Da área científica “Ciências, Matemática e Computação” 12 (doze) estados dispõem de instituições que ofertam especialização na área, são eles: DF, MS, MT, CE, MA, PI, ES, MG, RJ, SP, PR e RS.

As instituições com os maiores CI (Conceito Institucional) são: Instituto Militar de Engenharia, Universidade Federal de Lavras, Universidade Federal do Rio de Janeiro, Universidade Federal do Rio Grande do Norte e Universidade Federal Fluminense, todas elas com IC igual a 5 (cinco). Já a instituição com o menor IC é o Instituto Federal de Educação, Ciência e Tecnologia de Pernambuco, com IC igual a 2 (dois).

## 5. Conclusões

Portanto, as contribuições do trabalho relacionam-se ao fato de haver, no cenário atual, uma crescente necessidade das IFES em implementar e melhorar as práticas e metodologias da EaD. Afinal, a EaD é uma educação além do limite espaço-tempo, que visa transformar e evoluir o processo tradicional de aprendizagem, bem como flexibilizar os processos de ensino-aprendizagem, além de expandir as universidades. O trabalho pretende provocar interesse em instituições, pesquisadores e profissionais envolvidos com a implementação, administração e utilização de sistemas de informações gerenciais de apoio à decisão no contexto da EaD. Tais tecnologias se propõem em fornecer às IFES indicadores de qualidade, proporcionando tomadas de

decisões que visam à redução da evasão e retenção dos estudantes e, conseqüentemente, a melhoria da EaD.

Pode ser concluído a partir deste trabalho que o perfil da EaD no Brasil está em constante desenvolvimento e expansão, especificamente nas áreas científicas de conhecimento em “Educação” e “Ciências, Matemática e Comunicação”, dessa forma, viabiliza-se cada vez mais o acesso da população ao ensino superior. Ao explorar e analisar os dados da EaD constata-se o grande potencial do Brasil nesse contexto, podendo se destacar no cenário educacional mundial através de ações que promovam o processo de ensino e aprendizagem eficaz nos ambientes e-learning e nas instituições através do uso de tecnologias educacionais contemporâneas.

## Referências

- ABED - ASSOCIAÇÃO BRASILEIRA DE EDUCAÇÃO A DISTÂNCIA. Censo EaD: relatório analítico da aprendizagem a distância no Brasil. Curitiba: Ibpx, 2014.
- GOTTARDO, Ernani; KAESTNER, Celso Antônio Alves; NORONHA, Robinson Vida. Estimativa de Desempenho Acadêmico de Estudantes: Análise da Aplicação de Técnicas de Mineração de Dados em Cursos a Distância. Revista Brasileira de Informática na Educação, v. 22, p. 45-55, 2014.
- GUERCIO, H.; Marques, P.; STROELE, V.; SILVA, C. K. P.; BARRERE, E. Análise do Desempenho Estudantil na Educação a Distância Aplicando Técnicas de Mineração de Dados. In: Congresso Brasileiro de Informática na Educação, 2014, Dourados/MS. I Workshop de Mineração de Dados em Ambientes Virtuais de Ensino/Aprendizagem, 2014.
- INMON, William H. Building the Data Warehouse: Getting Started. 4a Edição. Editora: Wiley Publishing, Inc, 2005.
- KAMPFF, A. J. C. Mineração de dados educacionais para geração de alertas em ambientes virtuais de aprendizagem como apoio à prática docente. Tese (doutorado), Universidade Federal do Rio Grande do Sul, Programa de Pós-Graduação em Informática na Educação. Porto Alegre/RS/Brasil, 2009.
- MORAN, José Manuel. Educação a Distância no Brasil: situação e perspectivas, 2014. Disponível em: <<http://www2.eca.usp.br/moran>>. Acesso: 18 abr. 2015.
- RODRIGUES, R. L.; RAMOS, J. L. C.; SILVA, J. C. S.; GOMES, A. S. A literatura brasileira sobre mineração de dados educacionais. In: Congresso Brasileiro de Informática na Educação, 2014, Dourados. Workshop de Mineração de Dados em Ambientes Virtuais do Ensino/Aprendizagem, 2014.
- SANTANA, L. C.; MACIEL, A. M. A.; RODRIGUES, R. L. Avaliação do Perfil de Uso no Ambiente Moodle Utilizando Técnicas de Mineração de Dados. In: Simpósio Brasileiro de Informática na Educação, 2014, Dourados. Congresso Brasileiro de Informática na Educação, 2014.
- SILVA, L. A.; MORIN, A. H.; SATO, T. M. C. Práticas de Mineração de Dados no Exame Nacional do Ensino Médio. In: Congresso Brasileiro de Informática na Educação, 2014, Dourados. Workshop de Mineração de Dados em Ambientes Virtuais do Ensino/Aprendizagem, 2014. p. 651-660.