

UNIVERSIDADE FEDERAL DE SÃO CARLOS
CENTRO DE CIÊNCIAS EXATAS E DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

**MODELO E FRAMEWORK PARA O
DESENVOLVIMENTO DE FERRAMENTAS
ANALÍTICAS DE APOIO AO ENSINO,
APRENDIZAGEM E GESTÃO EDUCACIONAL**

GISLAINE CRISTINA MICHELOTI ROSALES

ORIENTADORA: PROFA. DRA. REGINA BORGES DE ARAÚJO

São Carlos - SP
Agosto/2014

UNIVERSIDADE FEDERAL DE SÃO CARLOS
CENTRO DE CIÊNCIAS EXATAS E DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

**MODELO E FRAMEWORK PARA O
DESENVOLVIMENTO DE FERRAMENTAS
ANALÍTICAS DE APOIO AO ENSINO,
APRENDIZAGEM E GESTÃO EDUCACIONAL**

GISLAINE CRISTINA MICHELOTI ROSALES

Tese apresentada ao Programa de Pós-Graduação em Ciência da Computação da Universidade Federal de São Carlos, como parte dos requisitos para a obtenção do título de Doutor em Ciência da Computação, área de concentração: Processamento de Imagens e Sinais
Orientadora: Profa. Dra. Regina Borges de Araújo

São Carlos - SP
Agosto/2014

**Ficha catalográfica elaborada pelo DePT da
Biblioteca Comunitária/UFSCar**

R788mf Rosales, Gislaine Cristina Micheloti.
Modelo e framework para o desenvolvimento de
ferramentas analíticas de apoio ao ensino, aprendizagem e
gestão educacional / Gislaine Cristina Micheloti Rosales. --
São Carlos : UFSCar, 2015.
208 f.

Tese (Doutorado) -- Universidade Federal de São Carlos,
2014.

1. Ciência da computação. 2. Educação a distância. 3.
Análise da aprendizagem. 4. Análise de dados. 5. Educação
- processo decisório. 6. Ciência de contexto. I. Título.

CDD: 004 (20^a)

UNIVERSIDADE FEDERAL DE SÃO CARLOS

CENTRO DE CIÊNCIAS EXATAS E DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

MODELO e FRAMEWORK PARA O DESENVOLVIMENTO DE FERRAMENTAS ANALÍTICAS DE APOIO AO ENSINO, APRENDIZAGEM E GESTÃO EDUCACIONAL

GISLAINE CRISTINA MICHELOTI ROSALES

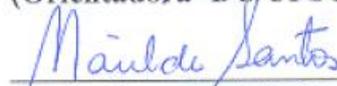
Tese apresentada ao Programa de Pós-Graduação em Ciência da Computação da Universidade Federal de São Carlos, como parte dos requisitos para a obtenção do título de Doutor em Ciência da Computação, área de concentração: Processamento de Imagens e Sinais.

Aprovado em 04 de setembro de 2014.

Membros da Banca:



Profa. Dra. Regina Borges de Araujo
(Orientadora- DG/UFSCar)



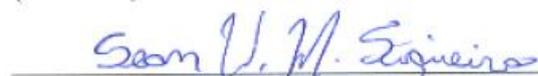
Profa. Dra. Marilde Terezinha Prado Santos
(DC/UFSCar)



Profa. Dra. Ellen Francine Barbosa
(ICMC/USP)



Profa. Dra. Liane Margarida Rockenbach Tarouco
(UFRGS)



Prof. Dr. Sean Wolfgang Matsui Siqueira
(UNIRIO)

São Carlos - SP
Agosto/2014

Dedicatória

*Dedico esta tese ao meu filho, Heitor, ao meu esposo, Fabrício,
e à minha mãe, Lourdes, que são a razão da minha alegria.*

AGRADECIMENTOS

À minha mãe, que por uma vida de amor e dedicação, não mediu esforços pela minha educação. Obrigada mãe, por me oferecer a oportunidade de estudar, por me incentivar a trilhar novas jornadas e por cobrar bons resultados ao final.

Ao meu amado esposo Fabrício, que foi meu ombro amigo durante as dificuldades que enfrentei. Em especial, pela paciência e compreensão reveladas, apesar dos momentos de ausência da minha parte ao longo dos anos de desenvolvimento desta tese.

À professora Regina Borges de Araújo, minha orientadora, pela oportunidade que me deu, pela competência científica, pela disponibilidade e pelo acompanhamento do trabalho, assim como pelas críticas, correções e sugestões feitas durante a orientação.

À professora Joice Lee Otsuka, pelas sugestões em artigos, pela disponibilidade sempre manifestada e pela amizade que construímos nesse período.

À Maria Angélica do Carmo Zanotto, pelo incansável apoio moral, pelas valiosas contribuições na área pedagógica, que me ajudaram a encontrar soluções que contribuíram para o desenvolvimento desta tese. Obrigada também pela permanente disponibilidade e pela amizade que se fortaleceu durante o desenvolvimento desta tese.

Aos amigos Janaína Cintra Abib, Cláudia Beatriz Berti, Rafaela Vilela, Fernando Vieira Duarte e Allan Oliveira, pela amizade, pelo carinho, pelos agradáveis momentos de convivência, pelo estimável apoio e pelas correções e sugestões no texto da tese.

Aos alunos de Iniciação Científica, em especial à Juliana Augusto Stocco, Wellington Camargo, Eurico Corte e Ubiratan Campos, pela dedicação ao trabalho e pelo comprometimento demonstrados.

A todos àqueles que se dispuseram a participar das pesquisas conduzidas para avaliação do trabalho.

RESUMO

O uso de novas tecnologias de informação e comunicação na área educacional, que vão além dos tradicionais Sistemas Gerenciadores da Aprendizagem (SGA), tem gerado um volume crescente de dados, que torna desafiadoras e complexas as análises de dados gerados para atender a tomada de decisão nos níveis de ensino, aprendizagem e gestão. Apesar das altas expectativas sobre a análise de dados no campo educacional, pesquisas atuais na área estão focadas mais especificamente sobre dados de alunos, seus processos e comportamentos de aprendizagem, até mesmo quando o foco da pesquisa é melhorar o ensino ou as ações em nível institucional. De modo a facilitar e estender o processo de análise dos dados para as áreas de ensino, aprendizagem e gestão para diferentes partes interessadas, este trabalho apresenta um Modelo Conceitual que deverá guiar a construção de aplicações analíticas educacionais cientes de contexto que apoiam a tomada de decisões educacionais nos níveis micro, meso e macro. O Modelo Conceitual propõe a coleta de dados educacionais a partir de diversas e heterogêneas fontes e de maneira descentralizada usando sensores lógicos e físicos. O Modelo suporta análises dos dados coletados em três níveis: análise descritiva, análise preditiva e análise prescritiva. O Modelo Conceitual foi implementado a partir de uma arquitetura de *framework* aberta, extensível e reusável, que oferece um caminho mais simples e unificado para a aquisição de comportamentos de usuários em aprendizagem *online*, a modelagem dos contextos coletados e análises. Para validação do Modelo Conceitual proposto, foram desenvolvidas três aplicações, a saber: ViTrackeR, para apoio à aprendizagem autorregulada provendo visualização de dados de rastreamento e recomendações personalizadas; ViMonitor, para apoio, em tempo real, às equipes de ensino e de gestão acadêmica fornecendo informações importantes sobre estudantes e tutores; e ViAssess, que provê suporte à segurança para aplicação de avaliações *online*. O Modelo Conceitual foi avaliado e validado em um ambiente real (por estudantes, tutores, professores e gestores). O *framework* foi avaliado por desenvolvedores de ferramentas analíticas educacionais e por pesquisadores especialistas no domínio desta pesquisa, obtendo resultados muito positivos. Os resultados das avaliações indicam que o Modelo Conceitual proposto suporta o desenvolvimento de aplicações analíticas educacionais nos três níveis de decisão, micro, meso e macro, e também suporta os três níveis de análises previstos: descritiva, preditiva e prescritiva.

Palavras-chave: educação a distância, análise da aprendizagem, análise de dados educacionais, tomada de decisão educacional, ciência de contexto.

ABSTRACT

The use of new information and communication technologies in education that go beyond the traditional Learning Management Systems (LMS), has generated a growing volume of data, making challenging and complex the analysis of data generated to meet the decision-making levels of teaching, learning and management. Despite high expectations for data analysis in education, current research in the area is focused more specifically on student data, processes and learning behaviors, even when the focus of the research is to improve the teaching or actions at the institutional level. In order to facilitate and extend the process of data analysis in the areas of teaching, learning and management for different stakeholders, this thesis presents a conceptual model that will guide the construction of educational analytical context aware applications that support educational decision making in the micro, meso and macro levels. The conceptual model proposes the collection of educational data from multiple, heterogeneous sources in a decentralized manner using logical and physical sensors. The Model supports analysis of data collected at three levels: descriptive analysis, predictive analysis and prescriptive analysis. The conceptual model was established from an open architecture framework, extensible and reusable, which offers a simpler and unified path for both the acquisition of user behaviors in online learning, and the modeling and analysis of the collected contexts. To validate the proposed conceptual model, three applications were developed, namely: ViTrackeR, to support self-regulated learning by providing visualization of data tracking and personalized recommendations; ViMonitor to support real-time, teams of teaching and academic management providing important information on students and tutors; and ViAssess, which provides support for secure assessments online. The conceptual model was evaluated and validated in a real environment (students, tutors, teachers and administrators). The framework was rated by both developers of educational and analytical tools and by expert researchers in the field of this research, obtaining very positive results. Evaluation results indicate that the proposed conceptual model supports the development of educational applications in the three analytical levels of decision making, micro, meso and macro, and also supports the three levels of analysis provided: descriptive, predictive and prescriptive.

Keywords: distance education, learning analysis, analysis of educational data, educational decision making, context awareness.

LISTA DE FIGURAS

Figura 2.1 – Contexto Turbulência em Aeronaves. Fonte: adaptado de (BAUER, 2003).	45
Figura 2.2 – Ontologia de contexto com foco no “estudante”. Fonte: adaptado de (YU, ZHOU e SHU, 2009).	47
Figura 2.3 – Uma árvore de decisão usada para classificar planta Iris. Fonte: adaptado de (GUAZZELLI, 2012).....	49
Figura 4.1 – Modelo Conceitual para o desenvolvimento de ferramentas analíticas educacionais. Fonte: elaborado pela autora.....	78
Figura 4.2 – Categorias de análises de dados. Fonte: elaborado pela autora.	86
Figura 5.1 – Arquitetura proposta do <i>Framework</i> para apoiar o desenvolvimento de aplicações educacionais analíticas. Fonte: elaborado pela autora.....	95
Figura 5.2 – Coleta distribuída de dados realizada por redes de sensores híbrida e tecnologias utilizadas. Fonte: elaborado pela autora.	99
Figura 5.3 – Representação da Modelagem de Contextos e Análise de Dados baseados em Ontologias. Fonte: elaborado pela autora.....	102
Figura 5.4 – Parte da Ontologia para Modelagem de Contextos em LA. (a) Representação gráfica da ontologia. (b) Classes. (c) Propriedades de objeto. (d) Propriedades de dados. Fonte: elaborado pela autora.	104
Figura 5.5 – Rede Bayesiana usada para modelar o comportamento do estudante em fórum de discussão. Fonte: elaborado pela autora.	106
Figura 5.6 – Representação gráfica do progresso acadêmico do estudante em uma barra de progresso multicolorida. Fonte: elaborado pela autora.	112
Figura 5.7 – Visualização do progresso acadêmico individual e exemplo de recomendação para o estudante. Fonte: elaborado pela autora.	113
Figura 5.8 – Visualização do progresso acadêmico em cada variável utilizada na rede Bayesiana, mostrando informações individuais e em comparação à média do grupo, ao valor mais alto e ao valor mais baixo encontrados. Fonte: elaborado pela autora.	114
Figura 5.9 – Cinco variáveis agregadas para três atividades de fóruns de discussão. Fonte: elaborado pela autora.	115
Figura 5.10 – Variável “acessos” observada em três atividades de fóruns distintas. Fonte: elaborado pela autora.	115

Figura 5.11 – <i>Downloads</i> e navegação por materiais <i>online</i> – o que era esperado e o que foi feito pelo aluno. Fonte: elaborado pela autora.	116
Figura 5.12 – Exemplo de recomendação feita pelo ViTrackeR: leitura de materiais teóricos antes de realizar um exercício. Fonte: elaborado pela autora.	117
Figura 5.13 – Exemplo de recomendação no ViTrackeR: manipular o objeto “equação de uma parábola”. Fonte: elaborado pela autora.	118
Figura 5.14 – Cinco variáveis observadas sobre três fóruns. Fonte: elaborado pela autora.	119
Figura 5.15 – Cinco variáveis observadas sobre três fóruns. Fonte: elaborado pela autora.	120
Figura 5.16 – Configuração de opções de monitoramento no ViAssess. Fonte: elaborado pela autora.	123
Figura 5.17 – Mensagem exibida pelo ViAssess ao identificar uma ação proibida durante uma avaliação <i>online</i> . Fonte: elaborado pela autora.	124
Figura 6.1 – Modelo de avaliação proposto para validação do Modelo Conceitual apresentado nesta tese. Fonte: elaborado pela autora.	126
Figura 6.2 – Quantidade de respostas para variáveis que medem a ajuda oferecida pelo ViTrackeR. Fonte: elaborado pela autora.	129
Figura 6.3 - Quantidade de respostas para itens 3.1 e 3.3 do questionário dos alunos. Fonte: elaborado pela autora.	130
Figura 6.4 - Quantidade de respostas para item 3.5. Fonte: elaborado pela autora.	130
Figura 6.5 – Respostas dos alunos sobre a utilidade da informação disponibilizada pelo ViTrackeR. Fonte: elaborado pela autora.	131
Figura 6.6 – Respostas dos alunos sobre a facilidade de uso do ViTrackeR e atividades. Fonte: elaborado pela autora.	131
Figura 6.7 – Respostas dos alunos sobre a motivação no uso do ViTrackeR. Fonte: elaborado pela autora.	131
Figura 6.8 – Diagrama de dispersão dos dados: aproveitamento escolar e tempo de estudo. Fonte: elaborado pela autora.	133
Figura 6.9 – Diagrama de dispersão dos dados: aproveitamento escolar e tempo em recursos (<i>Facebook</i> e/ou portais de notícias e entretenimento). Fonte: elaborado pela autora.	135
Figura 6.10 - Tempos (min) de estudo, utilização do <i>Facebook</i> e navegação em sites de notícias e/ou entretenimento para alunos com piores aproveitamentos. Fonte: elaborado pela autora.	137

Figura 6.11 - Tempos (min) de estudo, utilização do <i>Facebook</i> e navegação em sites de notícias e/ou entretenimento para alunos com melhores aproveitamentos na disciplina. Fonte: elaborado pela autora.	137
Figura 6.12 - Respostas dos instrutores referente ao item 2 do instrumento. Fonte: elaborado pela autora.	141
Figura 6.13 - Respostas dos instrutores referente ao item 3 do instrumento. Fonte: elaborado pela autora.	141
Figura G.1 – Idade dos alunos participantes. Fonte: elaborado pela autora.	203
Figura G.2 – Ano escolar. Fonte: elaborado pela autora.....	203
Figura G.3 – Tempo de Experiência. Fonte: elaborado pela autora.....	204

LISTA DE TABELAS

Tabela 3.1 – Resumo dos sistemas estudados: fontes de dados, usuários-alvo e Técnicas de Modelagem e Análise utilizadas.....	71
Tabela 6.2 – Frequências: variáveis aproveitamento e tempo de dedicação aos estudos.....	133
Tabela 6.3 – Resultado do coeficiente de <i>Spearman</i> entre tempo de dedicação aos estudos e aproveitamento.	134
Tabela 6.4 – Frequências para os itens aproveitamento, tempo gasto utilizando <i>Facebook</i> e tempo gasto navegando por portais e sites de notícias e entretenimento.	135
Tabela 6.5 – Resultado do coeficiente de <i>Spearman</i> entre aproveitamento escolar e tempo navegando por portais de notícias e entretenimento; e o tempo no <i>Facebook</i>	136
Tabela G.1 – Caracterização da amostra dos estudantes.	204
Tabela G.2 – Estatísticas: tópico 2 do instrumento de avaliação.	204
Tabela G.3 – Frequências: itens 3.1 e 3.3 do instrumento de avaliação.....	205
Tabela G.4 – Frequências: item 3.5 do instrumento de avaliação.....	205
Tabela G.5 – Frequências: itens 4.1, 4.3 e 4.5 do instrumento de avaliação.	205
Tabela G.6 – Frequências: itens 5.1 e 5.3 do instrumento de avaliação.....	205
Tabela G.7 – Frequências: item 5.5 do instrumento de avaliação.....	206

LISTA DE QUADROS

Quadro 1.1 – Principais procedimentos metodológicos utilizados na pesquisa. Fonte: elaborado pela autora.	28
Quadro 4.1 – Variáveis selecionadas para medir o sucesso/fracasso em fóruns de discussão. Fonte: elaborado pela autora.	79
Quadro 5.1 – Sensores Lógicos e Atuadores. Fonte: elaborado pela autora.	100
Quadro 5.2 – Sensores lógicos e físicos para aplicação de monitoramento de avaliações <i>online</i> . Fonte: elaborado pela autora.	122
Quadro 6.1 – Perfis dos instrutores. Fonte: elaborado pela autora.	141
Quadro 6.2 – Perfis dos gestores entrevistados. Fonte: elaborado pela autora.	144
Quadro 6.3 - Resumo das respostas dos gestores à questão 3. Fonte: elaborado pela autora.	146
Quadro 6.4 –Perfis dos desenvolvedores entrevistados. Fonte: elaborado pela autora.	150
Quadro 6.5 – Perfis dos especialistas entrevistados. Fonte: elaborado pela autora.	155

LISTA DE ABREVIATURAS E SIGLAS

A/D – Analógicos/Digitais

API - *Application Programming Interface*

AVA – Ambiente Virtual de Aprendizagem

BN – *Bayesian Network*

CAM - *Contextualized Attention Metadata*

CSCCL - *Computer-Suported Collaborative Learning*

CPA – Comissão Própria de Avaliação

DAML - *DARPA Agent Markup Language*

DARPA - *Defense Advanced Research Projects Agency*

DDDM - *Data-Driven Decision Making*

DL - *Description Logic*

DUMBO - *Dynamic, Ubiquitous, Mobile Meeting Board*

EaD - Educação a Distância

EDM – Educational Data Mining

FIPA - *Foundation for Intelligent Physical Agents*

GPS - *Global Positioning System*

HTML - *HyperText Markup Language*

IES – Instituição de Ensino Superior

JDL - *Joint Directory Laboratory*

KQML - *Knowledge Query Manipulation Language*

LAK - *Learning Analytics and Knowledge*

LMS - *Learning Managemet System*

MIT - *Massachussets Institute of Technology*

MOODLE - *Modular Object Oriented Dynamic Learning Environment*

OIL - *Ontology Inference Language ou Ontology Interchange Language*

OML - *Ontology Markup Language*

OUHK - *The Open University of Hong Kong*

OWL - *Web Ontology Language*

RB - Redes Bayesianas

RDF - *Resource Description Framework*

RDFS - *Resource Description Framework*

RDQL - *RDF Data Query Language*

RFID - *Radio-Frequency Identification*

SEAD – Secretaria de Educação a Distância

SGA - Sistema Gerenciador de Aprendizagem

SGBD - Sistema Gerenciador de Bancos de Dados

SHOE - *Simple HTML Ontology Extensions*

SOA – *Service Oriented Architecture*

SOS - *Sensor Observation Service*

SPARQL - *Query Language for RDF*

SPSS - *Statistical Package for the Social Sciences*

TPC - Tabela de Probabilidades Condicionais

TEL - *Technology Enhanced Learning*

TIDIA-AE - Tecnologia da Informação no Desenvolvimento da Internet Avançada -
Aprendizado Eletrônico

UAB - Universidade Aberta do Brasil

UFSCar - Universidade Federal de São Carlos

UML - *Unified Modeling Language*

VLE – *Virtual Learning Environment*

ViTrackerR – Ferramenta de Visualização de Dados e Recomendações

ViMonitor – Visualização e Monitoramento da Aprendizagem

ViAssess - *Avaliações Online Seguras*

XML - *Extensible Markup Language*

W3C - *World Wide Web Consortium*

WINDIS - *Wireless Network and Distributed Interactive Simulations*

SUMÁRIO

CAPÍTULO 1 - INTRODUÇÃO	20
1.1 Justificativas da Pesquisa e Problema	20
1.2 Solução proposta e Contribuições.....	23
1.3 Objetivos	25
1.4 Proposições de Pesquisa	26
1.5 Procedimentos Metodológicos	28
1.5.1 Abordagem e classificação da pesquisa	28
1.5.2 Método de pesquisa e instrumentos para coleta de dados	29
1.5.3 Amostra e seleção dos participantes.....	30
1.5.4 Análise dos dados	31
1.5.5 Descrição das atividades metodológicas e observações pertinentes.....	31
1.6 Estrutura e Organização do Trabalho	34
CAPÍTULO 2 - FUNDAMENTOS CONCEITUAIS	36
2.1 Considerações Iniciais.....	36
2.2 Aprendizagem Assistida por Tecnologias.....	37
2.2.1 Ambientes Pessoais de Aprendizagem	39
2.3 Contexto	41
2.3.1 Modelagem de Contextos.....	42
2.3.1.1 Abordagens para modelagem de contextos	43
2.4 Analítica.....	50
2.4.1 Análise da Aprendizagem.....	51
2.5 Dados em <i>Learning Analytics</i>	53
2.6 Considerações Finais	55
CAPÍTULO 3 - ESTADO DA ARTE EM LEARNING ANALYTICS	56
3.1 Considerações iniciais.....	56
3.2 <i>Learning Analytics</i>	58
3.2.1 Visualização	59
3.2.1.1 Avaliação e <i>Feedback</i>	59
3.2.1.2 Monitoramento e Tutoria	64

3.2.2	Recomendação	66
3.2.3	Personalização e Adaptação	69
3.2.4	Discussões sobre a captura de contextos em LA	72
3.3	Considerações Finais	73
CAPÍTULO 4 - MODELO CONCEITUAL PARA O DESENVOLVIMENTO DE FERRAMENTAS ANALÍTICAS EDUCACIONAIS.....		76
4.1	Considerações iniciais.....	76
4.2	Descrição do Modelo Conceitual proposto	78
4.2.1	Artefatos	78
4.2.1.1	Levantamento preliminar de Artefatos	81
4.2.2	Coleta e Pré-Processamento de Dados	83
4.2.2.1	Coleta dos Dados	83
4.2.2.2	Pré-processamento dos Dados	85
4.2.3	Análises.....	86
4.2.3.1	Técnicas para Análises de Dados	87
4.2.4	Decisão	89
4.3	Considerações Finais	90
CAPÍTULO 5 - FRAMEWORK PARA COLETA E ANÁLISES DE DADOS E APLICAÇÕES EDUCACIONAIS DESENVOLVIDAS.....		92
5.1	Considerações iniciais.....	92
5.2	Visão Geral do <i>Framework</i> de Coleta e Análises de Dados.....	94
5.2.1	Arquitetura do <i>Framework</i>	94
5.2.1.1	Fonte de Dados	96
5.2.1.2	Gerenciador de Dados	101
5.2.1.3	Modelagem de Contextos e Análises	101
5.2.1.4	Camada de Aplicações.....	107
5.2.2	Contribuições do <i>Framework</i> proposto.....	108
5.2.3	Privacidade dos Dados.....	109
5.3	Aplicações Desenvolvidas a partir do <i>Framework</i> proposto	110
5.3.1	ViTrackeR: Visualização de Dados e Recomendações	111
5.3.1.1	Progresso Acadêmico Sumarizado	112
5.3.1.2	Progresso Acadêmico Detalhado	114
5.3.1.3	Agente de Recomendações	116

5.3.2 ViMonitor: Visualização e Monitoramento da Aprendizagem	118
5.3.3 ViAssess: Avaliações <i>Online</i> Seguras.....	120
5.4 Considerações Finais	124
CAPÍTULO 6 - AVALIAÇÕES E APRESENTAÇÃO DOS RESULTADOS.....	125
6.1 Considerações iniciais.....	125
6.2 Avaliações realizadas com usuários das aplicações desenvolvidas	126
6.2.1 Avaliação realizada com Estudantes.....	126
6.2.1.1 Planejamento das Entrevistas e Instrumento de Coleta	127
6.2.1.2 Caracterização da amostra	128
6.2.1.3 Resultados obtidos: percepções dos estudantes após uso do ViTrackeR ..	129
6.2.1.4 Correlação de Dados.....	132
6.2.1.5 Discussões sobre os resultados obtidos nas entrevistas com estudantes: atendimento aos objetivos e às proposições de pesquisa	138
6.2.2 Avaliação realizada com instrutores (tutores e professores)	139
6.2.2.1 Planejamento das Entrevistas e Instrumento de Coleta	140
6.2.2.2 Caracterização da amostra	140
6.2.2.3 Resultados obtidos nas entrevistas com instrutores	141
6.2.2.4 Discussões sobre os resultados obtidos nas entrevistas com instrutores: atendimento aos objetivos e às proposições propostas	142
6.2.3 Avaliação realizada com gestores de projetos de EaD	143
6.2.3.1 Planejamento das Entrevistas e Instrumento de Coleta	143
6.2.3.2 Caracterização da amostra	144
6.2.3.3 Resultados obtidos nas entrevistas com gestores	145
6.2.3.4 Discussões sobre os resultados obtidos nas entrevistas com gestores: atendimento aos objetivos e às proposições propostas	148
6.3 Avaliação realizada com desenvolvedores de ferramentas de apoio à educação <i>online</i>	149
6.3.1 Planejamento das Entrevistas e Instrumento de Coleta	149
6.3.2 Caracterização da Amostra	150
6.3.3 Resultados obtidos nas entrevistas com desenvolvedores	151
6.3.4 Discussões sobre os resultados obtidos nas entrevistas com desenvolvedores	153
6.4 Avaliação realizada com especialistas em aplicações educacionais	154

6.4.1 Planejamento das Entrevistas e Instrumento de Coleta	154
6.4.2 Caracterização da Amostra	155
6.4.3 Resultados obtidos nas entrevistas com especialistas.....	156
6.4.4 Discussões sobre os resultados obtidos nas entrevistas com especialistas ..	157
6.5 Considerações Finais	158
CAPÍTULO 7 - CONSIDERAÇÕES FINAIS.....	159
7.1 Atendimento à questão de pesquisa, aos objetivos e às proposições	159
7.2 Limitações da Pesquisa.....	160
7.3 Trabalhos Futuros	161
REFERÊNCIAS.....	163
APÊNDICE A	185
APÊNDICE B	189
APÊNDICE C	196
APÊNDICE D	199
APÊNDICE E	201
APÊNDICE F.....	202
APÊNDICE G.....	203
APÊNDICE H	207

Capítulo 1

INTRODUÇÃO

“Technology represents a provocation to schools’ traditional ways of working. But investigating its role in tech-savvy schools clearly shows that, by building on two fundamental human needs – communication and curiosity – technology can be used to broaden students’ horizons.” (PERSSON, 2012, p. 21).

Este capítulo introdutório da tese apresenta na Seção 1.1 as justificativas da pesquisa e o problema em que se baseia o presente estudo; a solução proposta e as contribuições científicas e tecnológicas são apresentadas na Seção 1.2; os objetivos do trabalho são apresentados na Seção 1.3; as proposições de pesquisa são apresentadas na Seção 1.4; seguidas dos procedimentos metodológicos adotados, na Seção 1.5; e da estrutura e organização do trabalho na Seção 1.6.

1.1 Justificativas da Pesquisa e Problema

O conceito “big data” juntamente com o avanço das tecnologias analíticas têm sido apontados como meios potencialmente transformadores das economias e altamente significativos para aumentar a produtividade organizacional (MANYIKA et al., 2011, p. 13), melhorar a eficiência operacional e financeira (VAN BARNEVELD, ARNOLD e CAMPBELL, 2012, p. 2) e aumentar a vantagem competitiva das organizações (KIRON et al., 2011). No campo educacional conceitos como conteúdos abertos, dados abertos e recursos abertos, aliados à noção de transparência e facilidade de acesso aos dados têm sido indicados como tendências motivadoras para novas pesquisas na área nos próximos anos (VERBERT et al., 2012). Siemens (2013) aponta que sistemas educacionais atuais têm feito uso

limitado de dados disponíveis para melhorar o ensino, os quais estão normalmente, vinculados à aprendizagem e ao sucesso de aprendizes, estando atrás de outros setores que conseguem explorar melhor o potencial da análise de dados. Todavia, indica que tem havido uma recente explosão de interesse na adoção de ferramentas analíticas para solução de problemas atuais na educação, como a retenção e suporte ao aprendiz (SIEMENS, 2013, p. 2). Os relatórios *The Horizon Report* produzidos pela *The New Media Consortium* (NMC) em parceria com as principais organizações que atuam em pesquisas sobre ensino e aprendizagem com tecnologias (EDUCAUSE¹ e *International Society for Technology in Education*²) nas edições de 2011³, 2012⁴ e 2013⁵ apontaram *learning analytics* como uma tendência futura chave a ser explorada nos próximos anos (JOHNSON et al., 2011, 2012, 2013).

A análise de dados em contextos educacionais pode ser considerada como facilitadora para a tomada de decisões que garantam o sucesso em todos os níveis da educação (VAN BARNEVELD, ARNOLD e CAMPBELL, 2012), variando de ações específicas de estudantes individuais - por exemplo monitoramento da aprendizagem - a ações em nível de curso – por exemplo (re)planejamento do curso ou de conteúdos – e até mesmo no endossamento e indução de políticas educacionais institucionais. A necessidade de soluções para apoiar a tomada de decisão em diferentes níveis da educação foi apontada na conferência *Technological Ecosystems for Enhancing Multiculturality Conference*⁶ (TEEM) que discutiu o futuro da *Learning Analytics*, na trilha 5 – “A promised land for educational decision making? Present and future of learning analytics” (CONDE e HERNÁNDEZ-GARCÍA, 2013).

¹ EDUCAUSE é uma das mais importantes associações sem fins lucrativos comprometidos com o avanço do ensino superior. EDUCAUSE é formada por líderes, pesquisadores e profissionais que atuam com tecnologias da informação. Site oficial: www.educause.edu

² *International Society for Technology in Education* (ISTE) é uma organização sem fins lucrativos que realiza pesquisas direcionadas a educadores e líderes educacionais com o objetivo de capacitar alunos de educação online. Site oficial: <https://www.iste.org/>

³ <http://net.educause.edu/ir/library/pdf/HR2011.pdf>

⁴ <http://horizon.wiki.nmc.org/file/view/2012-Horizon.HE-Preview.pdf>

⁵ <http://www.nmc.org/pdf/2013-horizon-report-HE.pdf>

⁶ <http://teemconference.eu/2013/tracks/learning-analytics/>

As expectativas da comunidade científica sobre a análise de dados para atender aos diferentes perfis de usuários da educação *online*, tais como alunos, professores e gestores, são altas. Todavia, a análise de dados no domínio educacional em pesquisas atuais está focada especificamente sobre dados de alunos, seus processos e comportamentos de aprendizagem (VAN BARNEVELD, ARNOLD e CAMPBELL, 2012; GRELLER e DRACHSLER, 2012; SHUM e FERGUSON, 2012 e SIEMENS e LONG, 2011) mesmo quando o foco da pesquisa está voltado para o melhoramento do ensino ou das ações institucionais.

Diversas ferramentas têm sido propostas recentemente a fim de prover melhoria do ensino e/ou aprendizagem, cobrindo várias áreas de pesquisa, tais como recomendação e personalização (VERBERT et al., 2012; Wang et al. 2012); avaliação e *feedback* (MAELE et al., 2013; HWANG e CHANG 2011); e, monitoramento e tutoria (DYCKHOFF et al. 2012; FENG e HEFFERNAN, 2010). Embora, tais ferramentas apoiem o ensino e/ou aprendizagem, elas basicamente atuam de forma isolada em uma ou, no máximo, duas categorias: estudantes individuais - por exemplo, para monitoramento da aprendizagem; e/ou disciplina/curso – por exemplo, para fins de (re)planejamento de conteúdos.

Diante disso, observa-se que há a carência de modelos para suporte à tomada de decisão em educação mediada por tecnologias que apoiam o desenvolvimento de aplicações analíticas em diferentes níveis: **nível micro** - usuários individuais, como por exemplo para fins de aprendizagem *online*; **nível meso** - especialmente para fins de ensino e gestão em nível de disciplinas e curso; e, no **nível macro** - particularmente na gestão educacional e gestão de projetos de Educação a Distância (EaD). Assim, o foco de estudo desta pesquisa reside no seguinte questionamento:

Como oferecer suporte conceitual e tecnológico para apoiar a construção de aplicações analíticas que facilitam a tomada de decisão em educação mediada por tecnologias para o ensino, a aprendizagem e a gestão educacional?

1.2 Solução proposta e Contribuições

Esta pesquisa propõe um Modelo Conceitual para servir como base para o desenvolvimento de aplicações que apoiam a tomada de decisões no domínio de análise de dados educacionais. O Modelo Conceitual proposto busca facilitar e clarificar o entendimento sobre as etapas envolvidas no suporte às decisões nos múltiplos níveis de atuação (individual, disciplina e curso e institucional) e para os diferentes perfis de usuários envolvidos na educação *online*. As características inovadoras do Modelo proposto incluem: suporte aos múltiplos níveis de decisão educacional dirigida a dados (micro, meso e macro); apoio aos diferentes perfis de usuários de EaD (estudante, tutor, professor, coordenador de curso, coordenador pedagógico e gestor educacional); e, abrangência das principais etapas necessárias ao projeto e desenvolvimento de tecnologias para aplicações educacionais analíticas. Este Modelo Conceitual consiste na principal contribuição científica deste trabalho.

A segunda contribuição deste trabalho é o desenvolvimento de um *framework* ciente de contexto que implementa as seguintes etapas do Modelo proposto: coleta de dados relevantes ao domínio educacional, tratamento dos dados coletados, análise e entrega de informações resultantes das análises realizadas para aplicações interessadas. A coleta de dados realizada pelo *framework* é baseada em rede de sensores híbrida – composta por sensores físicos e sensores lógicos que podem, implicitamente, coletar dados sobre as interações e eventos sobre foco de atenção dos diferentes usuários diretamente envolvidos no processo de educação *online* (tais como estudantes, tutores e professores).

Além da captura de comportamentos de usuários convencionalmente monitorados em *browsers web*, sistemas gerenciadores da aprendizagem (SGA) ou bases de dados, o *framework* também coleta conteúdos gerados pelos usuários incluindo seu próprio ambiente pessoal – a partir do rastreamento de programas de aplicações locais, sistema operacional, redes de comunicação, etc. – e o ambiente físico que o rodeia. Conforme mencionado, além do monitoramento dos aprendizes, ações de instrutores também são monitoradas, tais como o tempo de resposta às dúvidas de alunos, assiduidade nos ambientes virtuais, quantidades de avaliações

contínuas disponibilizadas, demora/agilidade na correção de avaliações, substituição ou adição de materiais que possam auxiliar alunos em dificuldades, etc.

O *framework* disponibiliza modelos de contextos e raciocínio genéricos e extensíveis que foram desenvolvidos com o propósito de facilitar a análise de dados que, frequentemente, é realizada pelas aplicações. Desta forma, o *framework* facilita a construção de ferramentas analíticas educacionais cientes de contexto que podem atuar em um ou mais níveis de tomada de decisão e oferece um caminho simples e unificado tanto para aquisição dos comportamentos de usuários em ensino e aprendizagem *online* quanto para análises dos dados coletados, sendo essas as principais contribuições científicas do *framework*. Características peculiares do *framework* proposto, que consistem em suas principais contribuições tecnológicas, são: a fácil integração de novos nós sensores e atuadores, independência de SGA e independência de plataforma. Além da contribuição científica mencionada que motivou o desenvolvimento deste *framework*, outra motivação foi a necessidade de ferramentas que ajudam estudantes, instrutores e gestores a responderem suas questões individuais continuamente e eficientemente, uma vez que a maioria dos trabalhos na área de *learning analytics* são conceituais, conforme afirmam Johnson e colaboradores (2011).

A terceira contribuição deste trabalho é o desenvolvimento de três aplicações analíticas cientes de contexto construídas a partir do *framework* proposto. As aplicações foram desenvolvidas a fim de possibilitar a avaliação do Modelo Conceitual em cada um dos níveis de tomada de decisão mencionados. ViTrackeR, é uma ferramenta que auxilia a promoção da aprendizagem autorregulada e faz recomendações personalizadas individualmente. ViMonitor é uma ferramenta de visualização de dados de rastreamento especialmente dedicada a auxiliar, em tempo de execução, equipe docente e equipe de gestão acadêmica provendo informações importantes sobre o monitoramento da aprendizagem (estudantes) e sobre as participações e interação de tutores e professores. Esta ferramenta também provê visualizações que complementam a apresentação de informações do ViTrackeR. ViAssess é uma ferramenta destinada ao suporte de avaliações somativas *online* e provê mecanismos de segurança para a supervisão e controle de ações que poderiam facilitar a obtenção de melhores resultados por meios ilegais. Diferentemente de outros sistemas similares, as ferramentas propostas neste trabalho possuem melhor integração com o sistema de gestão de EaD, facilitando o

uso de informações geradas nos diferentes níveis da tomada de decisão que elas atuam. As aplicações desenvolvidas foram aplicadas em cursos reais e contribuíram, portanto, para a validação do Modelo proposto nesta tese.

1.3 Objetivos

As ferramentas analíticas de apoio aos processos educacionais, em geral, têm seus propósitos relacionados com o sucesso do aprendiz. O propósito fundamental deste trabalho é apoiar a obtenção do sucesso: (1) do aprendiz - durante seu processo de aprendizagem; (2) dos instrutores - em suas atividades de ensino; e, (3) dos gestores - em suas atividades de gestão.

Para alcançar esse propósito delineado, o principal objetivo desta tese é propor e validar um Modelo Conceitual para auxiliar a tomada de decisão educacional dirigida a dados em ensino e aprendizagem e em gestão educacional, atendendo aos diferentes perfis de usuários de EaD e integrando todas as etapas necessárias à análise educacional. O objetivo do Modelo Conceitual proposto é facilitar a compreensão sobre as etapas necessárias para a criação de ferramentas tecnológicas analíticas cientes de contexto que sirvam como suporte para a tomada de decisão educacional nos níveis micro, meso e macro.

Os seguintes objetivos secundários serviram para aplicação e validação do Modelo Conceitual proposto:

- implementação do Modelo proposto por meio do desenvolvimento de uma arquitetura de *software* genérica, utilizando o conceito de *framework*, que realiza a coleta descentralizada e distribuída de dados educacionais a partir de diversas e diferentes fontes usando redes de sensores e realiza análises de dados educacionais, em tempo de execução, a partir de modelagem preditiva e semântico conceitual. O *framework* desenvolvido possibilita a construção de aplicações para educação *online* independente do escopo de atuação e do perfil de usuário que ela atende. Entende-se por escopo de atuação da aplicação a subárea na qual ela está inserida, por exemplo: avaliação da aprendizagem, monitoramento da aprendizagem, recomendação de conteúdos e personalização da aprendizagem.

- desenvolvimento de três aplicações analíticas cientes de contexto integradas ao sistema de gestão de EaD que atuam em diferentes níveis de tomada de decisão educacional e atendem perfis diferentes de usuários em educação *online*.

1.4 Proposições de Pesquisa

Com base na revisão da literatura sobre sistemas de apoio à decisão educacional e sobre ferramentas analíticas para facilitar os processos de ensino e aprendizagem, o presente estudo parte da seguinte proposição de pesquisa:

Proposição Geral – O Modelo Conceitual proposto facilita a compreensão sobre as etapas necessárias à construção de aplicações analíticas educacionais e, por meio do *framework* que implementa essas etapas, possibilita o desenvolvimento de aplicações analíticas que apoiam a tomada de decisão educacional nos níveis micro, meso e macro.

A proposição geral proposta nesta tese é investigada a partir de outras cinco proposições específicas, conforme apresentado a seguir:

P1 – O Modelo Conceitual facilita o desenvolvimento de ferramentas analíticas que auxiliam na promoção da **aprendizagem autorregulada** de estudantes e melhoram a tomada de decisão em **nível micro**.

A *P1* é composta por cinco subproposições que melhor detalham o conceito de aprendizagem autorregulada. Essas subproposições, que vão de P1.A à P1.E, foram definidas a partir de estudos realizados com base na literatura existente de trabalhos tais como dos autores Pintrich (2000), Bandura (2001), Brockett e Hiemstra (1991), Hiemstra (1999) e Patterson, Lunyk-Child e Crooks (2002) e também com o suporte de professores do departamento de educação, em especial dos cursos de pedagogia e psicologia, e pedagogos da Secretaria de Educação a Distância (SEAD) da UFSCar, todos especialistas no assunto. Essas subproposições são apresentadas a seguir:

P1.A – A aplicação ajuda alunos a decidirem sobre o tempo de dedicação aos conteúdos de aprendizagem. (Gerenciamento do tempo de estudo)

P1.B – Informações apresentadas pelo sistema e recomendações personalizadas oferecidas pela aplicação auxiliam alunos a refletirem sobre sua própria aprendizagem. (Autorreflexão)

P1.C – A aplicação auxilia estudantes no monitoramento de seus estudos e de seu desempenho acadêmico continuamente. (Automonitoramento)

P1.D – Lacunas de aprendizagem podem ser identificados mais facilmente com apoio da aplicação, possibilitando aos estudantes a auto avaliação da aprendizagem individual e em relação a seus pares. (Autoavaliação)

P1.E – Informações e recomendações feitas pela aplicação auxiliam estudantes a tomarem decisões, especialmente em situações que necessitam de mudança de comportamento. (Autorreação)

P2 – O Modelo Conceitual proposto facilita o desenvolvimento de aplicações que atuam em **nível meso** e auxiliam instrutores (**tutores e professores**) a tomarem decisões sobre suas atividades docentes, durante e após o término de disciplinas.

P3 – O Modelo Conceitual proposto facilita o desenvolvimento de aplicações analíticas que apoiam a tomada de decisão educacional em **nível macro**, auxiliando **líderes e gestores** de projetos de EaD a **melhorarem estratégias educacionais** em longo prazo.

P4 – O **framework** proposto oferece um **caminho mais simples e unificado** para a **coleta distribuída** de dados a partir de diversas e variadas fontes e realiza análises sobre esses dados de forma satisfatória.

P5 – O **framework** proposto oferece **interfaces de programação que facilitam a construção de aplicações analíticas** sem a necessidade do desenvolvedor conhecer os detalhes de implementação das etapas de coleta e análise de dados.

1.5 Procedimentos Metodológicos

De maneira simplificada, o Quadro 1.1 mostra os principais procedimentos metodológicos utilizados na presente pesquisa.

Quadro 1.1 – Principais procedimentos metodológicos utilizados na pesquisa. Fonte: elaborado pela autora.

Item	Classificação
Abordagem	Qualitativa, exploratória e descritiva
Método de pesquisa	Estudo de caso
Instrumentos para coleta de dados	Questionários não-estruturados e questionários semiestruturados
Método de análise	<p>Estatística Descritiva foi aplicada para análise dos resultados obtidos nas entrevistas com alunos. As técnicas utilizadas foram: medidas de tendência central, medidas de dispersão e medidas de associação (coeficiente da correlação de <i>Spearman</i>).</p> <p>Análise de Conteúdo foi aplicada sobre resultados obtidos nas entrevistas com instrutores, gestores educacionais, desenvolvedores de ferramentas educacionais e pesquisadores com experiência na área desta tese.</p>

Fonte: elaborado pela autora.

1.5.1 Abordagem e classificação da pesquisa

A abordagem de pesquisa realizada neste trabalho é qualitativa, exploratória e descritiva, uma vez que o objetivo não visa à medição numérica para testar hipóteses em seu processo de interpretação, mas sim a exploração e o entendimento das várias opiniões e perspectivas dos usuários do sistema a partir da observação das variáveis. Nesta pesquisa considera-se a interpretação subjetiva dos indivíduos, múltiplas fontes de evidências e a proximidade da pesquisadora com o fenômeno estudado, características tais que evidenciam a pesquisa qualitativa segundo Martins (2010).

De acordo com GIL (1999), pesquisas podem ser classificadas em três categorias: exploratórias, descritivas e explicativas. Neste sentido, a presente pesquisa se desenvolveu em três etapas: a primeira, de caráter exploratório, proporcionou maior entendimento sobre os processos realizados em ensino e aprendizagem em ambientes virtuais, com o objetivo de, por meio de pesquisas bibliográficas a partir de artigos científicos, teses e livros, por meio de entrevistas realizadas com os agentes envolvidos no processo (estudante, tutores, professores e técnicos de sistema) e pela experiência da executora desta pesquisa e de colaborações de especialistas da área de domínio, coletar informações sobre a necessidade dos usuários e levantar requisitos que servissem de subsídios para o desenvolvimento do Modelo proposto, bem como do *framework* e das aplicações e posterior formulação dos instrumentos de coleta de dados aplicados à amostra selecionada.

A segunda etapa da pesquisa consistiu da criação do Modelo Conceitual proposto e sua aplicação na orientação do desenvolvimento do *framework* de coleta de dados e raciocínio e das três aplicações desenvolvidas no escopo deste trabalho a partir do *framework*.

A terceira etapa da pesquisa, de caráter descritivo, possibilitou o estudo, a análise e a interpretação da percepção dos usuários sobre aplicações desenvolvidas e sobre o Modelo proposto.

As etapas da pesquisa são escritas em detalhes na Seção 1.5.5 deste documento.

1.5.2 Método de pesquisa e instrumentos para coleta de dados

O método de pesquisa utilizado neste trabalho é o estudo de caso. De acordo com Miguel (2010) o estudo de caso pode ser visto como uma abordagem de caráter empírico que investiga um dado fenômeno dentro de um contexto real contemporâneo e possibilita profundo conhecimento sobre o mesmo. Além disso, o método escolhido é indicado quando o pesquisador busca responder questões do tipo “Como?” e “Por quê?” e baseia-se na análise de um limitado número de casos em que, no máximo, aplica-se uma limitada análise estatística (YIN, 2009).

A coleta de evidências em estudos de caso pode ser feita por meio entrevistas (estruturadas, semiestruturadas ou não-estruturadas), análise

documental e observação direta (YIN, 2001). De acordo com Martins (2010), os instrumentos mais indicados para coleta de dados em pesquisa qualitativa são as entrevistas semiestruturadas ou entrevistas não estruturadas desenvolvidas a partir de um modelo. Segundo o mesmo autor, instrumentos estruturados permitem ao pesquisador impor sua opinião sobre o problema de pesquisa, dificultando a observação da perspectiva dos participantes.

A fim de garantir o uso controlado do método científico e permitir certo grau de liberdade que possibilita alcançar os objetivos de investigação e ao mesmo tempo não ultrapassar os limites da ciência, neste trabalho houve predominância pelo uso de entrevistas semiestruturadas para realizar as entrevistas com os usuários do sistema. Entrevistas não-estruturadas também foram utilizadas para avaliação e validação do Modelo Conceitual pelos especialistas de domínio. Além disso, com o objetivo de manter a confiabilidade dos dados coletados, outras fontes de evidências foram adotadas em menor escala, conforme sugere Martins (2010): observações pessoais e documentos disponibilizados pela instituição de ensino, tais como registros de acessos, notas dos alunos, relatórios disponíveis no ambiente virtual e relatórios gerados pelo sistema desenvolvido no escopo deste trabalho. A utilização desses três tipos de fontes de evidências possibilita a técnica de triangulação dos dados, observando a convergência e divergência entre os mesmos. Os questionários aplicados aos participantes são apresentados nos Apêndices de B a F.

1.5.3 Amostra e seleção dos participantes

O método de amostragem selecionado foi o não probabilístico. Segundo Stevenson, 2001, nesse método não é possível estabelecer a variabilidade amostral com precisão, logo não é possível estimar o erro amostral. A principal característica desse tipo de amostragem é não fazer uso de formas aleatórias de seleção, mas sim selecionar os participantes por meio de critérios subjetivos do pesquisador (LAKATOS e MARCONI, 1989; GIL, 1999, MARÔCO, 2011).

A escolha do método de amostragem não probabilístico se justifica quando a população não está disponível para ser selecionada aleatoriamente (MARÔCO, 2011). Segundo os mesmos autores, a limitação de tempo e recursos (financeiros e pessoas) necessários para a realização de uma pesquisa com amostragem

probabilística também justificam a escolha do método. Neste trabalho, encontraram-se essas dificuldades mencionadas, o que direcionou a escolha do método de amostragem não probabilístico.

Para realização deste estudo, a seleção da amostra considerou a facilidade de acesso aos participantes. Outros fatores que contribuíram para o procedimento de amostragem foram a disponibilidade e a concordância dos participantes em participar da presente pesquisa.

As entrevistas foram realizadas com 22 estudantes, 2 tutores, 2 professores, 3 gestores educacionais, 2 desenvolvedores de ferramentas de apoio à EaD e 2 especialistas na área de pesquisa desta tese.

1.5.4 Análise dos dados

Neste estudo foram empregadas duas formas de análise das evidências, a análise de conteúdo e a análise descritiva. A análise de conteúdo pode ser utilizada como técnica de pesquisa para interpretar conteúdos textuais (BARDIN, 2002). De acordo com Mann (1995), a estatística descritiva usa técnicas estatísticas para descrever e sumarizar um conjunto de dados. Para análise descritiva dos dados, o comportamento das variáveis foi estudado a partir da análise de suas ocorrências. A fim de caracterizar a amostra em estudo, foram utilizadas as técnicas de medidas de tendência central que buscam caracterizar variáveis com mais frequência e medidas de dispersão das observações em torno das estatísticas de tendência central. Finalmente, o coeficiente de correlação de *Spearman* (FIGUEIREDO FILHO e SILVA JUNIOR, 2009) permitiu medir a associação que caracteriza a intensidade e a direção da variação comum entre algumas das variáveis observadas.

As análises estatísticas foram realizadas com auxílio do software *IBM Statistical Package for Social Sciences* (SPSS), versão 20. Os dados foram exportados para o *Microsoft Excel*, versão 2013, a fim de obter gráficos com melhor visualização.

1.5.5 Descrição das atividades metodológicas e observações pertinentes

Para atingir os objetivos propostos nesta tese, a abordagem metodológica inclui as seguintes atividades:

Fase 1: Estudos sobre o estado da arte

Inicialmente, foram realizados estudos sobre o estado da arte no desenvolvimento de tecnologias para enriquecer a aprendizagem *online* a fim de analisar os sistemas existentes, identificar aspectos que dificultam sua utilização e investigar e avaliar a arquitetura de *software* utilizada e o tipo de modelagem de contexto adotado.

O resultado dessa exploração, realizada a partir da revisão da literatura, possibilitou duas conclusões principais: os sistemas existentes não fazem, de maneira genérica e completa, a captura de contextos que atenda aos diferentes propósitos da análise de dados, mas limitam-se ao propósito da aplicação para o qual foi desenvolvida; e a modelagem de contextos, frequentemente, consiste em uma das camadas da arquitetura, estando intimamente relacionada ao propósito da aplicação ou, no máximo, ao conjunto de aplicações no mesmo escopo (tais como recomendações, monitoramento da aprendizagem ou avaliação da aprendizagem). A partir de tais conclusões foi identificado um conjunto de características importantes para a construção de uma arquitetura de *software* que suprisse às limitações identificadas: arquitetura aberta e extensível, possibilitando incorporação de novas fontes de dados; compartilhamento de contexto; e a separação entre a coleta dados e a análise, o que possibilita a utilização de diferentes modelagens de contextos e uso de diferentes técnicas de raciocínio.

Após essa investigação, foi realizada uma revisão da literatura na área de engenharia de *software* pela busca de estruturas que viabilizassem o desenvolvimento do sistema proposto. Esse estudo direcionou a escolha por *framework* por possibilitar atender, mais amplamente, as características identificadas.

Além disso, outra constatação ainda mais importante foi identificada durante esta fase da pesquisa, a carência de ferramentas tecnológicas que apoiam a tomada de decisão educacional em múltiplos níveis e, principalmente, a carência de um modelo conceitual genérico que facilita o entendimento das etapas necessárias ao desenvolvimento de tais ferramentas. Esta averiguação possibilitou o redirecionamento do trabalho priorizando a criação do Modelo Conceitual de apoio à tomada de decisão nos níveis micro, meso e macro, conforme mencionado anteriormente.

Considerando as constatações feitas sobre o estado da arte, novo estudo foi realizado a partir da literatura existente sobre metodologias, abordagens, modelos e arquiteturas de *software* para suporte à tomada de decisão dirigida a dados em múltiplos níveis, o que possibilitou identificar uma lacuna na área dada a carência de tais soluções.

Fase 2: Definição do Modelo Conceitual para o desenvolvimento de ferramentas analíticas educacionais e definição da arquitetura de *software* para implementação do Modelo

Partindo dos estudos realizados na fase 1, iniciou-se as atividades de definição do Modelo Conceitual proposto para apoiar o desenvolvimento de ferramentas analíticas educacionais e também a escolha da arquitetura de *software* para implementação desse Modelo.

Estudos a partir da revisão da literatura sobre arquiteturas de *software*; de discussões realizadas entre a pesquisadora, membros do grupo de pesquisa e a orientadora; e discussões com desenvolvedores de *software* possibilitaram a escolha da estrutura (*framework*) a ser utilizada para implementação do Modelo proposto nesta tese.

Dadas a escolha da estrutura de *software* e a criação do Modelo Conceitual para guiar o seu desenvolvimento, iniciaram-se as atividades de definição das camadas da arquitetura, a especificação e a implementação de acordo com as características identificadas anteriormente. Nesta fase da pesquisa também foram identificados e adotados padrões de projeto que facilitaram a adição de novas funcionalidades ao *framework*: *abstract factory* para manter a consistência do conjunto de sensores de acordo com a plataforma subjacente, *singleton* para garantir a criação de somente uma *factory* durante o ciclo de vida dos sensores em uma plataforma particular e *strategy* para isolar as diferentes implementações de sensores de acordo com a plataforma.

Fase 3: Desenvolvimento de aplicações a partir do *framework* proposto na fase 2

Três aplicações educacionais analíticas foram desenvolvidas a partir do *framework* proposto: ViTrackeR – sistema de recomendação, ViMonitor – sistema de visualização de dados, e ViAssess – sistema de segurança para avaliações formativas *online*. Cada aplicação levou em consideração a investigação realizada na fase 1 deste trabalho, sobre a modelagem de contextos, que permitiu a escolha das técnicas de modelagem e, conseqüentemente, das técnicas de inferência utilizadas.

Fase 4: Avaliação do Modelo Conceitual proposto, do *framework* e das aplicações

Como atividade final deste trabalho, uma avaliação foi conduzida com o objetivo de avaliar o Modelo Conceitual proposto nesta tese, avaliar o *framework* e também avaliar as aplicações desenvolvidas. As avaliações foram realizadas a partir da percepção de trinta e três participantes deste estudo e tiveram como objetivo principal a validação do Modelo proposto.

Esta fase da pesquisa envolveu a coleta de dados primários obtidos por meio de entrevistas com aplicação de questionários semiestruturados ou não-estruturados, análise dos dados obtidos e elaboração das discussões sobre os resultados obtidos. Os resultados das avaliações realizadas são apresentados no capítulo 7 deste documento.

1.6 Estrutura e Organização do Trabalho

A estrutura do trabalho está dividida em sete capítulos, sendo que o primeiro deles consistiu em apresentar a introdução contendo, as justificativas do trabalho, o problema de pesquisa desta tese, a solução proposta e contribuições, os objetivos, proposições de pesquisa e procedimentos metodológicos adotados.

O segundo capítulo apresenta os fundamentos conceituais utilizados no trabalho, e inclui aprendizagem assistida por tecnologias; contexto e técnicas de modelagem contextos; análises de dados e análise da aprendizagem. O capítulo três apresenta o estado da arte em *Learning Analytics* e aborda as principais pesquisas sobre ferramentas analíticas educacionais de visualização de dados, incluindo avaliação e *feedback* e monitoramento e tutoria; recomendações; e personalização e adaptação.

No capítulo quatro é apresentado o Modelo Conceitual proposto neste trabalho em detalhes, incluindo suas etapas: Artefatos; Coleta e Pré-Processamento de Dados; as fases de Análises dos Dados, abrangendo análise descritiva, análise preditiva e análise prescritiva e a camada de Decisão.

A implementação do Modelo Conceitual é apresentada no capítulo cinco, que contempla as principais contribuições do *Framework* e suas camadas, incluindo: a camada de Coleta de Dados, a camada de Gerenciamento dos Dados, a camada de Modelagens de Contextos e Análises e a camada de Aplicações. Questões sobre a privacidade dos Dados também são abordadas. Neste mesmo capítulo cinco também são apresentadas as aplicações (ViTrackeR, ViMonitor e ViAssess) desenvolvidas para avaliação do *Framework* proposto.

O sexto capítulo apresenta as avaliações conduzidas para validação do Modelo Conceitual proposto e os Resultados da pesquisa. No sétimo e último capítulo são apresentadas as considerações finais desta tese, incluindo o atendimento à questão de pesquisa, aos objetivos e às proposições propostas; as limitações da pesquisa; e, os trabalhos futuros.

Concluídas a contextualização do problema, da solução proposta e a apresentação dos elementos que delimitaram esta pesquisa, o capítulo subsequente consiste na descrição dos fundamentos conceituais utilizados no trabalho.

Capítulo 2

FUNDAMENTOS CONCEITUAIS

“In the twentieth century, the dominant approach to education focused on helping students to build stocks of knowledge and cognitive skills that could be deployed later in appropriate situations. This approach to education worked well in a relatively stable, slowly changing world in which careers typically lasted a lifetime. But the twenty-first century is quite different. The world is evolving at an increasing pace. When jobs change, as they are likely to do, we can no longer expect to send someone back to school to be retrained. By the time that happens, the domain of inquiry is likely to have morphed yet again.” (BROWN E ADLER, 2008, p. 30).

A Seção 2.1 apresenta as considerações iniciais; a Seção 2.2 apresenta conceitos envolvidos na aprendizagem assistida por tecnologias; na Seção 2.3 é apresentada a definição de contexto e as técnicas de modelagem utilizadas; a Seção 2.4 apresenta o conceito de Analytics, com ênfase em Learning Analytics (LA). A Seção 2.5 descreve sobre a importância dos dados em LA; e, por fim, a Seção 2.6 apresenta as técnicas de fusão de dados para tratamento de dados imperfeitos.

2.1 Considerações Iniciais

A tecnologia iniciou muitas mudanças em diferentes aspectos da vida humana. Ensino e aprendizagem são atividades humanas representativas que têm sido objeto dessas mudanças, uma vez que a forma como as pessoas ensinam, aprendem, colaboram e decidem são diretamente influenciadas pelas tecnologias.

Este cenário rico em tecnologias onde o processo de ensino e aprendizagem ocorre, assim como a gestão educacional, agrega diversos novos conceitos e novas abordagens para o suporte à tomada de decisão nesse processo, tais como

aprendizagem móvel e ubíqua, ambientes de aprendizagem pessoais, ecossistema de aprendizagem, *learning analytics* e *big data*. Na mesma direção, a noção de contexto também emerge como componente importante para ser considerado no desenvolvimento de tecnologias para o suporte às decisões em nível micro, nível meso e nível macro.

Este capítulo apresenta os principais conceitos envolvidos na área de análise de dados para apoio às decisões educacionais, incluindo a captura de contextos de usuários, técnicas de modelagens de dados, técnicas de tratamento para dados imperfeitos e métodos de análises dos dados.

2.2 Aprendizagem Assistida por Tecnologias

Aprendizagem Assistida por Tecnologias, do inglês *Technology Enhanced Learning* (TEL), objetiva o projeto, desenvolvimento e inovações técnico-sociais que suportam práticas de ensino e aprendizagem de indivíduos e organizações (MANOUSELIS et al., 2011). Desta forma, o termo TEL é usado para descrever a aplicação das tecnologias de informação e comunicação para o ensino e aprendizagem (KIRKWOOD e PRICE, 2014).

Inicialmente, TEL esteve relacionada apenas à entrega de conteúdos de aprendizagem baseada em computador, geralmente pelo uso de ambientes virtuais de aprendizagem (AVA). Esses AVA, também denominados sistemas gerenciadores da aprendizagem (SGA), ou *learning management systems* (LMS), suportam a entrega de conteúdos *online* estaticamente, seguindo um padrão de modularização dos cursos em unidades discretas (WILSON, SWITZER e PARRISH, 2007). Os SGA constituem-se em plataformas para armazenamento e disponibilização (na *web*) de conteúdos digitais de estudo e para o acompanhamento da aprendizagem a distância (RYANN, 2009). Exemplos de SGA são: Moodle⁷, Sakai⁸ e TIDIA-Ae⁹ e

⁷ Moodle: *Modular Object Oriented Dynamic Learning Environment*. <http://www.moodle.org/>.

⁸ <http://sakaiproject.org/>.

Blackboard¹⁰. Essa tecnologia possibilita que um curso baseado em papel seja convertido em formato digital e oferecido *online* (CHATTI et al., 2010), sem as adaptações necessárias e a criação de recursos adequados que satisfaçam as necessidades heterogêneas dos estudantes. De acordo com o mesmo autor, esses modelos de aprendizagem *top-down* centrados no SGA estão fadados ao fracasso, uma vez que a aprendizagem é mais do que apenas conteúdos estáticos e a tecnologia é apenas uma questão secundária. Neste sentido, novos conceitos são incorporados à TEL para que a educação a distância atenda desafios de aprendizagem distribuída, flexível, ubíqua, social e em rede.

A princípio, sistemas de aprendizagem, onde a colaboração é mediada por computadores, foram denominados como sistemas de *e-learning* (*electronic learning*). Mais tarde, com a introdução do conceito de mobilidade dos usuários e o desenvolvimento de tecnologias móveis com características de *desktop* e com acesso à internet de banda larga para o suporte da aprendizagem, o conceito de TEL evoluiu para *m-learning* (*mobile learning*) (XIA et al., 2013). Este novo contexto aliado ao acesso de conteúdos de qualquer lugar e a qualquer momento possibilitou experiências de aprendizagem em diferentes situações, e não apenas em ambientes escolares. Neste novo paradigma, a TEL suportada por tecnologias *wireless* e ubíqua é conhecida como *u-learning* (*ubiquitous learning*) (ROGERS et al., 2005; SYVÄNEN et al., 2005). Neste mesmo contexto, outros autores (PLYMALE, 2005; THOMAS, 2006; RYU, 2007) se referem à evolução da TEL como *p-learning* (*pervasive learning*).

Neste trabalho, o termo *Technology-Enhanced Learning*, ou mesmo sua forma abreviada TEL, é usado para descrever os diferentes pontos de vista, conforme mencionados anteriormente. A escolha do termo deve-se ao fato de TEL ser uma terminologia mais abrangente e por ignorar a ligação a uma única tecnologia existente, favorecendo a utilização de conceitos relacionados ambientes de aprendizagem pessoais e à ciência de contexto que serão abordados no decorrer do texto.

⁹ TIDIA-AE: *Tecnologia da Informação no Desenvolvimento da Internet Avançada - Aprendizado Eletrônico*. Projeto financiado pela Fapesp que destina-se à pesquisa e desenvolvimento de Ensino a Distância, com o objetivo de desenvolver ferramentas de código livre, sendo amparado por redes de alta velocidade.

¹⁰ <http://blackboard.grupoa.com.br/>.

2.2.1 Ambientes Pessoais de Aprendizagem

O termo ambientes de aprendizagem pessoais, ou *Personal Learning Environment* (PLE), foi utilizado pela primeira vez por Olivier e Liber (2001) para designar um modelo *peer-to-peer* centrado na aprendizagem pessoal e ao longo da vida (*lifelong learning*). Na visão de Siemens (2008) PLEs surgiram em resposta ao desenvolvimento de tecnologias centradas na interação social e na colaboração.

De acordo com Van Harmelen (2008), PLE são definidos como sistemas que auxiliam aprendizes a controlar e gerenciar sua própria aprendizagem com o objetivo de suprir falhas dos SGA tradicionais no suporte à aprendizagem centrada no estudante com os desafios apontados tais como ubiquidade, mobilidade, personalização, flexibilidade, dinamicidade e aprendizagem de natureza complexa.

Van Harmelen (2006) aponta algumas necessidades que fundamentam os PLE como a aprendizagem ao longo da vida, onde os aprendizes podem acessar uma interface padrão com diversos sistemas de *e-learning* de diferentes instituições; o atendimento às abordagens pedagógicas onde o controle da aprendizagem é do aprendiz; e o atendimento às necessidades individuais dos estudantes. Attwell e Costa (2008) abordam o conceito de PLE com enfoque na relevância da aprendizagem informal, ao longo da vida e na autonomia do estudante na construção de conhecimentos e organização da sua própria aprendizagem.

Pesquisas bibliográficas realizadas durante o desenvolvimento deste trabalho mostraram que muitas propostas de PLE foram radicais na proposição de novas abordagens para suporte à aprendizagem, onde sugeriram a substituição dos SGA tradicionais por novas ferramentas que oferecessem mais liberdade para o estudante na escolha e controle da aprendizagem, conforme visto em Jafari, Mcgee e Carmean (2006) e Chatti et al. (2010). Blackall (2005) também aborda o conceito nesta linha e destitui qualquer abordagem estruturada e não individual de qualquer ambiente de aprendizagem.

Embora muita pesquisa tenha sido gerada sobre o conceito de PLE, no presente trabalho os SGA são abordados como ferramentas importantes para aprendizagem dirigida a tecnologias tendo recebido melhorias incrementais e incorporação de novas ferramentas *web 2.0* que apoiam a aprendizagem centrada no aluno e a tomada de decisão no nível individual, de curso e institucional. Neste sentido, este trabalho trata PLE e SGA de maneira complementar, justificando-se

pelo fato de que a dissociação dos termos e escolha pela adoção de abordagens completamente não estruturadas exige novas arquiteturas e soluções tecnológicas ainda não estabelecidas e generalizadas atualmente, inclusive não disponíveis nas instituições de ensino brasileiras, o que inviabilizaria toda a pesquisa. Além do mais, vários desafios associados aos SGA têm limitado sua execução, especialmente em nível institucional, tais como a complexidade e o aumento no custo de treinamento e suporte necessários para uso de diversas tecnologias, maior dificuldade de gerenciar problemas de privacidade, confiabilidade do sistema e continuidade dos dados, conforme aponta (MOTT, 2010). Além disso, o mesmo autor aponta que a natureza “livre” de aplicativos baseados na *web* oferecem poucas alternativas de negociação quando há degradação de desempenho ou quando dados são expostos ou perdidos, por exemplo, por acidente, ou quando há interrupções de serviços ou descontinuidade de ferramentas.

Considerando o exposto, abordagens pedagógicas atuais para educação *online* não devem extinguir o SGA como ferramenta de interação entre os atores (professor, tutor, estudante), mas sim associá-la às inovações da *web 2.0*, o que permite a combinação de dados e serviços a partir de diferentes fontes a fim de oferecer aprendizagem mais personalizada, flexível e ubíqua. Para este fim, a coleta de contextos em ambientes TEL deve considerar diversas fontes, incluindo sensores físicos e lógicos, *softwares* e dispositivos móveis.

Os sensores físicos podem ser sensores de ambiente, que coletam presença, identificação, ruído, etc., ou podem ser sensores de dispositivo, que coletam dados tais como a identificação do usuário, localização, largura de banda disponível, configurações do dispositivo e mobilidade. Um sensor lógico é um componente de *software* (microagente), baseado em *desktop*, *web* ou dispositivos móveis, que pode capturar dados relevantes e específicos sobre o usuário, suas ações e interações, em um ambiente digital (ROSALES et al., 2012), tais como suas tarefas, progresso do trabalho, preferências, relacionamentos em redes sociais, horários de estudo, etc.

A captura de contextos a partir de sensores, físicos e/ou lógicos, permite às aplicações tomarem conhecimento sobre as ações do usuário de forma transparente para este, ou seja, a captura acontece a partir das ações naturais executadas pelo usuário, sem a entrada de dados no sistema de maneira explícita. Exemplos de contextos incluem: atividade que o usuário está fazendo, o tempo gasto para o desenvolvimento de atividades e testes, número de *posts* de fóruns gerados, lidos e

acessados, caminho de navegação, pausas, interação social online (tipos de comunidades frequentadas, tipo de conteúdo em foco, tempo gasto com interações sociais), entre outros. Na próxima Seção, a importância do contexto é descrita a partir do ponto de vista tecnológico e social.

2.3 Contexto

TEL é enriquecida pela adição de informações de contexto, onde sistemas adaptam seus serviços ao contexto do usuário. A necessidade de compreender contexto é importante para o projeto e desenvolvimento de ambientes TEL (KAENAMPORN PAN e O'NEILL, 2005). Além da sua implicação em aprendizagem, de maneira geral contexto é importante para o projeto e desenvolvimento de novas tecnologias ubíquas, estando associado à criação de novos caminhos de interação de usuários (DOURISH, 2004), o que inclui aprendizes, tutores, professores e gestores educacionais.

Neste sentido, a visão técnica de contexto mais amplamente aceita é definida e apresentada por (DEY e ABOWD, 1999), onde: contexto é *qualquer informação que pode ser usada para caracterizar a situação de uma entidade. Uma entidade é uma pessoa, lugar ou objeto que é considerado relevante para interação entre o usuário e uma aplicação, incluindo os próprios usuário e aplicação*. Para Dourish (2004) esta visão técnica de contexto é percebida como novos caminhos para conceituar a ação humana e a relação entre a ação e o sistema computacional.

Outras visões apresentam o contexto como entidade dinâmica e uma visão mais social. Em Syvänen et al. (2005), a noção de contexto é vista não como uma concha que envolve a aprendizagem em um dado tempo e local, mas sim como uma entidade dinâmica, construída pelas interações entre os aprendizes e seus ambientes.

Tipicamente, as definições para contexto são apresentadas de modo que possam facilitar a implementação de aplicações em seu escopo de atuação. No geral, os contextos são definidos de acordo com o domínio e tipo de aplicação ao qual estão sendo empregados. Contextos podem incluir informações sobre *usuários* (preferências, localização e direção, situação, etc.); os *dispositivos* utilizados no

ambiente de interação (disponibilidade, capacidade, etc.); e, o próprio *ambiente* (capacidade, tempo de espera, usuários conectados e tempo conectado).

Alguns contextos podem ser obtidos a partir da combinação de outros ou a partir da inferência realizada durante sua interpretação. Desta forma, contextos podem ser classificados em estáticos ou dinâmicos (HENRICKSEN, INDULSKA e RAKOTONIRAINY, 2002). Os contextos estáticos são aqueles que não sofrem alterações, tais como: identificação de um usuário (por exemplo registro de aluno - RA), CPF, etc. Contextos dinâmicos são todos os demais dados e informações do sistema, ou seja, aqueles que podem ser alterados. De acordo com (HENRICKSEN, INDULSKA e RAKOTONIRAINY, 2002), os contextos dinâmicos podem ser reclassificados como: *sentidos* – aqueles que são capturados diretamente por meio de sensores; *derivados* – aqueles que são derivados a partir de interpretações e inferências realizadas; ou *explícitos* – aqueles que são determinados explicitamente pelo usuário, tais como senha de acesso, dados da agenda, etc.

2.3.1 Modelagem de Contextos

Conforme mencionado anteriormente, uma aplicação pode desejar tanto os dados brutos, normalmente providos por sensores, quanto informações mais abstratas e elaboradas, provenientes de um processo de refinamento e interpretação (HENRICKSEN, INDULSKA e RAKOTONIRAINY, 2002).

Desta forma, a interpretação de contexto pode ser vista como sendo o conjunto de métodos e processos envolvendo a manipulação, a agregação, a derivação, a inferência e o refinamento sobre as informações contextuais, com o propósito de aumentar o nível de abstração dessas informações ou mesmo de compor novas informações contextuais derivadas, de modo a facilitar o entendimento de um determinado contexto pelas aplicações e auxiliá-las em seus processos de tomada de decisão. Conforme afirma (STRANG e LINNHOFF-POPIEN, 2004), a modelagem adequada dos contextos é a chave principal para um bom sistema ciente de contexto.

De acordo com Bettini, Pareschi, e Riboni (2008), a modelagem formal de contextos auxilia o desenvolvimento do sistema, principalmente porque possibilita formalização e organização do conhecimento do sistema separadamente à

interpretação dos dados, promovendo sua manutenção e evolução. Além disso, é possível tornar o sistema mais extensível a partir do reuso e compartilhamento de contextos entre aplicações cientes de contexto desenvolvidas para o mesmo domínio.

2.3.1.1 Abordagens para modelagem de contextos

As abordagens para modelagem de contextos incluem os seguintes modelos: baseado em chave-valor, esquemas de marcação, orientados a objetos, baseado em lógica, baseado em ontologias e com uso de situação lógica e modelagem preditiva (SCHILIT, ADAMS e WANT, 1994), (RYAN, 1999), (CHEN e KOTZ, 2000), (BETTINI, PARESCHI, e RIBONI, 2008), (HOAREAU e SATOH, 2009), (GUAZZELLI, 2012) e (SMITH, LANGE e HUSTON, 2012).

Modelagem Chave-Valor

Este tipo de modelagem usa o par *chave-valor* para definir uma lista de atributos e seus valores para descrever os contextos usados em aplicações cientes de contexto. A *chave* (ou identificação) consiste no atributo do ambiente e deve ter um valor associado a ela. Normalmente, esse valor é mensurado por sensores a partir da coleta realizada no ambiente.

O modelo chave-valor foi utilizado por Schilit, Adams e Want (1994) para gerenciar informações de localização. Este modelo é um dos mais simples e não permite a representação de contextos complexos e derivados. Apesar dos problemas com sua expressividade, este modelo pode ser mantido para sistemas distribuídos com grande quantidade de computadores e baixo acoplamento.

Exemplos de uso do modelo *chave-valor*, (SCHILIT e THEIMER, 1994):

<Room 31 temperature, 24>

<Disciplina 043 número_estudantes, 87>

Modelagem de Marcação

O modelo baseado em *esquemas de marcação (markup)* utiliza várias linguagens de marcação, principalmente *XML*. Inicialmente, este modelo foi proposto para suportar troca de mensagens entre *Clientes Móveis X Servidor* para oferecer serviços personalizados e rastreamento de usuários (onde? quando? velocidade?), (RYAN, 1999).

O modelo baseado em **linguagem de marcação** tem um funcionamento padrão que inicia com a solicitação de uma nova sessão, o seu uso e a finalização da sessão e funciona da seguinte forma, (RYAN, 1999):

i. solicita nova sessão:

```
<context session = "new"> </context>
```

ii. inicia-se nova sessão (questões de segurança à parte):

```
<context session = "123" /context>
```

iii. após toda a transmissão, finaliza a sessão:

```
<context session = "123" action = "close" />
```

De acordo com (STRANG e LINNHOFF-POPIEN, 2004) e (INDULSKA et al., 2003) o modelo baseado em marcações apresenta limitações para representação de contextos complexos e não é apropriado para representação de contextos genéricos.

Modelagem Orientada a Objetos

A modelagem orientado a objetos utiliza representação baseada no conceito de objetos e seus relacionamentos a partir do paradigma de orientação a objetivos (BAUER, 2003). Neste modelo, os objetos representam os contextos (temperatura, localização, identificação, etc.) e os relacionamentos existentes entre eles são descritos a partir do conceito de relacionamento convencional da orientação a objetos; as operações usadas para processamento dos contextos são encapsuladas pelos objetos. A principal linguagem adotada na modelagem de contextos nesta abordagem é a *unified modeling language* (UML).

Um exemplo de uso deste modelo de representação de contextos é apresentado a seguir, na Figura 2.1, para o monitoramento da turbulência e risco de colisão de aeronaves em (BAUER, 2003). No sistema proposto, o *contexto turbulência* apresenta o índice de turbulência (baseado em valores médios e pico) em uma faixa de 1 a 27 e é classificado como *light*, *moderate*, *severe*. Além de mostrar os índices, o sistema apresenta serviços (informação de vôo, alerta) que são os dados isolados. O controlador de seção, ou controlador de vôo, decide pela solução em parceria com o piloto que pode opinar. O *contexto colisão* é composto por dados relevantes sobre colisão com outra aeronave, tais como: sobreposição de rota e vôo mais baixo que o permitido.

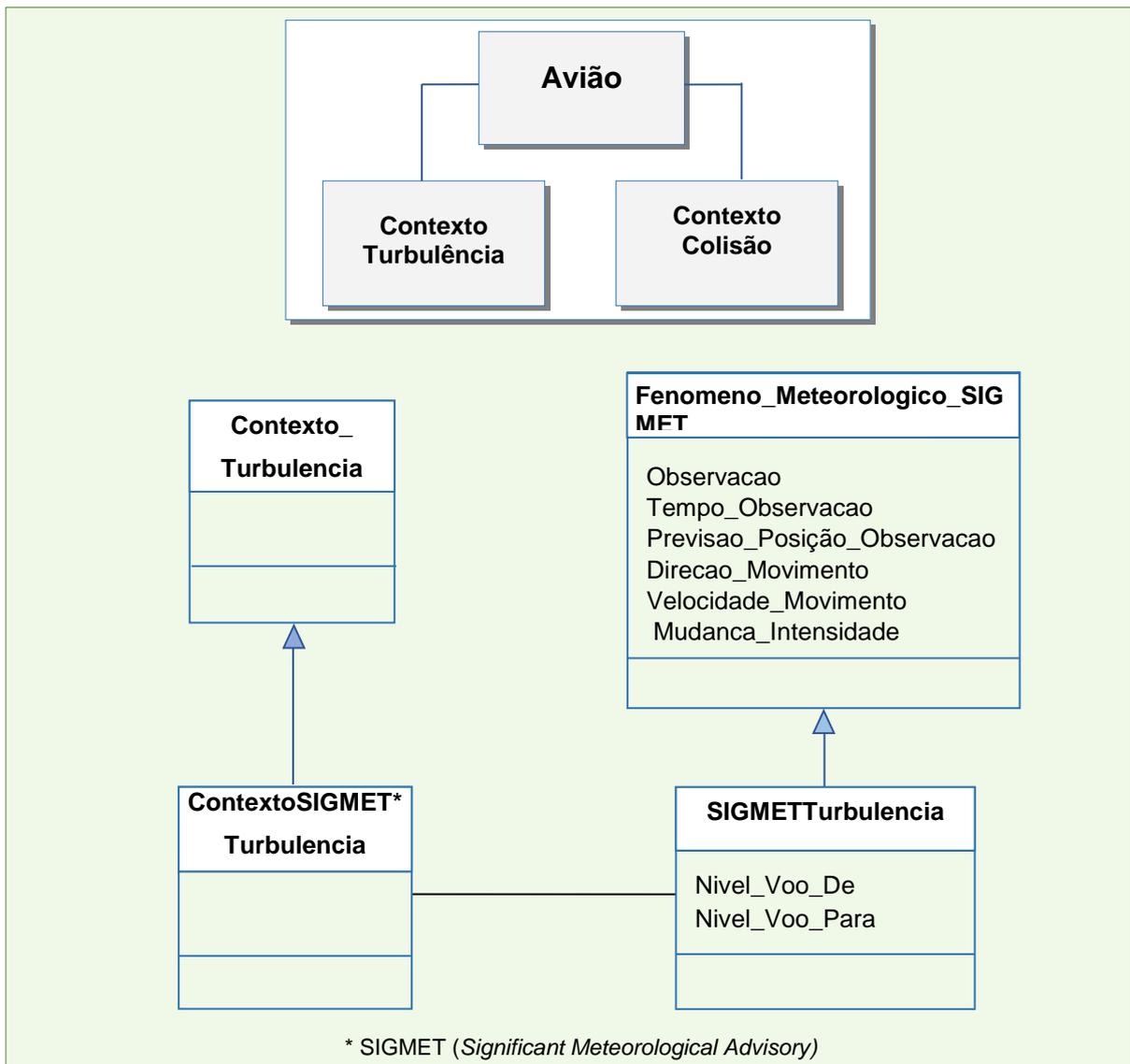


Figura 2.1 – Contexto Turbulência em Aeronaves. Fonte: adaptado de (BAUER, 2003).

A principal desvantagem do modelo orientado a objetos é que ele não permite caracterizar completamente os contextos só com uso de diagramas, é preciso da linguagem humana para detalhar as descrições. Além disso, detalhes da informação de contexto sobrecarregam o diagrama.

Modelagem Baseada em Lógica

A modelagem de contextos baseado em lógica utiliza fatos, expressões e regras (regras de primeira ordem, programação baseada em regras, etc.) para definir um modelo de contexto com alto nível de formalismo. Neste caso, o contexto é representado por fatos, expressões ou regras. Ranganathan et al. (2003), Karsiri e Mycroft (2003) e Henrichsen, Indulska e Rakotonirainy (2002) aplicam lógica de

primeira ordem para interpretação de informações de contexto. Já no trabalho de Sohn (2003), o autor propõe um *toolkit* para usuários finais desenvolverem aplicações cientes de contexto a partir da especificação de regras de primeira ordem.

A lógica define as condições em que uma expressão ou fato pode ser derivado de outras expressões ou fatos. Esse processo mostra como inferir a partir de um conjunto de regras utilizando um sistema formal. A maioria das abordagens utiliza Prolog.

Exemplos de modelagem baseado em lógica de primeira ordem são os seguintes (AKMAN e MEHMET, 1996):

- i. $ist(c,p) - is\ TRUE -$ quais as afirmações que podem ser feitas para que a proposição p seja VERDADEIRA no contexto C .
- ii. $c0: ist(contextof("Sherlock\ Holmes\ stories"), "Holmes\ is\ a\ detective") -$ É verdade que nas histórias de Sherlock Holmes, o Holmes é um detetive?

De maneira geral, o uso de programação baseada em regras tem provado ser intuitivo e bem sucedido para modelagem de contextos em aplicações cientes de contexto (HOAREAU e SATOH, 2009).

Modelagem Baseada em Ontologias

O uso de ontologias é fortemente indicado para a modelagem de contextos em *web* semântica (CHEN, FINN, e JOSHI 2003). As técnicas baseadas em ontologias suportam um vocabulário de predicados com semântica adicional para os contextos recuperados por sensores (LOKE, 2004), e a representação de dados nesses projetos pode ser compartilhada por diferentes sistemas ou podem permitir a recuperação de dados a partir de um armazenamento na *web*.

Com uso de ontologias, os contextos podem ser modelados a partir de preferências, características e perfis dos usuários. Além disso, novos contextos podem ser deduzidos a partir dos que já são conhecidos.

De acordo com (YU, ZHOU e SHU, 2009), ontologias possibilitam a descrição e modelagem de contextos em um caminho semântico, com características próprias da tecnologia de ontologias, tais como expressividade, extensibilidade, facilidade de compartilhamento e reuso, suporte à interpretação lógica e interoperabilidade entre serviços e aplicações.

O termo *ontologia*, quando empregado na representação do conhecimento, refere-se à descrição formal e explícita de conceitos, os quais são frequentemente concebidos como um conjunto de entidades, relacionamentos, instâncias, funções e axiomas (GRUBER, 1993).

Ambientes virtuais cientes de contexto devem adaptar suas características de acordo com as mudanças de contexto. Além disso, os espaços físicos de aprendizagem podem ser enriquecidos com diversos dispositivos heterogêneos que produzem informações de contexto. Desta forma, a construção de sistemas TEL deve prever aquisição de dados a partir de sensores físicos e/ou lógicos, agregação de contextos (tais como agregação de contextos relacionados a um estudante em particular) e interpretação em alto nível a partir de contextos brutos (YU, ZHOU e SHU, 2009).

Um exemplo de modelagem de contextos do estudante usando ontologias é apresentado na Figura 2.2 (YU, ZHOU e SHU, 2009).

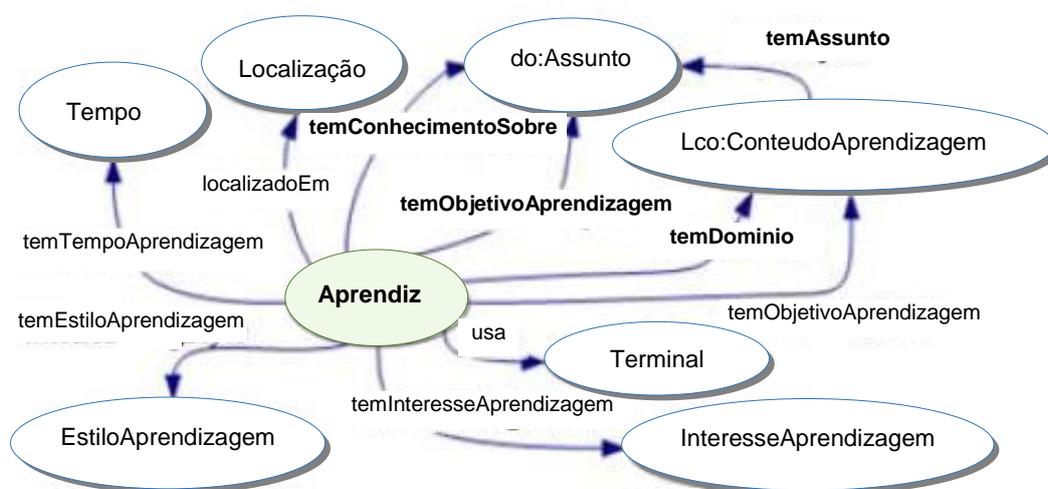


Figura 2.2 – Ontologia de contexto com foco no “estudante”. Fonte: adaptado de (YU, ZHOU e SHU, 2009).

De acordo com a representação ontológica apresentada na Figura 2.2, é possível descrever os contextos de aprendizagem do estudante “Michael” a partir do seguinte trecho em *OWL (Web Ontology Language)*.

```
<Aprendiz rdf:about="Michael">
  <temObjetivoAprendizagem>Data Mining</ TemObjetivoAprendizagem >
  <temDominio>CS201</ temDominio >
  <temTempoAprendizagem>19:00:00-22:00:00</ temTempoAprendizagem >
</ Aprendiz >
```

Outros modelos de representação de contextos foram propostos, incluindo modelos híbridos que unem mais de uma abordagem tradicional para melhorar a representação dos contextos. Dadas as características da modelagem baseada em ontologias, diversos autores, (STRANG e LINNHOFF-POPIEN, 2004), (HORROCKS, PATEL-SCHNEIDER e VAN HARMELEN, 2003), (YU, ZHOU e SHU, 2009), destacam o uso de ontologias como técnica promissora para modelagem de contextos.

Modelagem Preditiva

De acordo com (DICKY, 2012), modelagem preditiva é o nome dado ao conjunto de técnicas matemáticas com o objetivo comum de encontrar relações entre variáveis de entrada a fim de medir valores futuros da variável alvo. Em outras palavras, um modelo preditivo utiliza-se de técnicas para prever o futuro a partir do passado. No entanto, as relações não são perfeitas na prática (DICKY, 2012). Com isso, um grau de incerteza deve estar relacionado às previsões, normalmente um intervalo de predição tem algum nível de confiança atribuído, por exemplo 95%.

Um modelo preditivo pode usar aprendizagem supervisionada ou não supervisionada. No primeiro caso, dados são apresentados para um modelo preditivo com dados de entrada e dados de saídas desejáveis. O treinamento é repetido até que o modelo aprenda a função de mapeamento entre entradas fornecidas e saídas desejadas. Exemplos de modelos preditivos usando aprendizagem supervisionada são: redes neurais (*neural networks* - NN), máquinas de vetor (*support vector machines* - SVM), regressão logística e árvore de decisão. O modelo preditivo não supervisionado, são fornecidos apenas dados (contextos) de entrada. Neste caso, o modelo deve descobrir a relação entre os dados de entrada. *Clustering* é o tipo mais comum de modelo preditivo não supervisionado.

De acordo com Guazzelli (2012) árvore de decisão é representada por um conjunto de regras e é considerada a técnica de modelagem preditiva mais usada por ser genérica, fácil de entender e por explicar facilmente a lógica por trás da saída produzida. Nesta técnica um nó folha é alcançado após um conjunto de condições serem avaliadas e consideradas verdadeiras. A Figura 2.3 ilustra um exemplo de árvore de decisão para classificar a planta "Iris" a partir do tamanho da pétala. Neste exemplo, a seguinte regra pode ser empregada para indicar que a classificação da

planta é Iris-Setosa: “Se a o tamanho da pétala for menor que 2.6 então a planta é Iris-Setosa com probabilidade 1”.

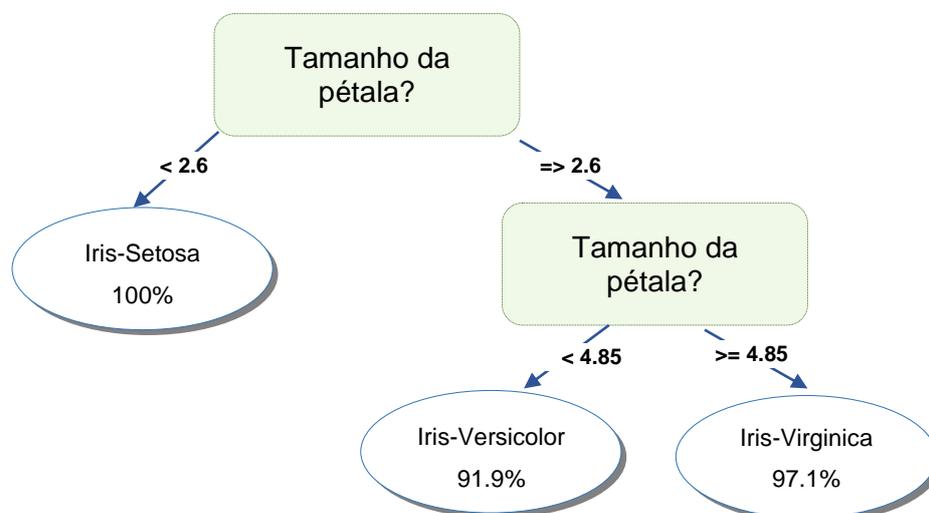


Figura 2.3 – Uma árvore de decisão usada para classificar planta Iris. Fonte: adaptado de (GUAZZELLI, 2012)

Análise preditiva é usada por instituições a fim de extrair valor a partir de dados históricos obtidos por sensores ou por pessoas. Dados de pessoas incluem transações estruturadas – por exemplo, compras online – ou não estruturadas – obtidas a partir de mídia social.

Abordagens para modelagem de contextos existentes diferem entre si em termos de poder de expressividade dos modelos de contexto, principalmente em relação ao suporte que oferecem e desempenho computacional para interpretação. Dentre as abordagens apresentadas, baseada em chave-valor, linguagem de marcação, orientada a objetos, baseada em lógica e ontologias, é possível combinar diferentes métodos e técnicas em modelagens híbridas para cumprir a todos os requisitos identificados para a aplicação.

Em vista do que foi apresentado, considera-se neste trabalho que dados de contexto em aprendizagem consistem de informações relevantes que indicam a situação acadêmica de um indivíduo e/ou de um grupo (turma, disciplina, curso, polo educacional) e da própria instituição. A análise dos dados de contexto em educação pode melhorar as compreensões sobre os processos de ensino e aprendizagem, munindo o próprio usuário, e demais agentes que atuam no processo educacional, de informações que auxiliam na tomada de decisão. Esse assunto é tratado na Seção seguinte.

2.4 Analítica

Analítica não é um conceito novo e tem suas origens na inteligência de negócios e mineração de dados (SHUM e FERGUSON, 2012). No primeiro caso, empresas utilizam ferramenta de monitoramento *web* (analítica da *web*) e análise preditiva há anos para melhorar o desempenho organizacional, prever tendências de consumo e o rumo em que os negócios podem tomar a fim de apoiar estratégias de *marketing* (LUAN, 2004). Norris, Baer e Offerman (2009) definem analítica como o conjunto de “processos de avaliação e análises de dados que possibilitam medir, melhorar e comparar o desempenho de indivíduos, programas, departamentos, instituições ou empresas, grupos de organizações e/ou toda a indústria”. Segundo Ravishanker (2011), analítica é a tomada de decisão dirigida a dados empregada para informar decisões sobre todos os níveis da empresa.

A mineração de dados é o campo da computação preocupado com grandes quantidades de dados para apoiar a descoberta de novas informações potencialmente úteis (PIATETSKY-SHAPIRO, 1995). Em educação, a mineração de dados se preocupa com o desenvolvimento de métodos para explorar dados em contextos educacionais a fim de entender melhor os estudantes e como eles aprendem (BAKER e YACEF, 2009).

Com isso, instituições de ensino estão voltando suas pesquisas para exploração de dados educacionais com diferentes objetivos, que buscam oferecer suporte às decisões de negócios e melhoramento do ensino e aprendizagem.

Particularmente em educação, há uma grande variedade de terminologias em torno de analítica. Nos trabalhos atuais o termo é definido conceitualmente (o que é) ou funcionalmente (o que faz) e possui diversas variações, tais como *Academic Analytics* (AA), *Learning Analytics* (LA), *Predictive Analytics* (PA) e *Action Analytics* (AcA). Não há consenso na definição e adoção desses termos. A rápida e crescente proliferação de trabalhos na área de analítica em educação incentivada pela explosão de dados disponíveis para análise possibilitou, e continua permitindo, o uso de diferentes definições para processos semelhantes. Esta Seção apresenta alguns desses termos.

Academic analytics possui definições que variam seu foco no nível de aprendiz – especialmente na predição de riscos e dificuldades acadêmicas

(CAMPBELL, DEBLOIS e OBLINGER, 2007; FRITZ, 2009 e BAEPLER e MURDOCH, 2010); instrutor – a fim de apoiar professores e tutores a determinarem necessidades de intervenções (CAMPBELL e OBLINGER, 2007 e CAMPBELL, 2007) ou instituição – com a finalidade de munir gestores de informações para apoiar a tomada de decisão frente aos desafios da universidade contemporânea (LUAN, 2004, GOLDSTEIN, 2005, DAWSON, MCWILLIAM e TAN, 2008 e RAVISHANKER, 2011). Apesar das diferentes definições e usos da terminologia *academic analytics*, o entendimento mais comum na comunidade acadêmica a relaciona com os negócios da instituição e pode ser compreendida como um processo para prover instituições de ensino com dados necessários para suporte à decisão operacional e financeira conforme apresenta Goldstein (2005).

Action analytics é definida por Norris, Baer e Offerman (2009) como um termo mais abrangente que AA e empregada para análises acadêmicas focadas na produtividade, desempenho acadêmico e administrativo. Para o autor, analítica da ação é mais sobre liderança e disseminação de mudanças culturais do que tecnologias.

Predictive analytics é compreendida como um processo de negócios e um conjunto de tecnologias que lidam com extração de informações para descobrir relacionamentos e padrões a partir de grandes volumes de dados - o passado - que pode ser usada para prever comportamentos e eventos – e o futuro (ECKERSEN, 2007; IBM, 2010a e IBM, 2010b).

Learning Analytics é o termo mais encontrado nos trabalhos relacionados à análise de dados educacionais e usualmente adotado como sinônimo de analítica.

Na Subseção seguinte são apresentadas algumas definições para LA encontradas na literatura.

2.4.1 Análise da Aprendizagem

A definição mais utilizada nos trabalhos atuais foi apresentada na 1ª Conferência Internacional sobre Aprendizagem Analítica e Conhecimento em 2011

(LAK11¹¹) e define LA como “a medição, coleta, análise e divulgação de dados sobre aprendizes em seus contextos, para propósitos de entender e aperfeiçoar a aprendizagem e os ambientes nos quais ela ocorre”, citado por Siemens e Long (2011, p. 32).

Johnson et al. (2011) descrevem LA como a “interpretação de uma ampla gama de dados produzidos e coletados a partir dos estudantes, a fim de avaliar o progresso acadêmico, prever o desempenho futuro e identificar possíveis problemas.” Elias (2011) considera LA como “um campo emergente no qual ferramentas analíticas sofisticadas são usadas para melhorar a aprendizagem e educação”. Para Siemens (2010), LA é “o uso de dados inteligentes, dados produzidos por aprendizes e modelos de análises para descobrir informações e conexões sociais a fim de prever e informar sobre aprendizagem”. Em Bach (2010) LA é vista como “o uso de modelagem preditiva e outras técnicas analíticas avançadas para ajudar a canalizar os recursos instrucionais, curriculares e de suporte para alcançar os objetivos de aprendizagem específicos”. Brown (2011) define LA como a “coleta e análise de dados associados com a aprendizagem de estudantes; para observar e entender comportamentos a fim de possibilitar intervenção apropriada”. Para Gilfus (2011), LA “é o uso de dados e modelos para prever o progresso e desempenho do estudante, e a possibilidade de agir sobre as informações”.

Apesar das diferentes definições, é possível afirmar que todas elas têm ênfase na compreensão de dados educacionais para o sucesso na aprendizagem. Além disso, as definições apresentadas não limitam sua admissão para os diferentes níveis da tomada de decisão, apesar de não terem sido criadas, diretamente, com essa finalidade.

Neste trabalho, LA é entendida como uma área da análise de dados que coleta e processa dados sobre aprendizes, tutores e professores, suas ações e interações em atividades de aprendizagem, dentro e fora do SGA, a fim de gerar conhecimento para auxiliar a decisão nos níveis individual, de curso e instituição.

¹¹ <https://tekri.athabascau.ca/analytics/> - 1st International Conference on Learning Analytics and Knowledge (LAK) 2011

2.5 Dados em *Learning Analytics*

Dados são elementos essenciais em LA. O SGA oferece um grande conjunto de dados sobre os usuários e suas interações, tais como leitura, escrita, *download* de materiais de estudo, *upload* de atividades, realização de avaliações, interações em discussões, notas, acessos, entre outras.

No entanto, a facilidade de acesso às tecnologias móveis e ubíquas envolve além do SGA ou outras bases de dados, dados obtidos a partir do ambiente digital do usuário, incluindo interações externas ao *browser*; dados, tais como a largura de banda e o tamanho da tela; e, dados do ambiente físico, tal como presença e ruído. A coleta de dados pode envolver sensores físicos e sensores lógicos. Com essa variedade de fontes, há dados de todos os tipos disponíveis para análise, por exemplo, tabelas estruturadas, páginas *web* semiestruturadas, textos não estruturados, arquivos multimídia (imagens, vídeos e áudios) e *hyperlinks*.

Esses dados podem estar em diferentes formatos, distribuídos em diversos meios de comunicação e locais da rede, o que torna o trabalho de integrar dados brutos a partir de múltiplas fontes heterogêneas em um desafio no desenvolvimento de sistemas de LA (CHATTI et al. 2012b). Verbert et al. (2012) afirmam que “embora uma enorme quantidade de dados tenha sido capturada a partir de ambientes de aprendizagem, é um processo difícil tornar esses dados disponíveis para propósitos de pesquisa” (p. 145).

Liu (2006) e Chatti et al. (2012b) indicam três passos que devem estar presentes na análise de dados: coleta e pré-processamento, análise e visualização, e pós-processamento. Preparação dos dados, ou o pré-processamento, é uma fase importante para remover impurezas, atributos irrelevantes e ruídos e também para preparar o dado para um método particular de análise, conforme destacam Liu (2006) e Chatti et al. (2012b). De acordo com Liu (2006), Romero e Ventura (2007) e Han e Kamber (2011), algumas das técnicas que podem auxiliar na preparação dos dados são: fusão dos dados, limpeza de dados, integração de dados, transformação de dados, redução de dados e modelagem dos dados.

A fase de análise dos dados, conforme já mencionada, consiste em explorar os dados obtidos na fase anterior a fim de descobrir padrões que podem ser úteis

para a tomada de decisão. Esta fase inclui também a visualização da informação, recomendações, personalização dos ambientes de aprendizagem e predição.

De acordo com Chatti et al. (2012b), a etapa de pós-processamento dos dados em LA envolve novas compilações dos dados a partir de fontes adicionais, novos atributos, novos indicadores e métricas, modificação das variáveis de análise ou uso de novas técnicas de análise.

Pesquisas relacionadas à fusão de dados estão, historicamente, bem estabelecidas para dados obtidos por meio de sensores de *hardware* (sensores de ambiente), os quais são gerados por sensores físicos calibrados com características bem definidas (KHALEGHI et al., 2013; WICKRAMARATHNE, 2012). No entanto, a maioria das técnicas de fusão existentes não é capaz de lidar com incertezas típicas de dados subjetivos coletados por meio de sensores lógicos, tais como dados de redes sociais, conteúdos de mensagens de fóruns e e-mails (WICKRAMARATHNE, 2012; YERVA, JEUNG, ABERER 2012; YERVA et al., 2012; GROSS et al., 2012; HALL et al., 2008; PRAVIA et al., 2008 e LAUDY e GOUJON, 2009). A fusão de dados obtidos por sensores lógicos pode ajudar a medir de forma mais confiável certos atributos de interesse, por exemplo, estado emocional (GROSS et al., 2012).

Há poucas pesquisas sobre a fusão de dados obtidos por sensores físicos e sensores lógicos para análise integrada. Na verdade, o conceito tem sido explorado recentemente a partir de dados de observação humana realizada sobre determinado fenômeno ou atributo de interesse, normalmente feitas por especialistas. Neste caso, as soluções propostas permitem humanos agirem como sensores (*soft sensors*) inserindo informações qualitativas, e portanto abertas a questionamentos, que serão fundidas com dados de sensores físicos com o objetivo de produzir melhores análises (HALL et al., 2008).

Recentemente, têm surgido algumas abordagens de fusão de dados a partir da coleta em múltiplas fontes. Yerva e colaboradores (2012) apresentam um *framework* que integra dados sociais e de sensores meteorológicos. Neste trabalho, a partir de informações fornecidas pelo usuário, via Twitter, por exemplo, o sistema prevê informações sobre o tempo no local onde o usuário deseja viajar.

Gross et al. (2012) apresentam uma arquitetura para realizar a fusão de dados de observações humanas e dados de sensores físicos em um *framework* único. Este trabalho inclui processamento de linguagem natural para elaboração de gráficos proposicionais a partir de observações linguísticas, conversão desses

gráficos para a forma atribuída, alinhamento e marcação de incertezas em observações humanas, entre outras.

Acharya e Kam (2011) fundem dados de sensores físicos com dados de sensores lógicos provenientes de relatórios contendo informações inseridas por especialistas de domínio. Os autores usaram representação probabilística de dados de sensores a partir da teoria de Dempster-Shafer (LUO E KAY, 1992) e operador de consenso para combinar opiniões humanas com incertezas e, então, usaram uma regra de fusão de probabilidade para gerar um sistema de fusão de dados de *hardware* e *software*. A solução proposta requer intervenção humana de especialistas para avaliar as análises obtidas e, entretanto, consiste de uma solução semiautomática.

2.6 Considerações Finais

Este capítulo apresentou os principais conceitos envolvidos na área de pesquisa deste trabalho, incluindo terminologias utilizadas; o conceito de contexto empregado às ferramentas analíticas educacionais; as técnicas de modelagem de contexto; os métodos para tratamento de dados, principalmente na fusão de dados imperfeitos e as analíticas empregadas sobre dados educacionais. O próximo capítulo apresenta o estado da arte em *Learning Analytics*.

Capítulo 3

ESTADO DA ARTE EM LEARNING ANALYTICS

“The rarely articulated implication of all of this data floating around is that un-augmented human cognition is no longer sufficient. Every day, every day, there is more to know, more ways to know it, and heightened expectations—by students, faculty members, alumni, foot-ball coaches, trustees, regulators, elected officials—that senior man-agers will do something efficacious with what they know. . . . The best tool in [their] battle against ignorance is advanced analytics”. (THORTON MAY, 2011).

Este capítulo apresenta o estado da arte em Learning Analytics, abordando as principais pesquisas sobre ferramentas analíticas educacionais nos últimos anos.

3.1 Considerações iniciais

A tomada de decisão dirigida a dados (*Data-Driven Decision Making - DDDM*), quando aplicada à educação envolve fazer uso dos dados, tais como os dados providos pelos SGA e por sensores, para o julgamento de educadores e gestores (SIEMENS E LONG, 2011). De acordo com Marsh, Pane e Hamilton (2006), o conceito de DDDM não é algo novo, ele foi introduzido nas décadas de 70 e 80 e desde então tem sido usado para identificar necessidades individuais, avaliar a eficácia de práticas de ensino e melhoria do planejamento estratégico. As fontes de dados usualmente empregadas nesses casos eram: questionários, entrevistas, observações e documentos institucionais.

Diversas publicações na área de educação e de gestão abordam o conceito de DDDM em educação, mas não enfatizam as pesquisas sob um viés tecnológico. Nesses trabalhos, a análise dos dados, por exemplo, não é automatizada, mas sim realizada por especialistas ou pelos próprios usuários da ferramenta (professores e gestores). Por exemplo, no trabalho de Marsh, Pane e Hamilton (2006), o conceito DDDM em educação é relacionado com professores e gestores, sistematicamente, coletarem e, eles próprios, analisarem dados que deverão guiar decisões que os ajudarão a melhorar o sucesso de aprendizes e escolas.

A partir de um estudo exaustivo sobre o uso de dados para tomada de decisões educacionais, Schidkamp e Lai (2013) afirmam que líderes de escolas e professores necessitam de, e devem cada vez mais usar, dados para tomar suas decisões, uma vez que evidências mostram que o uso do dado pode melhorar a aprendizagem do estudante. No entanto, as autoras sugerem que os dados podem ser obtidos a partir de métodos qualitativos, por exemplo, observações em salas de aula e entrevistas; ou quantitativos, por exemplo, resultados de avaliações.

Na mesma direção, Earl e Katz (2006) apresentam o pressuposto de que quando bem informados, líderes de escolas podem rever de forma mais eficaz as suas capacidades existentes, identificar os pontos fracos e oportunidades de melhorias. No entanto, para os professores, por exemplo, o pressuposto é que analisar resultados de suas avaliações lhes permitirá direcionar práticas de ensino para as necessidades individuais dos estudantes (MANDINACH e HONEY, 2008).

Ainda sob uma visão de gestão educacional, Macfadyen e Dowson (2012) indicam que a qualidade do ensino oferecido por uma instituição não é prevista pela dimensão de seus orçamentos institucionais, números ou valores de dólares investidos, ou pela proporção do corpo docente ou horas gastas em sala de aula, mas sim pelo que as instituições fazem com seus recursos. Os autores argumentam que os processos de tomada de decisão relacionados à organização dos recursos institucionais - humanos e material – e o planejamento para uso mais eficiente dos recursos existentes são características críticas para excelência de instituições. No entanto, os autores indicam que dada a quantidade imensa de dados disponíveis para análises, pode haver a necessidade da participação de pessoas especialistas que farão as análises sobre os dados. Tecnologias de análises de dados automatizadas não são mencionadas.

Todavia, considerando que *big data* é sinônimo de tomada de decisão dirigida a dados (DHIMAN e QUACH, 2012) e que em educação isso não é diferente, o emprego de tecnologias que automatizam os processos de coleta e análises de dados é primordial para apoiar a DDDM educacional. Neste contexto, os trabalhos publicados nesta linha estão inseridos na área de LA.

Conforme apresentado na Seção 2.4.1, a LA é uma área de pesquisa em crescimento e que oferece diversas oportunidades de investigação para os próximos anos (por exemplo, *Horizon Report*, JOHNSON et al. 2011, 2012 e 2013). Há diversos benefícios da análise de dados voltada para apoiar decisões em diferentes níveis, tais como: no nível micro ao ajudar alunos a refletirem sobre sua aprendizagem e melhorar suas decisões a fim de obterem sucesso (SIEMENS, 2013; GRELLER e DRACHSLER, 2012; DOLLÁR e STEIF, 2012); no nível meso ao ajudar instrutores a compreenderem melhor as dificuldades dos estudantes, repensar estratégias de ensino, prover melhores intervenções e facilitar a avaliação da aprendizagem (ARNOLD e PISTILLI, 2012; GOVAERTS et al., 2012; ALI, et al. 2012; DYCKHOFF et al., 2012); e, em nível macro ao orientar as decisões de instituições para, por exemplo, direcionarem melhor seus recursos (SMITH, LANGE, e HUSTON, 2012).

Desta forma, este capítulo apresenta o estado da arte em LA, classificando os principais trabalhos que geraram contribuições científicas para a área na última década em três categorias distintas: sistemas de visualização, sistemas de personalização e adaptação e sistemas de recomendação.

3.2 Learning Analytics

Recentemente, diversas aplicações têm sido propostas para capturar os comportamentos de usuários a fim de auxiliar a aprendizagem *online* em diferentes áreas, como testes de usabilidade em interação humano-computador, aprendizagem ubíqua, aprendizagem ciente de contexto, mineração de dados, recuperação da informação e captura de metadados de atenção contextualizados (*Contextualized Attention Metadata - CAM*) em TEL.

LA pode prover ferramentas educacionais poderosas para apoiar os usuários de educação *online* oferecendo melhores condições de sucesso para aprendizes, melhorar a efetividade das práticas pedagógicas de instrutores e melhorar a tomada de decisão em nível institucional. Neste sentido, as próximas seções apresentam o estado da arte em sistemas de visualização, sistemas de personalização e adaptação e sistemas de recomendação.

3.2.1 Visualização

Card, Mackinlay e Shneiderman (1999) definem visualização como “o uso de representações visuais de dados interativas e suportadas por computador para ampliar a cognição”. Visualização “promete nos ajudar a acelerar nossa compreensão e ação em um mundo de volumes de informação crescentes” (CARD, 2003, p. 542). Mazza e Dimitrova (2007) afirmam que a eficácia dos LMS pode ser melhorada a partir da integração de técnicas de visualização da informação para gerar representações gráficas para que professores possam entender as informações mais facilmente e rapidamente. Em um trabalho mais recente Ali e coautores (2012) mostraram que “visualização pode ser um meio eficaz para lidar com grandes quantidades de dados a fim de sustentar a carga cognitiva de educadores em um nível aceitável” (p. 488).

A visualização de informações resultantes das análises de dados em educação tem sido amplamente discutida nos trabalhos sobre avaliação e feedback e no suporte aos processos de tutoria e monitoria. O foco principal, geralmente associado ao desenvolvimento dessas ferramentas, está no apoio aos professores e instrutores.

3.2.1.1 Avaliação e *Feedback*

Arnold e Pistilli (2012) propõem o *Course Signals* que prediz e apresenta resultados de aprendizagem baseados em três tipos de dados: notas obtidas pelos alunos no curso, tempo para realização de tarefas e desempenho obtido anteriormente. O sistema detecta sinais de alerta que podem indicar que os aprendizes não estão desempenhando suas atividades satisfatoriamente e oferece *feedback* automático sobre o desempenho dos alunos. Um semáforo com as cores vermelha, laranja e verde é apresentado para o aluno. A cor vermelha indica que as

notas do aluno estão abaixo do que foi especificado pelo professor para aquele determinado momento na disciplina. Quando as notas do aluno estão acima do especificado, duas novas variáveis são utilizadas para calcular se o aluno está ou não no caminho certo: notas que ele tirou em outros cursos e o tempo na realização de tarefas. Caso o aluno esteja realizando as atividades adequadamente, a cor verde estará evidente, senão a cor laranja é usada para indicar que ele precisa melhorar suas atividades no curso. Este sistema é adotado na *Purdue University*¹² atualmente e utiliza exclusivamente três tipos de dados para análise e o ambiente virtual de aprendizagem como fonte de dados, conforme mencionado. Além dessas restrições, esta ferramenta utiliza o tempo em que o estudante esteve conectado a um determinado recurso e não o tempo que, de fato, ele levou para realizar uma tarefa. Neste caso, um aluno pode conectar-se no ambiente, acessar determinado recurso, mas redirecionar sua atenção para outras atividades, tais como assistir um filme, ou navegar em outros sites, e não estar usando, de fato, o recurso em que está conectado.

Uma proposta semelhante foi desenvolvida na *Carnegie Mellon University* como uma iniciativa para aprendizagem aberta implementando *feedback* instantâneo para atividades e avaliação do progresso do estudante (DOLLÁR e STEIF, 2012). O sistema batizado como *Online Learning Initiative* (OLI) é uma proposta para ensinar alguns conceitos de introdução à estatística. Como complemento, esta ferramenta oferece *feedback* aos alunos e dicas sobre como praticar habilidades individuais. Um *dashboard* é utilizado com as cores vermelha, laranja e verde que indicam em que medida o aluno atingiu os objetivos para o curso e atividades específicas. Para obter essas informações, o sistema analisa *logs* produzidos durante as interações dos alunos com a ferramenta, especialmente dados sobre o tempo que os alunos gastam em cada atividade, o tipo de proposta de atividades e as notas obtidas. Informações inseridas por meio de formulários *online* também são utilizados para as análises.

Student Activity Meter (SAM) é uma ferramenta de visualização do progresso do estudante no curso (GOVAERTS et al., 2012) destinada ao automonitoramento dos estudantes e ciência dos professores. A ferramenta realiza o rastreamento do tempo gasto pelos alunos durante o desenvolvimento das atividades do curso e

¹² <http://www.itap.purdue.edu/learning/tools/signals/>

apresenta, em *dashboard*, o número de estudantes no curso, o número de recursos utilizados e o tempo (mínimo, médio e máximo) gasto pelos estudantes para realização do curso. As fontes de dados são os *logs* do *Moodle* e rastreamento do tempo via *Twitter*. Esta solução tem um foco restrito aos recursos utilizados pelos estudantes e o tempo permanecido em cada um. Essas informações dão *insights* interessantes aos alunos e professores, tais como: maior tempo em determinados tipos de recursos podem indicar preferências dos estudantes por determinadas ferramentas; e muito tempo em determinada atividade pode indicar dificuldade pontual em determinado assunto. Todavia, esta ferramenta utiliza como dado de entrada para as análises o tempo conectado e não o tempo de estudo na ferramenta, efetivamente, sendo esta a principal limitação desta solução. Além disso, como a ferramenta tem por objetivo o automonitoramento dos estudantes, outras variáveis que possam apoiar esse objetivo poderiam ser consideradas nas análises, tais como status das atividades e dos materiais e prazos de entrega.

Learning Object Context Ontologies (LOCO-Analyst) sistema consiste de um *framework* ontológico desenvolvido para prover *feedback* aos instrutores sobre o desempenho dos estudantes em atividades (ALI, et al. 2012). O esforço do estudante é medido a partir do tempo gasto por ele na realização de atividades e pelos artefatos produzidos durante o curso. O professor pode visualizar os tipos de atividades que um estudante desenvolveu ou participou e interações sociais entre estudantes. Este sistema integra um conjunto de ontologias relacionadas à aprendizagem: ontologia do contexto da aprendizagem como componente central do *framework*; ontologia para definição do modelo do usuário (professores, autores de conteúdos e estudantes), que integra outras ontologias para definição de competências, desempenho e preferências; ontologia de objetos de aprendizagem *quiz* que representa formalmente os instrumentos de avaliação; e uma ontologia de domínio. As fontes de dados utilizadas neste *framework* consistem, exclusivamente, de *logs* e bases de dados. Esta solução é destinada ao apoio aos professores, especialmente para fins de intervenção, não cobrindo suporte aos demais usuários, tais como a aprendizagem autorregulada ou apoio às equipes de gestão pedagógica ou acadêmica.

Moodle dashboard é uma ferramenta similar à LOCO e possibilita aos professores a visualização de índices de frequência e resultados das autoavaliações

que podem dar uma indicação do progresso do estudante no aprendizado (PODGORELEC e KUHAR, 2011). A ferramenta também faz uma previsão do sucesso dos estudantes no curso usando árvore de decisão simples, com apenas uma variável - aproveitamento acadêmico (nota), a fim de indicar se o aluno está apto a ser aprovado ou não. Esta ferramenta foi desenvolvida para uso, particularmente, na plataforma Moodle. Estatísticas simples, tais como média, valor mais frequente e porcentagens, são aplicadas sobre os dados recuperados, restritamente, a partir da base de dados do SGA. Apesar de oferecer um nível mais básico de análises, uma característica positiva desta ferramenta é a simplicidade na apresentação das informações produzidas nas análises por meio de gráficos de linhas ou barras, o que facilita a interpretação das informações por parte dos professores. As principais limitações apresentadas nesta solução são duas: a primeira em relação ao conjunto de dados de entrada, cuja fonte única é o SGA; a falta de extensibilidade e reuso da ferramenta para outras plataformas, ou para inclusão de novas fontes de dados, ou mesmo para outros propósitos de aplicação é a segunda limitação apresentada.

Gradient's Learning Analytics System (GLASS) (LEONY et al., 2012) é uma ferramenta de visualização a partir de bases de dados de CAM (*Contextualized Attention Metadata*) *schema*, que armazenam eventos gerados durante o uso de aplicações enquanto o estudante está trabalhando em um ambiente de aprendizagem. A ferramenta provê duas visualizações: uma linha de tempo indicando a frequência de eventos de atividades e um gráfico de barras com eventos gerados pelos usuários. Dados de *logs* são analisados a fim de obter dados sobre as interações dos usuários com ferramentas do ambiente virtual e são armazenados em bases de dados para análises futuras. O sistema possibilita o uso de várias bases de dados. No entanto, outros tipos de fontes não são considerados, como as interações com ferramentas externas ao ambiente virtual de aprendizagem. Além disso, esta ferramenta oferece uma boa solução para recuperar dados complexos de *logs*, no entanto apenas fornece funcionalidades analíticas básicas para apresentação dos dados.

Na mesma direção, *Student Inspector* (SCHEUER e ZINN, 2007) é uma ferramenta de visualização baseada no uso de *logs* de dados de atividades realizadas pelos estudantes no ambiente de aprendizagem ActiveMath¹³. *Student Inspector* dá alguns *insights* sobre como os estudantes utilizam o sistema, incluindo dados sobre o aproveitamento acadêmico, obtido a partir de resultados de testes e erros típicos cometidos pelos alunos, dados sobre os tópicos e competências mais fracas e fortes e a sequência de atividades do aluno em uma seção de aprendizagem. Este sistema analisa dados de *logs*, metadados de objetos de aprendizagem e dados obtidos por outras bases relacionais.

iTree (NAKAHARA et al., 2005) é uma ferramenta desenvolvida para dispositivos móveis com objetivo de melhorar o engajamento de estudantes em atividades de fóruns. Esta ferramenta ilustra contribuições dos estudantes em fóruns na figura de uma árvore. O tamanho da árvore, tamanho da haste, número de ramos, folhas, frutos, bem como a cor do céu, indicam o grau de envolvimento do estudante nas atividades de fórum. A árvore é utilizada como um mecanismo de alerta e motivação para os estudantes melhorarem suas discussões. As variáveis utilizadas para construção da árvore e que afetam o seu crescimento são: número de *posts* dos estudantes, quantidade de leituras realizadas sobre os *posts*, quantidade de respostas aos *posts* e proporção total de *posts* no fórum. Esta ferramenta tem um propósito bastante limitado e, por este motivo, utiliza poucas entradas de dados.

Santos et al. (2013) apresentam o *StepUp!* que proporciona visualizações em dispositivos móveis sobre a interação social entre o usuário e seus pares, tempo gasto, artefatos produzidos, recursos utilizados pelo usuário e o progresso do estudante em atividades específicas do curso, como por exemplo a quantidade de postagens em *blogs*. O objetivo principal da ferramenta é mostrar para os estudantes como eles gastam o seu tempo. A ferramenta realiza rastreamento de ações do usuário dentro de navegadores *web* e em bancos de dados do ambiente virtual para obter dados como a identidade da pessoa que publicou uma mensagem em uma ferramenta blog, por exemplo, e o *timestamp*. Os mesmos dados são recuperados

¹³ <http://www.activemath.org/>

para interações via *Twitter* por meio de uma API. Assim como a maioria das ferramentas de visualização apresentadas anteriormente, este sistema tem um propósito específico e utiliza poucos dados de entrada para processamento e, portanto, gera um conjunto moderado de informação resultante das análises.

3.2.1.2 Monitoramento e Tutoria

Chen, Chang e Wang (2008) propõem uma ferramenta análoga a um tutor *online* que gera um conjunto de notificações aos estudantes a fim de ajudá-los no processo de aprendizagem. Um *website* é implementado para possibilitar o uso da ferramenta via dispositivo móvel. Comportamentos dos estudantes são monitorados a partir de registros sobre a aprendizagem em portfólios. Um modelo do estudante é construído a partir de dados, obtidos em *logs*, como resultados de avaliações e tarefas. Baseado no modelo do estudante e nos dispositivos de aprendizagem disponíveis, três módulos são desenvolvidos com objetivo de melhorar o desempenho do estudante: “consciência do status de aprendizagem”, uma agenda de lembretes e algumas notificações. Essas notificações são enviadas a um professor para aprovação antes de serem encaminhadas ao estudante via SMS. Nesta solução a maior limitação encontra-se no conjunto de dados utilizado como base para as análises, que consiste de poucas entradas para produzir resultados confiáveis, e também nos mecanismos de coleta de dados, que consiste da captura de dados básicos disponíveis e não suporta aquisição mais abrangente de contextos, tais como as interações dos estudantes fora do navegador ou em outras páginas.

Kosba, Dimitrova e Boyle (2005) apresentam o *Teacher ADVisor* (TADV) que baseia-se em intervenções geradas automaticamente. O sistema consiste de um *framework* que gera uma base de “conselhos” para professores a partir do rastreamento de dados coletados no ambiente virtual e do uso de fatores de certeza e teorias do conjunto *fuzzy*. Os conselhos destinados aos professores objetivam ajudá-los a melhorar seu *feedback* e guiar os estudantes durante o curso. A ferramenta de coleta de dados deste sistema considera apenas quantidade de postagens em fóruns de discussões e acessos aos materiais de estudo, consistindo, portanto, de um conjunto bastante limitado para análises do comportamento de estudantes em aprendizagem *online*. Considerando uma proposta de disciplina em que fóruns de discussão são pouco utilizados ou que o professor adota,

principalmente, materiais externos ao SGA como base para os estudos, esta ferramenta torna-se inviável.

Tool for Advanced Data Analysis in Education (TADA-Ed) é uma ferramenta desenvolvida no âmbito da área de mineração de dados educacionais e tem o objetivo de apoiar professores na descoberta de padrões pedagogicamente relevantes a partir da mineração sobre dados obtidos no rastreamento de tarefas *online* (MERCERON e YACEF, 2005). Estatísticas simples, como frequência de erros, e resultados dos algoritmos de mineração, como clusterização de alunos com base em seus erros, são apresentados aos professores. Neste sistema, professores devem selecionar as técnicas de mineração que desejam visualizar e analisar os resultados oferecidos pelo sistema após aplicação da técnica escolhida. Com isso, professores devem ter conhecimento suficiente sobre as técnicas de mineração, pois seus resultados podem ser complexos. A suposição de que os usuários conhecem as técnicas de mineração de dados suficientemente a ponto de selecioná-las e também interpretar seus resultados é a principal desvantagem desta proposta. Além disso, os dados utilizados para análises são, basicamente, quantidade de exercícios realizados pelo aluno e os acertos e erros cometidos, o que também consiste de uma limitação deste trabalho dado o objetivo proposto que centra-se na descoberta de padrões pedagogicamente relevantes.

Social Networks Adapting Pedagogical Practice (SNAPP) (BAKHARIA e DAWSON, 2011) é um projeto da *University of Wollongong*. Esta ferramenta foi desenvolvida para prover instrutores com visualizações sobre a evolução dos relacionamentos dos participantes em fóruns de discussão a fim de identificar padrões e diagnosticar problemas que possibilitam a intervenção. Os dados são apresentados aos professores em uma linha do tempo, que também ilustra eventos chave que ocorreram durante um período do desenvolvimento do curso. Esta ferramenta coleta dados sobre quem postou novas mensagens em fóruns e quem respondeu essas mensagens e, então, produz um diagrama de rede social sobre tais interações. Com isso, a proposta busca identificar alunos desconectados, o que pode indicar que esses alunos estão em risco, uma vez que têm pouco envolvimento na comunidade de aprendizagem *online*. Esta proposta é interessante para prover informações sobre o engajamento do estudante em fóruns de discussão. No entanto, limita-se a apenas este tipo de ferramenta e não considera se as postagens dos estudantes são relevantes ou não. Por exemplo, pode ocorrer que alunos com

maiores números de postagens enviem apenas mensagens que não contribuem para a discussão do grupo e, portanto, não atendem à proposta da atividade. Neste caso, a ferramenta pode fornecer um resultado de análise inconsistente.

Classroom View (FRANCE et al., 2006) é uma ferramenta destinada aos instrutores de cursos *online* para monitoramento e visualização das interações que acontecem em uma sala de aula virtual. O sistema, baseado em arquitetura multiagentes, coleta informações a partir do rastreamento de atividades de interação e apresenta os caminhos que o estudante percorreu durante uma seção de atividade e o número de vezes que um determinado caminho foi seguido. Essas informações podem ser usadas por professores para regulação de seções de treinamento.

Dyckhoff e colegas (2012) apresentam uma ferramenta de visualização que chamaram de *Learning Analytics Toolkit* (eLAT). O principal objetivo da ferramenta é processar grandes conjuntos de dados relacionados aos interesses de professores a fim de ajudá-los na reflexão sobre o seu ensino com uso de tecnologias e identificar oportunidades para intervenções e melhorias. Esta ferramenta possibilita professores explorarem e correlacionarem objetos de aprendizagem utilizados pelos estudantes, acessos realizados, frequência de *logins* e resultados de avaliações a fim de identificar alguns padrões, como por exemplo: se o aprendizado contínuo está ocorrendo e se isso está relacionado com os acessos realizados, momentos em que os alunos fazem exercícios, se acessam materiais antes ou depois das aulas. Esta ferramenta consiste de uma proposta mais abrangente que as citadas anteriormente, principalmente, por ampliar o conjunto de dados de entrada para as análises. Neste sistema, são considerados: dados pessoais sobre o estudante, dados numéricos sobre as interações realizadas, notas de exercícios e dados recuperados em calendários, tais como datas de avaliações, início de um curso, etc. No entanto, a única fonte de dados utilizada é a base de dados do SGA.

3.2.2 Recomendação

Várias pesquisas têm sido conduzidas enfatizando a importância de sistemas de recomendação em ambientes TEL, principalmente na indicação de recursos de aprendizagem apropriados a partir de uma variedade de opções (VERBERT et al., 2011), recomendação de conteúdos (YU, ZHOU e SHU, 2010) e a identificação de caminhos em uma rede de aprendizagem que são mais rápidos de completar ou

mais atraentes que outros (DRACHSLER et al., 2009). Nesta Seção, são apresentados alguns sistemas recomendadores.

Bozo, Alarcón e Iribarra (2010) apresentam um recomendador de objetos de aprendizagem de acordo com o contexto do professor, por exemplo, informações sobre o nível educacional, assunto, área, região e tipo de escola; e de acordo com os assuntos de aprendizagem, tais como o nível educacional e o tópico em questão. Essas informações necessárias para gerar as recomendações são explícitas sobre os recursos e usuários e nem sempre estão disponíveis, o que pode inviabilizar a proposta do trabalho.

Broisin et al. (2010) apresentam uma solução para recomendação de materiais para estudantes de acordo com seus objetivos e de suas atividades atuais. Para isso, o sistema captura anotações semânticas que os alunos realizam sobre os recursos de aprendizagem e dados das interações dos usuários com os recursos também são capturados e armazenados em um repositório CAM. Links para materiais recomendados são integrados à interface gráfica do ambiente de aprendizagem e e-mails contendo referências aos documentos recomendados são enviados para os interessados. Neste trabalho, é criado o modelo do usuário, que ajuda a determinar os perfis dos usuários do sistema, e também o modelo de documentos que classifica os recursos acessos pelos usuários, e assim fazer as recomendações. Este trabalho usa como base para coleta de dados uma proposta de federação de CAM, conforme apresentado em Botoianu et al. (2010). Embora esta abordagem objetiva um modelo de CAM genérico, requer implementação específica para cada aplicação rastreada (por exemplo, SGA e repositórios) e, portanto, apresenta baixo reuso da arquitetura de coleta.

Klasnja-Milicevic et al. (2011) apresentam um módulo de recomendação integrado ao sistema Protus, que consiste em um sistema de tutoria projetado para ajudar aprendizes no estudo da linguagem de programação Java, qual pode adaptar-se automaticamente conforme os interesses e nível de conhecimento dos aprendizes. O módulo de recomendação do Protus tem por objetivo recomendar uma lista de materiais úteis e que sejam de interesse dos aprendizes. Para isso, o sistema utiliza dados sobre o nível de conhecimento do estudante, preferências, hábitos, interesses e estilos de aprendizagem, principalmente a partir de mineração de *logs*. Baseado nesses dados, o sistema constrói um modelo do estudante

utilizando ontologias, que é alterado a cada nova sessão. O aluno pode ajudar a compor seu modelo editando seu estilo de aprendizagem preferido durante uma sessão. Por consistir de uma proposta com objetivo altamente específico de recomendação de materiais para uma disciplina em particular, esta proposta considera um conjunto consistente de dados para realizar as recomendações.

Ferreira et al. (2014) apresentam um modelo para recomendação ubíqua de conteúdos educacionais baseado em agentes para grupos de alunos criados dinamicamente (UbiGroup), à medida que o estudante se conecta ao ambiente. Esta proposta objetiva ajudar professores a pesquisarem e selecionarem materiais educacionais, na forma de objetos de aprendizagem, a partir do perfil do grupo de alunos e do contexto de aprendizagem. O perfil do grupo é composto por dados como a localização do estudante, atividades que ele está realizando, notas e preferências. O principal diferencial deste trabalho em relação a outros recomendadores de conteúdo na área educacional é que ele está inserido no domínio de computação ubíqua e considera algumas informações de contexto adicionais do usuário, tal como o sinal de GPS.

O projeto iWeaver (WOLF, 2003) é um recomendador de conteúdos que objetiva prover ambientes de aprendizagem gerenciáveis baseados na *web*. Para isso, o sistema oferece acesso a diferentes mídias para estudantes do SGA. O modelo do aprendiz é construído a partir do estilo de aprendizagem, escolhas navegacionais e ferramentas de aprendizagem utilizadas. O estilo de aprendizagem dos estudantes é obtido a partir de preenchimento de formulário. A coleta de dados é realizada a partir de *logs* do usuário e dados armazenados em bases educacionais. A principal limitação apontada na maioria dos trabalhos anteriores também está presente nesta proposta, que é o conjunto de dados de entrada restrito ao ambiente de aprendizagem virtual, ou no máximo aos dados de *logs* ou dados inseridos pelo usuário por meio de formulários.

K-InCa (NABETH et al., 2005), que é um acrônimo para, *Knowledge Intelligent Conversational Agents*, objetiva auxiliar pessoas a aprenderem e adotarem práticas de gerenciamento de conhecimento compartilhado. Esta ferramenta usa ontologias no processo de modelagem e inferência a partir dos contextos e foi proposta para o suporte a intervenções que ajudam aprendizes a passarem, progressivamente, de um estágio de *expertise* para outro. K-InCa é implementado sob arquitetura

multiagentes a partir de um conjunto de agentes inteligentes cognitivos que implementa categorias de intervenções, geradas a partir do estilo cognitivo, preferências e contexto do usuário. A aquisição dos contextos é feita por um agente a partir da leitura de HTML, a fim de coletar interações dos usuários com o sistema, tais como tipos de atividades que ele participa.

Yu, Zhou e Shu (2010) propõem a *Semantic Learning Space*, uma infraestrutura para recomendação de conteúdos. O recomendador usa dados, tais como o assunto estudado e o tempo de estudo, para propor as recomendações aos estudantes. Se um aluno gastou muito tempo em um determinado assunto, o sistema recomendará outro tópico; por exemplo, se o aluno ficou um longo tempo estudando matemática, será sugerida uma pausa para estudar inglês. Conhecimento a priori do usuário sobre determinado assunto, objetivos, capacidade do terminal (por exemplo, resolução de vídeo e formatos suportados) e as condições da rede (por exemplo, largura de banda e tempo de resposta) também são dados utilizados para gerar as recomendações. Esta ferramenta usa o conceito de sensores virtuais para obter o contexto do usuário, mas não provê uma solução automática para aquisição dos contextos e sim a entrada de dados é feita explicitamente via interface HTML, a partir do preenchimento de formulários.

3.2.3 Personalização e Adaptação

De acordo com Brusilovsky, Kobsa e Nejdí (2007), a tarefa de personalização pode ser visualizada como um problema de predição, onde o sistema tenta prever o nível de interesse do usuário, ou a utilidade para ele, em categorias de conteúdos, páginas ou itens, e ranqueia isso de acordo com valores preditivos. Sistemas recomendadores de conteúdos e recomendadores de conexões sociais são tipos de personalização e, por este motivo, serão apresentados nesta Seção de personalização.

De acordo com Greller e Drachsler (2012) LA oferece a promessa de “aprendizado mais personalizado” que deve possibilitar que os aprendizes tenham experiências de aprendizagem mais efetivas. A personalização da aprendizagem possibilita conceber cursos considerando que os aprendizes não iniciam, necessariamente, o curso no mesmo estágio e não caminham, necessariamente, no

mesmo ritmo de aprendizagem. Nesta Seção, são apresentados alguns sistemas de personalização da aprendizagem no escopo de *learning analytics*.

Paraskakis (2005) propõe um sistema que objetiva facilitar a aprendizagem ao longo da vida e treinamentos vocacionais, por meio do acesso a materiais de aprendizagem que melhor atendam às necessidades individuais do estudante. Três módulos são implementados pelo sistema: gerenciamento de conteúdos, integração de conteúdos e acesso à banda larga multimodal. O sistema seleciona objetos de aprendizagem e os personaliza a fim de possibilitar o acesso de usuários via diferentes larguras de banda, conforme o contexto em que se encontra. Dados como agenda, tarefas, perfil pessoal e experiências anteriores, dispositivo disponível e interesses pessoais ajudam a compreender o contexto do usuário.

Japanese Polite-Expressions Learning Assisting System (JAPELAS) é um sistema desenvolvido para o suporte à aprendizagem da língua japonesa (YIN et al., 2010). Ele busca facilitar o aprendizado de expressões polidas e é direcionado, principalmente, a estudantes estrangeiros para facilitar a boa comunicação com nativos. Considerando que no Japão as expressões mudam de acordo com o contexto (por exemplo, hiponímia, distância social e a formalidade na conversa), JAPELAS usa agentes para prover um vocabulário apropriado ao contexto do usuário. Este sistema usa duas fontes de dados: bases de dados para obter o perfil do aprendiz (por exemplo, nome, idade, gênero, ocupação e interesses) e localização do usuário (por exemplo, se ele está em uma loja, uma sala privada ou em casa). Os dados sobre o perfil do estudante são informados na inicialização do sistema e sua localização é detectada por *tags* RFID e GPS. Esses dados são, então, utilizados para determinar a formalidade da conversa (salas de reunião são incluídas em situações informais).

Schmidt e Winterhalter (2004) apresentam o *Learning In Process* (LIP), um sistema cujo objetivo é entregar conteúdos de aprendizagem sob demanda a partir de um repositório de objetos de aprendizagem e metadados semânticos. O sistema utiliza o perfil do usuário, estilo de aprendizagem, tempo de aprendizagem estimado e nível de competência para compor o contexto do usuário. A aquisição de contextos neste sistema é feita a partir de *web browsers*, interações do usuário com aplicações (via *plugins*) e dados armazenados em bases relacionais (*queries*).

Tabela 3.1 – Resumo dos sistemas estudados: fontes de dados, usuários-alvo e Técnicas de Modelagem e Análise utilizadas.

Sistema	Fontes de Dados				Usuários -Alvo		Técnicas de Modelagem e Análise utilizadas			
	Log	Formulários	Bases de Dados	Sensores Físicos	Estudante	Professor	Semântico	Preditivo	Regras de Decisão Simples	Fuzzy
Course Signals (ARNOLD e PISTILLI, 2012)	X				X			X		
OLI (DOLLÁR e STEIF, 2012)	X	X				X		X		
SAM (GOVAERTS et al., 2012)	X	X			X	X			X	
LOCO-Analyst (ALI, et al. 2012)	X					X	X			
Moodle Dashboard (PODGORELEC e KUJAR, 2011)	X					X			X	
GLASS (LEONY et al., 2012)	X				X	X		X		
Student Inspector (SCHEUER e ZINN, 2007)	X					X		X		
iTree (NAKAHARA et al., 2005)	X				X					X
StepUp! (SANTOS et al., 2013)	X	X			X	X				X
(CHEN, CHANG e WANG, 2008)	X				X	X	X			X
Teacher ADVisor (KOSBA, DIMITROVA e BOILE, 2005)	X					X				X
TADA-Ed (MERCERON e YACEF, 2005)	X					X		X		
SNAPP (BAKHARIA e DAWSON, 2011)	X					X	X			
Classroom View (FRANCE et al., 2006)	X					X		X		
eLAT (DYCKHOFF et al., 2012)	X					X			X	
(BOZO, ALARCÓN e IRIBARRA, 2010)	X					X	X			
(BROISIN et al., 2010)	X				X		X			
Protus (KLASNJA-MILICEVIC et al., 2011)	X				X		X			
UbiGroup (FERREIRA et al., 2014)	X			X		X	X			
iWeaver (WOLF, 2003)	X				X				X	
K-InCa (Nabeth et al., 2005)	X				X		X			
Semantic Learning Space (YU, ZHOU e SHU, 2010)	X	X		X	X	X	X			
(PARASKAKIS, 2005)	X	X		X	X				X	
JAPELAS (YIN, 2010)	X		X	X	X				X	
LIP (SCHMIDT e WINTERHALTER, 2004)		X	X		X		X			

3.2.4 Discussões sobre a captura de contextos em LA

A captura de contextos em LA, onde há grande diversidade de fontes e tipos de dados, é um processo trabalhoso e requer muitos esforços de aplicações individuais. É comum a aquisição de contextos estar associada à lógica da aplicação e ter a coleta restrita a apenas uma parte dos contextos dos usuários conforme apresentado nos sistemas mencionados.

Chatti et al. (2012a) publicaram um estudo sobre fontes de dados que têm sido usadas por sistemas na área de aprendizagem assistida por tecnologias. Eles descobriram que 85% das ferramentas analisadas usam apenas a base de dados do SGA, enquanto as demais usam dados coletados em arquivos de *log* ou então dados publicamente disponíveis. Em muitos casos, a descoberta de contexto do usuário não é automática, mas requer entrada de dados explícita por meio de formulários. Abordagens implícitas geralmente implicam apenas na recuperação de dados gerados no *browser*.

Diversos trabalhos foram propostos na literatura com o objetivo de expandir a coleta de contextos para além das atividades realizadas dentro do SGA ou em *logs*. *TaskTracer* (DRAGUNOV et al., 2005 e JENSEN et al., 2010) coleta a interação do usuário a partir de *plugins* específicos às aplicações do *MS Office*, o que representa a maior limitação desta solução. SWISH (OLIVER et al., 2006) tem uma abordagem similar ao *TaskTracer*, mas realiza o monitoramento a partir de eventos *window* gerados por qualquer aplicação no *PC Microsoft Windows*. SWISH usa *listeners* para os eventos produzidos por cada janela do sistema.

Wolpers et al. (2007) e Muñoz-Merino et al. (2010) apresentam um *framework* que usa um conjunto de ferramentas e *addons* para coleta de dados do comportamento do usuário a partir de aplicações no servidor ou *desktop* do usuário. Nesta solução, cada aplicação monitorada tem seu próprio formato XML, em que os dados coletados são salvos. Esses arquivos XML são transformados em um XML *Schema* via *plugins*. A principal desvantagem desta solução é que apenas atividades que lidam com documentos acessados pelo usuário são observadas, demais atividades, tais como ouvir uma música, jogar ou navegar na web não são monitoradas. Botoianu e colegas (2010) propoem uma federação de metadados de atenção contextualizada (*contextualized attention metadata* - CAM) para obter dados produzidos por qualquer ferramenta ou sistema computacional. Apesar da proposta

objetivar um modelo genérico de coleta de dados, requer uma implementação específica para cada aplicação rastreada (por exemplo, é totalmente dependente de SGA e do repositório de objetos de aprendizagem utilizado) e, portanto, apresenta baixo reuso da arquitetura de coleta.

Com isso, apesar de diversas soluções serem propostas para a recuperação de dados educacionais além do SGA, a investigação de arquiteturas de *software* orientadas para a aquisição de dados educacionais ainda é um desafio na comunidade científica.

3.3 Considerações Finais

Conforme apresentado neste capítulo, pesquisas sobre *Learning Analytics* oferecem a promessa de prever e melhorar o sucesso e retenção dos estudantes (OLMOS e CORRIN, 2012; SMITH, LANGE, e HUSTON, 2012) em parte porque possibilita que estudantes, professores ou instituições tomem decisões que influenciarão no sucesso da aprendizagem; no sucesso de práticas de ensino; no engajamento; na detecção antecipada de riscos de abandono e evasão; nas necessidades e oportunidades de intervenção; e no redirecionamento de recursos.

A análise de dados no domínio educacional em pesquisas atuais está focada especificamente sobre dados de alunos, seus processos e comportamentos de aprendizagem (VAN BARNEVELD, ARNOLD e CAMPBELL, 2012), (GRELLER e DRACHSLER, 2012) e (SIEMENS e LONG, 2011) mesmo quando o foco da pesquisa esteja direcionado para o melhoramento do ensino ou das ações institucionais. Geralmente, a coleta de dados sobre outros usuários envolvidos no processo de ensino-aprendizagem, tais como de tutores e professores, não é realizada e, tais dados nem mesmo são considerados nas análises. Neste caso, a tomada de decisão em nível macro, por exemplo, não está bem apoiada.

Além disso, a maioria dos trabalhos apresentados foca na aquisição de dados a partir de meios convencionais, preferencialmente bases de dados, *logs*, sistemas de posicionamento (RFID e GPS) e sensores físicos, tais como câmeras e microfones, o que limita a obtenção de resultados mais refinados a partir das

análises. A coleta de dados externos ao *browser*, quando é feita, possui restrições quanto ao reuso e extensibilidade da solução apresentada.

Na perspectiva de apoio à tomada de decisão, trabalhos atuais em LA focam em apresentar dados, especialmente, a partir de análises descritiva (visualização) e preditiva (personalização, adaptação e recomendação). Um nível mais avançado de análise de dados, a análise prescritiva, que utiliza-se das análises descritiva e preditiva para indicar sugestões aos usuários, não é alcançado, o que consiste em outra limitação dos trabalhos atuais.

Alguns trabalhos abordam a DDDM em educação sob uma perspectiva do negócio da instituição, tal como na visão apresentada por Luan (2004), que indica que diversas questões críticas em negócios são semelhantes em educação. Por exemplo, enquanto empresas de negócios estão interessadas em questões do tipo “Quem são meus clientes fiéis?” ou “Quais são os clientes susceptíveis a desertar para o concorrente?” em educação as questões seriam “Quais alunos propensos a frequentar mais aulas?” ou “Que tipo de curso vai atrair mais alunos?”. Na visão deste autor, técnicas de mineração de dados podem ser aplicadas à educação da mesma maneira que é feita na área de negócios. Sob um mesmo ponto de vista, em um relatório especial da eSchool News, produzido pela IBM para divulgação do software estatístico SPSS, Natsu (2010) relata que a análise de dados pode ajudar líderes de educação a cortar custos, prover eficiência nos processos e melhorar o ensino e a aprendizagem. No relatório, a autora apresenta um caso de sucesso no emprego de análise estatística educacional da Universidade da Califórnia que economizou \$167 milhões durante 5 anos em dez campi e cinco centros médicos. Van Barneveld, Arnold e Campbell (2012) entendem que a aplicação da análise de dados em educação visa a adoção de práticas para garantir o sucesso organizacional em todos os níveis, abordando questões sobre retenção, admissões, angariação de fundos e eficiência operacional. Esses trabalhos adotam um modelo único tanto para análise de negócios e quanto para análise educacional, onde a tomada de decisão segue um princípio organizacional não abordando particularidades nem da área de negócios e nem da área educacional, o que torna as propostas apresentadas extremamente genéricas.

Dadas as considerações apresentadas, conclui-se a necessidade de novas soluções que direcionem a construção de aplicações para o apoio à tomada de decisão educacional em diferentes níveis de decisão, tais como o nível micro, o nível

meso e o nível macro; a partir de um grande conjunto de dados capturados em diversas fontes, que extrapolem os meios tradicionais de coleta de dados e considere recursos atualmente empregados no ensino e aprendizagem, tais como interações em redes sociais e interações com aplicações externas ao SGA; e que alcance um nível mais avançado na análise de dados do que apenas análises descritivas e preditivas. Neste sentido, o próximo capítulo apresenta a proposta de um Modelo Conceitual que busca vencer as limitações apresentadas. O Modelo Conceitual proposto foi implementado em uma arquitetura de *Framework* e utilizado, na prática, por usuários em situações de cursos reais a partir de três aplicações educacionais analíticas construídas com base no *framework* desenvolvido.

Capítulo 4

MODELO CONCEITUAL PARA O DESENVOLVIMENTO DE FERRAMENTAS ANALÍTICAS EDUCACIONAIS

Este capítulo apresenta o Modelo Conceitual proposto nesta tese para apoiar e guiar o desenvolvimento de ferramentas analíticas educacionais. O Modelo é proposto com o objetivo de servir como base para a construção de aplicações que apoiam a tomada de decisões no domínio educacional nos níveis micro, meso e macro. A Seção 4.1 apresenta as considerações iniciais do capítulo; a Seção 4.2 descreve o Modelo Conceitual proposto em detalhes, incluindo todas as suas etapas: Artefatos, etapa descrita na Subseção 4.2.1; a Coleta e Pré-Processamento dos Dados, descrita na Subseção 4.2.2; a etapa de Análises realizadas sobre os dados, descrita na Subseção 4.2.3, abrangendo análise descritiva, análise preditiva e análise prescritiva; e, por fim, a etapa de Decisão, descrita na Subseção 4.2.4.

4.1 Considerações iniciais

A crescente variedade de fontes geradoras de dados e velocidade de produção desses dados na área de educação gera um volume de dados cada vez maior. Avanços em métodos computacionais para análise automática desses dados prometem melhoria dos processos de aprendizagem na educação formal (SIEMENS e BAKER, 2012), no ensino e na gestão educacional. Pesquisas sobre tecnologias, métodos e modelos para análises desses dados têm sido desenvolvidas por duas

comunidades distintas, o grupo que pesquisa mineração de dados educacionais, em inglês *Educational Data Mining* (EDM), e a comunidade que atua em análise da aprendizagem, em inglês *Learning Analytics and Knowledge* (LAK).

De acordo com Siemens e Baker (2012), apesar de haver similaridades e objetivos comuns, especialmente voltados para a melhoria da qualidade das análises de dados educacionais em grande escala, essas comunidades têm orientações tecnológicas, ideológicas e metodológicas distintas. A primeira distinção observada é com relação ao tipo de descoberta que é priorizada. EDM tem foco maior na descoberta automatizada, enquanto LAK tem foco maior no julgamento humano e a descoberta automatizada é apenas uma ferramenta para alcançar este objetivo. Outra diferença entre as duas comunidades é o tipo de personalização e adaptação suportado. A maioria dos modelos EDM são utilizados para adaptação automatizada, por exemplo sistemas de tutoria inteligente; enquanto que modelos LAK são, geralmente, utilizados com o objetivo de informar e capacitar alunos e instrutores apresentando-lhes evidências. A terceira, e principal, diferença entre essas duas comunidades de pesquisa está relacionada ao entendimento sobre os fenômenos, determinado pelos paradigmas holístico e reducionista (SIEMENS e BAKER, 2012). Na comunidade LAK pesquisadores atuam, tipicamente, na tentativa de compreender os sistemas integralmente, em sua totalidade, e são encontrados muitos *frameworks* holísticos que objetivam apoiar as necessidades de diferentes partes interessadas (*stakeholders*) a partir de informações extraídas dos dados. A comunidade EDM tem forte atuação na descoberta com modelos, onde a maioria dos *frameworks* são reducionistas e objetivam a generalização dos modelos (por exemplo *cross-validation*).

O Modelo Conceitual proposto neste trabalho foi concebido a partir de uma visão holística sobre os processos de aprendizagem, ensino e gestão e tem como objetivo fornecer evidências para o julgamento humano, a partir da análise de dados brutos, a fim de apoiar a tomada de decisão em diferentes níveis. Conhecimentos já produzidos na área de EDM, especialmente a adoção de técnicas de mineração de dados, são utilizados a fim de atingir melhores resultados nas análises.

A Seção seguinte descreve o Modelo Conceitual proposto neste trabalho.

4.2 Descrição do Modelo Conceitual proposto

O Modelo Conceitual proposto neste trabalho é constituído de quatro dimensões principais: artefatos, coleta de dados, análises e decisão (Figura 4.1). As quatro dimensões propostas deverão orientar a construção de aplicações de LA de maneira mais fácil e útil e deverão possibilitar que projetistas responsáveis por processos educacionais analíticos implementem soluções para seus usuários partindo de uma visão holística sobre os resultados alcançados.

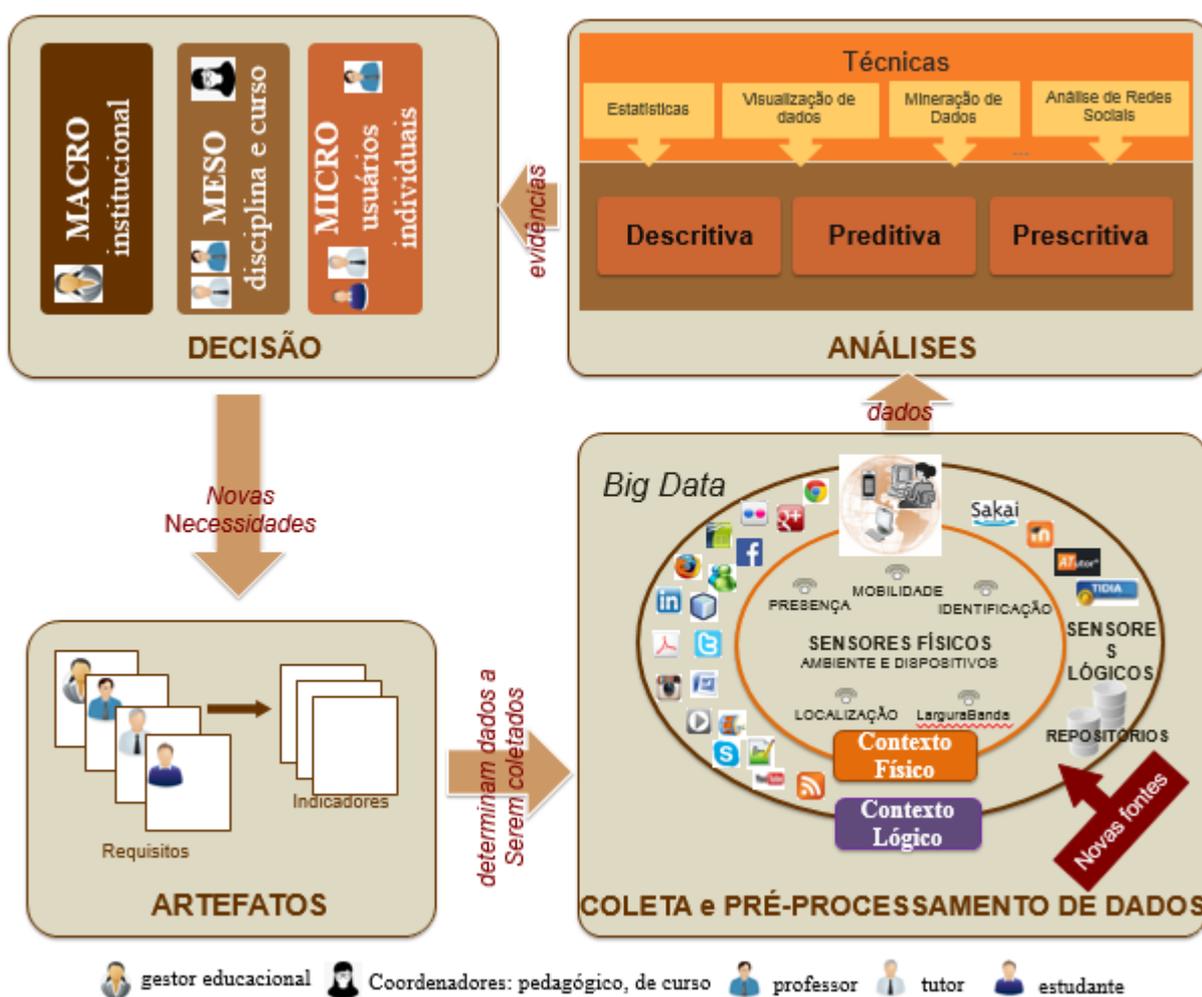


Figura 4.1 – Modelo Conceitual para o desenvolvimento de ferramentas analíticas educacionais. Fonte: elaborado pela autora.

4.2.1 Artefatos

Os artefatos consistem da primeira dimensão do Modelo Conceitual proposto. Artefatos podem ser classificados em requisitos e indicadores. Os requisitos

compreendem as necessidades dos usuários do sistema, são mapeados por meio de variáveis e medidos por meio de indicadores.

Um exemplo de requisito é a medição do sucesso/fracasso em fóruns de discussão. Fóruns de discussão são as ferramentas mais populares em aprendizagem *online*. Embora trabalhos anteriores (SCHELLENS e VALCKE, 2006; MACFADYEN e DAWSON, 2010; GRESS et al., 2010; STRIJBOS, 2011) conduzidos na área de aprendizagem colaborativa suportada por computador (Computer-Suported Collaborative Learning - CSCL) tenham muito bem apresentado variáveis e/ou indicadores de sucesso para fóruns de discussão, esses trabalhos não consideraram a possibilidade de obter dados não disponíveis tradicionalmente no SGA, tal como a porcentagem em que as mensagens de um determinado fórum foram lidas pelo estudante. Partindo desses trabalhos e dos estudos conduzidos pela pesquisadora com especialistas e usuários do sistema para levantamento de requisitos e indicadores, as variáveis para o requisito “medição do sucesso/fracasso em fóruns de discussão” foram mapeados e estão sumarizados no Quadro 4.1. Um total de dez variáveis foram identificadas para fóruns, a partir das quais sete são extraídas do SGA e três (indicadas no Quadro 4.1 por MR, MF e SL) são extraídas a partir de dados obtidos por sensores. A coleta de dados por meio de sensores será abordada na Seção 4.2.2.1. A descrição sobre os estudos conduzidos no escopo deste trabalho para o levantamento de requisitos e indicadores preliminares serão abordados na Seção 4.2.1.1.

Quadro 4.2 – Variáveis selecionadas para medir o sucesso/fracasso em fóruns de discussão.
Fonte: elaborado pela autora.

Variáveis observadas	Significado
Access (AC)	Número de acessos realizados no fórum.
MessagesAmount (MA)	Número de <i>posts</i> submetidos.
ReadMessages (RM)	Porcentagem de mensagens lidas no fórum.
MouseFocus (MF)	Quantidade de tempo em que o mouse permaneceu sobre mensagens do fórum.
Scroll (SL)	Porcentagem de rolagem da página verticalmente.
MessagesRelevance (MR)	Relevância de mensagens submetidas. Mensagens muito curtas, tais como “Eu concordo”, “Eu discordo”, são ignoradas.
Keywords (KW)	Número de palavras-chave em mensagens.
WordsAmount (WA)	Tamanho da mensagem em termos de quantidade de palavras.
Replica (RE)	Número de réplicas postadas.
Rejoinder (RJ)	Número de tréplicas postadas.

Outros exemplos de necessidades que podem ser mapeadas em requisitos e indicadores são relacionados a seguir:

- Conhecer o comportamento do aprendiz e suas preferências: pode ser utilizado para oferecer recomendações aos estudantes a fim de, por exemplo, ajudá-los a melhorar seu aproveitamento escolar. Variáveis relevantes neste exemplo podem incluir: tipos de recursos acessados (texto, vídeo, simulação, etc.) a partir do SGA, *download* de materiais de aprendizagem a partir do SGA; tipos de ações realizadas em paralelo aos estudos (enquanto navega pelo curso no SGA ou enquanto lê/vê um material de aprendizagem baixado localmente), tais como participação em grupos sociais, ouvir música, jogar; preferências por mídias e tipos de materiais, etc.

- Monitorar e/ou bloquear o uso de determinados recursos durante avaliações *online*: pode ser utilizado para prover mecanismos para o suporte às avaliações *online* seguras. Variáveis observadas podem ser: o uso de dispositivos de armazenamento secundário; uso de câmera e microfone; uso de *software* de comunicação instantânea, tais como *Skype* ou *Hangout*; uso de e-mail; navegação por sites; uso de editores de texto, ou outros aplicativos, etc.

- Determinar as condições do dispositivo e da rede do usuário: pode ser utilizado para, por exemplo, traçar um contexto mais amplo do usuário a fim de fazer recomendações, ou, ainda, com o objetivo de prover melhores soluções de aprendizagem adaptativa personalizada. Variáveis relevantes neste exemplo são: tamanho da tela; largura de banda, latência da rede; carga energética disponível, etc.

O Modelo Conceitual proposto neste trabalho determina que novos requisitos podem ser identificados continuamente a partir das necessidades indicadas pelos usuários durante o uso do sistema. Essa tarefa deve ser, preferencialmente, semiautomática, a fim de facilitar a extensão dos artefatos produzidos. Conforme mencionado, os requisitos são mapeados em indicadores mensuráveis quantitativamente e traduzem as necessidades dos usuários. Requisitos e indicadores também determinam quais dados serão coletados para satisfazê-los e, portanto, quais variáveis serão observadas.

A identificação de variáveis a serem observadas para cada requisito identificado, especialmente a determinação dos indicadores, são questões desafiadoras do ponto de vista pedagógico. Primeiro porque medir, independentemente do quê se deseja mensurar, por exemplo uma ação do usuário

ou o seu comportamento, é uma tarefa subjetiva e que depende, fundamentalmente, da intervenção humana. Neste sentido, as variáveis identificadas para cada requisito, bem como sua mensuração, devem ser extensíveis e flexionáveis, conforme interesse do usuário que está utilizando o sistema. Para isso, é necessário fornecer ao usuário uma lista extensível de requisitos e variáveis relacionadas a cada um para que possa selecionar o que é de seu interesse.

Tomando como exemplo o requisito “medição do sucesso/fracasso do estudante em fóruns de discussão”, o professor deve poder selecionar quais serão seus indicadores, conforme a proposta de fórum desejada. Neste caso, o professor deverá selecionar quais variáveis serão consideradas para determinar o sucesso ou o fracasso do estudante, bem como o peso de cada uma delas. Com isso, é possível que o professor estabeleça os seus requisitos e indicadores para cada atividade, assim como para uma unidade de aprendizagem ou para a disciplina.

4.2.1.1 Levantamento preliminar de Artefatos

Determinar corretamente os requisitos e os indicadores é crucial para produzir resultados eficazes em LA. Com o propósito de flexibilizar a utilização do Modelo proposto, principalmente para fins de reusabilidade, foi realizado, como parte deste trabalho de doutorado, um estudo que comparou ferramentas existentes nas seguintes plataformas de SGA: Moodle, SAKAI e TIDIA-AE. A partir deste estudo, concluiu-se que grande parte das ferramentas encontradas nessas plataformas possuem funcionalidades iguais ou similares. Com isso, foi possível selecionar o seguinte grupo de ferramentas presentes nas principais plataformas de SGA a fim de padronizar o mapeamento de variáveis observadas e indicadores: *link* interno para materiais de estudo (texto, html e vídeo); *link* externo para *websites*; fórum; chat; tarefa; quiz; wiki e questionário.

Um conjunto preliminar de variáveis, e indicadores, compatível com as ferramentas selecionadas foi identificado a partir de contribuições teóricas feitas por autores de diversos campos da educação e psicologia do ensino e da aprendizagem e também a partir do suporte direto de professores e especialistas em educação que estiveram envolvidos neste projeto.

Adicionalmente, foi conduzido um estudo com 17 estudantes de cursos de graduação a distância, 9 tutores, 4 professores, 2 coordenadores de curso, 2 coordenadores pedagógicos, 4 *designers* instrucionais e 3 profissionais técnicos que

realizam manutenção de SGA. Este estudo teve como objetivo validar as variáveis e indicadores identificados preliminarmente e também levantar novos requisitos a partir das necessidades de usuários considerando três aplicações com propósitos distintos: (1) organização dos estudos e monitoramento do sucesso/fracasso para fins de autorregulação da aprendizagem – estudo dirigido com alunos; (2) acompanhamento e avaliação (*assessment*) dos processos de ensino e aprendizagem, principalmente, para fins de intervenção – estudo dirigido com tutores, professores, *designers* instrucionais e coordenadores de curso e pedagógicos; e, (3) monitoramento e aplicação de avaliações *online* seguras – conduzido com professores, *designers* instrucionais e técnicos de Tecnologias da Informação (TI). Essas três aplicações serão apresentadas no capítulo 5 desta tese como ViTrackeR, ViMonitor e ViAssess, respectivamente.

Para este levantamento de requisitos, foram utilizados questionários durante entrevistas com os participantes. Durante as seções de entrevistas, foram apresentados exemplos de indicadores para os usuários, que foram pouco explicados, a fim de coletar opiniões e novas ideias. Os seguintes itens foram apontados como importantes para a maioria dos estudantes: o monitoramento dos prazos de entrega de atividades; e, o acompanhamento contínuo do desempenho escolar. Tutores, professores e coordenadores indicaram como pontos essenciais a serem monitorados: se o aprendizado contínuo está ocorrendo; e, a identificação de situações que necessitam de intervenção a fim de apoiar esse processo. Com relação ao monitoramento de avaliações *online*, professores e técnicos de TI indicaram como pontos principais a serem observados: atividades paralelas do estudante durante uma avaliação; e, a possibilidade de bloquear algumas ações do estudante durante avaliações, tais como o uso de *pen drive*.

Com este estudo, obteve-se um conjunto amplo de requisitos e indicadores que foram, parcialmente, implementados no *framework* desenvolvido, cuja descrição geral é apresentada no capítulo 5 deste trabalho.

A coleta de dados educacionais é a base para o desenvolvimento de ferramentas de LA que deverão servir ao apoio dos processos de ensino, aprendizagem e gestão. A Seção seguinte apresenta a segunda dimensão do Modelo Conceitual proposto nesta tese, que consiste na coleta e no pré-processamento de dados.

4.2.2 Coleta e Pré-Processamento de Dados

O SGA consiste na principal fonte de dados centralizados sobre as atividades e interações dos usuários, tais como acesso aos materiais de aprendizagem e testes realizados, além de dados estatísticos, como a quantidade de interações e o número de tentativas de respostas. Outras bases de dados, públicas e privadas, podem ser usadas para prover dados de acordo com o interesse de aplicações, por exemplo, silos tradicionais de dados disponíveis via arquiteturas orientadas a serviço – (*Service Oriented Architecture* - SOA). Além disso, o crescente uso de informações e tecnologias de comunicação em educação, móveis e ubíquas, amplia o conjunto de dados produzidos pelos usuários e suas aplicações, o que deixa trilhas de dados disponíveis que podem aumentar a qualidade e profundidade das análises (Siemens, 2013). Recursos tais como *players* multimídia, mensageiros instantâneos, jogos, páginas *web*, e-mails e programas de compartilhamento de arquivos são frequentemente usados durante a aprendizagem. Com isso, os dados podem estar em diferentes formatos, distribuídos ao longo do espaço, do tempo e da mídia. A dimensão *Coleta e Pré-processamento de Dados*, presente na Figura 4.1, ilustra o ambiente ciente de contexto no âmbito educacional, onde determinado contexto pode compreender informações sobre o mundo físico e informações sobre o mundo lógico.

4.2.2.1 Coleta dos Dados

A captura de contextos é passo essencial para apoiar usuários durante seus processos de ensino, aprendizagem e gestão. Tipicamente, ciência de contexto provê serviços a partir de sensores distribuídos em ambientes físicos inteligentes. Partindo desta abordagem, neste trabalho, a coleta de dados em ambientes de aprendizagem abertos envolve o uso de rede de sensores híbrida, composta por sensores físicos e sensores lógicos (ROSALES et al. 2010), e podem aumentar o volume de dados úteis, que refletem as atividades distribuídas do usuário, conduzindo a resultados mais sólidos de análise da aprendizagem

Os sensores físicos podem ser sensores de ambiente, ou sensores de dispositivo. Dados obtidos por sensores físicos podem ser muito úteis para análises educacionais. Por exemplo, se um determinado aluno não usa restaurantes

universitários, isso pode indicar que ele não é tão inserido no campus e, talvez, possa ser um candidato potencial para evadir da instituição.

Conforme apresentado no capítulo 2 desta tese, um sensor lógico é um componente de *software* (microagente), que pode capturar dados relevantes e específicos, tais como suas tarefas, progresso do trabalho, preferências, relacionamentos em redes sociais, horários de estudo, etc. Esta abstração utilizada para a coleta de dados a partir de ambientes digitais possibilita a cobertura de todos dados abundantemente espalhados nos ambientes com os quais o usuário interage.

A abordagem desenvolvida nesta tese para a coleta de contextos tem a capacidade de coletar dados distribuídos em ambientes de rápida mudança de informação e a partir de múltiplas e heterogêneas fontes a fim de criar um conjunto de dados educacionais útil.

A coleta de dados a partir do dispositivo do usuário deve ser transparente para ele e deve ser feita sem sua intervenção explícita e sem perturbar suas atividades de estudo. Além disso, a coleta de dados deve cobrir amplamente os comportamentos dos usuários a fim de produzir análises confiáveis e para servir, satisfatoriamente, aos diversos propósitos de aplicações de LA. Para isso, sensores lógicos devem rastrear pistas de dados e obter um conjunto significativo e de granularidade fina a partir do monitoramento das interações do usuário dentro e fora do *browser*, a partir de interações com aplicações e a partir do monitoramento de arquivos do sistema.

Com o potencial de coleta mencionado, onde a proliferação de dispositivos, móveis ou não, é crescente e a disponibilidade de sensores, lógicos ou físicos, é igualmente progressiva, é possível obter um conjunto extenso e profundo de dados sobre os usuários. Esses dados podem incluir: o local onde o aluno está, com quem ele está, qual dispositivo ele está usando, o que ele está fazendo (tomando sorvete, assistindo um filme no cinema, estudando em casa via SGA, acessando as redes sociais, ouvindo música, assistindo uma vídeoaula), quais os ruídos presentes no ambiente, entre outros. Contudo, monitorar e medir um conjunto imenso de dados que não sejam relevantes pode ser algo indesejável, principalmente pelo fato de produzir um conjunto massivo de dados que não serão utilizados e que poderão, inclusive, dificultar a geração de análises confiáveis. Desta forma, a identificação correta dos artefatos, incluindo as variáveis associadas a cada requisito e os

indicadores que serão utilizados para fazer as medições, é algo crucial para o bom funcionamento de uma ferramenta educacional analítica e deve ser cuidadosamente planejada durante o projeto e desenvolvimento de sistemas desta natureza.

4.2.2.2 Pré-processamento dos Dados

As fontes de dados indicadas no Modelo Conceitual proposto incluem: (i) o ambiente digital pessoal do usuário; (ii) o ambiente físico no qual ele se encontra e com o qual ele interage; (iii) os repositórios de dados. Considerando essa variedade de fontes e considerando, ainda, a diversidade de tecnologias utilizadas na coleta, problemas relacionados à qualidade do dado, abordados no capítulo 2 desta tese, podem estar presentes. Por exemplo, a aquisição de dados a partir de diferentes bases (institucionais, públicas ou privadas, e SGA) pode apresentar ambiguidade e conflito de dados, uma vez que os mesmos dados podem estar duplicados em diferentes fontes, por exemplo, informações pessoais do estudante. A coleta a partir do ambiente digital do usuário também pode apresentar problemas na qualidade do dado obtido. Por exemplo, a ausência do atributo 'gênero' pode dificultar a identificação do tipo de música que o estudante gosta de ouvir enquanto estuda matemática.

Desta forma, uma fase de pré-processamento dos dados deve ser prevista com a finalidade de transformar dado imperfeito em um formato adequado que possa ser utilizado como entrada para um método particular de análise. Várias técnicas, principalmente do campo de mineração de dados, podem ser utilizadas nesta fase e incluem: limpeza de dados, integração de dados, transformação de dados, redução de dados e fusão de dados (LIU, 2006; ROMERO e VENTURA, 2007 e HAN e KAMBER, 2011).

Dados coletados nesta segunda dimensão do Modelo Conceitual proposto podem ser utilizados para encontrar padrões significativos e para criar modelos descritivos, preditivos e prescritivos para apoiar a aprendizagem, o ensino e a gestão educacional. Alguns exemplos incluem a identificação de estudante academicamente em risco de fracasso; recomendações que auxiliem o estudante a melhorar seu aproveitamento acadêmico; personalização de ambientes de aprendizagem de acordo com o ritmo e estilo de aprendizagem do usuário; monitoramento contínuo da aprendizagem e do aproveitamento escolar para fins de intervenção e monitoramento da atuação de equipes de ensino e intervenções.

A Seção seguinte apresenta os tipos de análises previstas no Modelo proposto.

4.2.3 Análises

A terceira dimensão do Modelo Conceitual proposto nesta tese consiste da análise dos dados obtidos na fase anterior, com os objetivos de facilitar a compreensão sobre o que aconteceu e o que está acontecendo (análise descritiva); prever situações baseando-se em dados históricos (análise preditiva) e oferecer *insights* para decisões mais inteligentes baseadas em dados sobre o que deverá ser feito (análise prescritiva). Com isso, a análise de dados educacionais prevista no Modelo Conceitual contempla três níveis, ou categorias, de análises, conforme ilustrado na Figura 4.2.



Figura 4.2 – Categorias de análises de dados. Fonte: elaborado pela autora.

A análise descritiva proporciona melhor entendimento sobre o que aconteceu, tendo como base dados históricos (DAVENPORT, 2012). Resultados de análises descritivas são, geralmente, apresentados por meio de relatórios, tabelares ou gráficos, que possibilitam facilitar a compreensão sobre as razões para o sucesso ou fracasso no passado.

Modelos de análise descritiva evoluíram para modelos preditivos, que utilizam dados históricos sobre um evento e integra novos conjuntos de dados a fim de fazer previsões sobre o futuro ou indicar a probabilidade de que uma situação ocorra (NYCE, 2007). A análise preditiva, conforme visto no capítulo 3 deste documento, teve valor expressivo em trabalhos desenvolvidos na área de LA nos últimos anos.

Um nível mais elevado de análise em relação às categorias anteriores, ainda escassamente explorado em educação *online*, e pouco explorado na área de gestão de negócios, é a análise prescritiva. Este tipo de análise parte da previsão obtida na categoria 2 de análise e prescreve recomendações, ou ações, para influenciar o que acontecerá no futuro (SCHAFFHAUSER, 2014). Um exemplo para este tipo de análise em educação, apresentado por Schaffhauser (2014) é um sistema de monitoramento da aprendizagem que poderia recomendar material adicional ou *sites* da *web* para um aluno com baixo desempenho em um determinado tópico. Apesar de ser um conceito atual e cujas discussões sobre seu emprego na área de educação estar nos seus primeiros rudimentos, neste trabalho esta categoria de análise está presente no Modelo Conceitual proposto e poderá ser empregada para resolver diversos problemas atuais da educação *online*, tais como a retenção dos alunos, a melhoria da aprendizagem e o re/planejamento pedagógico.

4.2.3.1 Técnicas para Análises de Dados

Diversas técnicas podem ser aplicadas para investigar dados e encontrar padrões escondidos em conjuntos de dados educacionais. O Modelo Conceitual proposto contempla quatro grupos de técnicas de análises que podem ser utilizados sobre os dados coletados em situações de ensino e aprendizagem: estatísticas, visualização de dados, mineração de dados e análise de redes sociais. Outras técnicas podem ser adicionadas conforme necessidade e interesse de aplicações.

O primeiro conjunto de técnicas indicado no Modelo é de **Estatísticas**. A maioria dessas estatísticas são disponibilizadas pelo próprio SGA por meio de relatórios de dados. Esses dados, embora úteis, muitas vezes são difíceis de serem interpretados, sobretudo pela grande quantidade de informações exibidas, geralmente, em formato tabelar.

De acordo com Card, Mackinlay e Shneiderman (1999), a **Visualização de Dados** consiste no uso de representações de dados visuais, suportadas por computadores para ampliar a cognição. Desta forma, o conjunto de técnicas de

visualização de dados proposto no Modelo Conceitual consiste em representar os resultados, obtidos a partir dos métodos estatísticos de análises, em formatos de maior representatividade visual a fim de facilitar a interpretação e análise das informações. Diversas representações gráficas, como histogramas, gráficos lineares, retangulares, de setores, barra de progresso, *scatterplot*, mapas e representações 3D, podem ser usadas para representar as informações em formato claro e compreensível. Visualização de dados possibilita oferecer a observação sobre um conjunto maior de variáveis em uma mesma representação visual. É desejável que as ferramentas educacionais analíticas ofereçam a opção de selecionar diferentes possibilidades de representações visuais para o usuário a fim de possibilitar o uso da representação que ele está mais familiarizado ou que tenha maior facilidade para interpretar.

Liu (2006) define a **Mineração de Dados** como o processo de descoberta de padrões úteis ou conhecimento a partir de fontes de dados, tais como bases de dados, textos, imagens e web. Há diversos métodos que podem ser utilizados para minerar dados, a maioria se enquadra nas seguintes categorias: aprendizagem supervisionada (ou classificação e previsão), aprendizagem não supervisionada (ou agrupamento) e regras de associação (LIU, 2006). Técnicas de mineração de dados são previstas no Modelo Conceitual proposto a fim de ampliar o suporte e atendimento às diferentes necessidades de usuários, viabilizando a construção de ferramentas analíticas a partir de modelos de análise preditivos e prescritivos.

A **Análise de Redes Sociais** possibilita representar visualmente, por meio de grafos, as relações sociais e conexões a partir do estudo quantitativo das ligações entre os indivíduos (D'ANDREA; FERRI e GRIFONI, 2009). A análise de redes sociais pode ser utilizada, por exemplo, para melhorar recomendações.

As técnicas de análises de dados podem ser utilizadas individualmente ou em conjunto, dependendo do objetivo da ferramenta de análise. O Modelo Conceitual proposto nesta tese considera esses diversos grupos de técnicas apresentado com o propósito de apoiar a construção de ferramentas de análise em qualquer uma das categorias apresentadas: descritiva, preditiva e prescritiva. Para isso, a implementação de uma ferramenta analítica educacional deve vencer o desafio de projetar e desenvolver análises a partir da técnica, ou técnicas, mais apropriada para atender aos objetivos da aplicação e minimizar o tempo entre a análise e a ação, apoiando a tomada de decisão de estudantes, de instrutores e/ou de gestores. A

ferramenta deve, ainda, possibilitar a extensão incremental do volume de dados para análises e garantir desempenho e escalabilidade.

A dimensão de Análises presente no Modelo Conceitual proposto provê evidências resultantes das análises realizadas para apoiar a tomada de decisão educacional, que consiste da quarta dimensão do Modelo, apresentado na Seção seguinte.

4.2.4 Decisão

Conforme exposto no capítulo 3 deste documento, sobre o estado da arte em tomada de decisão educacional dirigida a dados, a utilização de dados a fim de melhorar práticas educacionais não é algo novo. Trabalhos publicados no início da década anterior relatam que a análise de dados para avaliar os processos educacionais pode direcionar melhores decisões (FELDMAN e TUNG, 2001; LACHAT, 2002; PARDINI, 2000; PROTHEROE, 2001, SYMONDS, 2003). Todavia, o grande volume de dados disponíveis nos mais variados meios e formatos heterogêneos, e as inúmeras análises possíveis sobre esses dados, gera grande quantidade de trabalho e esforço humano, difícil de mensurar, e de extrema complexidade de ser realizado em um universo de educação *online* que pode reunir dados de centenas ou milhares de estudantes e outros usuários do ambiente. Sendo assim, o uso de tecnologias computacionais para coleta e análise de dados é fundamental para fornecer acesso às informações, facilitar a compreensão e análise sobre os processos educacionais e possibilitar melhores decisões individuais (nível micro), em nível de disciplina e de curso (nível meso) e em nível de gestão educacional (nível macro).

O **nível micro** de decisão do Modelo Conceitual reflete decisões tomadas por usuários em relação às suas ações individuais. As decisões tomadas nesse nível não causam consequências, diretas, aos demais usuários. Exemplos de decisões neste nível: o aluno decide se seguirá a agenda de estudo proposta pelo sistema; e o tutor decide se fará a correção das atividades da turma A antes de corrigir as atividades da turma B, conforme sugerido pela ferramenta. Decisões em nível micro são tomadas por estudantes, tutores e professores, conforme ilustrado na Figura 4.1.

Em **nível meso**, as decisões são tomadas a partir da análise de um conjunto de variáveis e suas relações no tocante a uma disciplina ou a um curso. Exemplos

de decisões neste nível são: o tutor pode decidir intervir em um fórum em que o nível de participação dos alunos está baixo; o professor pode decidir alterar os critérios de avaliação de uma atividade; o coordenador de curso pode decidir encaminhar os tutores de determinado polo de apoio presencial para capacitação. Neste nível, as decisões são tomadas por tutores, professores, coordenadores de curso e coordenadores pedagógicos.

As decisões em **nível macro** são tomadas por gestores de projetos de EaD e estão relacionadas à gestão institucional. Um gestor pode decidir, por exemplo, oferecer maior número de vagas em determinado curso – essa decisão é comumente tomada por gestores em instituições privadas. Outras possibilidades neste nível de gestão são: monitoramento da retenção e do grau de evasão; monitoramento do grau de conclusão; observação de estratégias de intervenção em níveis de curso e polo; monitoramento de equipes e avaliação de políticas institucionais para fins de replanejamento.

De acordo com Marsh, Pane e Hamilton (2006), a tomada de decisão educacional pode ser dividida em duas categorias: aquelas que implicam na utilização dos dados para informar, identificar ou esclarecer (por exemplo, identificação de necessidades) e decisões que envolvem o uso de dados para agir (por exemplo, mudança curricular e realocação de recursos). Neste segundo caso, uma vez que a ação foi executada, novos dados podem ser coletados para servirem como retroalimentação ao sistema conduzindo a um ciclo contínuo de coleta de dados e análises a fim de gerar novos resultados para apoio às decisões, representado na Figura 4.1 do Modelo Conceitual como *feedback*.

4.3 Considerações Finais

Com o objetivo de orientar a construção de tecnologias de apoio à decisão, o Modelo Conceitual proposto amplia a coleta de dados para além dos dados internos (bases institucionais e SGA), considerando também as interações do usuário em outros ambientes (físicos e digitais) e o uso de diferentes tecnologias (de *software* e de *hardware*). O Modelo também contempla uma variedade de técnicas (Estatísticas, Visualização de Dados, Mineração de Dados e Análise de Redes Sociais) que

possibilitam a implementação das três categorias de análises (descritiva, preditiva e prescritiva) a fim de facilitar as interpretações sobre os resultados, realizar previsões e medir o impacto de decisões futuras.

O Modelo Conceitual proposto neste trabalho foi implementado a partir de um *framework* que contempla todas as dimensões propostas do Modelo. O *framework* é apresentado na Seção seguinte.

Capítulo 5

FRAMEWORK PARA COLETA E ANÁLISES DE DADOS E APLICAÇÕES EDUCACIONAIS DESENVOLVIDAS

“Education has always been considered as potentially one of the most productive breeding grounds for technology. On the other hand, technology can be considered as a catalyst for teaching and learning and can help in enhancing positive educational change.” (GIBSON, ALDRICH e PRENSKY, 2007).

Este capítulo do trabalho apresenta um framework ciente de contexto que consiste de dois componentes principais: a captura de contextos e a modelagem e raciocínio de contextos. A abordagem desenvolvida tem a capacidade de coletar dados distribuídos em ambientes de rápida mudança de informação e a partir de diversas fontes e processá-los em modelos de contextos extensíveis para atender às necessidades de serviço e aplicações de LA com diferentes propósitos.

5.1 Considerações iniciais

A captura de contextos em LA, onde há grande diversidade de fontes e tipos de dados, é um processo complexo. Conforme visto no capítulo 3, várias das soluções e aplicações existentes para a coleta de dados educacionais possuem limitações que restringem seu uso. As principais limitações incluem: aquisição de contextos fortemente acoplada à lógica da aplicação; coleta restrita a apenas uma parte dos contextos dos usuários; fontes de dados usadas são frequentemente associadas apenas aos clássicos arquivos de *log* e/ou bases de dados dos SGA;

dados de contexto são limitados aos eventos gerados dentro do *browser* (ações realizadas pelo estudante fora do *browser* não são monitoradas); coleta a partir de *plugins* específicos às aplicações proprietárias; coleta a partir de preenchimento de formulários, e não de maneira automática; coleta de atividades restritas apenas a acesso, pelo usuário, a documentos e conteúdos do SGA (atividades como ouvir música, jogar ou navegar na *web*, usualmente, não são monitoradas); solução dependente de SGA e repositório de objetos de aprendizagem utilizado. Neste sentido, embora muitas soluções estejam sendo desenvolvidas, a construção de arquiteturas de *software* orientadas para a aquisição de dados educacionais ainda é um desafio na comunidade científica, especialmente no que se refere ao estudo de mecanismos para a ampla e automática recuperação de contextos do usuário com o objetivo de melhorar os serviços oferecidos a eles.

A fim de facilitar o desenvolvimento de aplicações analíticas educacionais para o apoio ao ensino, à aprendizagem e à gestão educacional, em especial as tarefas de coleta e análises de dados, o Modelo Conceitual proposto neste trabalho foi implementado a partir de um *framework*. Segundo Gamma et al., (1994), um *framework* é um conjunto de classes relacionadas, as quais permitem a reutilização em um projeto com classes específicas do *software*.

O *framework* desenvolvido neste trabalho, e apresentado neste capítulo, classifica-se como um *framework* caixa cinza, uma vez que traz funcionalidades prontas para serem utilizadas, sem a necessidade do contato direto com o código, mas também estão presentes características como generalidade e facilidade de extensão (FAYAD, SCHIMIDT e JONHSON, 1999). O *framework* desenvolvido possui arquitetura aberta, extensível, robusta e independente de plataforma a partir da abstração de redes de sensores híbridos, que desacopla o processo de captura de contextos da lógica da aplicação. O rastreamento de dados é feito a partir de múltiplas e heterogêneas fontes, disponíveis em diferentes formatos, incluindo: (i) bases de dados de informação contextual (SGA, públicas ou privadas), (ii) ambiente de trabalho digital (*personal workspace*) do usuário e (iii) ambiente físico. Com isso, contextos capturados podem descrever um comportamento do usuário, uma atividade do usuário, um ambiente (físico ou lógico) do usuário, o estado físico do usuário ou essas informações interligadas, tal como um comportamento do usuário em uma determinada atividade em um dado ambiente e em condições específicas. Intenções de interação também são monitoradas, tais como tentativas de respostas

não confirmadas e cliques realizados sobre objetos em uma página. As intenções de interação consistem de um tipo especial de dados, que, a princípio, não foi armazenado em bases, mas que pode ser útil para compreender, por exemplo, o comportamento do usuário.

5.2 Visão Geral do *Framework* de Coleta e Análises de Dados

A fim de estender a coleta de dados para além das fontes tradicionais, que são SGA e bases de dados institucionais, sensores lógicos são depositados em três pontos principais: dispositivo do usuário, SGA e ambiente físico. No primeiro caso, eventos gerados pelos usuários e suas aplicações, incluindo as intenções de interação, são monitoradas por sensores lógicos implementados com dois tipos de tecnologias: agentes e *plug-ins* de *browser*. No segundo caso, interfaces simples de serem manipuladas são oferecidas para consultas às bases de dados. Por fim, sensores físicos monitoram a ocorrência de eventos em ambientes físicos, tais como salas de aula, laboratório ou biblioteca.

5.2.1 Arquitetura do *Framework*

O *framework* desenvolvido possui arquitetura aberta e modular e visa disponibilizar um conjunto de classes que auxiliam o desenvolvedor na implementação de aplicações analíticas educacionais. O *framework*, continuamente, coleta dados distribuídos em diferentes fontes, transfere esses dados para aplicações e serviços interessados e armazena dados brutos para acessos futuros. O *framework* também realiza a modelagem de contextos e análises sobre os dados, obtidos a partir de diversas fontes, a fim de minimizar a complexidade na implementação das aplicações. A Figura 5.1 mostra a arquitetura multicomponente do *framework* proposto.

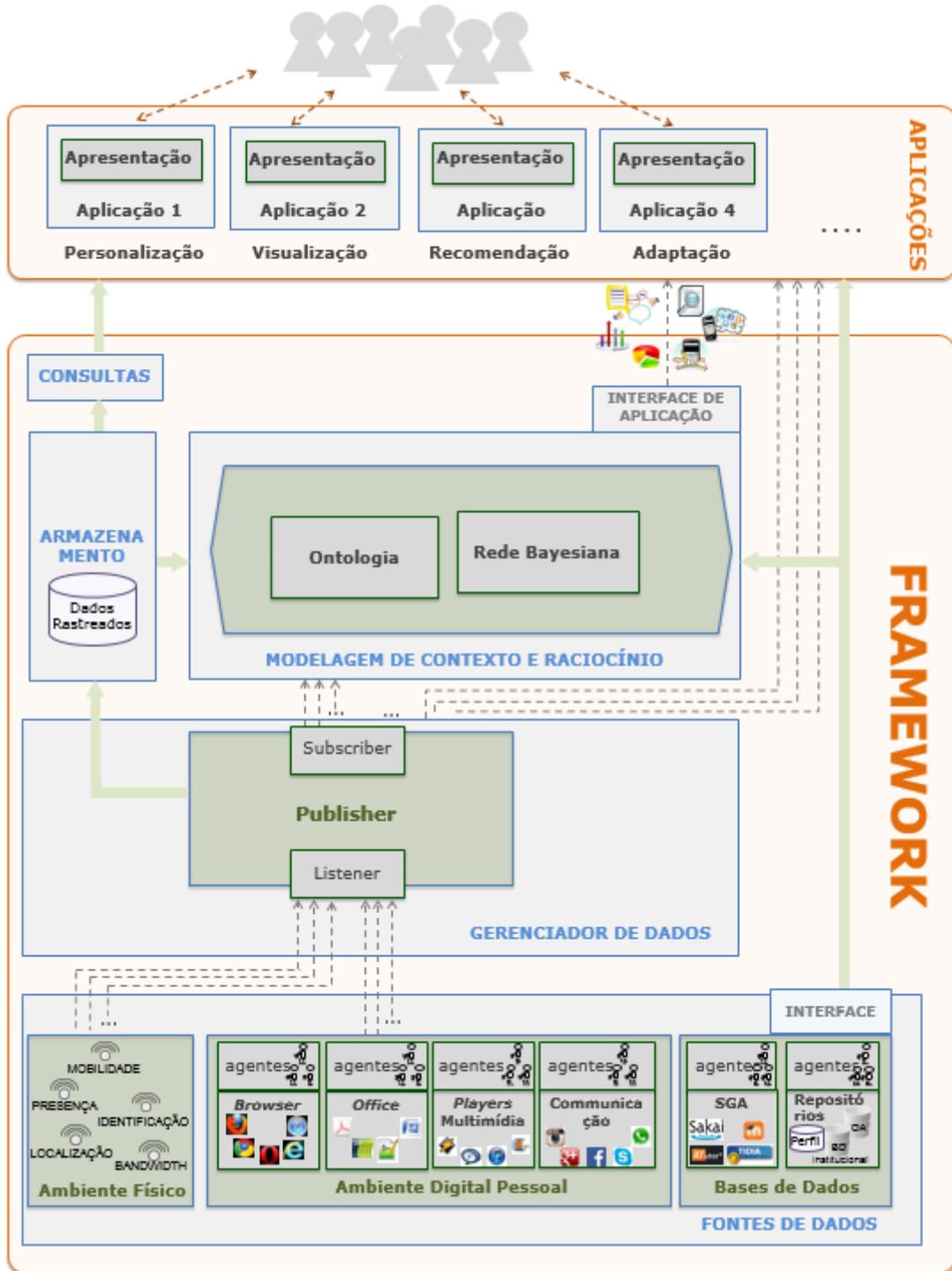


Figura 5.1 – Arquitetura proposta do *Framework* para apoiar o desenvolvimento de aplicações educacionais analíticas. Fonte: elaborado pela autora.

As funcionalidades do *framework* são apresentadas seguindo uma arquitetura em camadas, as quais são descritas nas subseções seguintes.

5.2.1.1 Fonte de Dados

O *framework* foi implementado a partir de padrões de *design* que descrevem uma solução reusável para captura de contextos a partir de três tipos de fontes de dados distintas: ambiente físico; ambiente digital pessoal do usuário; e, bases de dados de informações contextuais (por exemplo, SGA, bases públicas ou privadas).

A coleta realizada a partir do ambiente digital pessoal do usuário utiliza nós sensores lógicos, onde cada nó coleta, paralela e independentemente uns dos outros, um evento específico. A coleta realizada por sensores lógicos possibilita amplo rastreamento dos comportamentos do usuário a partir de interações executadas em páginas *web* e eventos de aplicações locais. Dados coletados a partir do rastreamento das interações via *browser* incluem: a duração da interação e sequência de eventos naquela interação; duração de intervalos de tempo de inatividade (ou seja, sem interações) durante as visitas às páginas do SGA ou demais páginas na *web*; cliques em objetos ativos da página; uso de barra de rolagem, entre outros. O rastreamento do ambiente local, externo ao *browser*, inclui: interações com objetos baixados a partir de uma disciplina no SGA e duração da interação; execução de aplicações paralelamente enquanto estuda no SGA, relacionadas ou não com a unidade de estudo atual, tais como mensageiros instantâneos, tocadores multimídia, etc.; e, o uso de dispositivos de entrada e saída, tais como *pen drive*, câmera e memória secundária. Características do dispositivo físico (por exemplo, tamanho de tela) e da rede de comunicação (por exemplo, largura da banda, latência) também são obtidos por sensores lógicos.

Sensores lógicos são implementados como um sistema multiagentes, onde cada sensor lógico é considerado um micro agente artificial com existência própria e independente dos demais. Tais sistemas oferecem uma organização estruturada e provê características de trabalho concorrente e distribuído com padrões de interação sofisticados (BRADSHAW, 1997). Assim, micro agentes (sensores lógicos) são projetados para agirem em paralelo e tem características de sentir e coletar um tipo de dado específico. A especificação e divisão de funcionalidades entre micro agentes provê requisitos importantes em arquitetura distribuída, por exemplo, modularidade, flexibilidade, capacidade de modificação e extensibilidade.

O *framework* também suporta mecanismos de atuação no lado cliente. Atuadores são agentes responsáveis por realizar uma ação apropriada no ambiente,

conforme interesse das aplicações. Por exemplo, o atuador *lockUserActuator* bloqueia o computador do usuário. Outro exemplo de atuador implementado é o *showOnBrowserActuator*, responsável por criar um elemento `<div>` para exibir parâmetros CSS e HTML na janela do *browser*.

Com isso, a coleta realizada pelo *framework* proposto produz um conjunto de dados detalhado e de granularidade final que pode refletir fielmente as atividades distribuídas do usuário e, então, conduzir a análises de dados educacionais mais sólidas e precisas.

5.2.1.1.1 Pré-processamento dos Dados

Dada a diversidade de fontes e o grande volume de dados produzidos por diferentes interfaces consideradas na fase de aquisição realizada pelo *framework*, torna-se evidente a complexidade do tratamento dos dados obtidos a fim de eliminar, ou reduzir, imperfeições. Considera-se que não há como resolver todos os problemas que podem estar relacionados à imperfeição dos dados em único algoritmo de fusão sem acarretar outros problemas decorrentes da complexidade da infraestrutura de *software* necessária.

A centralização de todos esses dados obtidos (navegação, interação, comunicação do usuário, dados do ambiente físico e dados de repositórios) em um único repositório de dados a fim de facilitar o processo de fusão é algo impraticável e, portanto, uma solução desconsiderada. Isso não se deve a questões de armazenamento físico, que não é um problema dadas as excelentes características de arquiteturas de *hardware* de armazenamento atuais, mas sim às questões de gerenciamentos que se fazem necessárias e novas soluções que sejam eficientes em resolver o problema da fusão de todos esses dados de maneira integrada.

Conforme apresentado na Seção 2.6.2 deste documento, algumas pesquisas sobre abordagens para fusão de dados obtidos por fontes heterogêneas estão sendo propostas atualmente. No entanto, não há soluções automatizadas, por enquanto, que resolvam todos os problemas de imperfeição em dados de fontes híbridas, especialmente pela dificuldade de lidar com incertezas de dados subjetivos, tais como redes sociais, conteúdos de mensagens de fóruns e e-mails. Soluções atuais necessitam de intervenção humana para validação ou complementação de dados e não são totalmente automáticas.

Abordagens diferentes podem ser usadas para fusão de dados no ambiente físico, no ambiente digital e em repositórios. O tratamento de dados para eliminar imperfeições, apesar de extremamente importante, não é o foco deste trabalho. Partindo deste ponto de vista e considerando, ainda, que eventuais falhas no funcionamento do sistema proposto nesta tese, ocasionadas por imperfeições nos dados coletados, não comprometem a segurança de pessoas, também não promovem danos materiais, financeiros ou naturais, a questão da qualidade do dado não é implementada em todos os níveis de abstração do *framework* proposto, mas apenas no nível da rede física. Essa decisão foi tomada tendo em vista que o tempo necessário para proposição de uma solução que resolva os problemas com dados imperfeitos considerando a heterogeneidade de tipos de dados, fontes, e interfaces de coleta, inviabilizaria o desenvolvimento do *framework* proposto, e portanto, desta pesquisa.

5.2.1.1.2 Aspectos tecnológicos

O *framework* foi escrito em Java. Por razões de implementação, a coleta de dados no ambiente digital do usuário é dividida em duas subredes de sensores lógicos: sensores para rastreamento de eventos dentro do *browser*; e sensores para rastreamento fora do *browser*. No primeiro caso, sensores lógicos são implementados usando tecnologias JavaScript e AJAX e são embarcados em *browsers* via *plugin*. Dados coletados no *browser* são publicados no *sink*, conforme ilustra a Figura 5.2. Um *container servlet*, Jetty¹⁴, é usado para manipular solicitações assíncronas. Neste trabalho, *sink* é um nó receptor responsável por gerenciar os mecanismos de *publish/subscribe* (p/s) localmente e também pela comunicação entre a rede e serviços.

No segundo caso, sensores para rastreamento fora do *browser* são implementados usando *Java Agent Development Framework* (JADE) (BELLIFEMINE, POGGI, e RIMASSA, 1999) em conformidade com a *Foundation for Intelligent Physical Agents* (FIPA¹⁵) que suporta o desenvolvimento de sistemas multiagentes. Um sensor lógico é implementado como um micro agente coletor de

¹⁴ Jetty Web Server, <http://www.eclipse.org/jetty/about.php>.

¹⁵ Foundation for intelligent physical agents, <http://www.fipa.org/>.

dados. Um agente tem um modelo de computador multitarefa, onde tarefas (ou comportamentos) do agente são executados simultaneamente.

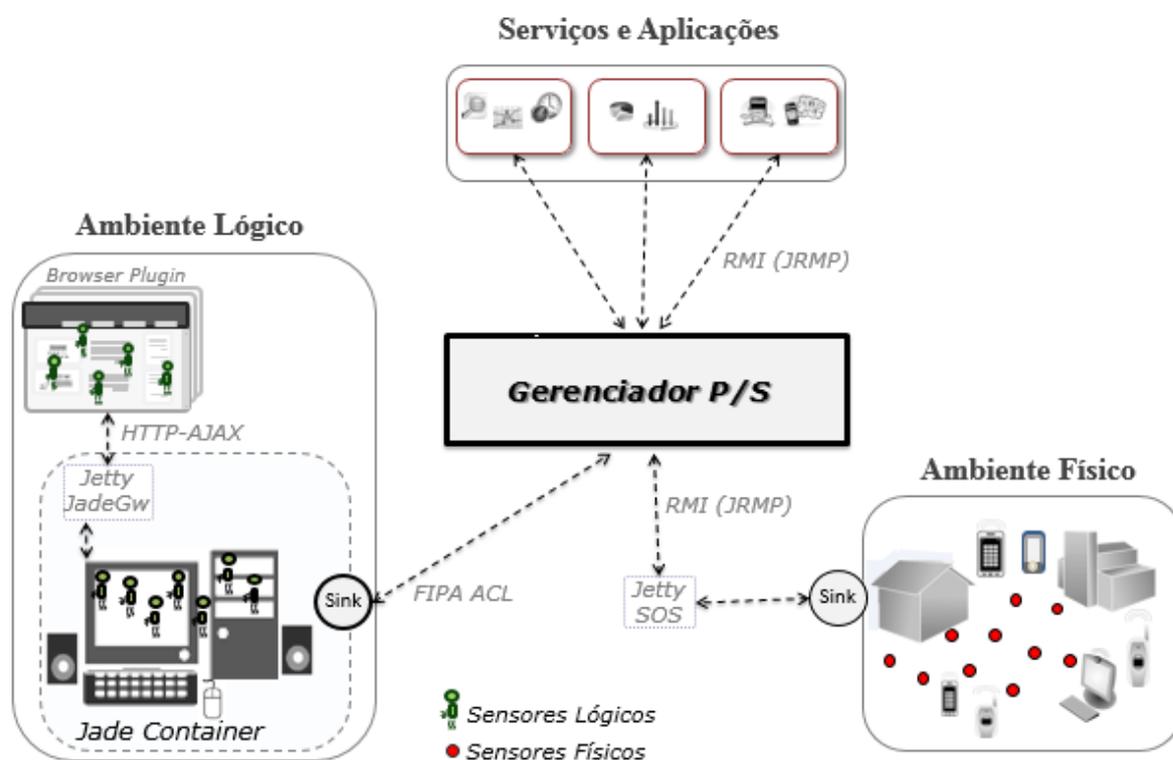


Figura 5.2 – Coleta distribuída de dados realizada por redes de sensores híbrida e tecnologias utilizadas. Fonte: elaborado pela autora.

O Quadro 5.1 apresenta alguns exemplos de sensores lógicos desenvolvidos no escopo deste trabalho a partir da identificação de requisitos. Sensores físicos também são usados para coleta de dados convencionais, tais como a localização do usuário. *Sensor Observation Service* (SOS¹⁶) é usado para consultas aos dados dos sensores na rede física em tempo real. O padrão PostgreSQL é usado como base de dados para armazenamento persistente.

Outro aspecto tecnológico relacionado ao *framework* desenvolvido é que ele facilita a adição de novas funcionalidades, caracterizando-o como uma solução reusável. Os seguintes padrões de *design* foram utilizados com o propósito de reduzir a reescrita de código: (i) *abstract factory* para manter um conjunto consistente de sensores de acordo com a plataforma subjacente, por exemplo em sistemas Windows deve-se garantir que nenhum sensor implementado para Linux

¹⁶ Sensor observation service, <http://www.opengeospatial.org/standards/sos>.

será criado; (ii) *singleton*, para assegurar a criação de uma *factory* somente durante o ciclo de vida dos sensores em uma plataforma, ou máquina, em particular; e (iii) *strategy* para isolar as diferentes implementações de sensores de acordo com a plataforma em uso. A implementação do *framework* possibilita a fácil integração de novos nós sensores. Para isso, algumas funcionalidades são disponibilizadas para facilitar seu uso, tais como as seguintes opções de manipulação de sensores: *create_sensor* – cria um sensor com parâmetros básicos para serem estendidos a fim de incorporar novas fontes de dados; *update_sensor* – possibilita alterar parâmetros de sensores já existentes, incluir ou excluir parâmetros; e *delete_sensor* – possibilita a exclusão de sensores criados que não serão utilizados pelo sistema.

As decisões de projeto foram influenciadas pelo desenvolvimento extensível de novos nós sensores e atuadores, rápido desempenho, independência de plataforma e solução reusável.

Quadro 5.3 – Sensores Lógicos e Atuadores. Fonte: elaborado pela autora.

	Nome do Sensor	Dado Coletado
Dentro do Browser (DB)	quiz_sensor	Cliques realizados em formulários.
	whereAtLMS_sensor	Localização do usuário no SGA.
	login_sensor	Login no SGA.
	focusChange_sensor	Recebimento e perda de foco (seleção de objetos).
	scrolling_sensor	Uso de barras de rolagem.
	resizing_sensor	Redimensionamento da janela.
	keystrokes_sensor	Pressionamento de teclas.
	accessedMaterials_sensor	Materiais teóricos acessados no SGA.
	accessedActivities_sensor	Atividades acessadas no SGA.
Fora do Browser (FB)	USBsensor	Uso de portas USB.
	AppMonitorSensor	Aplicações em execução no dispositivo do usuário.
	ResponseTimeSensor	Tempo de resposta em comunicação entre dois pontos.
	BandwidthSensor	Velocidade de <i>download</i> e <i>upload</i> .
	InstantCommunicationSensor	Ferramentas de comunicação instantânea sendo usadas.
	CameraSensor	Uso da câmera.
	MicrophoneSensor	Uso do microfone.
	Atuador	Atuação
FB	LockUserActuator	Bloqueia o computador do usuário – função semelhante à opção de <i>hibernate</i> .
DB	showOnBrowserActuator	Cria um elemento <div> para exibir parâmetros recebidos.

5.2.1.2 Gerenciador de Dados

A camada de recepção e gerenciamento dos dados, identificada por Gerenciador de Dados na Figura 5.1, é responsável por gerenciar o envio e recebimento dos dados entre aplicações e redes de sensores em ambas as direções. Dados são encaminhados, em tempo real, para aplicações e também armazenados em base de dados para acesso futuro.

O padrão de comunicação *publish/subscribe* foi utilizado por oferecer escalabilidade e topologia de rede dinâmica. O mecanismo *publish/subscribe* segue uma arquitetura fracamente acoplada dirigida a eventos, ou seja, dados são gerados e consumidos independentemente. Assim, *subscribers* podem ser vistos como consumidores de eventos (tais como aplicações e sensores), os quais registram interesse em receber eventos e, então, recebem notificações assincronamente. *Publishers* são geradores de eventos, ou seja, cada evento coletado é publicado para todos os *subscribers* interessados.

Os dados coletados por sensores são persistidos em uma base de dados relacional para acesso posterior e processamentos. Considerando o grau de variabilidade e o número de campos de dados opcionais, foi escolhido XML como formato de dados, que facilita sua compreensão e uso por aplicações e serviços.

5.2.1.3 Modelagem de Contextos e Análises

Neste trabalho, a fim de atender aos diferentes propósitos de aplicações (tais como recomendação, visualização e segurança de avaliações *online*) a modelagem de contextos e o raciocínio, a partir da aplicação de técnicas de análises sobre os contextos, são fases dissociadas do processo de coleta, o que possibilita a utilização de técnicas direcionadas ao foco da aplicação, facilidade de extensão e reuso.

A fim de reduzir a complexidade das aplicações analíticas cientes de contexto, melhorar a capacidade de evolução do sistema e possibilitar o reuso para aplicações de diferentes propósitos, a camada de modelos de contextos proposta no *framework* usa duas abordagens para modelagem dos dados: um modelo gráfico probabilístico (Rede Bayesiana) para análise dos dados, com a finalidade de encontrar relações entre variáveis de entrada e medir valores futuros - predição; e, um modelo semântico-conceitual, baseado em ontologias, para garantir descrição

semântica dos contextos com alta expressividade. Um conjunto de regras de decisão é utilizado para propósitos de facilitar e melhorar a visualização dos dados para compor os dois modelos de contexto a fim de alcançar um nível de análise prescritiva.

A abordagem adotada possibilita o compartilhamento de contextos, permitindo uma interpretação comum de informações de contextos em diferentes aplicações.

5.2.1.3.1 O Modelo Semântico-Conceitual - Ontologias

Modelos ontológicos especificam seus conceitos, relações e fatos de maneira uniforme (CHO e HONG, 2008), (HERVÁS, BRAVO, e FONTECHA, 2010) com apoio à interoperabilidade, descrição semântica e compartilhamento de contextos a partir de uma interpretação comum.

O modelo ontológico proposto neste trabalho tem a função principal de descrever a informação em um caminho semântico e também proporcionar suporte ao raciocínio lógico, provendo ainda características próprias da tecnologia, tais como expressividade, extensibilidade, facilidade de compartilhamento e reuso. Para isso, um serviço de inferência de informação contextual e interfaces de consultas às ontologias, ilustrados na Figura 5.3, são disponibilizados pelo *framework*.

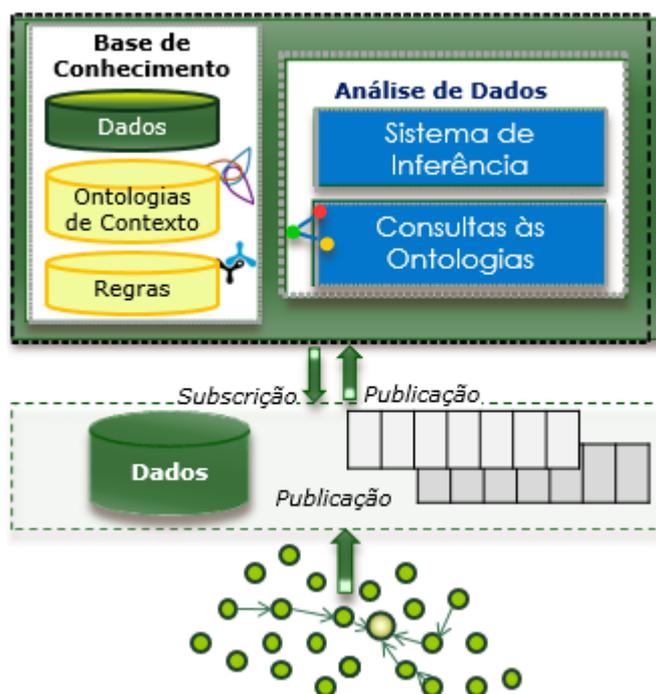


Figura 5.3 – Representação da Modelagem de Contextos e Análise de Dados baseados em Ontologias. Fonte: elaborado pela autora.

As interfaces de consulta às ontologias utilizam a linguagem SPARQL de consulta para RDF (SEABORNE, 2006). O sistema de análise de dados provê um conjunto de inferências baseados em lógica de primeira ordem para a interpretação dos eventos e geração de informações. A *Web Ontology Language* (OWL) é utilizada para descrever os contextos a fim de facilitar o processo de inferência baseado em regras (ZHAO et al., 2008). Isso possibilita que sensores em diferentes plataformas tenham um vocabulário comum e um conjunto comum de conceitos quando interagindo entre si.

A base de conhecimento é composta por dados brutos, para acesso futuro quando requisitado por aplicações; o repositório Ontologias de Contexto, presente na Figura 5.3, consiste das ontologias desenvolvidas; e, um conjunto de regras, que podem ser aplicadas sobre os dados. O *framework* Jena¹⁷ foi utilizado para a manipulação de ontologias e aplicação de inferência. Ontologias foram usadas para modelar as interações do estudante em aprendizagem *online*, conforme requisitos e indicadores previamente determinados. Foram desenvolvidas duas ontologias principais: uma ontologia de interações e comportamento do usuário no SGA; e, uma ontologia de interações e comportamento do usuário com outros recursos, ferramentas e aplicações. No primeiro caso, são mapeadas todas as ferramentas comuns à maioria dos SGA e suas relações, conforme estudo apresentado na Seção 4.2.2.1. Na segunda ontologia foram mapeados requisitos e indicadores previamente determinados relacionados às interações do usuário fora do SGA. A Figura 5.4 ilustra apenas parte da ontologia desenvolvida para modelagem dos contextos de interações dentro do SGA.

¹⁷ Jena. Java Framework for building Semantic Web Applications licensed by The Apache Software Foundation. <http://jena.sourceforge.net/>

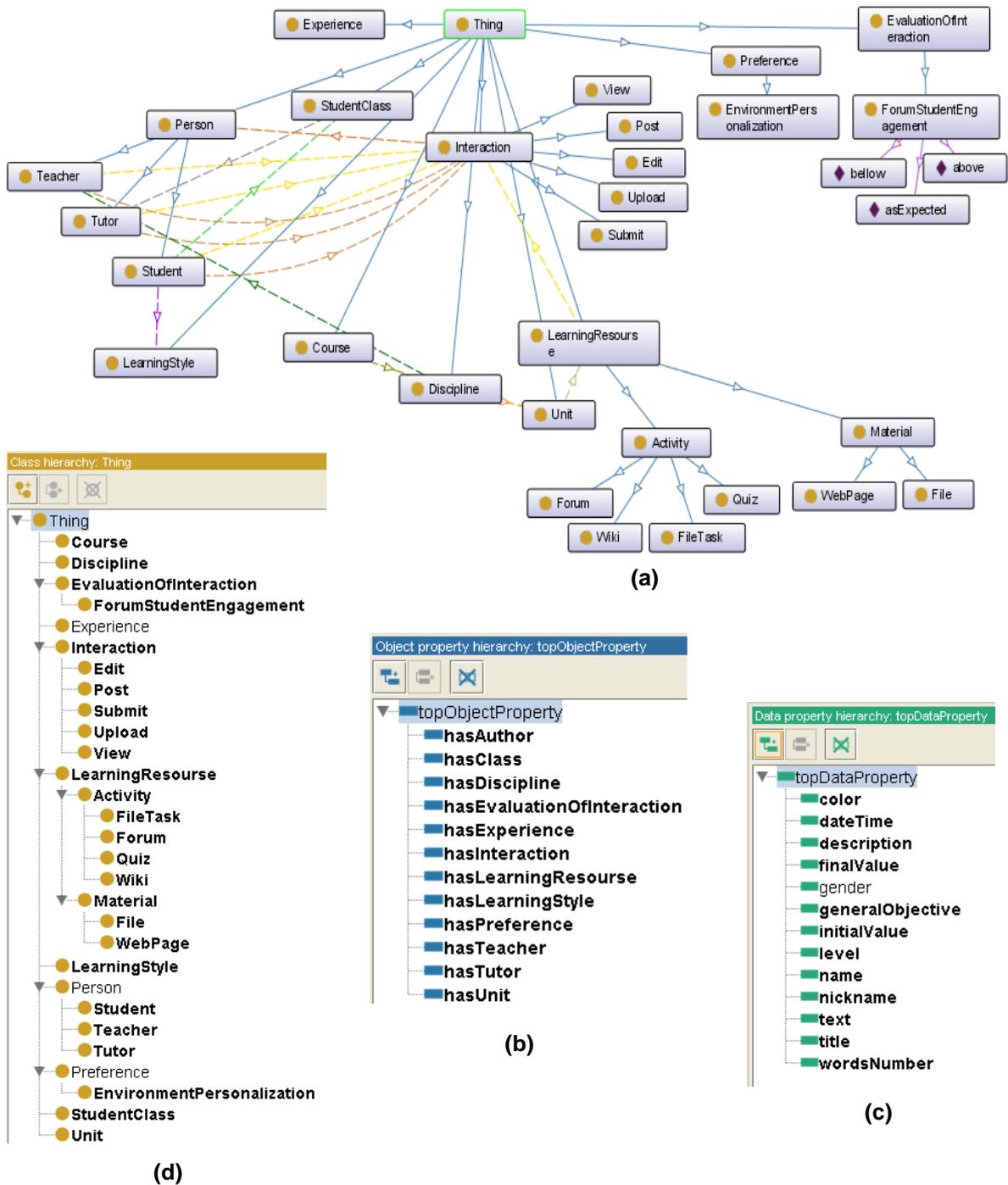


Figura 5.4 – Parte da Ontologia para Modelagem de Contextos em LA. (a) Representação gráfica da ontologia. (b) Classes. (c) Propriedades de objeto. (d) Propriedades de dados. Fonte: elaborado pela autora.

Conforme apresentado no capítulo 3, ontologias consistem de um método de modelagem de contextos bastante adotado em LA, especialmente para fins de adaptação e personalização. Todavia, deve-se pesar a escolha entre alta a expressividade de modelos ontológicos e sua complexidade de gerenciamento, especialmente na representação de dados dinâmicos. De acordo com Bettini e

Brdiczka (2010), ontologias não são adequadas para processar informações em ambientes dinâmicos de constante mudança da informação de maneira escalável. Esse modelo foi disponibilizado no *framework* para uso de aplicações onde o tempo de raciocínio em ambientes de rápida mudança de informação não é algo crítico, tais como aplicações de monitoramento da aprendizagem *online* para fins de avaliação, onde a maioria dos dados utilizados como entrada para as análises são obtidos a partir de bases de dados de contextos (coletados por sensores e armazenados para acesso futuro).

A Seção seguinte apresenta o modelo de contexto baseado em Redes Bayesianas implementado e disponibilizado no *framework*.

5.2.1.3.2 O Modelo Gráfico Probabilístico – Redes Bayesianas

Redes Bayesianas (RB), desenvolvidas na década de 1980 por Judea Pearl (1988), são um modelo probabilístico inspirado pela causalidade (esquema causa-efeito) e provê um modelo gráfico no qual cada nó representa uma variável randômica em um domínio. Em nosso caso elas representam comportamentos dos estudantes, e cada arco representa um relacionamento de influência causal ou dependência entre essas variáveis. A Figura 5.5 ilustra a Rede Bayesiana que modela as relações entre as variáveis associadas com comportamentos dos usuários em fóruns de discussão. Essas variáveis foram identificadas no levantamento de requisitos, conforme apresentado na Seção 4.2.1.1 e no Quadro 4.1.

Atualmente, de acordo com Brusilovsky and Millán (2007), as RB são consideradas uma das melhores técnicas disponíveis e comumente usada como recurso poderoso para diagnósticos (inferências sobre as possíveis causas de um evento) e predição (estado/evolução futuro das variáveis dadas as evidências).

Cada nó da RB tem uma Tabela de Probabilidades Condicionais (TPC) que especifica a probabilidade de cada estado possível de um nó dada a combinação de estados de seus pais, conforme mostra a TPC do nó *MsgRel* ilustrado na Figura 5.5; ou então a distribuição da probabilidade a priori para nós sem pais (ou nós independentes), como a TPC do nó *Access* (Figura 5.5).

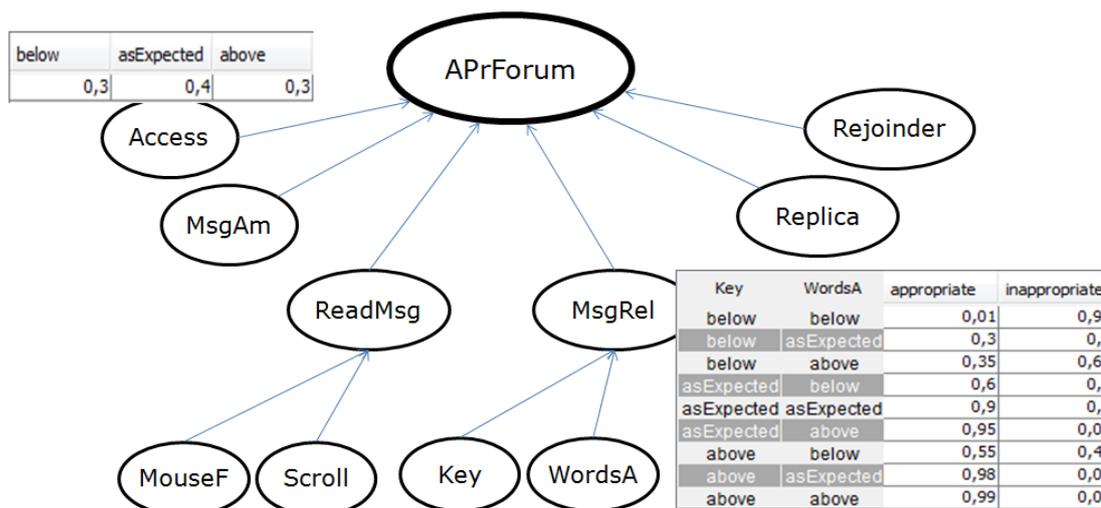


Figura 5.5 – Rede Bayesiana usada para modelar o comportamento do estudante em fórum de discussão. Fonte: elaborado pela autora.

Por ser complexa e extensa, a Figura 5.5 ilustra apenas parte do modelo desenvolvido, neste caso restrito às atividades de fóruns. Essa rede é usada para determinar o progresso acadêmico em fóruns de discussão e permite a modelagem de parâmetros qualitativos (estrutura da RB – nós e arcos) e informações quantitativas (parâmetros da RB – probabilidade a priori, condicionais ou distribuições). Este modelo inclui os aspectos relevantes do mundo real, tais como comportamentos observados e relações entre eles, e as probabilidades condicionais especificadas representam a força desses relacionamentos. Cada comportamento observado é classificado em uma das seguintes categorias: abaixo do esperado, dentro do esperado e acima do esperado; ou então como: apropriado e inapropriado. Os parâmetros de cada categoria podem ser definidos pelo professor durante o planejamento da disciplina. Ele pode fazer isso para cada comportamento observado. Entretanto, para reprimir a rejeição do sistema pelo excesso de configurações, essas configurações são opcionais e, caso elas não sejam feitas pelo professor, o sistema usa valores padrão pré-definidos. Por exemplo, o valor que determina o número de acessos esperado do estudante em uma atividade de fórum.

É possível fazer inferências a partir da RB para prever o progresso acadêmico de estudantes e prover recomendações em relação à sua participação em atividades, na disciplina e no curso. Para o exemplo do progresso acadêmico em fóruns de discussão, obtém-se valores do nó *APrForum* a partir dos valores dos nós

independentes. O progresso acadêmico do estudante em determinado fórum é dado pelo maior valor calculado de probabilidade – positivo ou negativo.

Para calcular a inferência probabilística é usado um conjunto de variáveis proposicionais $V(V_1, V_2, \dots, V_n)$ e um subconjunto de variáveis E de V com valores definidos, usado como evidência, $E = \varepsilon$ (positivo e negativo; ou como: abaixo do esperado, dentro do esperado ou acima do esperado; ou como: apropriado e inapropriado). O cálculo da probabilidade condicional, baseado no teorema de Bayes, é dado por $p(V_i = v_i | E = \varepsilon)$, onde a variável V_i tem um valor v_i dada a evidência.

Sendo assim, para determinar o progresso do estudante é preciso calcular a probabilidade $p(APrForum = positive)$, ou seja, $p(APrForum = positivo | AC, MA, RM, MR, RE, RJ)$ e $p(APrForum = negativo)$, ou seja, $p(APrForum = negativo | AC, MA, RM, MR, RE, RJ)$. Assumindo que a probabilidade $p(APrForum = positivo)$ é 94.99%, $p(APrForum = negativo)$ é 5.01%, pode-se inferir que o estudante está tendo um progresso positivo na atividade.

Para implementação do modelo probabilístico foi utilizada a *engine* de inferência Bayesiana implementada no pacote de API Weka para calcular o progresso acadêmico em cada atividade avaliativa, na unidade (ou tópico), em disciplinas e, então, o seu progresso global em um determinado curso.

5.2.1.4 Camada de Aplicações

A camada do topo da arquitetura é composta de aplicações analíticas educacionais cientes de contexto que manipulam os dados de contexto de acordo com sua conveniência. Essas aplicações podem ter diferentes propósitos, conforme já mencionado. O *framework* foi desenvolvido para facilitar a construção dessas aplicações e pode ser utilizado para reduzir o trabalho de desenvolvimento, especialmente na coleta de dados distribuída e nas análises de dados.

As aplicações devem implementar a camada de apresentação de dados, que envolve a entrega e formatação da informação para o usuário por meio de componentes de interface que realizam controle da aparência visual nas saídas de informações.

5.2.2 Contribuições do *Framework* proposto

A modelagem e a interpretação de contextos para atender vários propósitos da LA é uma questão crítica, observada e deduzida a partir da investigação de trabalhos propostos nos principais eventos e documentos das principais organizações que atuam com pesquisas em LA, tais como a 4th *International Conference on Learning Analytics na Knowledge* em Indianápolis, EUA (LAK 2014); a 13th *International Conference on Advanced Learning Technologies* (ICALT, 2013), em Beijing, China; *International Conference WWW/Internet* (2013), no Texas, EUA; EDUCAUSE, *International Society for Technology in Education* (ISTE) e outros. Conforme apresentado no capítulo 3 sobre o estado da arte de LA, a modelagem de contextos, geralmente, limita-se a atender às necessidades da aplicação em foco, por exemplo, recomendação, personalização ou visualização, e há poucos trabalhos que buscam atender aos vários propósitos da LA.

Uma proposta de modelagem e inferência/raciocínio de contexto genérica que atenda aos diferentes propósitos da LA talvez seja algo inviável, uma vez que a aplicação prática e usabilidade são, frequentemente, inversamente proporcionais à generalidade do modelo (BOLCHINI et al., 2007), ou seja, quanto mais expressiva e poderosa, menos prática e útil. Além disso, sem conhecer o propósito da aplicação, qualquer tipo de interpretação pode gerar uma reação inapropriada por parte da aplicação. Considerando isso, o *framework* desenvolvido no escopo deste trabalho entrega dados em um formato útil para que aplicações e serviços interessados apliquem técnicas de modelagem e raciocínio conforme seu interesse. Adicionalmente, o *framework* implementa um modelo híbrido de modelagem de contextos que une uma abordagem preditiva e uma abordagem semântico-conceitual a fim de melhorar a representação dos contextos.

Um conjunto de regras de decisão também foi implementado a fim de facilitar a visualização de informações graficamente em diferentes formatos e escalas. Com isso, é possível oferecer para o usuário mais do que somente uma recomendação, mas também informações que o ajudam a compreender as variáveis associadas àquela recomendação, ou seja, o “porquê” da recomendação. Por exemplo, um sistema pode indicar para um aluno que seu desempenho acadêmico é de 40% de sucesso em uma determinada disciplina, ou seja, o semáforo está vermelho; e pode

recomendar que ele gaste mais tempo estudando materiais teóricos disponibilizados pelo professor. Para que o aluno possa refletir sobre o “porquê” seu sucesso é de 40% apenas, o sistema pode ainda oferecer um conjunto de informações gráficas relacionadas às variáveis avaliadas no modelo preditivo utilizado para gerar as recomendações. Desta forma, um nível de análise prescritiva é alcançado.

Com isso, há duas contribuições principais trazidas na proposta do *framework* desenvolvido. A primeira contribuição está relacionada à coleta descentralizada de contextos que desacopla tarefas mencionadas anteriormente e produz um conjunto de dados detalhado e que pode refletir mais fielmente as atividades distribuídas de usuários. Além disso, não apenas dados de estudantes são coletados e modelados, mas também dos demais usuários envolvidos nas atividades de ensino e aprendizagem, tais como tutores e professores. A segunda contribuição está relacionada aos modelos de contextos disponíveis no *framework* que consistem de modelagens genéricas e extensíveis com objetivo de facilitar a implementação de aplicações de LA. Desta forma, atividades de coleta de contextos, modelagem e inferência, atualmente implementadas pelas próprias aplicações, são agregadas em um *framework* único.

5.2.3 Privacidade dos Dados

Observações extensivas sobre o comportamento dos usuários é desejável e benéfico para ele e possibilita a provisão de serviços personalizados que são de seu interesse. Todavia, a observação dos comportamentos dos usuários, com a abordagem descrita neste trabalho, traz algumas preocupações sobre a privacidade das informações coletados.

A fim de assegurar a privacidade dos usuários, foram considerados alguns aspectos durante o projeto e desenvolvimento do *framework*. A primeira preocupação foi desenvolver um *software* legítimo com alto consentimento do usuário. Antes de instalar a ferramenta no lado cliente, um termo de consentimento é apresentado aos usuários com orientações explícitas sobre as funcionalidades do *software*, os impactos que ele pode causar no dispositivo do usuário e os tipos de dados que serão coletados. Neste mesmo termo de consentimento, são apresentadas formas de contato com a pesquisadora e com a instituição de ensino responsáveis pelos dados, de modo que o usuário pode solicitar a visualização dos

dados coletados a qualquer momento, questionar os procedimentos utilizados e solicitar exclusão de qualquer informação pessoal.

Durante a instalação do *software*, usuários podem selecionar quais sensores lógicos deverão ser instalados. Além disso, a ferramenta (por meio das aplicações analíticas que serão apresentadas na Seção seguinte) oferece opções de pausar e/ou desativar a coleta de dados. Com isso, o usuário pode determinar quanto e em que medida os dados pessoais serão capturados.

É importante destacar que os sensores lógicos não fazem cópias de documentos, total ou parcialmente, e também não monitoram conteúdos inseridos/lidos pelos usuários em seus e-mails, mensageiros instantâneos, formulários *online* ou qualquer outro tipo de troca de mensagens. Com isso, buscou-se proteger o consentimento informado do usuário, evitar interpretações destrutíveis e oferecer possibilidades de poder sobre a coleta de dados, autonomia e integridade.

A fim de validar o *framework* proposto, três aplicações analíticas educacionais cientes de contexto foram desenvolvidas e são apresentadas na Seção seguinte.

5.3 Aplicações Desenvolvidas a partir do *Framework* proposto

A fim de possibilitar a validação do *framework* proposto, foram desenvolvidas três aplicações analíticas educacionais cientes de contexto: ViTrackeR, que auxilia na promoção da aprendizagem autorregulada de estudantes, provendo a visualização de dados de rastreamento e recomendações personalizadas que objetivam facilitar a tomada de decisões em nível individual; ViMonitor, que consiste em uma ferramenta particularmente dedicada a auxiliar, em tempo real, a equipe de ensino e a equipe pedagógica, provendo informações importantes para a tomada de decisões em nível de disciplina e curso a partir das interações e participação de estudantes e tutores em disciplinas *online*; e, ViAssess, que provê suporte refinado para aplicação de avaliações *online* seguras. Essas aplicações são descritas nas seções seguintes.

5.3.1 ViTracker: Visualização de Dados e Recomendações

ViTracker é uma implementação construída a partir do *framework* proposto, que integra dados do SGA e dados de sensores para produzir análises refinadas sobre o engajamento do estudante em um curso e para oferecer visualização e exploração da informação processada. Esta aplicação, automaticamente, determina o progresso acadêmico individual do estudante e, proativamente, o auxilia a partir de sugestões de ações que podem ser tomadas para melhorar seu progresso. A ferramenta também possibilita a visualização de informações relevantes que podem ajudar o estudante a tomar decisões durante seu processo de aprendizagem. Soluções anteriores que suportam os processos de ensino e aprendizagem focam, principalmente, no diagnóstico de problemas e, geralmente, são identificados ao final dos cursos, quando a recuperação da aprendizagem é mais difícil, e em muitos casos, impraticável.

A aprendizagem autorregulada é um processo de construção ativo, por meio do qual estudantes estabelecem metas para sua aprendizagem e em seguida monitoram, regulam e controlam sua cognição, motivação e comportamento, guiados por seus objetivos e características contextuais do ambiente (PINTRICH, 2000). Conforme proposto por Bandura (2001), a criação de um ambiente educacional favorável para a aprendizagem autorregulada, onde há espaço para a escolha e controle do estudante, propicia autonomia relacionada à sua própria aprendizagem. Neste sentido, aplicações que atuam no suporte da aprendizagem autorregulada proveem a avaliação de lacunas de aprendizagem do estudante, reflexão, autoavaliação e gerenciamento da informação. Desta forma, ViTracker foi desenvolvida com o objetivo de prover assistência personalizada aos estudantes em aprendizagem *online*. Esta aplicação utiliza a modelagem preditiva usando redes Bayesianas, conforme apresentado na Seção 5.2.1.3.2.

ViTracker possibilita aos estudantes o monitoramento de informações a partir de uma perspectiva individual ou em grupo. Neste sentido, a ferramenta provê dois módulos com funcionalidades distintas: visualização do progresso acadêmico sumarizado e visualização do progresso acadêmico detalhado.

5.3.1.1 Progresso Acadêmico Sumarizado

A visualização do progresso acadêmico sumarizado é apresentada durante a navegação do estudante no SGA e permite que ele acompanhe seu progresso acadêmico em determinado componente avaliativo (atividade, tópico ou disciplina) por meio de um gráfico de barras multicolorido (Figura 5.6).



Figura 5.6 – Representação gráfica do progresso acadêmico do estudante em uma barra de progresso multicolorida. Fonte: elaborado pela autora.

O gráfico apresentado na Figura 5.6 convenientemente provê observação sobre aumentos e quedas em *scores* de avaliação. O gráfico consiste de três zonas: segura, de alerta e de perigo. A porção vermelha indica zona de perigo, a porção laranja indica zona de alerta e a verde é usada para a zona segura.

Para esta apresentação de dados sumarizados, ViTrackeR usa o atuador *showOnBrowserActuator*. Este atuador aceita parâmetros HTML, incluindo os resultados das inferências via representação gráfica, e parâmetros CSS para formatação da saída, tais como tamanho e posicionamento dos elementos na tela, incluindo formatação da própria interface de apresentação.

O módulo de visualização do Progresso Acadêmico Sumarizado utiliza indicadores de comportamento provendo diferentes níveis de análise, permitindo que estudantes possam observar, refletir e analisar seu desempenho acadêmico em relação aos demais alunos da turma. A Figura 5.7 ilustra a barra de navegação utilizada pelo sistema para exibir as informações ao estudante enquanto navega no SGA. A barra de progresso e o valor numérico da probabilidade de sucesso prevista (progresso acadêmico) são exibidos em (A).

O estudante pode facilmente acessar informações detalhadas sobre cada comportamento observado (variáveis utilizadas no modelo Bayesiano) que levou ao valor previsto no componente avaliativo clicando em um *link* em formato de lupa (B). Ao selecionar esta opção, a Figura 5.8 será exibida. Em (C) o estudante visualiza uma recomendação feita pelo sistema; e, em (D) o estudante tem as opções de desativar o recebimento de informações temporária ou permanentemente.

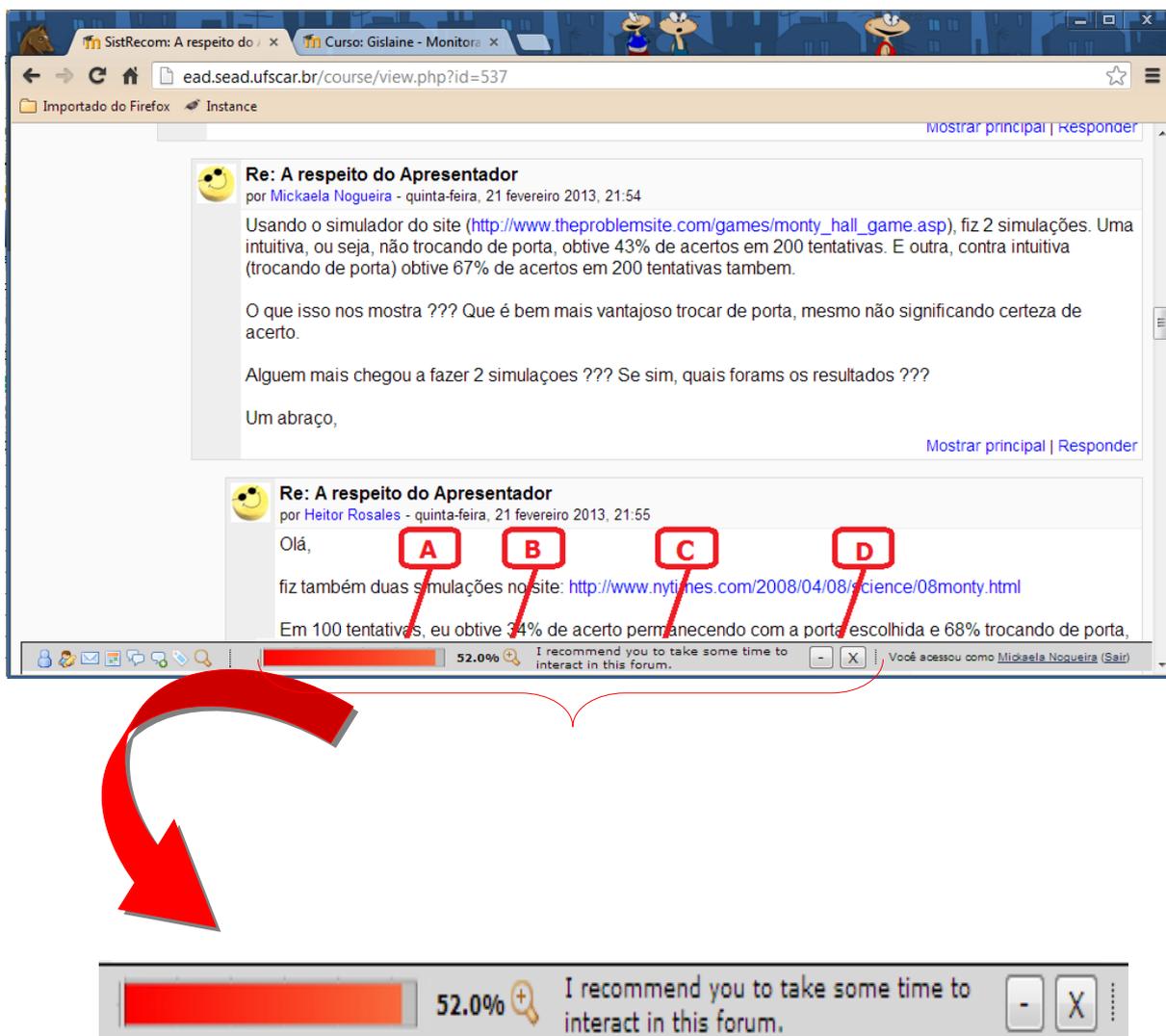


Figura 5.7 – Visualização do progresso acadêmico individual e exemplo de recomendação para o estudante. Fonte: elaborado pela autora.

A Figura 5.8 apresenta informações sobre cada variável utilizada para gerar a probabilidade prevista no gráfico de barras multicolorido. Nesta opção, o estudante pode visualizar seu comportamento em relação à média da turma, ao maior e menor valores encontrados.

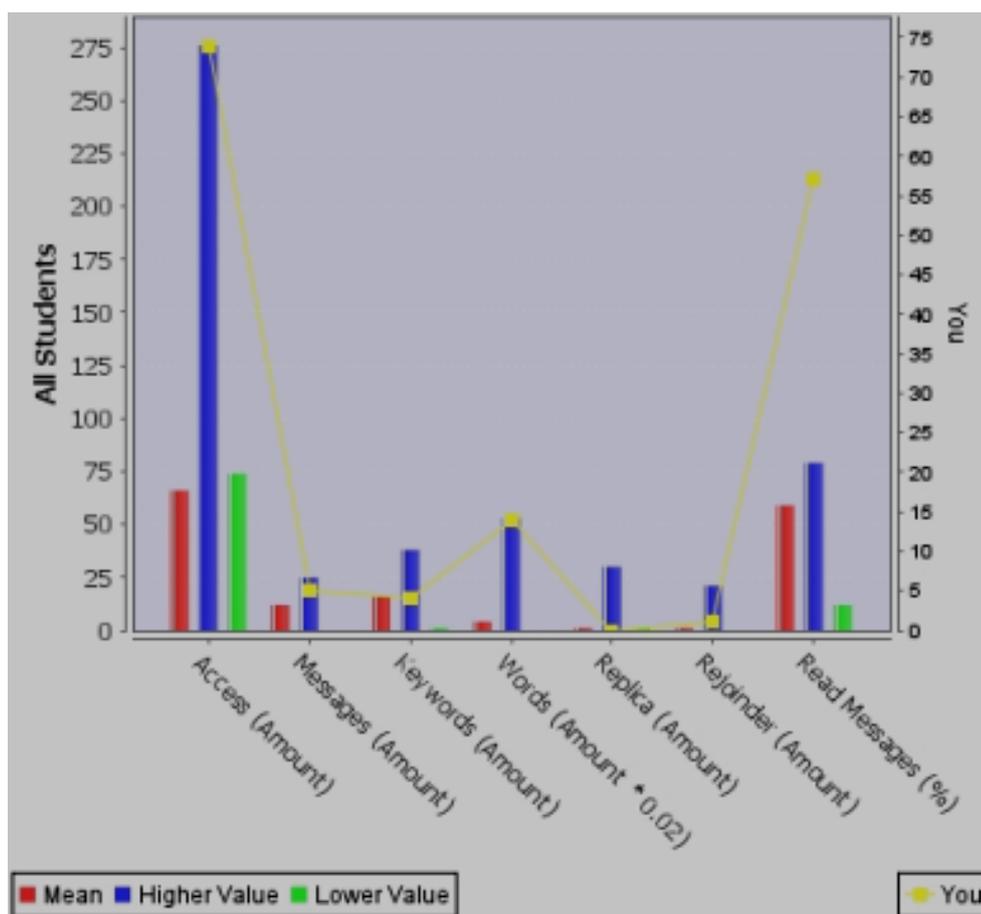


Figura 5.8 – Visualização do progresso acadêmico em cada variável utilizada na rede Bayesiana, mostrando informações individuais e em comparação à média do grupo, ao valor mais alto e ao valor mais baixo encontrados. Fonte: elaborado pela autora.

5.3.1.2 Progresso Acadêmico Detalhado

O módulo de Progresso Acadêmico Detalhado é uma opção implementada pelo ViTrackeR que permite aos estudantes, via interface gráfica do usuário, aplicar filtros para diretamente acessar dados descritivos. Indicadores selecionados podem ser restritos a um simples recurso do SGA ou estendido a múltiplos recursos. Gráficos podem ser gerados com um número arbitrário de variáveis representando informações que possibilitam a análise de uma variável individual ou um conjunto de variáveis observadas.

A Figura 5.9 mostra um exemplo usando o gráfico do tipo *spider* para representar cinco variáveis agregadas em três atividades de fórum. Tanto as variáveis, quanto as atividades, são selecionadas pelo estudante. Os raios de cada variável são calculados automaticamente, de acordo com o seu valor correspondente.

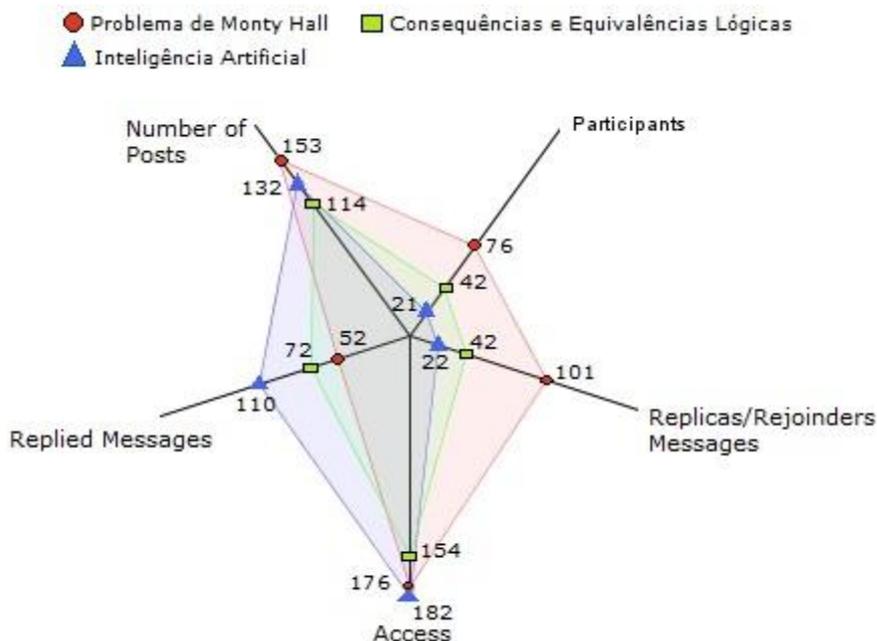


Figura 5.9 – Cinco variáveis agregadas para três atividades de fóruns de discussão. Fonte: elaborado pela autora.

Na Figura 5.10 o estudante observa a quantidade de acessos (variável observada) realizados em três fóruns distintos.

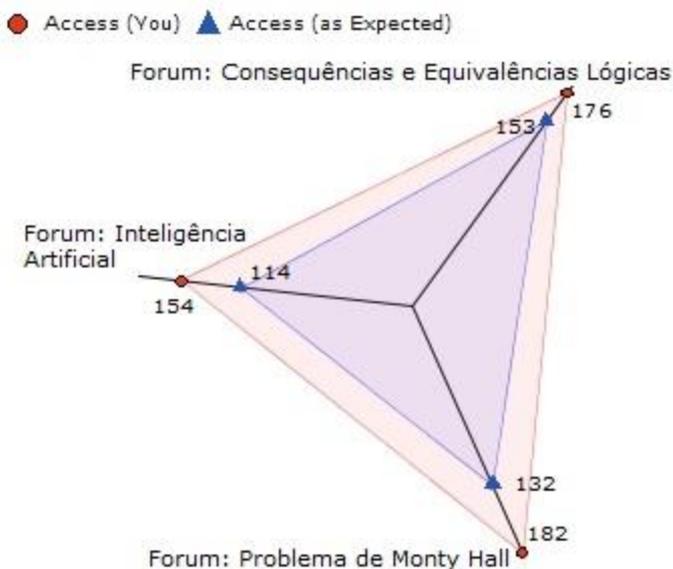


Figura 5.10 – Variável “acessos” observada em três atividades de fóruns distintas. Fonte: elaborado pela autora.

Na Figura 5.11, o aluno observa seis tópicos, ou unidades de estudo, as quais podem compor uma disciplina relacionada a duas variáveis: arquivos baixados e navegação por materiais de aprendizagem *online*.

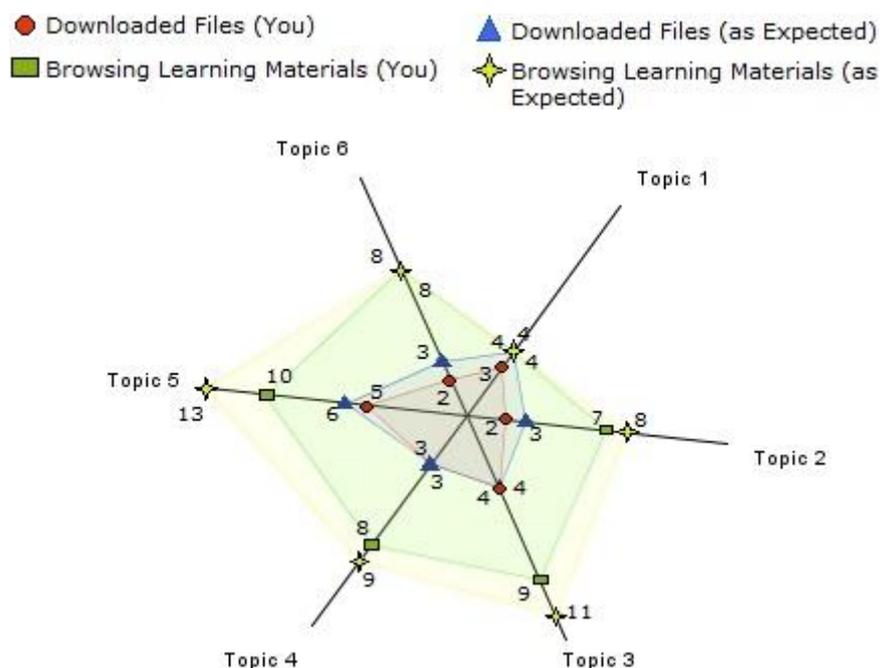


Figura 5.11 – Downloads e navegação por materiais online – o que era esperado e o que foi feito pelo aluno. Fonte: elaborado pela autora.

5.3.1.3 Agente de Recomendações

ViTrackeR implementa um agente que provê recomendações personalizadas e dinâmicas, baseadas no progresso acadêmico individual. Essas recomendações são relacionadas, por exemplo, aos materiais que já foram lidos/assistidos ou àqueles que o aluno ainda não estudou; às atividades nas quais o estudante mostrou baixo aproveitamento acadêmico e às ações que podem ser feitas para obter um desempenho melhor (análise prescritiva). O agente de recomendações também apresenta lembretes dos próximos eventos agendados e próximos prazos. Alguns exemplos de recomendações feitas pelo agente desenvolvido como parte do ViTrackeR são apresentados a seguir:

- Eu recomendo que você faça o teste “Estrutura de dados: análise da complexidade” antes das 23h55 desta data.
- Eu recomendo que você assista ao vídeo “Inferência Lógica” do tópico “5 – Inferência Lógica e Sistemas de Derivação” antes de fazer este exercício.
- Eu recomendo que você revise cuidadosamente as mensagens antes de submetê-las na wiki “Inclusão de elementos em árvore B”.
- Eu recomendo que você reserve alguns minutos para ler as mensagens postadas no fórum “Métodos de ordenação”.

Um exemplo de recomendação feita pelo agente de recomendações, é apresentado na Figura 5.12. Neste exemplo, um estudante iniciou a resolução de um questionário de correção automática disponível no tópico 2 da disciplina de pré-cálculo. ViTrackeR detecta que o estudante está errando as respostas do questionário e também que o aluno ainda não leu os materiais teóricos da disciplina. Neste caso, o agente de recomendações recomenda a leitura dos materiais antes do aluno continuar a responder o questionário (Figura 5.12).

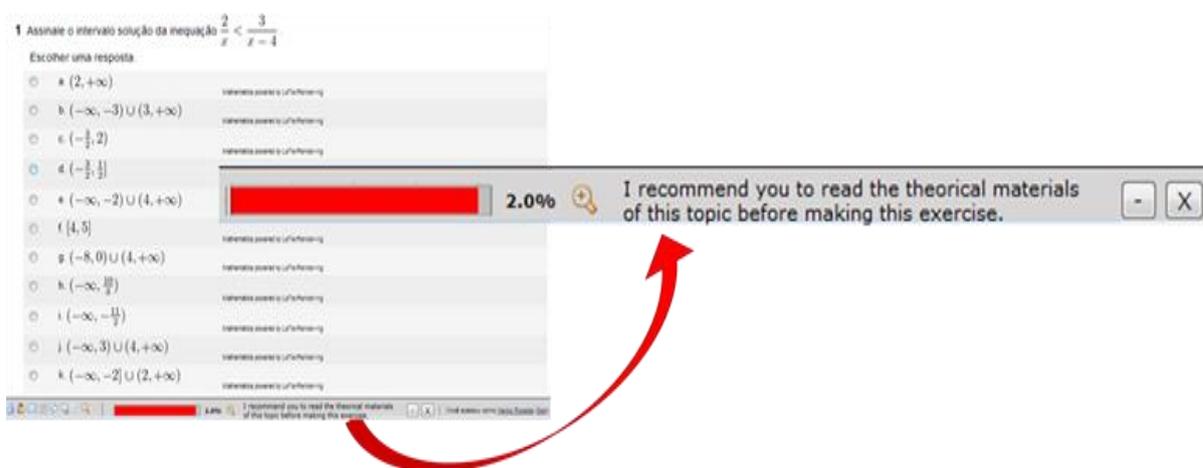


Figura 5.12 – Exemplo de recomendação feita pelo ViTrackeR: leitura de materiais teóricos antes de realizar um exercício. Fonte: elaborado pela autora.

Outro exemplo de recomendação feita pelo ViTrackeR é apresentado na Figura 5.13. Neste exemplo, um estudante que apresenta progresso acadêmico abaixo do esperado no tópico 3 de uma determinada disciplina está lendo os materiais teóricos disponibilizados no ambiente. Esses materiais consistem de textos e de um objeto interativo com o qual o usuário pode manipular, *online*, pontos de um sistema de coordenadas Cartesiano. ViTrackeR detecta que este estudante já acessou todos os materiais teóricos e que alguns materiais foram acessados por duas vezes, mas o aluno não interagiu com o recurso disponível. Neste caso, o agente de recomendações sugere ao aluno que manipule o objeto.

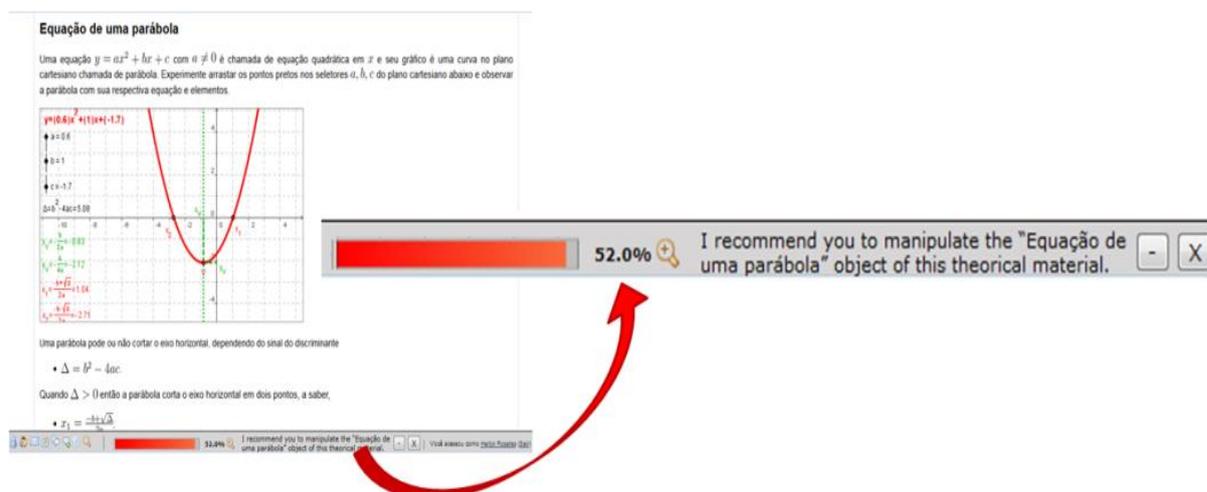


Figura 5.13 – Exemplo de recomendação no ViTracker: manipular o objeto “equação de uma parábola”. Fonte: elaborado pela autora.

A possibilidade de desativar o recebimento de informações – progresso acadêmico e recomendações, permite aos estudantes que são mais autorregulados em sua aprendizagem, e portanto com maior autonomia, a autoavaliação em relação aos demais alunos da turma; reflexões sobre seu desempenho acadêmico e aplicação de suas próprias estratégias de aprendizagem.

5.3.2 ViMonitor: Visualização e Monitoramento da Aprendizagem

ViMonitor é uma ferramenta destinada aos professores e tutores para o monitoramento contínuo da aprendizagem e visualização dos resultados de análises a partir de dados obtidos pela rede de sensores. Esta ferramenta usa OWL para descrever o contexto dos usuários e gerar informação útil e contextualizada a partir da observação de seus comportamentos. Esta ferramenta gera um conjunto de relatórios e visualizações gráficas para suporte a tomada de decisões de tutores e professores durante todo o desenvolvimento da disciplina. Professores podem usar informações providas pelo sistema para avaliar seu trabalho docente e adaptar suas práticas de ensino conforme a necessidade dos alunos.

A equipe de ensino pode, via interface gráfica do usuário, selecionar requisitos a partir de uma lista disponibilizada pelo sistema para obter informação sobre as atividades dos estudantes e dos tutores via representação gráfica em várias formas e escalas.

Requisitos (variáveis) podem ser restritos a um recurso simples do SGA, tal como uma atividade em *chat*, ou estendido a múltiplos recursos ou usuários. Por

exemplo, é possível visualizar cinco variáveis diferentes - quantidade de participantes; número total de postagens; mensagens de respostas ao tópico; mensagens do tipo réplica e do tipo tréplica; e acessos, agregadas para três fóruns, conforme ilustra a Figura 5.14. Neste gráfico, o professor pode analisar e comparar o nível de participação dos estudantes em três atividades distintas e pode ter um *insight* sobre o nível de interação social do grupo, por exemplo, a quantidade de mensagens respondidas, a quantidade de réplicas e de tréplicas podem refletir o interesse do grupo em dar continuidade às discussões de um fórum específico em vez de iniciar novos tópicos com discussões isoladas.

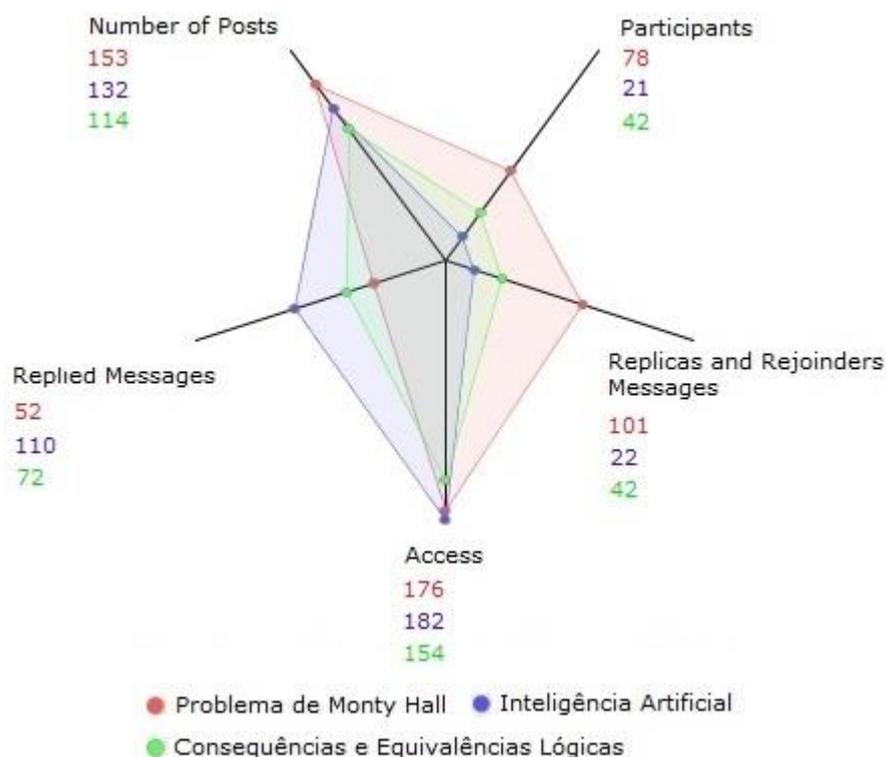


Figura 5.14 – Cinco variáveis observadas sobre três fóruns. Fonte: elaborado pela autora.

O exemplo apresentado na Figura 5.15 classifica o tempo de resposta em questionários em três categorias: abaixo do esperado, dentro do esperado e acima do esperado. Neste exemplo, o professor está interessado no tempo despendido pelo aluno lendo e refletindo sobre as questões do questionário antes de submeter a resposta. Esta informação possibilita, por exemplo, que o professor analise se há questionários nos quais os estudantes submetem suas respostas sem nem mesmo ler as questões, o que pode sugerir que as respostas foram “chutadas”. O gráfico apresentado neste exemplo mostra que no questionário “Diagrama de Classes” o

tempo despendido na leitura e reflexão das questões foi classificado como abaixo do esperado para 28% dos estudantes, dentro do esperado para 67% e acima do esperado para 5% dos alunos. Por configuração padrão, a métrica utilizada para este cálculo considera apenas o tempo de leitura às questões e não o tempo necessário para reflexões. Neste caso, o sistema considera que o usuário deverá levar, pelo menos, 10 segundos para a leitura de um conjunto de 200 caracteres. Desta forma, a cada 200 caracteres em que o aluno levar 10 segundos, o sistema classificará o tempo de leitura como dentro do esperado. Valores superiores a esse tempo, classificam o tempo de leitura como acima do esperado e valores inferiores classificam o tempo de leitura como abaixo do esperado. O professor pode alterar os indicadores e aumentar ou reduzir o tempo (em segundos) necessário para leitura e reflexão das questões, conforme a proposta e complexidade da atividade.

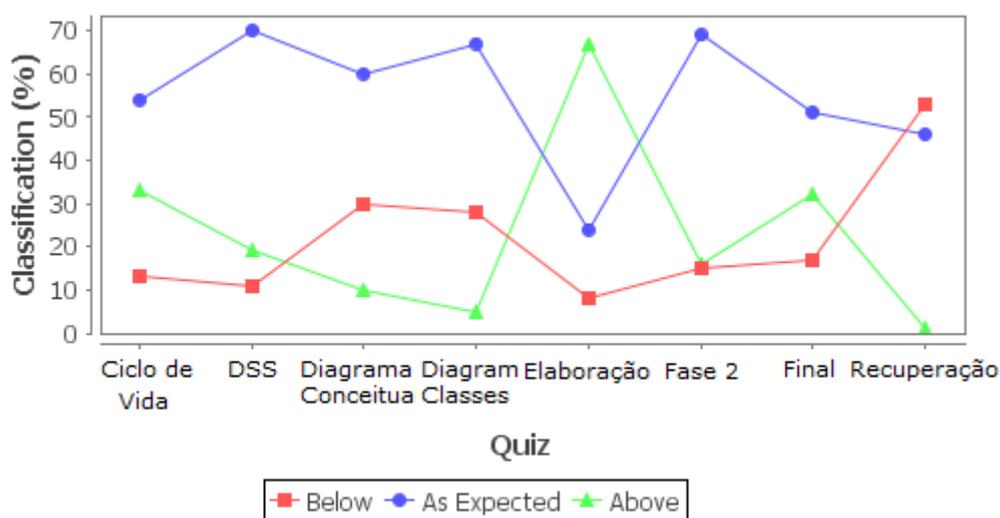


Figura 5.15 – Cinco variáveis observadas sobre três fóruns. Fonte: elaborado pela autora.

Para o ensino *online*, informações sobre as atividades dos estudantes em tempo real podem ser importantes, uma vez que o atraso na tomada de decisões podem levar ao rompimento no processo de aprendizagem, insucesso do estudante e até mesmo ao risco de abandono.

5.3.3 ViAssess: Avaliações *Online* Seguras

No Brasil, a legislação que regulamenta os cursos de educação a distância, estabelece a obrigatoriedade e prevalência das avaliações presenciais sobre outras formas de avaliação (BRASIL, 2005). Em consonância com a legislação, o

documento de Referenciais de Qualidade para Educação Superior a Distância apresentado pela SEED/MEC (BRASIL, 2007 p. 17) destaca a necessidade de garantia de segurança nas avaliações presenciais:

As avaliações da aprendizagem do estudante devem ser compostas de avaliações a distância e avaliações presenciais, sendo estas **últimas cercadas das precauções de segurança e controle de frequência, zelando pela confiabilidade e credibilidade dos resultados** [grifo meu].

Dados do relatório analítico da aprendizagem a distância no Brasil do Censo EAD.br de 2010 apontam que na educação a distância, a prova escrita presencial é aplicada por 82% das instituições, tanto durante o curso quanto no final, dentre outras formas de avaliação presencial e a distância, como provas práticas, apresentação de trabalhos, etc. (ABED, 2010).

As provas escritas presenciais ainda são predominantemente baseadas na mídia impressa, envolvendo um complexo processo de preparação, impressão e distribuição. Esse processo em geral tem alto custo, é lento (exige a elaboração da prova com semanas de antecedência) e requer a manipulação das provas por outras pessoas além do professor, podendo comprometer o sigilo.

A aplicação de avaliações presenciais *online* em um SGA parece sanar muitos desses problemas, reduzindo o custo e o tempo de distribuição, possibilitando que o próprio professor disponibilize a prova no SGA no momento da avaliação. Além disso, torna possível a elaboração de avaliações utilizando e integrando diferentes recursos (vídeos, áudio, gráficos em 3D, etc.).

No entanto, as avaliações *online*, mesmo quando realizadas presencialmente, ainda são muito vulneráveis, pois estando *online*, é difícil assegurar que o aluno não consulte materiais não permitidos pelo professor, não troque informações, mensagens ou até mesmo tenha sua prova realizada por terceiros. Outra dificuldade reside em prover mecanismos que permitam ao professor e tutores um acompanhamento das atividades do estudante durante e após a realização dos exames.

Desta forma, a ferramenta ViAssess foi desenvolvida com o propósito de prover mecanismos para avaliações *online* seguras a partir da supervisão das

interações do usuário em seu ambiente digital. Dados coletados por sensores em tempo real sobre ações e interações do usuário, e de suas aplicações, são submetidos a um processo de inferência a partir de um conjunto de funções com parâmetros alterados dinamicamente que permitem à aplicação identificar ações ilícitas do aluno durante a realização de uma avaliação *online*.

A partir do levantamento de requisitos junto aos profissionais da educação a distância, conforme apresentado na Seção 4.2.1, foi possível identificar um conjunto de sensores lógicos que foram implementados (Quadro 5.2) pelo *framework* e utilizados pelo ViAssess. Alguns sensores físicos também foram identificados como necessidades apontadas durante o levantamento de requisitos. A sigla SF do Quadro 5.2 indica sensores físicos, SL sensores lógicos.

Quadro 5.2 – Sensores lógicos e físicos para aplicação de monitoramento de avaliações *online*. Fonte: elaborado pela autora.

	Nome do Sensor	Tipo de Monitoramento Realizado
SL	usb_sensor	Monitora acesso à porta USB com dispositivo de armazenamento secundário. Exemplos: <i>pen drive</i> , disco secundário.
SL	parallelApplications_sensor	Monitora arquivos acessados pelo usuário. Exemplos: texto.doc, video.mpeg.
SL	accessedMaterials_sensor	Monitora materiais do SGA acessados pelo usuário. Exemplos: livro eletrônico, sites recomendados.
SL	instantCommunication_sensor	Monitora ferramentas de comunicação instantânea utilizadas. Exemplo: skype, MSN, gtalk
SF	userIdentification_sensor	Sensor biométrico. Exemplo: digital do usuário, face do usuário.
SL	login_sensor	Identifica o usuário que realiza <i>login</i> no SGA.
SL	accessedActivities_sensor	Monitora atividades do SGA acessadas pelo usuário. Exemplos: fórum, <i>chat</i> , <i>wiki</i> .
SF	mobileSign_sensor	Identifica sinais de dispositivos móveis no espaço coberto pelos sensores. Exemplos: celular e tablet.
SL	accessedSites_sensor	Monitora sites acessados pelo usuário

A partir de uma interface gráfica (Figura 5.16), professores podem configurar o sistema para receber informações de monitoramento durante avaliações *online*. Um professor pode, por exemplo, selecionar as seguintes opções a serem monitoradas: acesso aos materiais e atividades da disciplina no SGA; navegação em

páginas web fora do SGA; uso de ferramentas de comunicação instantânea (Skype, MSN e Hangout) e uso de dispositivos de armazenamento secundário (pendrives e cartões).

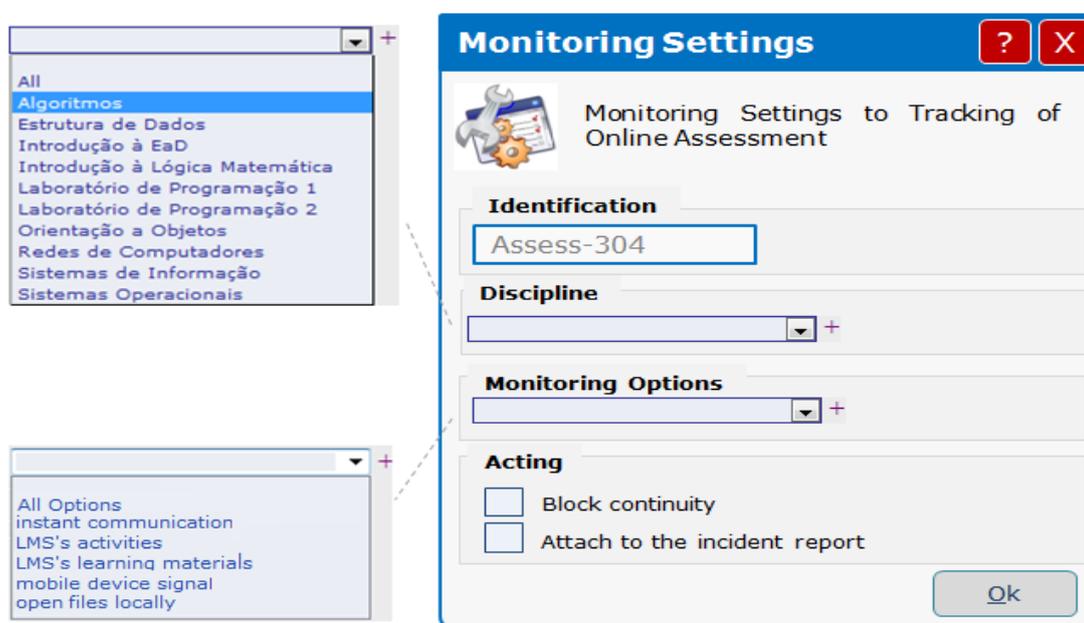


Figura 5.16 – Configuração de opções de monitoramento no ViAssess. Fonte: elaborado pela autora.

Ao identificar um evento de risco, ou seja, uma ação que o aluno não deveria executar, a aplicação recebe uma notificação e emite ações a determinados atuadores da rede de sensores, com o intuito de garantir a credibilidade da avaliação *online*. A Figura 5.17 ilustra a interface exibida no dispositivo do usuário/estudante pelo atuador após a aplicação identificar uma tentativa de acesso não autorizado aos materiais da disciplina durante a realização de uma avaliação. Neste caso, o sistema só será desbloqueado com a inserção de uma senha gerada aleatoriamente logo após o bloqueio. Essa senha é enviada via SMS ao celular cadastrado (celular do supervisor da avaliação, por exemplo). Logo, não há como o aluno proceder a realização da prova sem que pelo menos o supervisor da avaliação tenha ciência da tentativa ilícita de acesso.

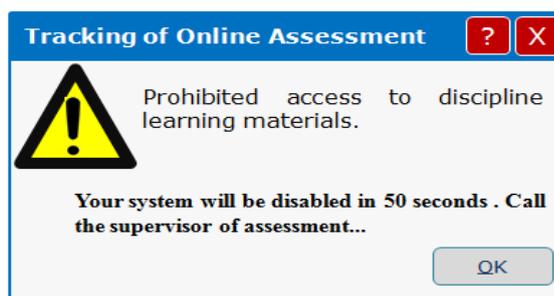


Figura 5.17 – Mensagem exibida pelo ViAssess ao identificar uma ação proibida durante uma avaliação *online*. Fonte: elaborado pela autora.

5.4 Considerações Finais

Este capítulo apresentou um *framework* desenvolvido a partir da abstração de redes de sensores para coleta de dados educacionais e análises. Sensores podem ser depositados em ambientes físicos e lógicos para captura de eventos de interesse para o monitoramento refinado de diferentes aspectos e visões do processo de ensino, aprendizagem e gestão educacional. A separação entre a coleta de dados e o processamento de contextos (modelagem e inferência), adicionado às decisões de projeto mencionadas, objetivam a aquisição de um conjunto de dados extensivo e independente do domínio da aplicação a fim de alcançar análises confiáveis, que melhor respondam às necessidades reais dos usuários em vários contextos.

A fim de oferecer bom equilíbrio entre poder de expressividade e compartilhamento semântico e representação formal com alto nível de abstração, e também garantir modelos mais simples para aplicações que não necessitam de tais características, neste trabalho optou-se por implementar e disponibilizar no *framework* para uso de aplicações duas modelagens de contexto, modelo ontológico e modelo Bayesiano. Para as análises estatísticas também foi utilizado um conjunto de regras de decisão, que facilita a visualização de informações com maior representatividade visual e a geração de gráficos dinâmicos em tempo real.

O *framework* desenvolvido nesta tese foi avaliado a partir de seu uso para a construção de três aplicações com objetivos e propósitos diferentes: ViTrackeR, ViMonitor e ViAssess. Essas aplicações foram avaliadas individualmente a partir da sua utilização por usuários em cursos reais. Os resultados obtidos nessas avaliações são apresentados no capítulo seguinte.

Capítulo 6

AVALIAÇÕES E APRESENTAÇÃO DOS RESULTADOS

“We need better evaluation to acknowledge and esteem excellence when it occurs, as it does.” (JOHN HATTIE, 2008, p. 146).

Neste capítulo, são apresentados e discutidos os resultados obtidos nas avaliações realizadas nesta tese. As avaliações objetivaram a validação do Modelo Conceitual proposto. A Seção 6.1 apresenta as considerações iniciais do capítulo. A Seção 6.2 apresenta os resultados obtidos na avaliação conduzida junto aos usuários das aplicações desenvolvidas. A Seção 6.3 apresenta a avaliação do framework na perspectiva de desenvolvedores de ferramentas educacionais. A Seção 6.4 apresenta resultados e discussões sobre a avaliação do Modelo Conceitual proposto na perspectiva de pesquisadores experientes na área da presente tese. Por fim, a Seção 6.5 apresenta as limitações desta pesquisa, seguida das considerações finais.

6.1 Considerações iniciais

O Modelo Conceitual proposto nesta tese foi avaliado a partir de três tipos de fontes de dados diferentes, conforme Figura 6.1: (1) usuários das aplicações analíticas desenvolvidas; (2) desenvolvedores de ferramentas educacionais; e, (3) pesquisadores e especialistas de domínio.

De acordo com Eddy et al. (2012), Refsgaard e Henriksen, (2002) e Sargent (2011), a validação de um modelo julga a sua acurácia, ou seja, a comprovação de que o modelo é uma representação razoável do sistema real e que reproduz o

comportamento do domínio de aplicação com fidelidade suficiente para satisfazer aos objetivos de análise. A validade aparente (*face validity* ou *expert review*) indica o grau em que o modelo, seus pressupostos e aplicações correspondem às evidências atuais, a partir do julgamento realizado por especialistas do domínio (SECOLSKY, 1987; NEVO, 1985). O conceito de validade aparente foi aplicado a partir da entrevista com pesquisadores especialistas de domínio.

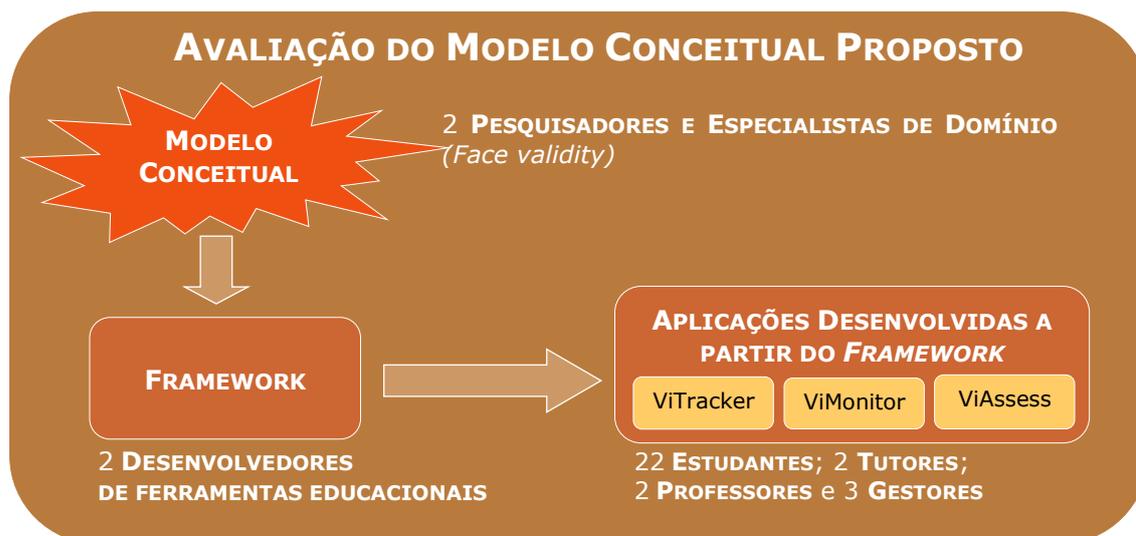


Figura 6.1 – Modelo de avaliação proposto para validação do Modelo Conceitual apresentado nesta tese. Fonte: elaborado pela autora.

Os resultados das avaliações realizadas são apresentados nas seções 6.2, 6.3 e 6.4 a seguir.

6.2 Avaliações realizadas com usuários das aplicações desenvolvidas

A avaliação realizada com usuários das aplicações se divide em: (1) avaliação pelos estudantes, (2) avaliação pelos instrutores (tutores e professores) e (3) avaliação pelos gestores educacionais e são apresentadas a seguir.

6.2.1 Avaliação realizada com Estudantes

A avaliação conduzida a partir do ponto de vista dos estudantes teve como objetivos diretos: (1) avaliar se as análises oferecidas pelo ViTracker ofereceram

suporte para apoiar a aprendizagem autorregulada - por meio do automonitoramento, da autorreação e da auto-motivação; e (2) avaliar a utilidade da informação apresentada pela ferramenta, a facilidade na execução e a facilidade de uso.

Esta avaliação buscou verificar a veracidade da seguinte proposição:

P1 - As ferramentas analíticas, desenvolvidas a partir do modelo proposto, auxiliam na promoção da **aprendizagem autorregulada** de estudantes e melhoram a tomada de **decisão em nível micro**.

6.2.1.1 Planejamento das Entrevistas e Instrumento de Coleta

Conforme procedimentos metodológicos adotados neste trabalho, descritos na Seção 1.5, o método de amostragem utilizado foi o não probabilístico, uma vez que os participantes foram selecionados considerando a facilidade de acesso aos mesmos e o interesse em participar do estudo.

Os alunos participantes da pesquisa são estudantes do curso de Análise e Desenvolvimento de Sistemas do Instituto Federal de Educação, Ciência e Tecnologia do Campus Araraquara (IFSP). O ViTrackeR foi integrado ao SGA, Moodle, utilizado pelos alunos, o que possibilitou o uso da aplicação em um ambiente real. Três disciplinas foram selecionadas para aplicação da pesquisa.

Inicialmente, foi feita uma apresentação aos alunos sobre a proposta do ViTrackeR. A apresentação também abordou a necessidade de instalar os sensores no computador pessoal do aluno e a descrição de todos os dados que poderiam ser coletados. Em seguida, os alunos participantes assinaram um termo de consentimento de instalação do sistema e coleta de dados a partir de seu computador pessoal, apêndice A.

Conforme mencionado na Seção 1.5, neste trabalho predominou-se o uso de entrevistas para realizar a coleta de dados com usuários das aplicações. As entrevistas com estudantes foram realizadas a partir do uso de questionário semiestruturado em duas etapas. A primeira etapa, aplicada antes do uso do ViTrackeR, teve a finalidade de identificar a experiência prévia dos alunos com ferramentas de organização de estudos, pertencentes ou não ao SGA.

A segunda etapa foi aplicada após o uso do ViTracker, ao final das disciplinas, a fim de coletar informações sobre a percepção dos alunos em relação à ferramenta e avaliar o atendimento aos objetivos relacionados à sua avaliação. A avaliação foi

feita quantitativamente usando escala de Likert de cinco pontos onde as respostas foram codificadas das seguintes formas: -2, -1, 0, 1, 2 ou então 0, 1, 2, 3, 4. No primeiro caso, -2 representa discordância em maior grau e 2 representa concordância em maior grau. O valor 0 indica ponto neutro. No segundo caso, onde os pontos variam de 0 a 4, o aluno indica o grau em que cada variável analisada está presente e o valor 0 indica ausência.

A fim de dar maior grau de liberdade para as respostas dos estudantes, todos os tópicos avaliados continham campos abertos para observações ou sugestões. O questionário aplicado aos alunos pode ser visualizado no Apêndice B. As entrevistas duraram, em média, 1 h e 15 minutos. Os alunos tiveram um computador disponível para responder às questões durante a entrevista, enquanto a pesquisadora anotava as observações que os alunos faziam em cada resposta.

Para avaliação das respostas, foram empregadas duas formas de análise das evidências, a análise de conteúdo e a análise descritiva. A análise de conteúdo foi utilizada para interpretar conteúdos textuais (BARDIN, 2002). A estatística descritiva possibilitou descrever e sumarizar o conjunto de dados. O comportamento das variáveis foi estudado a partir da análise de suas ocorrências e o coeficiente de correlação de *Spearman* permitiu medir a associação que caracterizou a intensidade e a direção da variação comum entre algumas das variáveis observadas.

A Seção 6.2.1.2 apresenta as características da amostra estudada.

6.2.1.2 Caracterização da amostra

Participaram desta pesquisa, vinte e dois estudantes, sendo 63,5% homens e 36,5% mulheres. A idade dos participantes variou entre 18 anos e 50 anos (Figura G.1 ilustrativa disponível no Apêndice G).

Dos respondentes, 68,2% estavam no primeiro ano do curso, 27,3% dos alunos estava no segundo ano e apenas um aluno (4,5%) estava no terceiro ano do curso (Figura F.2). Aproximadamente 50% dos alunos possuía menos que seis meses de experiência na utilização do Moodle, sete alunos possuíam de seis meses a um ano de experiência, quatro alunos possuíam de um a três anos e apenas um aluno possuía mais que seis anos de experiência (Figura F.3).

6.2.1.3 Resultados obtidos: percepções dos estudantes após uso do ViTracker

Na aplicação da segunda etapa do instrumento de coleta, os estudantes foram questionados acerca de suas percepções sobre os seguintes tópicos: ajuda oferecida pelo sistema; as recomendações realizadas pelo ViTracker; utilidade da informação; e, facilidade de uso da ferramenta e se seu uso trouxe motivação para estudar. A as seções seguintes apresentam os resultados obtidos.

6.2.1.3.1 Sobre a ajuda oferecida pelo ViTracker

As percepções dos alunos sobre a ajuda oferecida pelo ViTracker são ilustradas na Figura 6.2. Das respostas obtidas, 77% dos alunos considerou que o sistema ajudou muito a gerenciar o tempo de estudo (item 2.1 do instrumento de coleta). No quesito autogerenciamento dos estudos (item 2.3), 72% dos alunos respondeu que o sistema ajudou muito a decidir o que executar primeiro. A mesma porcentagem dos alunos indicou que o sistema ajuda muito no auto monitoramento da aprendizagem (item 2.5). Por fim, 72% dos alunos indicou que o sistema ajudou muito a atingir as metas de aprendizagem (item 2.7) estabelecidas.

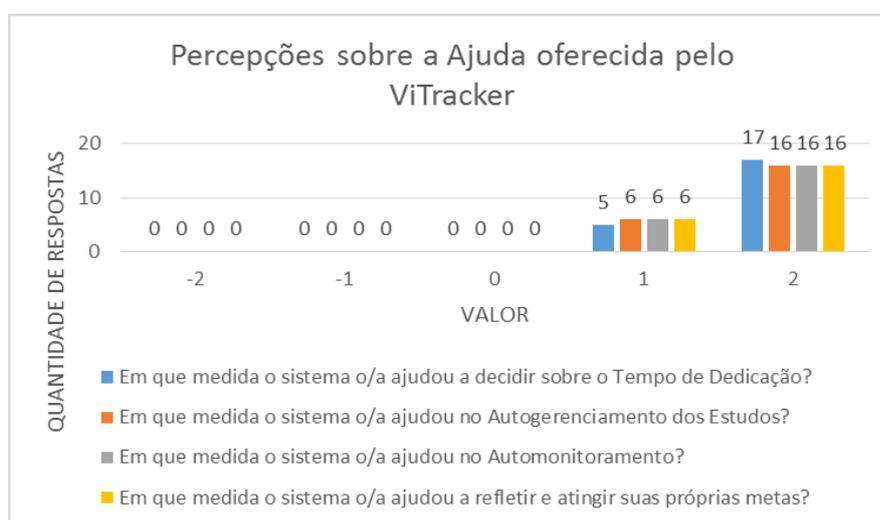


Figura 6.2 – Quantidade de respostas para variáveis que medem a ajuda oferecida pelo ViTracker. Fonte: elaborado pela autora.

Estatísticas obtidas a partir dos dados coletados são apresentadas na Tabela G.2 do Apêndice G. Não houve respostas indicando que o sistema atrapalhou as atividades de aprendizagem. A partir dessas evidências, pode-se concluir que 100% dos alunos avaliou a ajuda oferecida pelo ViTracker como positiva.

6.2.1.3.2 Sobre as recomendações feitas pelo ViTrackerR

Os valores obtidos na avaliação das recomendações apresentadas pelo ViTrackerR são apresentados nas Figuras 6.3 e 6.4. Observa-se que 59% dos alunos mudou muito sua atitude no ambiente (item 3.1) após receber recomendações do sistema; em torno de 50% dos alunos seguiu as recomendações completamente (item 3.3), e 82% dos alunos indicou que as recomendações foram extremamente úteis (item 3.5).

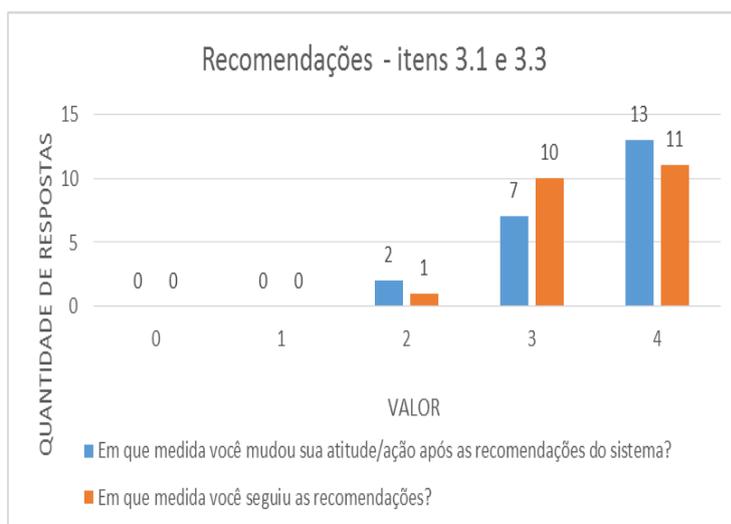


Figura 6.3 - Quantidade de respostas para itens 3.1 e 3.3 do questionário dos alunos. Fonte: elaborado pela autora.

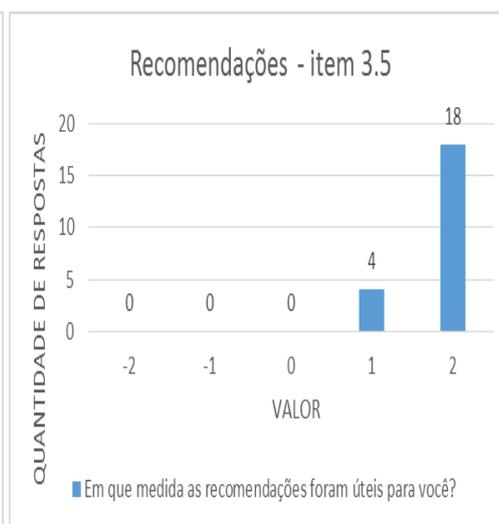


Figura 6.4 - Quantidade de respostas para item 3.5. Fonte: elaborado pela autora.

Durante a entrevista, alguns alunos foram questionados sobre o motivo de não terem seguido completamente as recomendações, já que as consideravam completamente úteis. O principal motivo apontado foi a limitação de tempo para executar o que o sistema recomendava. As frequências obtidas a partir dos dados coletados neste tópico são apresentadas nas Tabelas G.3 e G.4 do Apêndice G.

6.2.1.3.3 Sobre utilidade das informações disponibilizadas pelo ViTrackerR

A Figura 6.5 ilustra as respostas obtidas no tópico 4 do instrumento de coleta, que possibilita observar que mais de 85% dos alunos avaliou que as informações oferecidas pelo sistema são extremamente úteis. As frequências das respostas são indicadas na Tabela G.5 (Apêndice G). Nenhum aluno considerou que as informações foram inúteis ou extremamente inúteis.

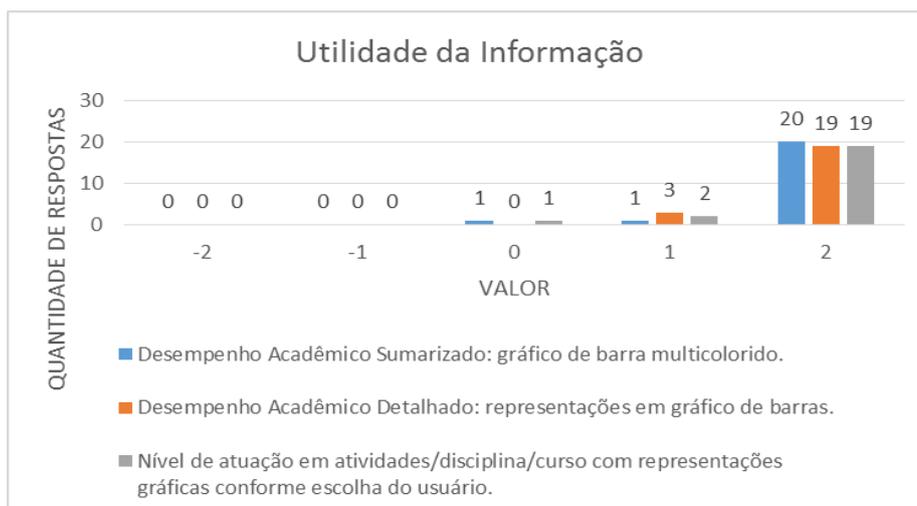


Figura 6.5 – Respostas dos alunos sobre a utilidade da informação disponibilizada pelo ViTrackerR. Fonte: elaborado pela autora.

6.2.1.3.4 Sobre a facilidade de uso do ViTrackerR

A avaliação dos alunos sobre a facilidade de uso do ViTrackerR (Figuras 6.6 e 6.7) indica que 63,5% dos alunos considerou que o sistema é extremamente fácil de usar (item 5.1), 77% dos alunos respondeu que o sistema facilitou extremamente suas atividades no ambiente virtual (item 5.3) e 100% dos alunos foram motivados ou muito motivados pelo sistema (item 5.5). As Tabelas G.6 e G.7 (Apêndice G) indicam as frequências obtidas nos itens avaliados.

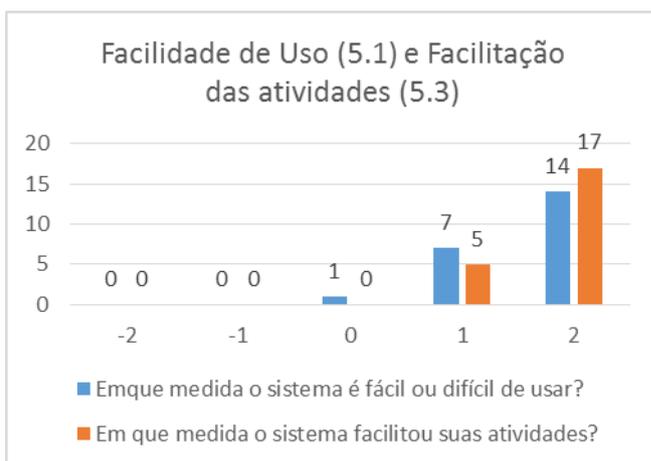


Figura 6.6 – Respostas dos alunos sobre a facilidade de uso do ViTrackerR e atividades. Fonte: elaborado pela autora.

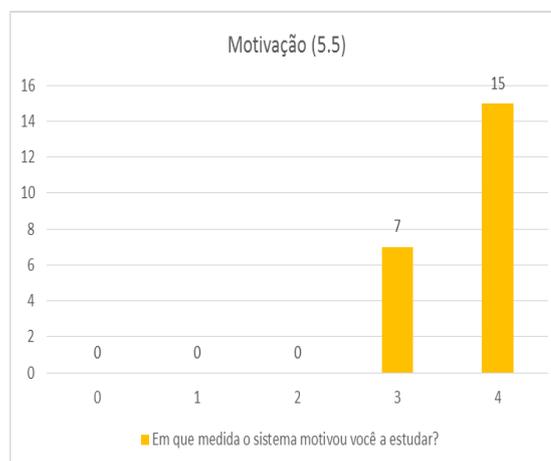


Figura 6.7 – Respostas dos alunos sobre a motivação no uso do ViTrackerR. Fonte: elaborado pela autora.

Nenhum aluno indicou dificuldade ao utilizar o sistema ou indicou que suas atividades foram dificultadas com o uso do sistema.

6.2.1.4 Correlação de Dados

Dados coletados pelos sensores foram analisados em conjunto com dados obtidos pelo ViTrackeR e dados de relatórios do SGA. As análises realizadas, por meio da triangulação desses dados, possibilitaram a identificação de possíveis causas para algumas situações observadas.

As principais observações feitas estão relacionadas ao desempenho final do aluno, medido pelo aproveitamento escolar e pela frequência na disciplina. O aproveitamento final é associado a uma nota final que varia entre zero e dez. As correlações foram analisadas a partir do coeficiente de correlação de postos de *Spearman*. A escolha deste coeficiente se deve ao fato dos valores não seguirem uma distribuição normal, necessitando, portanto, de testes não-paramétricos. A escolha desta técnica também foi motivada pelo fato da amostra ser pequena.

O coeficiente de correlação de *Spearman* (ρ) é uma medida de associação não paramétrica que mede a intensidade da relação entre duas variáveis pelo menos ordinais (Marôco, 2011). A expressão simplificada, após manipulação algébrica, é a seguinte:

$$\rho = 1 - \frac{6 \sum d_i^2}{(n^3 - n)}$$

Onde d_i = representa a diferença entre cada posto de valor correspondentes de x e y , ou seja: (postos de x_i dentre os valores de x) - (postos de y_i dentre os valores de y); e n = é o número dos pares dos (x_i, y_i) .

O coeficiente varia da seguinte forma: $-1 \leq \rho \leq 1$. Quanto mais próximo o ρ estiver dos extremos, maior será a evidência de associação entre as variáveis. Se o ρ for igual a -1 , indica correlação negativa perfeita. Quanto mais próximo de -1 estiver, significa que as variáveis variam em sentido contrário, isto é, conforme uma variável aumenta, a outra diminui. O valor 1 indica correlação positiva perfeita. Quanto mais próximo de 1 estiver o coeficiente de correlação, indica que conforme uma variável aumenta, a outra também aumenta. O valor zero, indica que não há dependência entre as variáveis. Foram realizadas duas análises, descritas a seguir:

Análise 1: a primeira análise buscou identificar se o aproveitamento final do aluno teve associação com o tempo de estudo. O tempo de estudo para esta análise

foi calculado em minutos e equivale ao tempo em que o aluno esteve conectado à disciplina e realizando alguma das seguintes atividades:

- 1) navegando pela página principal da disciplina,
- 2) navegando em uma das atividades (teóricas ou práticas) da disciplina,
- 3) acessando localmente um material baixado (*downloaded*) a partir da disciplina,
- 4) com uma das páginas da disciplina aberta, porém executando outras atividades não relacionadas à disciplina por um período menor que 20 minutos (quando o *login* do usuário no SGA expira).

A Figura 6.8 mostra o diagrama de dispersão dos dados obtidos.

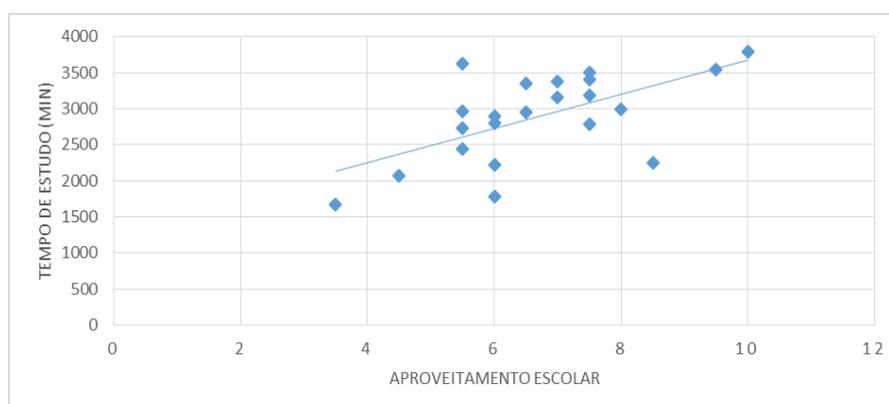


Figura 6.8 – Diagrama de dispersão dos dados: aproveitamento escolar e tempo de estudo. Fonte: elaborado pela autora.

A partir da análise do gráfico apresentado na Figura 6.8 e das frequências obtidas nos itens analisados (Tabela 6.2), observa-se que o valor mínimo para o aproveitamento foi de 3.5, o valor máximo foi de 10.0. Sobre o tempo de dedicação aos estudos, o valor mínimo, dado em minutos, foi de 1670,40 minutos, equivalente a 27,84 horas e o tempo máximo foi de 3787,20 minutos, equivalente a 63,12 horas.

Tabela 6.1 – Frequências: variáveis aproveitamento e tempo de dedicação aos estudos.

		Aproveitamento	Tempo de Dedicação ao Estudo
N	Válido	22	22
	Ausente	0	0
Mediana		6,500	2962,800
Desvio Padrão		1,532	599,939
Mínimo		3,500	1670,400
Máximo		10,000	3787,200

O coeficiente de *Spearman* entre tempo de dedicação aos estudos e aproveitamento, ilustrado na Tabela 6.3, é de 0.555, portanto, uma correlação positiva de moderada magnitude. Desta forma, conforme aumenta o tempo de dedicação aos estudos, o aproveitamento também aumenta (moderadamente).

Tabela 6.2 – Resultado do coeficiente de *Spearman* entre tempo de dedicação aos estudos e aproveitamento.

			Tempo de Dedicação ao Estudo	Aproveitamento
rô de <i>Spearman</i>	Tempo de Dedicação ao Estudo	Coeficiente de Correlação	1,000	0,555**
		Sig. (2 extremidades)	.	0,007
		N	22	22
	Aproveitamento	Coeficiente de Correlação	0,555**	1,000
		Sig. (2 extremidades)	0,007	.
		N	22	22

** . A correlação é significativa no nível 0,01 (2 extremidades).

Análise 2: esta análise buscou identificar o grau de correlação entre o aproveitamento do aluno e a utilização de recursos paralelos durante os estudos. Para isso, foram realizadas as seguintes análises:

- 1) Correlação entre o aproveitamento do aluno e o tempo gasto usando o site de rede social *Facebook* (usou-se a média obtida durante o período da disciplina);
- 2) Correlação entre o aproveitamento do aluno e o tempo gasto em navegação por sites classificados como notícias e/ou entretenimento;

A Figura 6.9 ilustra a dispersão dos dados brutos associados ao aproveitamento escolar e o tempo gasto pelo aluno utilizando o *Facebook* e o tempo gasto navegando por sites e portais de notícias e/ou entretenimento. Observa-se que o aluno com menor aproveitamento, 3.5, apresentou elevados índices de utilização de ambos os recursos paralelos analisados.

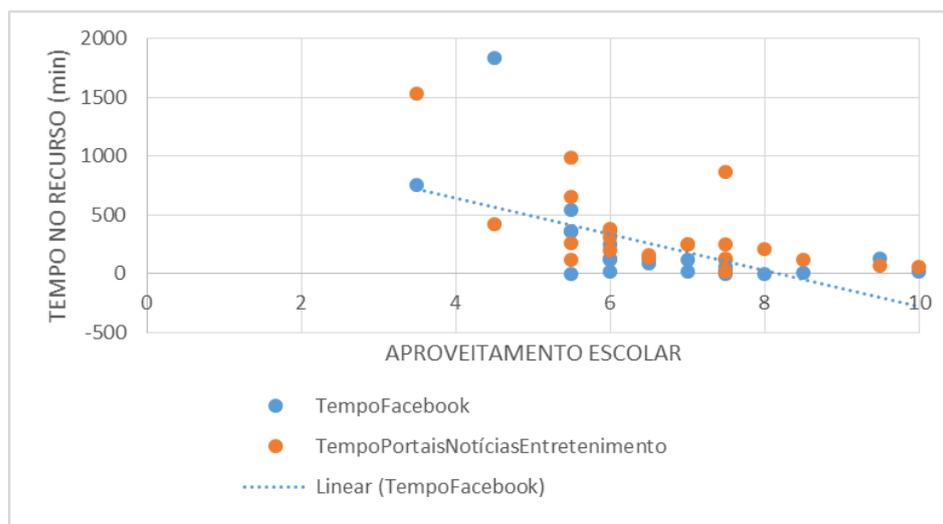


Figura 6.9 – Diagrama de dispersão dos dados: aproveitamento escolar e tempo em recursos (Facebook e/ou portais de notícias e entretenimento). Fonte: elaborado pela autora.

A Tabela 6.4 apresenta as frequências obtidas para os itens analisados.

Tabela 6.3 – Frequências para os itens aproveitamento, tempo gasto utilizando Facebook e tempo gasto navegando por portais e sites de notícias e entretenimento.

		Aproveitamento	Tempo no Facebook	Tempo em Portais de Notícias e/ou Entretenimento
N	Válido	22	22	22
	Ausente	0	0	0
Mediana		6,500	122,265	254,505
Desvio Padrão		1,532	405,998	322,186
Mínimo		3,500	0,000	58,900
Máximo		10,000	1835,400	1532,100

O valor mínimo obtido para a variável *Tempo no Facebook* indica que há aluno(s) que não usou/usaram o *Facebook* durante suas atividades de estudo no SGA. O valor máximo para esta mesma variável é de 1835,4 minutos, o que equivale a 30,59 horas de utilização do *Facebook* paralelamente aos estudos no SGA. As frequências do tempo gasto pelo aluno navegando em portais e sites de notícias e/ou entretenimento indica que todos os alunos utilizaram esse recurso paralelamente aos estudos em certo grau. O índice de correlação de *Spearman* obtido na análise entre o aproveitamento e o tempo gasto acessando o *Facebook* foi de -0.677, uma correlação negativa de magnitude moderada. Assim, índices mais elevados na utilização do *Facebook*, indicam piores aproveitamentos. A mesma situação é confirmada no tempo gasto na navegação de sites de notícias e/ou

entretenimento, $\rho = -0.627$. A Tabela 6.5 apresenta esses resultados calculados a partir do SPSS.

Tabela 6.4 – Resultado do coeficiente de Spearman entre aproveitamento escolar e tempo navegando por portais de notícias e entretenimento; e o tempo no Facebook.

			Aproveitamento	Tempo em Portais de Notícias e Entretenimento	Tempo no Facebook
Coeeficiente de Correlação	Aproveitamento	Coeeficiente de Correlação	1,000	-0,627**	-0,677**
		Sig. (2 extremidades)	.	,002	0,001
		N	22	22	22
	Tempo em Portais de Notícias e/ou Entretenimento	Coeeficiente de Correlação	-0,627**	1,000	0,515*
		Sig. (2 extremidades)	0,002	.	0,014
		N	22	22	22
	Tempo no Facebook	Coeeficiente de Correlação	-0,677**	0,515*	1,000
		Sig. (2 extremidades)	0,001	0,014	.
		N	22	22	22

** . A correlação é significativa no nível 0,01 (2 extremidades).

* . A correlação é significativa no nível 0,05 (2 extremidades).

Isolando os extremos, os dois alunos com aproveitamentos mais baixos e os dois alunos com melhores aproveitamentos na turma, observa-se que, para esses quatro alunos, o tempo gasto com sites de entretenimento e o tempo gasto no Facebook enquanto o aluno estudava, muito provavelmente, teve influência no aproveitamento final obtido e, portanto, no seu sucesso/fracasso na disciplina. A Figura 6.10 ilustra que o aluno com aproveitamento 3.5 teve um tempo total de estudo usando o SGA de 27.84 horas. Durante este tempo, este aluno passou 12.54 horas navegando no Facebook e 25.54 horas navegando em sites de notícias e/ou entretenimento simultaneamente. O aluno com aproveitamento 4.5 apresentou 35.44 horas de estudo, das quais 30,59 horas ele também esteve no Facebook e 7.12 horas navegando por sites e portais de entretenimento.

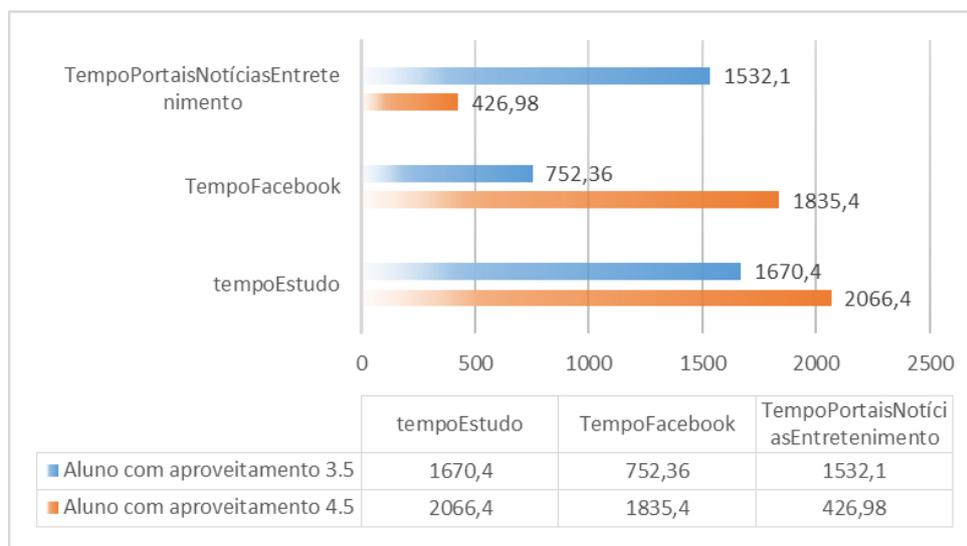


Figura 6.10 - Tempos (min) de estudo, utilização do Facebook e navegação em sites de notícias e/ou entretenimento para alunos com piores aproveitamentos. Fonte: elaborado pela autora.

A Figura 6.11 ilustra a mesma informação apresentada na Figura 6.10, porém apenas para os dois alunos com melhores aproveitamentos. Observa-se que o tempo total de estudo no SGA do aluno com aproveitamento 10 foi de 63,12 horas, das quais apenas 19 minutos ele também esteve no Facebook e, aproximadamente, 1 hora ele esteve navegando por portais de notícias e entretenimento. Situação semelhante se repete para o aluno com aproveitamento 9.5.

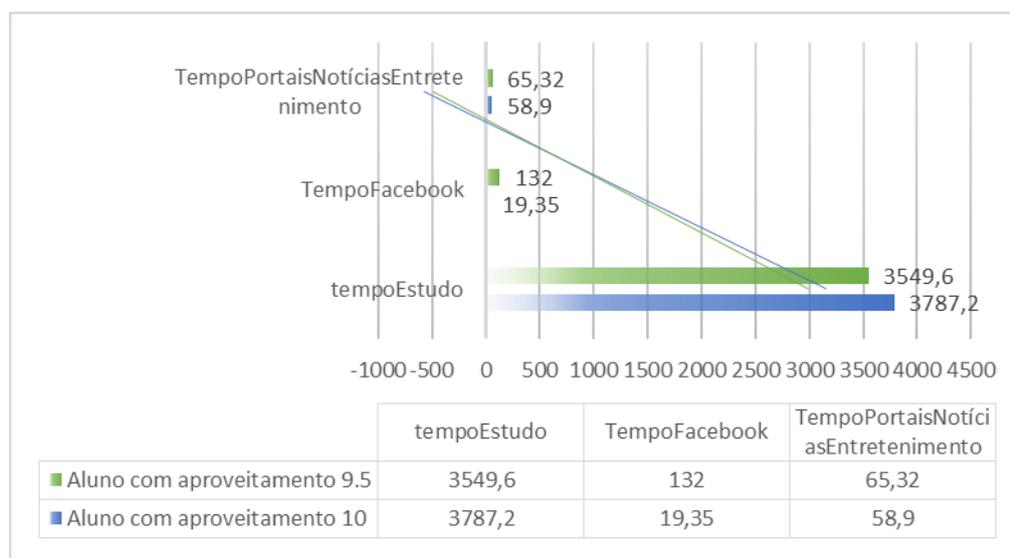


Figura 6.11 - Tempos (min) de estudo, utilização do Facebook e navegação em sites de notícias e/ou entretenimento para alunos com melhores aproveitamentos na disciplina. Fonte: elaborado pela autora.

6.2.1.5 Discussões sobre os resultados obtidos nas entrevistas com estudantes: atendimento aos objetivos e às proposições de pesquisa

Os objetivos propostos para a avaliação na perspectiva de estudantes consistiram em: (1) avaliar se as análises oferecidas pelo ViTrackeR ofereceram suporte para apoiar a aprendizagem autorregulada; e, (2) avaliar a utilidade da informação apresentada pela ferramenta, a facilidade na execução e a facilidade de uso. Analogamente aos objetivos propostos, a proposição identificada por P1 consiste da seguinte afirmação:

P1 - As ferramentas analíticas, desenvolvidas a partir do modelo proposto, auxiliam na promoção da aprendizagem autorregulada de estudantes e melhoram a tomada de decisão em nível micro.

A proposição P1 se desdobra em outras cinco subproposições *P1.A*, *P1.B*, *P1.C*, *P1.D*, *P1.E*, descritas na Seção 1.4 deste documento. A *P1.A* pode ser validada a partir das respostas obtidas no item 2.1 do instrumento aplicado aos alunos, uma vez que 100% dos respondentes indicou que o sistema ajudou, ou ajudou muito, a decidir sobre tempo de dedicação aos estudos. Em atendimento à subproposição *P1.B*, 100% dos alunos indicou, no item 2.7 do instrumento, que o sistema ajudou, ou ajudou muito, a refletir sobre sua aprendizagem durante o desenvolvimento da disciplina. Igualmente, 100% dos alunos avaliou que a aplicação ajudou, ou ajudou muito, a monitorar seus estudos e atividades, validando, portanto, a subproposição *P1.C*. Sobre a autoavaliação da aprendizagem, presentes na proposição *P1.D*, respostas referentes ao item 4 do instrumento indicam que 97% dos alunos apontou que as informações oferecidas pelo sistema foram úteis e, portanto, os ajudou na autoavaliação da aprendizagem. As recomendações oferecidas pelo ViTrackeR foram avaliadas no item 3 do instrumento de coleta e validam a subproposição *P1.E*, uma vez que mais de 95% dos respondentes seguiu as recomendações propostas e 91% dos alunos mudou seu comportamento em decorrências das recomendações.

Com isso, considera-se que a proposição *P1* é verdadeira quando aplicada ao grupo de estudantes participantes desta pesquisa, cujas características foram apresentadas na Seção 6.2.1.2, uma vez que o ViTrackeR, construído a partir do entendimento das etapas especificadas no Modelo Conceitual proposto, e a partir do *framework* que implementa as etapas essenciais propostas no Modelo, auxiliou na

promoção da aprendizagem autorregulada dos estudantes e auxiliou a tomada de decisão em nível individual.

Voltando-se agora ao atendimento aos objetivos propostos, observa-se que a partir das percepções dos vinte e dois alunos que utilizaram o ViTrackeR, o primeiro objetivo - suporte à aprendizagem autorregulada - é atendido satisfatoriamente, o que pode ser comprovado pelas avaliações realizadas que validaram a **P1** e suas subproposições. O segundo objetivo de avaliação na perspectiva dos estudantes consistiu em avaliar a utilidade da informação, a facilitação das atividades e a facilidade de uso da ferramenta. De acordo com a análise das respostas obtidas pelos estudantes nos itens 4 e 5 do instrumento de coleta evidencia-se que mais de 95% dos estudantes considerou que as informações são úteis, ou muito úteis, e a mesma porcentagem foi obtida na avaliação da facilidade de uso do sistema. Dos respondentes, 100% dos alunos considerou que o sistema facilitou suas atividades. Desta forma, considera-se que o segundo objetivo proposto para esta avaliação também foi satisfeito.

A Seção seguinte apresenta o planejamento das avaliações realizadas com instrutores, os resultados obtidos e discussões.

6.2.2 Avaliação realizada com instrutores (tutores e professores)

Tutores e professores são usuários potenciais de ferramentas analíticas desenvolvidas a partir do *framework* proposto e, portanto, representam uma fonte importante na avaliação do Modelo Conceitual. O objetivo desta avaliação é verificar a efetividade das aplicações construídas para apoio à tomada de decisão no nível meso, desenvolvidas com fins de melhorar a tomada de decisão para usuários deste nível. Para isso, a seguinte proposição é apresentada:

P2 – O Modelo Conceitual proposto facilita o desenvolvimento de aplicações que atuam em nível meso e auxiliam instrutores (tutores e professores) a tomarem decisões sobre suas atividades docentes durante e após o término de disciplinas.

Esta avaliação foi conduzida a partir do uso do ViMonitor e também a partir do ViAssess. Não foi possível permitir o uso dessas duas aplicações pelos participantes desta pesquisa durante o período de oferta das disciplinas, o que possibilitaria uma avaliação mais pontual e específica. No entanto, resultados das análises obtidos pela pesquisadora durante a oferta das disciplinas em que o estudo foi aplicado,

foram apresentados aos participantes e permitiram que os mesmos opinassem sobre a efetividade do ViMonitor. Uma simulação do uso do ViAssess foi realizada junto aos professores a fim de possibilitar a participação dos mesmos nesta pesquisa.

6.2.2.1 Planejamento das Entrevistas e Instrumento de Coleta

A escolha intencional da amostra quando se opta pelo uso de estudo de caso é uma situação típica (Miles e Huberman, 1994). A escolha da amostra considerou a atuação dos participantes como instrutores em EaD e a aceitação em participar da presente pesquisa. Convites foram realizados por email, onde também foi feita uma apresentação geral do projeto e os objetivos da pesquisa foram elucidados. No mesmo email também foram apresentados o protocolo a ser seguido na entrevista e o tempo previsto de duração da mesma. As entrevistas levaram em média três horas de duração e foram todas gravadas, conforme a aceitação dos participantes. Também foram realizadas anotações durante as sessões.

Inicialmente, foi feita uma apresentação do Modelo Conceitual proposto e em seguida, foram apresentados exemplos de análises realizadas durante o uso do ViTracker pelos alunos do IFSP e exemplos obtidos pelo ViMonitor. Por último, foi realizada a entrevista com os participantes.

O instrumento de coleta de dados utilizado nas entrevistas com instrutores tem formato semiestruturado (disponível no Apêndice C). O instrumento consiste de 16 questões e é composto pelos seguintes tópicos: (1) informações pessoais; (2) gerenciamento pessoal, avaliação e supervisão dos alunos; e, (3) apoio oferecido para a tomada de decisões no nível meso.

A Seção seguinte apresenta as características da amostra selecionada.

6.2.2.2 Caracterização da amostra

Participaram desta pesquisa quatro instrutores, sendo dois homens e duas mulheres. Os perfis dos entrevistados são apresentados no Quadro 6.1.

Quadro 6.5 – Perfis dos instrutores. Fonte: elaborado pela autora.

	Instrutor 1	Instrutor 2	Instrutor 3	Instrutor 4
Formação acadêmica	Graduação, mestrado e doutorado (em andamento) em Ciência da Computação.	Graduação em Ciência da Computação. Mestrado em engenharia elétrica.	Graduação em pedagogia. Graduação em direito. Mestrado e Doutorado em Educação Escolar.	Graduação em pedagogia. Mestrado em Educação.
Tempo de experiência em educação <i>online</i>	6 anos como tutor.	6 anos como tutor.	4 anos como tutor e 3 anos como professor.	6 anos como professor.

Os resultados obtidos nas entrevistas com os instrutores são apresentados na Seção seguinte.

6.2.2.3 Resultados obtidos nas entrevistas com instrutores

Ao serem questionados sobre em que medida os instrutores aceitariam que uma ferramenta os ajudasse no gerenciamento de suas atividades docentes no SGA (Q 2.1); no acompanhamento dos alunos (Q 2.3) e na avaliação dos alunos (Q 2.6), 100% dos participantes indicou que aceitaria ou aceitaria completamente a ajuda oferecida pelo sistema, conforme ilustra a Figura 6.12.

As respostas indicam que 100% dos participantes aceitaria, ou aceitaria completamente, a ajuda do sistema (Q 3.1) e também adaptaria, ou adaptaria completamente, suas estratégias (Q 3.2), conforme ilustra a Figura 6.13.

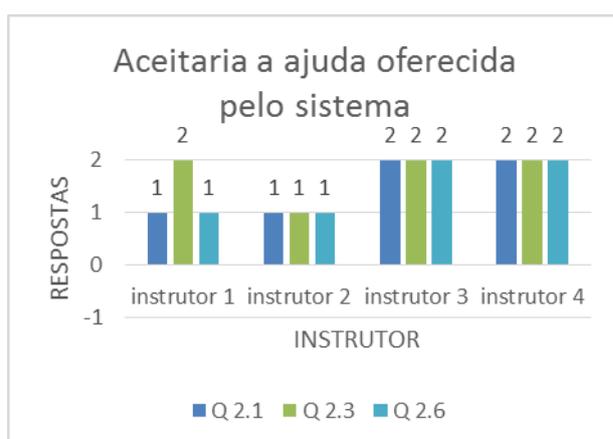


Figura 6.12 - Respostas dos instrutores referente ao item 2 do instrumento. Fonte: elaborado pela autora.

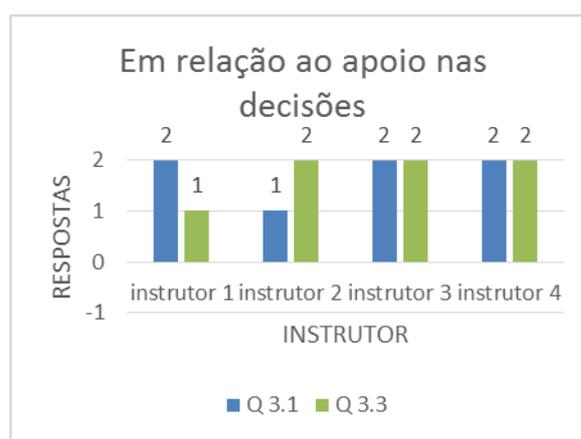


Figura 6.13 - Respostas dos instrutores referente ao item 3 do instrumento. Fonte: elaborado pela autora.

6.2.2.4 Discussões sobre os resultados obtidos nas entrevistas com instrutores: atendimento aos objetivos e às proposições propostas

Comentários realizados durante a participação desta pesquisa indicam que os entrevistados concordam que as visualizações produzidas pelo ViMonitor auxiliam na supervisão e acompanhamento de alunos continuamente e possibilitam intervenções durante o processo, prevenindo o fracasso escolar.

O instrutor 3 fez a seguinte observação: “...uma dificuldade que sinto é acompanhar todas as mensagens em fóruns de discussão, se já dei retorno aos alunos, quais já participaram e quais necessitam de incentivo”.

Os instrutores também opinaram sobre a facilidade que a ferramenta oferece na identificação de lacunas e dificuldades de aprendizagem. Todos os entrevistados concordaram que as informações fornecidas pelo ViMonitor ajudam nessa identificação e podem apoiar, inclusive, decisões sobre mudança nas estratégias de ensino durante o desenvolvimento das unidades de aprendizagem. O instrutor 2 comentou que gostaria de medir quanto os materiais complementares postados pelo professor ajudaram no sucesso do aluno. Uma análise do comportamento do estudante e da sua curva de aprendizagem podem dar indicações sobre isso. O instrutor 2 também comentou que a ferramenta ViMonitor “...pode auxiliar no aprendizado do tutor, ou seja, em como ser um bom tutor”, pois orienta suas ações no ambiente, conforme o número de atividades aumenta.

Com relação ao ViAssess, os entrevistados indicaram interesse em tal aplicação. As opções de monitoramento de segurança para realização de avaliações *online* são interessantes, especialmente, para disciplinas em que o uso do computador é desejável.

De acordo com as opiniões expressas pelos instrutores, as aplicações ViMonitor e ViAssess auxiliam instrutores na tomada de decisões sobre as atividades docentes durante a disciplina e também auxiliam no replanejamento da mesma para uma oferta futura. Desta forma, a proposição **P2** pode ser considerada verdadeira para esta amostra estudada, uma vez que as aplicações mencionadas foram construídas a partir do *framework*, o que consiste da implementação do Modelo proposto. A próxima Seção apresenta o planejamento, resultados obtidos e discussões sobre as avaliações realizadas com gestores para avaliação.

6.2.3 Avaliação realizada com gestores de projetos de EaD

A avaliação realizada com gestores de projetos de educação a distância teve como objetivo: avaliar se modelo proposto, implementado a partir do *framework*, e tendo como ilustrações resultados obtidos por meio de análises realizadas nas três aplicações desenvolvidas (ViTracker, ViMonitor e ViAssess), pode facilitar a tomada de decisão do gestor de projetos de EaD. Partindo do objetivo desta avaliação, a seguinte proposição foi avaliada:

P3 – O Modelo Conceitual proposto facilita o desenvolvimento de aplicações analíticas que apoiam a tomada de decisão educacional em nível macro, auxiliando líderes e gestores de projetos de EaD a melhorarem estratégias educacionais em longo prazo.

6.2.3.1 Planejamento das Entrevistas e Instrumento de Coleta

Para este estudo, a seleção intencional levou em consideração as experiências em cargos de liderança de equipes atuantes em educação *online* e, especialmente, experiências na gestão de projetos de educação a distância. A experiência da pesquisadora condutora deste estudo e os relacionamentos construídos durante sua atuação em EaD possibilitaram a identificação de três potenciais interessados em participar da pesquisa. Outro fator que contribuiu para a escolha dos gestores foi a concordância deles em participar do presente estudo.

Esta tese e os objetivos da pesquisa foram apresentados de maneira resumida por e-mail. As entrevistas levaram em média três horas de duração. Duas entrevistas foram gravadas e foram realizadas anotações durante as sessões. Uma entrevista não pode ser gravada pela indisponibilidade do equipamento de gravação no momento. Para esta última entrevista, foram anotadas as respostas e observações do respondente.

As entrevistas tiveram um caráter informal com a intenção de deixar o participante à vontade para realizar interrupções com o objetivo de tirar dúvidas, fazer sugestões ou apenas para discutir pontos que considerasse relevantes. Primeiramente, foi feita uma apresentação dos objetivos da entrevista, seguida da apresentação do Modelo Conceitual proposto nesta tese. Em seguida, foram apresentados exemplos de análises realizadas a partir de dados da coleta realizada

durante o uso do ViTrackeR pelos alunos para avaliação da ferramenta e também exemplos usando dados randômicos que possibilitaram a ilustração de análises, possíveis de serem obtidas a partir dos dados coletados, que a pesquisadora julgou serem úteis para gestores. As simulações foram realizadas a partir de dados randômicos, dada a impossibilidade de uso do sistema em um grande conjunto de disciplinas e usuários que pudessem produzir dados suficientes para as análises no nível de gestão, tais como acompanhamento da retenção, monitoramento da evasão, monitoramento do aproveitamento de alunos de diferentes disciplinas, cursos e polos e monitoramento de equipes.

O instrumento de coleta de dados utilizado nas entrevistas com gestores, disponível no Apêndice D, é não-estruturado e consistiu de seis questões principais. Para avaliação das respostas dos gestores foi realizada a análise de conteúdo sobre os dados obtidos. A Seção seguinte apresenta as características da amostra selecionada.

6.2.3.2 Caracterização da amostra

Participaram desta avaliação três gestores, sendo todos homens. Todos os entrevistados são, também, pesquisadores na área de educação a distância. Os perfis dos entrevistados são apresentados no Quadro 6.2.

Quadro 6.2 – Perfis dos gestores entrevistados. Fonte: elaborado pela autora.

	Gestor 1	Gestor 2	Gestor 3
Formação acadêmica	Especialização em gestão e liderança universitária; Especialização em docência do ensino superior; Mestrado em filosofia da educação; Doutorado em educação.	Especialização em gestão acadêmica e universitária; Mestrado em bioengenharia; Doutorado em engenharia de produção.	Especialização em didática; Especialização em informática na educação; Mestrado em educação com foco em EaD; Doutorado em educação; Pós-doutorado em Educação a Distância.
Tempo de experiência	10 anos totais, sendo 8 anos na função de gestor.	12 anos totais, todos na função de gestor.	19 anos totais, sendo 13 anos na função de gestor.
Áreas de pesquisa	Fundamentos da educação – teoria crítica; Escola de Frankfurt e Nacionalidade emancipatória.	Gestão educacional privada	Trabalho Docente; Tecnologias, Linguagens, Cognição e Educação a Distância.
Quantidade de alunos sob sua gestão	22.000	3.500	3.600

	Gestor 1	Gestor 2	Gestor 3
Quantidade de cursos sob sua gestão	38	Não informado	11
Categoria Administrativa da Instituição	Privada	Privada	Pública Federal

Os resultados obtidos nas entrevistas com os gestores são apresentados na Seção seguinte.

6.2.3.3 Resultados obtidos nas entrevistas com gestores

Com relação à primeira questão do instrumento de coleta de dados, **Q1**, os participantes da pesquisa foram questionados sobre sua percepção em relação ao Modelo proposto, se de fato ele pode apoiar a tomada de decisões no nível institucional, uma vez implementado. As seguintes respostas foram obtidas:

Gestor 1:

Sim. Todas as informações e análises apresentadas que podem ser obtidas a partir do Modelo proposto, mesmo as informações em nível de estudante, nível de tutores e nível de professores, consigo levar para nível macro de decisão e observo que essas informações são extremamente úteis para facilitar as decisões que devem ser tomadas continuamente em nível institucional.

Gestor 2: “Sim. O Modelo pode obter dados diferenciados e pode fornecer um conjunto de informações que consiste no *core business* da instituição. ”

Gestor 3: “Sim. O Modelo prevê a coleta de grandes quantidades de dados de diferentes tipos, muitos difíceis de serem obtidos por observação e análise humana, que podem melhorar a gestão acadêmica em nível institucional. ”

Com relação à segunda questão, **Q2**, se as visualizações gráficas fornecidas são adequadas e se podem apoiar as decisões em nível institucional, todos os entrevistados apontaram que a possibilidade de ter acesso às informações fornecidas por meio uma ferramenta conforme apresentado, agrupadas e em escala, pode contribuir significativamente para analisar diversas situações do seu dia a dia. Na terceira questão, **Q3**, os gestores foram questionados sobre sua opinião em

relação ao apoio que o Modelo pode oferecer em onze itens avaliados, cujas respostas foram resumidas no Quadro 6.3.

Quadro 6.3 Resumo das respostas dos gestores à questão 3. Fonte: elaborado pela autora.

	Gestor 1	Gestor 2	Gestor 3
Q3.1 Práticas pedagógicas	“Sim, sobretudo para avaliar nos diferentes níveis – unidade, curso, polo, grupo de tutores, grupo de turmas, etc. “	Sim.	Sim.
Q3.2 Monitoramento da retenção	“Sim, sem dúvida. “	Sim.	“Sim. Prefiro me referir a este item como não-evasão”.
Q3.3 Detecção de evasão	“Sim, especialmente na detecção antecipada – o risco de evasão. “	Sim.	Sim.
Q3.4 Monitoramento do grau de conclusão	Sim.	Sim.	Sim.
Q3.5 Observação de estratégias de intervenção	Sim.	“Acredito que informações sobre estratégias de intervenção sejam mais úteis para o nível meso de tomada de decisão, onde atuam os coordenadores de curso e coordenadores pedagógicos. ”	Sim.
Q3.6 Monitoramento de equipes	Sim.	Sim.	Sim.
Q3.7 Reclassificação de pessoal	“Sim. Por exemplo, dependendo de dados sobre o engajamento e atuação do tutor/professor, poderia me ajudar a decidir sobre manter ou não alguém no quadro. ”	Sim.	Sim.
Q3.8 Identificação de necessidade de capacitação	Sim.	Sim.	Sim.
Q3.9 (re)planejamento pedagógico	Sim.	Sim.	Sim.
Q3.10 Avaliação de currículos	“Sim, no sentido de mapear a motivação no semestre, os desafios, etc.”	Sim.	Sim.

	Gestor 1	Gestor 2	Gestor 3
Q3.11 Replanejamento de políticas institucionais	“Sim, permitir, por exemplo, identificar se o Modelo prevê quantidade adequada de atividades, de tutores, de encontros presenciais, etc. “	Sim.	“Acredito que as informações que o sistema (Modelo implementado) tem potencial de fornecer podem apoiar o endossamento de políticas institucionais. ”

A questão 4 do instrumento utilizado durante a entrevista, **Q4**, questionou os três gestores sobre outros tipos de aplicações que poderiam ser implementadas com o objetivo de auxiliar a tomada de decisão em nível institucional, tendo como base o Modelo Conceitual proposto. As seguintes aplicações foram indicadas: “...uma aplicação que estivesse integrada à CPA de forma complementar, apoiando nas análises de dados para primeira dimensão do SINAES, em especial para avaliação do Plano de Desenvolvimento Institucional. “; a medição do contato com tutores, satisfação, cruzamento de dados atc.”; e, “...uma aplicação que cruzasse dados baseados na coleta distribuída, conforme proposto pelo Modelo, com dados de respostas de alunos a um conjunto de questionamentos feitos para coletar a opinião dos usuários. A proposta seria o gestor inserir perguntas no sistema que seriam apresentadas aos estudantes, eventualmente e automaticamente, tais como: “Você gosta desse tutor?”, “Você está com algum problema?”, “Você gostou dos materiais desta disciplina?”, dentre outras.

Ao serem questionados se o Modelo avança no apoio à gestão educacional, tendo como comparação as tecnologias atuais de conhecimento do gestor, **Q5**, os gestores responderam, de imediato, que sim.

O gestor 1 indicou que “atualmente as IES, principalmente as que atuam com EaD, têm tentado investir em recursos analíticos, tais como o *Blackboard*, que é o *top* do mercado atualmente. ” No entanto, ele comentou que não viu nada ainda que considera os dados privados do estudante e suas intenções e apontou que isso consiste em um passo avançado em relação ao que já existe.

O gestor 2 respondeu que ficou “...encantado em conhecer o projeto...” e que desconhece ferramentas que realizam análises de dados e apresentam evidências no nível de detalhes, e tão úteis, quanto as que são possíveis a partir do Modelo proposto.

O gestor 3 indicou que os dados de análises disponíveis atualmente, normalmente consistem apenas em dados do SGA e são difíceis de serem tabulados, trabalhados e que é difícil tirar conclusões a partir deles, principalmente pela dificuldade de obtê-los e organizá-los a fim de fazer cruzamentos. Observou, ainda, que muitas análises são realizadas manualmente em sua instituição.

Na **Q6** os gestores foram questionados se estariam dispostos que sua equipe de TI integrasse as aplicações desenvolvidas no escopo deste trabalho ao seu SGA. Os três gestores responderam que sim à esta questão.

A Seção seguinte apresenta as discussões sobre os resultados obtidos nesta avaliação junto aos gestores.

6.2.3.4 Discussões sobre os resultados obtidos nas entrevistas com gestores: atendimento aos objetivos e às proposições propostas

A avaliação conduzida com gestores experientes na gestão de projetos de EaD, onde uma das instituições selecionadas é pioneira em educação *online* no país, serviu para validar o Modelo Conceitual proposto na perspectiva desse grupo de participantes, e ratificar que o Modelo pode ser utilizado para o desenvolvimento de aplicações que apoiam a tomada de decisão no nível macro, conforme foi proposto no objetivo desta avaliação.

As respostas às questões e as observações feitas pelos gestores confirmaram a utilidade da informação obtida a partir de análises sobre os diferentes tipos de dados obtidos por meio da coleta distribuída e que tais informações podem facilitar e melhorar estratégias educacionais que fazem parte do cotidiano do gestor de projeto de EaD. Todos os gestores indicaram que o Modelo Conceitual proposto pode apoiar a tomada de decisões no nível institucional.

A apresentação deste trabalho para os gestores possibilitou identificar novas aplicações que podem ser construídas com base no Modelo Conceitual e *framework* propostos. Os dois gestores de instituição privada fizeram observações semelhantes ao indicarem que uma ferramenta que apoie o trabalho da CPA seria de grande utilidade para a gestão em nível macro.

Unanimemente, os gestores indicaram interesse em utilizar o sistema apresentado em suas IES. Entende-se que o sistema apresentado consiste do *framework*, implementado a partir do Modelo proposto e também das três aplicações (viTracker, ViMonitor e ViAssess) desenvolvidas no escopo deste trabalho.

Com o exposto, a proposição P3, que parte da afirmação de que o Modelo Conceitual proposto, por meio do *framework*, possibilita o desenvolvimento de aplicações analíticas que apoiam a tomada de decisão educacional em nível macro, auxiliando líderes e gestores de projetos de EaD a melhorarem estratégias educacionais em longo prazo, pode ser considerada verdadeira para a amostra selecionada, cuja caracterização foi feita na Seção 6.2.3.2 deste trabalho.

A Seção seguinte apresenta o planejamento, os resultados e as discussões a partir das entrevistas com desenvolvedores de ferramentas educacionais.

6.3 Avaliação realizada com desenvolvedores de ferramentas de apoio à educação *online*

A avaliação conduzida junto aos desenvolvedores de ferramentas para o apoio à educação *online* teve como objetivo avaliar o *framework* implementado a partir do Modelo Conceitual proposto nesta tese. Duas proposições foram propostas a fim de possibilitar a avaliação do *framework* na perspectiva de desenvolvedores:

P4 – O *framework* proposto oferece um caminho mais simples e unificado para a coleta distribuída de dados a partir de diversas e variadas fontes e realiza análises sobre esses dados de forma satisfatória.

P5 – O *framework* proposto oferece interfaces de programação que facilitam a construção de aplicações analíticas sem a necessidade do desenvolvedor conhecer os detalhes de implementação das etapas de coleta e análise de dados.

6.3.1 Planejamento das Entrevistas e Instrumento de Coleta

Assim como nas avaliações anteriores, a escolha da amostra foi intencional. Levou-se em consideração as experiências dos potenciais interessados e sua atuação no desenvolvimento de ferramentas de apoio à educação *online* e a concordância em participar deste estudo. O contato inicial foi por e-mail, onde foi feita uma explanação do projeto e uma breve apresentação do *framework* e objetivos da avaliação. Com um dos desenvolvedores participantes deste projeto, em um primeiro encontro foi feita a apresentação do projeto de maneira geral e,

detalhadamente, do *framework* proposto. Este desenvolvedor colocou-se à disposição para utilizar o *framework* e construir uma aplicação a partir dele a fim de possibilitar melhor avaliação do mesmo. A fim de agilizar a utilização do *framework*, alguns e-mails foram trocados entre o desenvolvedor e a pesquisadora para auxílio durante a utilização do *framework*. Todo o código fonte, assim como a documentação do *framework*, foi entregue a este participante. Em um segundo encontro foi realizada a entrevista com este desenvolvedor. As duas entrevistas tiveram duração média de, aproximadamente, 4 horas.

Com o segundo desenvolvedor, foi agendado somente um encontro, dada a indisponibilidade de tempo do participante para utilizar o *framework*. Neste encontro foi realizada a apresentação do Modelo de maneira geral, e do *framework* em detalhes, foram apresentados a arquitetura de *software* e trechos do código-fonte exemplificando a coleta e o processo de análise dos dados para possibilitar a avaliação por parte deste desenvolvedor. Em seguida foi realizada a entrevista. Este encontro teve duração de 4 horas e 30 minutos.

O instrumento de coleta de dados utilizado em ambas as entrevistas consiste de seis questões não estruturadas (disponível no Apêndice E). Para avaliação das respostas dos desenvolvedores foi realizada a análise de conteúdo sobre os dados obtidos.

6.3.2 Caracterização da Amostra

Participaram desta avaliação dois desenvolvedores experientes na construção de ferramentas de apoio à educação *online*, sendo ambos homens. Os perfis dos entrevistados são apresentados no Quadro 6.4.

Quadro 6.4 – Perfis dos desenvolvedores entrevistados. Fonte: elaborado pela autora.

	Desenvolvedor 1	Desenvolvedor 2
Formação acadêmica	Graduação em Sistemas, Especialização em Gestão de Projetos; Especialização em Administração Acadêmica e Escolar; Mestrado em Tecnologia e Inovação.	Graduação em Sistemas de Informação; Mestrado em Ciência da Computação.

	Desenvolvedor 1	Desenvolvedor 2
Experiência profissional	10 anos de experiência com programação, todos voltados para o desenvolvimento de ferramentas educacionais para EaD.	17 anos de experiência com programação; 13 anos de experiência no desenvolvimento de ferramentas educacionais.
Funções exercidas em equipes de desenvolvimento de ferramentas educacionais	Programador; Analista de requisitos; Líder de projetos.	Programador; Analista de sistemas; Coordenador de equipe de desenvolvimento.

Os resultados obtidos nas entrevistas com os desenvolvedores são apresentados na Seção seguinte.

6.3.3 Resultados obtidos nas entrevistas com desenvolvedores

A avaliação com os desenvolvedores de ferramentas para apoio à educação *online* serviu para avaliar o *framework* proposto e foi organizada em cinco categorias de questões, que podem ser visualizadas no Apêndice E.

A primeira questão, **Q1**, buscou identificar se a implementação do *framework* abrange todas as etapas necessárias que justificam a escolha da abstração e, portanto, facilitam a construção de aplicações educacionais analíticas.

O primeiro participante respondeu afirmativamente que sim e comentou que a contextualização da coleta foi muito bem planejada e que os dados obtidos em diversas fontes, além do SGA, consistem em um aporte para análises mais específicas que podem levar a novas evidências, diferentes das que são tradicionalmente feitas, e que, certamente, serão de interesse da instituição. O segundo entrevistado respondeu que as etapas presentes no *framework* contemplam todos os passos necessários para implementação de aplicações educacionais analíticas.

Com relação à segunda questão, **Q2**, buscou-se identificar se a implementação do *framework* satisfaz a coleta de dados distribuída, o tratamento dos dados e as análises, conforme prevê o Modelo Conceitual. A percepção de ambos os desenvolvedores é que o *framework* implementou todas as etapas indicadas satisfatoriamente. Um dos desenvolvedores comentou que o potencial de

coleta de dados é além do que ele esperava para uma ferramenta analítica, já que está acostumado apenas com relatórios gerados a partir do SGA e, no máximo, com alguns relatórios com cruzamentos de dados, também obtidos a partir do SGA. Outro comentário realizado é que o código está fácil de manipular e possibilita sua utilização sem problemas.

Ao serem questionados se o *framework* facilita a construção de ferramentas analíticas para a tomada de decisão nos níveis de ensino, aprendizagem e gestão educacional, referente à **Q3**, ambos responderam que sim. O desenvolvedor 1 indicou que as interfaces foram bem definidas e as tecnologias possibilitam o reuso do código. O segundo desenvolvedor indicou que, de maneira ampla, pode afirmar que sim. No entanto, para dar uma opinião mais específica sobre o código, precisaria utilizá-lo.

Na **Q4** os participantes foram questionados se o *framework* oferece um caminho unificado para a coleta distribuída de dados e para as análises. O desenvolvedor 1 indicou que conseguiu utilizar a mesma estrutura de *software* (*framework*) para realizar a coleta e a interpretação dos dados a fim de obter análises mais aprofundadas convenientemente. Comentou, ainda, que não teve dificuldades em utilizar o código. O desenvolvedor 2 também comentou que a proposta do *framework* apresentado oferece um caminho unificado para desenvolver diferentes aplicações que apoiam as decisões em diferentes níveis educacionais.

Com relação às interfaces de programação disponíveis no *framework*, ambos os desenvolvedores indicaram que elas são fáceis de usar e que não há necessidade de conhecer detalhes de implementação para utilizar os serviços de coleta e análise dos dados. O desenvolvedor 1 comentou que o conhecimento do código será necessário somente se houver interesse em estender funcionalidades a partir da estrutura que já está disponível, para implementação de novos tipos de fontes de dados, por exemplo. O desenvolvedor 2 observou que a IES que realiza o *front-end* tem como tarefa apenas a apresentação dos resultados, e não precisa implementar a coleta de grande volume de dados e as análises sobre eles.

Na próxima Seção são apresentadas as discussões sobre os resultados obtidos nesta avaliação.

6.3.4 Discussões sobre os resultados obtidos nas entrevistas com desenvolvedores

Partindo do objetivo proposto para esta avaliação e dadas as respostas obtidas nas entrevistas realizadas, é possível considerar que o *framework* foi avaliado positivamente na perspectiva de dois desenvolvedores experientes no desenvolvimento de ferramentas educacionais.

- Ambos os desenvolvedores, ao serem questionados, indicaram que a arquitetura proposta do *framework* contempla, de maneira satisfatória, todas as etapas necessárias para análise de dados em educação *online*, as quais foram todas implementadas adequadamente em termos de codificação, documentação e resultados. O desenvolvedor 1, que teve contato próximo com a documentação e os códigos do *framework* ao utilizar o mesmo com o propósito de criar uma aplicação, pode fazer uma avaliação mais pontual do mesmo. Este desenvolvedor comentou durante a entrevista que o código não apresentou erros lógicos e teve facilidade em sua utilização, comentou ainda que a documentação disponibilizada orientou adequadamente o desenvolvimento da aplicação que pode ser construída de maneira rápida e descomplicada. A única inconveniência indicada foi que os *plug-ins* de *browser* implementados e disponibilizados são para os navegadores *Chrome* e para *Internet Explorer*, e o desenvolvedor não utilizava nenhum deles. Para testar a aplicação que construiu, o desenvolvedor teve que instalar um deles.

Os desenvolvedores indicaram, de maneira geral, que ficaram satisfeitos com a proposta do *framework* e que consideram que essa ferramenta será muito útil para apoiar as decisões individuais, em nível de curso e também no nível de gestão. Eles comentaram que a maioria das necessidades atuais das instituições é na análise de dados para a gestão em nível macro. Os gestores necessitam de informações que somente a análise de dados do SGA não pode fornecer, especialmente sobre o comportamento dos usuários, tais como a motivação e o engajamento.

Outra observação feita pelos participantes foi que as soluções disponíveis hoje para análise de dados educacionais, em especial para educação superior *online*, além de terem custo elevado para a instituição, não contemplam a coleta e análises de dados externos ao SGA, ou no máximo, cobrem a coleta de dados em bases institucionais. Os entrevistados mencionaram, ainda, que as soluções

disponíveis apresentam limitações em relação à proposta deste trabalho e atendem, apenas em parte, às necessidades do gestor, do professor e do aluno. Opinaram que o *framework* proposto neste trabalho pode prover um melhor acompanhamento dos processos no nível de aprendizagem, de ensino e também de gestão educacional, e, seguramente, pode facilitar a tomada de decisão nesses níveis.

Dadas as respostas obtidas nas entrevistas realizadas com os desenvolvedores e as opiniões manifestas, é possível considerar que **as proposições** de pesquisa **P4**, sobre a coleta distribuída de dados e as análises disponibilizadas pelo *framework* serem satisfatórias, e **P5**, sobre facilidade proporcionada pelo *framework* para desenvolvimento de aplicações educacionais analíticas, **são verdadeiras** para a amostra estudada.

A próxima Seção apresenta o planejamento da avaliação conduzida com especialistas no domínio desta tese e também os resultados obtidos.

6.4 Avaliação realizada com especialistas em aplicações educacionais

Com o objetivo de dar sustentação à validação do Modelo Conceitual proposto nesta tese, foi conduzida uma avaliação do Modelo a partir da avaliação da percepção de dois especialistas que desenvolvem pesquisas sobre ferramentas de apoio à educação *online*. As opiniões desses especialistas são apresentadas nas subseções que seguem.

6.4.1 Planejamento das Entrevistas e Instrumento de Coleta

Para a avaliação do Modelo Conceitual proposto nesta tese junto aos especialistas de domínio foi feita uma seleção intencional baseada no julgamento da pesquisadora, onde foram considerados potenciais interessados em participar deste estudo. A seleção dos interessados considerou um perfil de pesquisador com experiência acadêmica de, pelo menos, 10 anos; sua atuação em projetos de pesquisas relacionadas ao tema deste trabalho; sua participação como líder em

projetos de pesquisa e implantação de ferramentas de apoio à educação *online* e também a disponibilidade e interesse em participar deste estudo.

Em um contato inicial por e-mail, foi feita uma apresentação geral do projeto e dos objetivos desta avaliação. Após concordar em participar, as entrevistas foram agendadas com dois pesquisadores que atenderam ao perfil desejado. As entrevistas tiveram um caráter informal com interrupções previstas para esclarecimentos e discussões pertinentes. As sessões duraram, em média, 2 horas e 30 minutos e foram gravadas com a permissão dos participantes. Também foram realizadas anotações durante as entrevistas. Um dos entrevistados desconhecia o projeto e para facilitar sua avaliação, foram enviados, antecipadamente, artigos publicados pela pesquisadora e a apresentação em *slides* que foi utilizada durante a entrevista. O segundo participante já tinha conhecimento sobre o projeto dada uma colaboração pontual em uma parceria no início do desenvolvimento do mesmo.

O instrumento de coleta de dados utilizado consiste de seis questões não estruturadas, que são apresentadas na próxima Seção. Para avaliação das respostas dos gestores foi realizada a análise de conteúdo sobre os dados obtidos. O instrumento de coleta de dados utilizado encontra-se no Apêndice F.

6.4.2 Caracterização da Amostra

Participaram desta avaliação dois pesquisadores doutores que desenvolvem, ativamente, pesquisas sobre ferramentas de apoio à educação *online*, sendo ambos do gênero feminino. Os perfis dos entrevistados são apresentados no Quadro 6.5.

Quadro 6.5 – Perfis dos especialistas entrevistados. Fonte: elaborado pela autora.

	Pesquisador 1	Pesquisador 2
Formação acadêmica	Graduação, mestrado e doutorado em Ciência da Computação.	Graduação engenharia elétrica, mestrado e doutorado em Ciência da Computação.
Tempo de experiência	19 anos com pesquisas relacionadas ao desenvolvimento de ferramentas para apoio à educação <i>online</i> ; e, 6 anos como gestora de projetos e de equipes de implantação de soluções para apoio à educação <i>online</i> .	19 anos com pesquisas relacionadas à Educação a Distância; 11 anos como gestora de projetos de pesquisa e de equipes de implantação de ferramentas computacionais para suporte à EaD.

	Pesquisador 1	Pesquisador 2
Áreas de pesquisa	Educação a Distância; Aprendizagem Colaborativa Apoiada por Computador; Agentes de interface/sistemas multiagentes; Usabilidade; <i>Design</i> de sistemas aprendizagem colaborativa; Acessibilidade; Recursos educacionais abertos interativos e <i>Design</i> e avaliação de jogos educacionais.	Educação a Distância; Teoria das Inteligências Múltiplas aplicada ao ensino na Educação a Distância; Tecnologias Educacionais e Acessibilidade.

Os resultados obtidos nas entrevistas com os pesquisadores são apresentados na Seção seguinte.

6.4.3 Resultados obtidos nas entrevistas com especialistas

A avaliação conduzida junto aos especialistas de domínio foi realizada com o objetivo de validar o Modelo Conceitual proposto nesta tese e foi organizada por meio de entrevista contendo três questões não estruturadas.

As respostas apresentadas para a **Q1** indicam que ambos os entrevistados consideram que o Modelo proposto nesta tese é adequado para apoiar o desenvolvimento de aplicações educacionais analíticas. O pesquisador 1 comentou que a coleta de dados a partir das diferentes fontes utilizadas, assim como os serviços de análise disponíveis, torna o Modelo mais completo que os tradicionalmente propostos. O pesquisador 2 mencionou que considera que o contexto do usuário pode ser melhor definido a partir da coleta sob a abstração de redes de sensores híbrida e que as fontes previstas consistem de um conjunto amplo para aquisição de dados, possibilitando uma cobertura maior para a captura de contextos. O pesquisador 2 observou ainda que o Modelo prevê diferentes tecnologias para análises dos dados e possibilita a prospecção de dados a partir da detecção de relacionamentos entre variáveis, oferecendo um suporte mais adequado para a tomada de decisão nos diferentes níveis mencionados, especialmente atendendo às necessidades do gestor de EaD, que necessita de suporte tecnológico que o ajude a sistematizar suas práticas de gestão.

Na **Q2** os participantes foram questionados sobre as etapas presentes no Modelo Conceitual, se elas são adequadas e suficientes para atender ao seu propósito de apoiar o desenvolvimento de ferramentas analíticas para suporte à tomada de decisão nos diferentes níveis e para os diferentes perfis. Ambos

entrevistados responderam que sim. O pesquisador 1 comentou que as etapas são satisfatórias e adequadas e considera que o Modelo é completo e atende aos perfis de aluno, tutor, professor e gestor educacional. Este pesquisador mencionou, ainda, que a implementação do Modelo a partir do *framework* satisfaz adequadamente quesitos de extensibilidade ao permitir a adição de novos componentes e a extensão dos componentes existentes. O pesquisador 2 comentou que o Modelo prevê as etapas essenciais à análise de dados educacionais com algumas vantagens que mencionou: ampla detecção de contextos dos usuários; diferentes modelos de análises, possibilitando o uso do Modelo, implementado a partir do *framework*, para atendimento às aplicações de LA com diferentes atuações, tais como personalização e avaliação; e, por conseguinte, apoia a tomada de decisão em diferentes níveis.

Sobre a validação do Modelo Conceitual proposto nesta tese, **Q3**, o pesquisador 1 concordou que a validação proposta é adequada, especialmente, por contemplar a avaliação do *framework* a partir da perspectiva de desenvolvedores. O pesquisador 2 também considerou que a avaliação conduzida a partir das percepções dos participantes (*stakeholders*) em relação ao sistema, somada à avaliação feita pelos desenvolvedores sobre o *framework* e também as opiniões de pesquisadores especialistas no domínio podem ser utilizadas para validar o Modelo Conceitual proposto neste trabalho. As discussões a respeito dos resultados obtidos são apresentadas na Seção seguinte.

6.4.4 Discussões sobre os resultados obtidos nas entrevistas com especialistas

Pesquisadores no domínio deste trabalho julgaram se o Modelo Conceitual, implementado a partir do *framework* proposto, e as aplicações desenvolvidas satisfazem aos objetivos de análise propostos. De acordo com as opiniões expressas e julgamentos feitos, o Modelo pode ser considerado satisfatoriamente validado e apoia a construção de ferramentas educacionais analíticas direcionadas para aprendizagem, para o ensino e/ou para a gestão educacional.

Ambos pesquisadores manifestaram-se satisfeitos com o Modelo quando questionados e também afirmaram que as etapas previstas apoiam, convenientemente, a construção de aplicações educacionais analíticas que podem

atender diferentes perfis de usuários e diferentes níveis de tomada de decisão em educação.

Dado o exposto, considera-se que na opinião dos participantes desta pesquisa, o Modelo Conceitual proposto nesta tese, implementado a partir do *framework* de coleta de dados por meio de redes de sensores híbrida e métodos análises a partir de modelos preditivo e semântico, pode ser utilizado por instituições de ensino para facilitar a compreensão sobre as etapas envolvidas na coleta e análise de dados educacionais, especialmente por apoiar a tomada de decisão em diferentes níveis (individual, curso/disciplina e gestão).

Com isso, é verdade afirmar que um sistema implementado a partir do Modelo Conceitual e *framework* propostos pode apoiar na previsão do sucesso/fracasso de estudantes; no monitoramento do desempenho acadêmico e aproveitamento escolar enquanto a disciplina está em andamento, permitindo intervenções e correções; em recomendações e orientações individuais definidas a partir da identificação do comportamento e preferências dos usuários; e na avaliação da aprendizagem.

6.5 Considerações Finais

Este capítulo descreveu em detalhes os procedimentos metodológicos utilizados nesta pesquisa. Foram apresentados o planejamento das entrevistas com cada uma das fontes de dados utilizadas, os resultados obtidos nas entrevistas e discussões sobre tais resultados.

As entrevistas foram realizadas com participantes que atuam nos diferentes níveis de decisão educacional. No nível micro, as avaliações com estudantes possibilitaram a verificação da proposição P1; no nível meso, as entrevistas realizadas com instrutores, possibilitaram a investigação da P2; no nível macro, gestores puderam atestar a veracidade da proposição P3. Desenvolvedores avaliaram o *framework* proposto, oferecendo suporte para corroborar a veracidade das proposições P4 e P5. Por fim, uma validação aparente foi conduzida com especialistas na área de pesquisa deste trabalho.

O capítulo seguinte apresenta as conclusões desta tese.

Capítulo 7

CONSIDERAÇÕES FINAIS

Este capítulo apresenta as considerações finais deste trabalho. A Seção 7.1 apresenta o atendimento à questão de pesquisa que delineou esta tese, aos objetivos e às proposições. A Seção 7.2 apresenta as limitações da pesquisa. A Seção 7.3 apresenta trabalhos futuros.

7.1 Atendimento à questão de pesquisa, aos objetivos e às proposições

Como questão de pesquisa desta tese, propôs-se a seguinte indagação: como oferecer suporte conceitual e tecnológico para apoiar a construção de aplicações analíticas que facilitam a tomada de decisão em educação mediada por tecnologias nos níveis de ensino, aprendizagem e gestão educacional?

De modo correspondente à questão de pesquisa, o objetivo principal deste estudo consiste em propor e validar um modelo para auxiliar a tomada de decisão educacional dirigida a dados nos níveis de ensino e aprendizagem com tecnologias e de gestão educacional, atendendo aos diferentes perfis de usuários de EaD e integrando todas as etapas necessárias à análise educacional. A avaliação conduzida neste trabalho buscou analisar as percepções de diversos tipos de fontes, incluindo usuários das aplicações desenvolvidas a partir do *framework* (estudantes, instrutores e gestores); desenvolvedores com experiência no desenvolvimento de ferramentas para apoio à EaD que avaliaram o *framework*, inclusive em termos de código; e, especialistas na área de pesquisa deste trabalho, que puderam opinar sobre a proposta geral deste trabalho e, em especial, sobre o Modelo Conceitual

proposto. Conforme apresentado no Capítulo 6 deste documento, os resultados obtidos nas avaliações em todas as esferas investigadas respaldam o objetivo geral proposto. Sendo assim, admite-se que o objetivo geral proposto neste trabalho foi alcançado a partir dos resultados obtidos nas avaliações conduzidas para as amostras selecionadas.

No que se refere ao atendimento às proposições de pesquisa, a proposição geral deste estudo foi desmembrada em outras cinco proposições: P1, P2, P3, P4 e P5. Conforme discussões apresentadas no Capítulo 6 deste documento, as cinco proposições provenientes da proposição geral foram declaradas verdadeiras para as amostras estudadas. Com isso, a proposição geral deste trabalho, que propõe um Modelo Conceitual para facilitar a compreensão sobre as etapas necessárias à construção de aplicações analíticas educacionais e, por meio do framework que implementa essas etapas, possibilita o desenvolvimento de aplicações analíticas que apoiam a tomada de decisão educacional nos níveis micro, meso e macro, pode ser considerada válida para este estudo.

Retomando a questão de pesquisa desta tese, o suporte conceitual e o suporte tecnológico que apoiem a construção de aplicações analíticas que facilitam a tomada de decisão em educação mediada por tecnologias nos níveis de ensino, aprendizagem e gestão educacional, podem ser obtidos, respectivamente, por meio do Modelo Conceitual proposto nesta tese, assim como do *framework* proposto que implementou todas as etapas previstas no Modelo, para o estudo de caso conduzido com as amostras selecionadas e nas condições indicadas.

7.2 Limitações da Pesquisa

Esta tese consiste de uma pesquisa de natureza exploratória e descritiva, não tendo, portanto, o propósito de esgotar o tema de pesquisa, mas sim de investigá-lo a partir da análise e interpretação dos fatos com o fito de validar o Modelo Conceitual proposto e também de coletar sugestões que poderão ser consideradas para novas versões do *framework* e das aplicações desenvolvidas no escopo deste trabalho. Por este motivo, uma possível limitação observada está associada ao fato desta pesquisa validar o Modelo Conceitual proposto apenas a partir dos estudos

realizados sob as três facetas apresentadas: (1) nas perspectivas dos usuários das três aplicações desenvolvidas, (2) nas perspectivas de desenvolvedores de ferramentas educacionais e também (3) na avaliação por pesquisadores experientes na área (*face validity*). No entanto, devido à falta de recursos de tempo, de orçamento e de equipes de desenvolvimento e testes, não foi realizada uma validação que poderia oferecer maior consistência ao Modelo Conceitual proposto, tal como a combinação de outras validações, por exemplo validação de consistência externa (*cross-validation*) e também a validade interna (verificação). Desta forma, ainda não é possível mensurar quantitativamente a capacidade de generalização do Modelo.

Outra limitação encontrada foi a impossibilidade de utilizar as três aplicações (ViTrackeR, ViMonitor e ViAssess) desenvolvidas em ambientes reais para uma amostra maior - um conjunto maior de disciplinas ministradas e em diferentes cursos, e também por um período maior que um semestre letivo. Estudos futuros estão planejados para superar essas limitações apresentadas.

7.3 Trabalhos Futuros

Como trabalho futuro desta tese pretende-se validar o modelo proposto a partir de testes em novos conjuntos de dados (*cross-validation*), utilizando um grupo maior de usuários e cursos. Também está prevista a avaliação da escalabilidade e desempenho do sistema à medida que aumenta, incrementalmente, o volume de dados para análises. Como terceiro passo, a coleta de dados deverá ser expandida para outros canais informais de aprendizagem, a fim de suportar o rastreamento a partir de aprendizagem em nuvem e uso de múltiplos dispositivos, simultaneamente, pelo mesmo usuário.

Como trabalhos futuros, o *framework* deverá implementar novas técnicas de análises de dados automatizadas que possibilitem a identificação de tendências, correlações, conexões e o comportamento dos dados continuamente durante o desenvolvimento da disciplina. Além disso, está prevista uma reavaliação dos modelos estatísticos propostos a fim de incluir o tratamento de mudanças repentinas no comportamento dos usuários. Neste caso, o modelo estatístico deverá ser capaz

de explicar mudanças imprevistas no comportamento humano, por exemplo, um estudante que teve um bom aproveitamento acadêmico durante todo o curso, frequência satisfatória e participou ativamente nas interações; de repente, este estudante abandona o curso. Esse aluno pode representar uma mutação no comportamento dos dados, o que é difícil de ser previsto.

REFERÊNCIAS

ABED. Censo EAD.BR. **Relatório analítico da aprendizagem a distância no Brasil**. São Paulo: Pearson Education do Brasil, 2010. Disponível em: <http://www.abed.org.br/censoead/CensoEaDbr0809_portugues.pdf>. Acesso em: outubro 2012.

ACHARYA, S.; KAM, M. Evidence combination for hard and soft sensor data fusion. In: INTERNATIONAL CONFERENCE ON INFORMATION FUSION, 14., 2011, Chicago. **Proceedings of the 14th International Conference on Information Fusion**. Chicago: ISIF, 2011. P. 1008-1015.

AKMAN, V.; MEHMET, S. Steps toward Formalizing Context. **AI Magazine**, v. 17, n. 3, p. 55-72, 1996.

ALI, L.; HATALA, M.; GASEVIC, D.; JOVANOVIC, J. A qualitative evaluation of evolution of a learning analytics tool. **Computers & Education**, v. 58, n.1, p. 470–489, 2012.

ARNOLD, K. E.; PISTILLI, M. D. Course Signals at Purdue: Using learning analytics to increase student success. In: INTERNATIONAL CONFERENCE ON LEARNING ANALYTICS & KNOWLEDGE, 2., 2012, New York. **Proceedings of the 2nd International Conference on Learning Analytics & Knowledge**. New York: ACM, 2012. p. 267–270.

ATTWELL, G; COSTA, C. **Integrating personal learning and working environments**. Pontydysgu - Bridge to Learning, 2008. Disponível em: <<http://www.pontydysgu.org/research/working-and-learning/>>. Acesso em: fevereiro 2014.

BACH, C. **Learning Analytics: Targeting Instruction, Curricula and Student Support**. Office of the Provost. Drexel University, 2010. Disponível em: <http://www.iiis.org/CDs2010/CD2010SCI/EISTA_2010/PapersPdf/EA655ES.pdf>. Acesso em: maio 2014.

BAEPLER, P.; MURDOCH, C. J. Academic Analytics and Data Mining in Higher Education. **International Journal for the Scholarship of Teaching and Learning**. v. 4, N. 2, Article 17, 2010.

BAKER, R. S. J. D.; YACEF, K. The state of educational data mining in 2009: A review and future visions. **Journal of Educational Data Mining**, v. 1, n. 1, p. 3-17, 2009.

BAKHARIA, A.; DAWSON, S. SNAPP: A Bird's-Eye View of Temporal Participant Interactions. In: INTERNATIONAL CONFERENCE ON LEARNING ANALYTICS & KNOWLEDGE, 1., New York. **Proceedings of the 1st International Conference on Learning Analytics & Knowledge**, NY, USA: ACM, 2011, p. 168-173.

BANDURA, A. Social cognitive theory: an agentic perspective. **Annual Review of Psychology**. Palo Alto: Annual Review, v. 52, p.1-26, 2001.

BARDIN, L. **Análise de Conteúdo**. São Paulo: Edições 70, 2002.

BAUER, J. **Identification and Modeling of Contexts for Different Information Scenarios in Air Traffic**. Ph.D. Thesis, Technology University of Berlin, Berlin, 2003.

BELLIFEMINE, F.; POGGI, A.; RIMASSA, G. JADE – A FIPA-compliant agent framework. In: CONFERENCE ON PRACTICAL APPLICATIONS OF INTELLIGENT AGENTS AND MULTI-AGENT SYSTEMS (PAAM), 1999, **Proceedings of Conference on Practical Applications of Intelligent Agents and Multi-agent Systems (PAAM)**, 1999, v. 99, p. 97-108.

BETTINI, C.; BRDICZKA, O.; HENRICKSEN, K.; INDULSKA, J.; NICKLAS, D.; RANGANATHAN, A.; RIBONI, D. A survey of context modelling and reasoning techniques. **Pervasive and Mobile Computing**, v. 6, n. 2, 2010, p. 161–180.

BETTINI, C.; PARESCHI, L.; RIBONI, D. Efficient profile aggregation and policy evaluation in a middleware for adaptive mobile applications. **Pervasive and Mobile Computing**, v.4 n.5, 2008, p.697-718.

BLACKALL, L. **Die LMS die! You too PLE!**. Learn Online, 2005. Disponível em: <<http://teachandlearnonline.blogspot.com/2005/11/die-lms-die-you-too-ple.html>>. Acesso em: maio 2014.

BOLCHINI, C.; CURINO, C. A.; QUINTARELI, E.; SCHREIBER, F. A.; TANCA, L. Data-oriented Survey of Context Models. **ACM SIGMOD RECORD**, New York, Vol. 36, No.4, 2007, p.19-26.

BOZO, J., ALARCÓN, R., IRIBARRA, S. Recommending Learning Objects According to a Teachers Context Model. In: EC-TEL, 2010, Berlin Heidelberg. **Proceedings of EC-TEL**, Berlin Heidelberg. LNCS 6383, 2010. p. 470-475.

BRADSHAW, J. M. **Software Agents**. [S.l.], MIT Press, 1997. 490 p.

BRASIL. **DECRETO Nº 5.622**, DE 19 DE DEZEMBRO DE 2005. Regulamenta o art. 80 da Lei nº 9.394, de 20 de dezembro de 1996, que estabelece as diretrizes e bases da educação nacional. Ministério da Educação. Disponível em: <http://www.planalto.gov.br/ccivil_03/_Ato2004-2006/2005/decreto/D5622.htm>. Acesso em: 05 2012.

BRASIL. **Referenciais de Qualidade para EAD de 24 de maio de 2007**. Secretaria de Educação a Distância. Disponível em: <portal.mec.gov.br/seed/arquivos/pdf/referenciaisead.pdf> 2007.

BROCKETT, R.G.; HIEMSTRA, R. **Ministério da Educação: Perspectives on theory, research and practice**. New York: Routledge, 1991.

BROISIN, J.; BRUT, M.; BUTOIANU, V.; SEDES, F.; VIDAL, F. A Personalized Recommendation Framework based on CAM and Document Annotations. In: WORKSHOP ON RECOMMENDER SYSTEMS FOR TECHNOLOGY ENHANCED LEARNING, 1., 2010, **Proceedings of the 1st Workshop on Recommender Systems for Technology Enhanced** . 2010. p. 2839–2848.

BROWN, M. **Learning Analytics: The Coming Third Wave**. EDUCAUSE Learning Initiative Brief, 2011. Disponível em: <<http://www.educause.edu/Resources/LearningAnalyticsTheComingThir/227287>>. Acesso em: maio 2014.

BROWN, J. S.; ADLER, R. P. **Minds on Fire: Open education, the Long Tail, and Learning 2.0**. EDUCAUSE Review, 2008 vol.43, no.1:16-32. Disponível em: <<http://net.educause.edu/ir/library/pdf/ERM0811.pdf>>. Acesso em: julho 2014.

BRUSILOVSKY, P.; KOBASA, A.; NEJDL, W. **The Adaptive Web: Methods and Strategies of Web Personalization**. Springer-Verlag Berlin Heidelberg, LNCS v. 4321, 2007.

BRUSILOVSKY, P.; MILLÁN, E. **User models for adaptive hypermedia and adaptive educational systems**. In Brusilovsky, P., Kobsa, A. and Nejd, W. (Eds.): *The Adaptive Web*, LNCS 4321, p. 3–53, Springer-Verlag Berlin, 2007.

CAMPBELL, J. P.; DEBLOIS, P. B.; OBLINGER, D. G. **Academic Analytics: A New Tool for a New Era**. *EDUCAUSE Review*, 2007. pp. 41–57. Disponível em: <<http://net.educause.edu/ir/library/pdf/erm0742.pdf>>. Acesso em: abril 2014.

CAMPBELL, J. P. **The Grand Challenge: Using Analytics to Predict Student Success**. *EDUCAUSE Midwest Regional Conference*, 2007. Disponível em: <<http://net.educause.edu/ir/library/pdf/MWR07085.pdf>>. Último acesso em: abril 2014.

CAMPBELL, J. P.; OBLINGER, D. G. **Academic Analytics**. *EDUCAUSE white paper*, 2007. Disponível em: <<http://net.educause.edu/ir/library/pdf/PUB6101.pdf>>. Acesso em: abril 2014.

CARD, S.; MACKINLAY, J.; SHNEIDERMAN, B. **Readings in information visualization: using vision to think**. San Diego, CA: Morgan Kaufmann Publishers, 1999.

CARD, S. Information visualization. In: A. Sears, e J. A. Jacko (Eds.), **The human-computer interaction handbook: Fundamentals, evolving technologies, and emerging applications**. New York, NY: Lawrence Erlbaum Assoc Inc, 2003.

CHATTI, M. A.; AGUSTIAWAN, M. R.; JARKE, M.; SPECHT, M. Toward a personal Learning Environment Framework. **International Journal of Virtual and Personal Learning Environments**, 1(4), 66-85, 2010.

CHATTI, M. A.; AGUSTIAWAN, M. R.; JARKE, M.; SPECHT, M. Toward a personal learning environment framework. In: **Design, implementation, and evaluation of virtual learning environments**. IGI Global, 2012a, p 20–40.

CHATTI, M. A.; DYCKHOFF, A. L.; SCHROEDER, U.; THÜS, H. A Reference Model for Learning Analytics. **International Journal of Technology Enhanced Learning (IJTEL)**, Vol.4, No.5/6, 2012b, p.318 – 331.

CHEN, G.; KOTZ, D., **A Survey of Context-Aware Mobile Computing Research**. Technical report, Dartmouth College, 2000.

CHEN, H.; FINN, T.; JOSHI, A. An Ontology for Context-Aware Pervasive Computing Environments. **Knowledge Engineering Review**, 18, 3, pp. 197-207, 2003.

CHEN, G. D.; CHANG, C. K.; WANG, C. Y. Ubiquitous learning website: scaffold learners by mobile devices with informationaware techniques. **Computers & Education** 50(1):77–90 28, 2008.

CHO, D. J.; HONG, M. W. A design of ontology context model in ubiquitous learning environments. In: INTERNATIONAL CONFERENCE ON COMPUTERS. 12., 2008, Wisconsin. **Proceedings of the 12th WSEAS international conference on Computers**, Wisconsin, USA, 2008. p. 844–848.

CONDE, M. A.; HERNÁNDEZ-GARCÍA, A. A promised land for educational decision-making?: present and future of learning analytics. In: INTERNATIONAL CONFERENCE ON TECHNOLOGICAL ECOSYSTEM FOR ENHANCING MULTICULTURALITY, 1., 2013, New York. **Proceedings of the First International Conference on Technological Ecosystem for Enhancing Multiculturality**. ACM New York, NY, USA 2013. p. 239-243.

DAVENPORT, T. **The three 'tives' of business analytics; predictive, prescriptive and descriptive**. CIO Enterprise Forum, 2012. Disponível em: <http://www.enterprisecioforum.com/en/video/three-tives-business-analytics-predictiv?q4851921=1>. Acesso em: maio 2014.

D'ANDREA, A.; FERRI, F.; GRIFONI, P. An Overview of Methods for Virtual Social Network Analysis. In: **Abraham, Ajith et al. Computational Social Network Analysis: Trends, Tools and Research Advances**. Springer. 2008, p. 8.

DAWSON, S.; MCWILLIAM, E.; TAN, J. P.-L. **Teaching Smarter: How Mining ICT Data Can Inform and Improve Learning and Teaching Practice**. Hello! Where Are You in the Landscape of Educational Technology? Proceedings ascilite Melbourne, 2008. Disponível em: <http://www.ascilite.org.au/conferences/melbourne08/procs/dawson.pdf>. Acesso em: maio 2014.

DEY, A. K.; ABOWD, G. D. **Towards a better understanding of context and context-awareness**. GVU Technical Report GITGVU- 99-22, College of Computing, Georgia Institute of Technology, 1999. Disponível em: <ftp://ftp.cc.gatech.edu/pub/gvu/tr/1999/99-22.pdf>. Acesso em: agosto 2011.

DHIMAN, K.; QUACH, B. Google's Go and Dart: parallelism and structured web development for better analytics and applications. In: CONFERENCE OF THE

CENTER FOR ADVANCED STUDIES ON COLLABORATIVE RESEARCH, 12., 2012, Riverton. **Proceedings of the 2012 Conference of the Center for Advanced Studies on Collaborative Research**. Riverton, NJ, USA 2012. p. 253-254.

DICKEY, D. A. **Introduction to Predictive Modeling with Examples** . SAS Global Forum - Statistics and Data Analysis, 2012. Disponível em: <http://support.sas.com/resources/papers/proceedings12/337-2012.pdf>. Acesso em: março 2014.

DOLLÁR, A.; STEIF, P. S. Web-based statics course with learning dashboard for instructors. In: COMPUTERS AND ADVANCED TECHNOLOGY IN EDUCATION, 2012, Napoli. **Proceedings of Computers and Advanced Technology in Education (CATE 2012)**, Napoli, Italy, 2012.

DOURISH, P. What we talk about when we talk about context. **Personal Ubiquitous Computing**. Vol. 8, No. 1, 2004, p. 19-30.

DRACHSLER, H.; HUMMEL, H.; VAN DEN BERG, B.; ESHUIS, J.; WATERINK, W.; NADOLSKI, R.; BERLANGA, A.; BOERS, N.; KOPER, R. Effects of the Isis Recommender System for Navigation Support in Self-Organized Learning Networks. **Educational Technology and Society**, vol. 12, no. 3, p. 115-126, 2009.

DRAGUNOV, A. N.. DIETTERICH, T. G.. JOHNSRUDE, K.. MCLAUGHLIN, M.. LI, L.; HERLOCKER, J. TaskTracer: A Desktop Environment to Support Multi-tasking Knowledge Workers. In: INTERNATIONAL CONFERENCE ON INTELLIGENT USER INTERFACES, 10., 2005, New York. **Proceedings of the 10th International Conference on Intelligent User Interfaces**. New York, USA: ACM Press, 2005. p.75-82.

DYCKHOFF, A. L.; ZIELKE, D.; BÜLTMANN, M.; CHATTI, M. A.; SCHROEDER, U. Design and Implementation of a Learning Analytics Toolkit for Teachers. **Educational Technology & Society**, 15 (3), 58–76, 2012.

EARL, L.; KATZ, S. **Leading schools in a data rich world**. Thousand Oaks, CA: Corwin Press, 2006.

ECKERSEN, W. W. **Predictive Analytics: Extending the Value of Your Data Warehousing Investment**. TDWI Best Practices Report, 2007. Disponível em: <http://www.teradata.com/t/assets/0/206/308/35d4edc6-3408-4cbd-8946-a803e49d8288.pdf>. Acesso em: março 2014.

EDDY, D. M.; HOLLINGWORTH, W.; CARO, J. J.; TSEVAT, J.; MCDONALD, K. M.; WONG, J. B. Model Transparency and Validation: A Report of the ISPOR-SMDM Modeling Good Research Practices TaskForce-7. In: **International Society for Pharmacoeconomics and Outcomes Research (ISPOR)**. Pp. 843 – 850, 2012.

ELIAS, T. **Learning Analytics: Definitions, Processes and Potential**, 2011. Disponível em: [em: <http://learninganalytics.net/LearningAnalyticsDefinitionsProcessesPotential.pdf>](http://learninganalytics.net/LearningAnalyticsDefinitionsProcessesPotential.pdf). Acesso em: maio 2014.

FAYAD, M. E.; SCHIMIDT, D. C.; JONHSON, R. E. Application Frameworks. In: FAYAD, M. E.; SCHIMIDT, D. C.; JONHSON, R. E., **Building Application Frameworks: Object-Oriented Foundations of Framework Design**. John Wiley & Sons, 1999. p.3-27.

FELDMAN, J.; TUNG, R. **Using data-based inquiry and decision making to improve instruction**. ERS Spectrum 19(3), 10–19. 2001. Disponível em: [em: <http://www.naesp.org/sites/default/files/Student%20Achievement_blue.pdf>](http://www.naesp.org/sites/default/files/Student%20Achievement_blue.pdf) Acesso em: abril 2014.

FENG, M.; HEFFERNAN, N. Can We Get Better Assessment From A Tutoring System Compared to Traditional Paper Testing? Can We Have Our Cake (Better Assessment) and Eat It too (Student Learning During the Test)? In: BAKER, R.S.J.D., MERCERON, A., PAVLIK, P.I. JR. (EDS.), 2010, Pittsburgh. **Proceedings of the 3rd International Conference on Educational Data Mining**: Pittsburgh, PA, USA, 2010. p. 41-50.

FERGUSON, R. **The State Of Learning Analytics in 2012: A Review and Future Challenges**. Technical Report KMI-12-01. Knowledge Media Institute. The Open University, UK, 2012. Disponível em: [em: <http://oro.open.ac.uk/36374/1/IJTEL40501_Ferguson%20Jan%202013.pdf>](http://oro.open.ac.uk/36374/1/IJTEL40501_Ferguson%20Jan%202013.pdf). Acesso em: maio 2014.

FERREIRA, L. G. A., GLUZ, J. C., BARBOSA, J. VICARI, R. An Ubiquitous Content Recommendation Model for Groups of Learners. In: WORKSHOP MULTIAGENT SYSTEM BASED LEARNING ENVIRONMENTS, 2014, Honolulu. **Proceedings of Multiagent System Based Learning Environments**. Honolulu – Hawaii, USA. 2014.

FIGUEIREDO FILHO, D. B.; SILVA JUNIOR, J.; A. Desvendando os Mistérios do Coeficiente de Correlação de Pearson (r). **Revista Política Hoje**, Vol. 18, n. 1, 2009.

FRANCE, L.; HERAUD, J.-M.; MARTY, J.-C.; CARRON, T.; HEILI, J. Monitoring virtual classroom: visualization techniques to observe student activities in an e-learning system. In: INTERNATIONAL CONFERENCE ON ADVANCED LEARNING TECHNOLOGIES, 6., 2006. **Proceedings of the sixth international conference on advanced learning technologies**, IEEE, 2006. p. 716–720.

FRITZ, J. (2009). Using Course Activity Data to Raise Awareness of Underperforming College Students. In: T. Bastiaens et al. (Eds.). **Proceedings of World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education**. Chesapeake, VA: AACE, 2009 p. 2586-2589.

GAMMA, E.; HELM, R.; JOHNSON, R.; VISSIDES, J. **Design Patterns: Elements of Reusable Object-Oriented Software**. Hardcover, Addison-Wesley Professional; 395 p., 1994.

GANDON, F. Engineering an ontology for a multi-agents corporate memory system. In: SYMPOSIUM ON THE MANAGEMENT OF INDUSTRIAL AND CORPORATE KNOWLEDGE. **Proceedings of the International Symposium on the Management of Industrial and Corporate Knowledge**. 2001. p. 209-228.

GENNARI, J.; MUSEN, M. A.; FERGERSON, R. W.; GROSSO., W. E.; CRUBEZY, M.; ERIKSON, H.; NOY, N. F.; TU, S. W. The evolution of Pretégé: An environment for knowledge-based systems development. **International Journal of Human-Computer Studies**, v. 58, n. 1, p. 89-1213, 2003.

GIBSON, D.; ALDRICH, C.; PRENSKY, M. **Games and Simulation in Online Learning: Research and Development Frameworks**. Library of Congress Cataloging-in-Publication Data, 2007.

GILFUS ANALYTICS. **Learning Analytics**, 2011. Disponível em: <<http://www.gilfusacademicanalytics.com/capabilities/learning-analytics>>. Acesso em: maio 2014

GRELLER, W.; DRACHSLER, H. Translating Learning into Numbers: A Generic Framework for Learning Analytics. *Educational Technology & Society*, 15, 3, 42-57, 2012.

GOLDSTEIN, P. J. **Academic Analytics: The Use of Management Information and Technology in Higher Education**. Key Findings, Boulder, CO: EDUCAUSE Center for Applied Research, 2005, Disponível em: <http://net.educause.edu/ir/library/pdf/ecar_so/ers/ers0508/EKF0508.pdf>. Acesso em: janeiro 2014.

GOVAERTS, S.; VERBERT, K.; DUVAL, E.; PARDO, A. The student activity meter for awareness and self-reflection. In: THE ACM SIGCHI CONFERENCE ON HUMAN FACTORS IN COMPUTING SYSTEMS, AUSTIN, TX, USA, 5–10 May 2012. **Proceedings of the 2012 ACM annual conference extended abstracts on human factors in computing systems extended abstracts**, ACM, 2012. p. 869–884.

GROSS, G. A.; NAGI, R.; SAMBHOOS, K.; SCHLEGEL, D. R.; SHAPIRO, S. C.; TAUER, G. Towards Hard+Soft Data Fusion: Processing Architecture and Implementation for the Joint Fusion and Analysis of Hard and Soft Intelligence Data. In: INTERNATIONAL CONFERENCE ON INFORMATION FUSION, 15., 2012. **Proceedings of International Conference on Information Fusion**, 2012. p. 955-962.

GRUBER, T. A translation approach to portable ontology specification. **Knowledge Acquisition Journal**. 199–220, 1993.

GUAZZELLI, A. **Predicting the future**. Part 2: Predictive modeling techniques. Copyright IBM Corporation, 2012. Disponível em: <<http://www.ibm.com/developerworks/library/ba-predictive-analytics2/ba-predictive-analytics2-pdf.pdf>> Acesso em: junho 2014.

GIL, A. C. **Método e Técnicas de Pesquisa Social**. São Paulo: Editora Atlas, 1999.

GRESS, C. L. Z.; FIOR, M.; HADWIN, A. F.; WINNE, P. H. Measurement and assessment in computer-supported collaborative learning. **Computers in Human Behavior**, Vol. 26, No. 5, p. 806-814, 2010.

HALL, D.; LLINAS, J.; MCNEESE, M.; MULLEN, T. A framework for dynamic hard/soft fusion. In: Information Fusion, 11., 2008. **Proceedings of International Conference on Information Fusion**, Pennsylvania, Pittsburg, 2008. p. 1–8.

HAN, J.; KAMBER, M. **Data Mining: Concepts and Techniques**. San Francisco, CA: Elsevier. 2011.

HATTIE, J. **Visible Learning: A Synthesis of Over 800 Meta-Analyses Relating to Achievement**. Routledge; 392 p., 2008.

HENRICKSEN, K.; INDULSKA, J.; RAKOTONIRAINY, A. Modeling Context Information in Pervasive Computing Systems. In: INTERNATIONAL CONFERENCE ON PERVASIVE COMPUTING, 1., 2002, London. **Proceedings of the First International Conference on Pervasive Computing**, London, UK: Springer-Verlag,

2002. p.167-180.

HERVÁS, R.; BRAVO, J.; FONTECHA, J. A context model based on ontological languages: a proposal for information visualization. **Journal of Universal Computer Science**, v. 16, p. 1539–1555, 2010.

HIEMSTRA, R. **Self-Directed Learning**. In: Rothwell, W. and Sensening, K. (ed.). *The SourceBook for Self-Directed Learning*. Amherest, MA, p 9 – 19, 1999.

HOAREAU, C.; SATOH, I. Modeling and Processing Information for Context-Aware Computing: A Survey. In: *NEW GENERATION COMPUTING*, OHMSHA, **Ltd. and Springer**, p. 177-196, 2009.

HORROCKS, I.; PATEL-SCHNEIDER, P. F.; VAN HARMELEN, F. From SHIQ and RDF to OWL: The making of a web ontology language. **Journal of Web Semantics**, p. 7–26, 2003.

HWANG, G-J.; CHANG, H-F. A formative assessment-based mobile learning approach to improving the learning attitudes and achievements of students. **Computer & Education**. Vol. 56, pp 1023–1031, 2011.

IBM CORPORATION. **Real World Predictive Analytics: Putting Analysis into Action for Visible Results**. 2010a. Disponível em: <<http://public.dhe.ibm.com/common/ssi/ecm/en/ytw03112usen/YTW03112USEN.PDF>>. Acesso em: março 2014.

IBM CORPORATION. **Predictive Analytics: Defined**. About SPSS Inc., 2010b. <<http://www.spss.com.hk/corpinfo/predictive.htm>>. Acesso em: abril 2014.

INDULSKA, J.; ROBINSON, R.; RAKOTONIRAINY, A.; HENRICKSEN, K. Experiences in using cc/pp in context-aware systems, In: M.-S. Chen, P. K. Chrysanthis, M. Sloman, A. B. Zaslavsky (eds.), **Mobile Data Management**, vol. 2574 of *Lecture Notes in Computer Science*, Springer, p. 247-261, 2003.

JAFARI, A.; MCGEE, P.; CARMEAN, C. Managing courses, defining learning: What faculty, students, and administrators want. *EDUCAUSE Review*, 2006. Disponível em: <<https://net.educause.edu/ir/library/pdf/erm0643.pdf>>. Acesso em: junho 2014.

JENSEN, C.; LONSDALE, H.; WYNN, E.; CAO, J.; SLATER, M.; DIETTERICH, T. G. The life and times of files and information: a study of desktop provenance. In: *SIGCHI*

CONFERENCE ON HUMAN FACTORS IN COMPUTING SYSTEMS PAGES, 2010, New York. **Proceedings of the SIGCHI Conference on Human Factors in Computing Systems Pages**. ACM: New York, NY, USA, 2010. p. 767-776.

JOHNSON, L.; SMITH, R.; WILLIS, H.; LEVINE, A.; HAYWOOD, K. **The 2011 Horizon Report**. Austin, Texas: The New Media Consortium, 2011.

JOHNSON, L.; ADAMS, S.; CUMMINS, M. **NMC Horizon Report: 2012 K-12 Edition**. Austin, Texas: The New Media Consortium, 2012.

JOHNSON, L.; ADAMS, S.; CUMMINS, M.; ESTRADA, V.; FREEMAN, A.; Ludgate, H. **The NMC Horizon Report: 2013 Higher Education Edition**. Austin, Texas: The New Media Consortium, 2013.

KAENAMPORN PAN, M.; O'NEILL, E. History as part of context. In: UK-UBINET WORKSHOP, 3., 2005. **Proceedings of the 3rd Uk-UbiNet Workshop**. University of Bath, UK, 2005, p. 9-11.

KHALEGHI, B.; KHAMIS, A.; KARRAY, F. O., RAZAVI, S. N. Multisensor data fusion: A review of the state-of-the-art. **Information Fusion**, v. 14, 28-44, 2013.

KARSIRI, E.; MYCROFT, A. Applying Bayesian Networks to Sensor-Driven Systems. In: INTERNATIONAL SYMPOSIUM ON WEARABLE COMPUTERS, 10., 2003. **Proceedings of the 10th IEEE International Symposium on Wearable Computers (ISWC'03)**, IEEE, 2003. p. 149-150.

KOSBA, E.; DIMITROVA, V.; BOYLE, R. Using student and group models to support teachers in web-based distance education. In: INTERNATIONAL CONFERENCE ON USER MODELING, 10., 2005, Edinburgh. **Proceedings of the 10th International Conference on User Modeling**. Edinburgh, UK, 2005. p 124–133.

KIRON, D.; SHOCKLEY, R.; KRUSCHWITZ, N.; FINCH, G.; HAYDOCK, M. **Analytics: The widening divide**. MIT Sloan Management Review, 2011. Disponível em: <<http://sloanreview.mit.edu/reports/analytics-advantage/>>. Acesso em: junho 2014.

KIRKWOOD, A.; PRICE, L. Technology-enhanced learning and teaching in higher education: what is 'enhanced' and how do we know? A critical literature review. **Learning, Media and Technology**, v. 39, n.1, p. 6–36, 2014.

KLASNJA-MILICEVIC, A., VESIN, B., IVANOVIC, M., e BUDIMAC, Z. E-Learning personalization based on hybrid recommendation strategy and learning style identification. **Computers & Education**, 56, 885–899, 2011.

LACHAT, M.A. Data-driven high school reform: The breaking ranks model. **Center for Resource Management**, Hampton, 2002.

LAKATOS, E.; MARCONI, M.A. **Metodologia do trabalho científico**. São Paulo: Atlas, 1989.

LAUDY, C.; GOUJON, B. Soft data analysis within a decision support system. In: INTERNATIONAL CONFERENCE ON INFORMATION FUSION, Seattle, 2009. **Proceedings of International Conference on Information Fusion (ICIF'09)**. Seattle, WA, 2009. p. 1889–1896.

LEONY, D.; PARDO, A.; de la FUENTE VALENTÍN, L.; SÁNCHEZ DE CASTRO, D.; DELGADO KLOOS, C. GLASS: a learning analytics visualization tool. In: INTERNATIONAL CONFERENCE ON LEARNING ANALYTICS AND KNOWLEDGE LAK '12, 2., 2012, New York. **Proceedings of the 2nd International Conference on Learning Analytics and Knowledge LAK '12**. ACM: New York, NY, USA, 2012. p. 162-163.

LIU, B. Web Data Mining. Exploring Hyperlinks, Contents, and Usage Data. **Berlin Heidelberg: Springer**, 2006.

LOKE, S. W. Representing and reasoning with situations for context-aware pervasive computing: a logic programming perspective. **The Knowledge Engineering Review**, 19, 3, pp. 213-233, 2004.

LUO, R. C.; Kay, M. G. Data fusion and sensor integration: State-of the-art 1990s. In Abidi, M. A. and Gonzalez, R. C., editors, **Data Fusion in Robotics and Machine Intelligence**, chapter 3, pages 7–135. **Academic Press, Inc.**, San Diego, CA, USA, 1992.

MACFADYEN, L. P.; DAWSON, S. Mining LMS data to develop an “early warning system” for educators: A proof of concept. **Computers & Education**, Vol. 54, No. 2, p. 588-599, 2010.

MACFADYEN, L. P.; DAWSON, S. Numbers Are Not Enough. Why e-Learning Analytics Failed to Inform an Institutional Strategic Planning. **Educational Technology & Society**, 15 (3), p. 149–163, 2012.

MAELE, J. V., BATEN, L., BEAVEN, A., RAJAGOPAL, K. e-Assessment for Learning: Gaining Insight in Language Learning with Online Assessment Environments. *Computer-Assisted Foreign Language Teaching and Learning: Technological Advances*, IGI Global, p. 245-261, 2013.

MANDINACH, E. B.; HONEY, M. **Data driven school improvement: Linking data and learning**. New York: Teachers College Press. 304 p., 2008.

MANN, P. S. **Introductory Statistics**, 2nd Edition, Wiley, 2005.

MANOUSELIS, N.; DRACHSLER, F.; VUORIKARI, R.; HUMMEL, H.; KOPER, R. Recommender Systems in Technology Enhanced Learning. **Recommender Systems Handbook**: L.R.P.B. KANTOR, F. RICCI, e B. SHAPIRA, eds., Springer, 2011. p. 387-415.

MANYIKA, J.; CHUI, M.; BROWN, B.; BUGHIN, J.; DOBBS, R.; ROXBURGH, C.; BYERS, A. H. **Big data: The next frontier for innovation, competition, and productivity**. McKinsey Global Institute, 2011. Disponível em: <http://www.mckinsey.com/insights/mgi/research/technology_and_innovation/big_data_the_next_frontier_for_innovation>. Acesso em: janeiro 2014.

MARÔCO, J. **Análise Estatística com o SPSS Statistics**. Report Number Análise e Gestão da Informação. LDA 5ª edição, 2011.

MARSH, J. A.; PANE, J. F.; HAMILTON, L. S. **Making Sense of Data-Driven Decision Making in Education**. Rand Corporation, 2006. <http://www.rand.org/content/dam/rand/pubs/occasional_papers/2006/RAND_OP170.pdf>. Acesso maio 2014.

MARTINS, R. A. Abordagens quantitativa e qualitativa. In: MIGUEL, Paulo A.M (Org.). **Metodologia de pesquisa em engenharia de produção e gestão de operações**. Rio de Janeiro: Elsevier, 2010. p. 45-61.

MAZZA, R.; DIMITROVA, V. CourseVis: A graphical student monitoring tool for supporting instructors in Web-based distance courses. **International Journal of Human-Computer Studies**, 65(2), p. 125–139, 2007.

MERCERON, A., YACEF, K. TADA-Ed for Educational Data Mining. **Interactive Multimedia Electronic Journal of Computer-Enhanced Learning**, v. 7 n. 1, 2005.

MIGUEL, Paulo A. C. Adoção do estudo de caso na engenharia de produção. In: MIGUEL, Paulo A.M. (Org.). **Metodologia de pesquisa em engenharia de produção e gestão de operações**. Rio de Janeiro: Elsevier, 2010. p.129-143.

MILES, M. B.; HUBERMAN, A. M. **Qualitative Data Analysis: an expanded sourcebook**. 2ª ed. London: Sage Publications, 1994. 338 p.

MOTT, J. **Envisioning the Post-LMS Era: The Open Learning Network**. EDUCAUSE, 2010. Disponível em: <<http://www.educause.edu/ero/article/envisioning-post-lms-era-open-learning-network>>. Acesso em: maio 2014.

MUÑOZ-MERINO P. J.; PARDO, A.; KLOOS, C. D.; MUÑOZ-ORGANERO, M.; WOLPERS, M. NIEMANN, K.; FRIEDRICH, M. CAM in the semantic web world. In: INTERNATIONAL CONFERENCE ON SEMANTIC SYSTEMS. I-SEMANTICS, 6., 2010, New York. **Proceedings of the 6th International Conference on Semantic Systems. I-SEMANTICS '10**. New York, NY, USA: ACM, 2010.

NABETH, T.; ANGEHRN, A.; RAZMERITA, L.; RODA, C. InCA: a cognitive multi-agents architecture for designing intelligent & adaptive learning systems. **Computer Science and Information Systems**, v. 2, n. 2, p. 99-114, 2005.

NAKAHARA, J.; HISAMATSU, S.; YAEGASHI, K.; YAMAUCHI, Y. iTree: does the mobile phone encourage learners to be more involved in collaborative learning? In: CONFERENCE ON COMPUTER SUPPORT FOR COLLABORATIVE LEARNING: (CSCL), 2005. **Proceedings of the 2005 conference on computer support for collaborative learning: learning 2005: the next 10 years!**. International Society of the Learning Sciences, 2005. p 470–478.

NATSU, J. **Advanced Analytics: Helping Educators Approach the Ideal**,” eSN **Special Report**, eSchool News, 2010, 17–23, Disponível em: <http://www-935.ibm.com/services/ie/gbs/pdf/Smarter_Education_Advanced_Analytics.pdf>. Acesso em: maio 2014.

NEVO, B. Face validity revisited. **Journal of Educational Measurement**, v. 22, n. 4, p. 287–293, 1985.

NORRIS, D.; BAER, L.; OFFERMAN, M. **A National Agenda for Action Analytics**. In: NATIONAL SYMPOSIUM ON ACTION ANALYTICS, St. Paul, 2009. **Proceedings of National Symposium on Action Analytics**. St. Paul, 2009.

NYCE, C. **Predictive Analytics White Paper**, American Institute for Chartered Property Casualty Underwriters/Insurance Institute of America, 2007, p. 1. Disponível em: <<http://www.theinstitutes.org/doc/predictivemodelingwhitepaper.pdf>>. Acesso em: julho 2014.

OLIVER, N.; SMITH, G.; THAKKAR, C.; SURENDRAN, A. C. SWISH: semantic analysis of window titles and switching history. In: INTERNATIONAL CONFERENCE ON INTELLIGENT USER INTERFACES, 10., 2006. **Proceedings of International Conference on Intelligent User Interfaces**. 2006. p. 194–201.

OLIVIER, B.; LIBER, O. Lifelong Learning: the need for portable personal learning environments and supporting interoperability standards. Bristol. **The JISC Centre for Educational Technology Interoperability Standards, Bolton Institute**, 2001.

PARASKAKIS, I. Ambient Learning: a new paradigm for e-learning. In: INTERNATIONAL CONFERENCE ON MULTIMEDIA AND INFORMATION & COMMUNICATION TECHNOLOGIES IN EDUCATION (m-ICTE), 3., 2005. **Proceedings of the 3rd International Conference on Multimedia and Information & Communication Technologies in Education**. Spain, 2005. p. 26-30.

PARDINI, P. Data, well done. **Journal of Staff Development**, v. 21, n. 1, p. 12–18, 2000.

PATTERSON, C.; LUNYK-CHILD, O. CROOKS, D. A New Perspective on Competencies for Self-Directed Learning. **NCBI: PubMed**. Ontario: Hamilton. v. 41, n. 1, 2002.

PEARL, J., 1988. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. **Morgan Kaufmann Publishers**, San Mateo, USA.

PERSSON, M. Ericsson Group. **The tools of education – soon at a museum near you**. Disponível em: <http://www.ericsson.com/res/thecompany/docs/publications/business-review/2012/issue1/the_tools_of_education.pdf>. Acesso em: julho 2014.

PIATETSKY-SHAPIRO, G. Guest editor's introduction: Knowledge discovery in databases – from research to applications. **Journal of Intelligent Information Systems**, v. 4, n. 1, p. 5-6, 1995.

PINTRICH, P. R.. *The role of goal orientation in self-regulated learning*. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), **Handbook of self-regulation**. San Diego, CA, p. 451–502, 2000.

PLYMALE, W. O. **Pervasive computing goes to school**. *EDUCAUSE Review*, Vol. 40, No.1, 2005, p. 60–61. Disponível em: <<https://net.educause.edu/ir/library/pdf/ERM0519.pdf>>. Acesso em: julho 2014.

PODGORELEC, V.; KUCHAR, S. Taking Advantage of Education Data: Advanced Data Analysis and Reporting in Virtual Learning Environments. **Electronics and Electrical Engineering**, v. 114, n. 8, p. 111-116, 2011.

PRAVIA, M.; PRASANTH, R.; ARAMBEL, P.; SIDNER, C.; CHONG, C.-Y. Generation of a fundamental data set for hard/soft information fusion. In: **FORMATION FUSION**, 11., 2008. **Proceedings of International Conference on Formation Fusion, 2008**. p. 1–8.

PROTHEROE, N. Improving teaching and learning with data-based decisions: Asking the right questions and acting on the answers. **ERS Spectrum**, v. 19, n. 3, p. 4–9, 2001.

RANGANATHAN, A.; McGRATH, R. E.; CAMPBELL, R. H.; MICKUNAS, M. D. Use of ontologies in a pervasive computing environment. **The Knowledge Engineering Review**, 18, 3, pp. 209-220, 2003.

RAVISHANKER, G. **Doing Academic Analytics Right: Intelligent Answers to Simple Questions**. Research Bulletin 2. Boulder, CO: EDUCAUSE Center for Analysis and Research, 2011. Disponível em: <from <https://net.educause.edu/ir/library/pdf/ERB1102.pdf>>. Acesso em: maio 2014.

REFSGAARD, J. C.; HENRIKSEN, H. J. MODELLING GUIDELINES – A THEORETICAL FRAMEWORK. In: **HarmoniQuA – State-of-the-art Report on QA guidelines**, 2002.

ROGERS, Y.; PRICE, S.; RANDELL, C.; FRASER, D. S.; WEAL, M.; FITZPATRICK, G. Ubi-learning integrating indoor and outdoor learning experiences. **Communications of the ACM**, v. 48, n. 1, p. 55–59, 2005.

ROMERO, C.; VENTURA, S. Educational Data Mining: a Survey from 1995 to 2005. **Expert Systems with Applications**, v. 33, n.1, p. 135-146, 2007.

ROSALES, G. C. M.; ARAÚJO, R. B.; OTSUKA, J. L.; ROCHA, R. V. Using Logical Sensors Network to the Accurate Monitoring of the Learning Process in Distance Education Courses. In: INTERNATIONAL CONFERENCE ON ADVANCED LEARNING TECHNOLOGIES, Athens, 2011. **Proceedings of the IEEE International Conference on Advanced Learning Technologies**. Athens, Georgia, 2011.

ROSALES, G.C.M.; ARAUJO, R. B.; OTSUKA, J. L.; Logical Sensor Networks to the monitoring of learning events. In: INTERNAL WORKSHOP OF INSTITUTO NACIONAL DE CIÊNCIA E TECNOLOGIA EM SISTEMAS EMBARCADOS CRÍTICOS (INCT-SEC 2011), 7-9 December, Águas de Lindóia, Brazil, 2011. **Proceedings of INCT-SEC**. São Paulo: USP, 2011.

ROSALES, G.C.M.; ARAUJO, R. B.; OTSUKA, J. L.; BEDER, D.; Data Collection Framework to the Learning Monitoring Using Logical Sensor Network, In: INTERNATIONAL CONFERENCE ON ADVANCED LEARNING TECHNOLOGIES, 12., 2012. **Proceedings of 12th International Conference on Advanced Learning Technologies**, IEEE, 2012. p.194-196.

RYAN, N. **ConteXtML: Exchanging contextual information between a mobile client and the FieldNote server**. Online Version, August 1999. Disponível em: <<http://www.cs.kent.ac.uk/projects/mobicomp/fnc/ConteXtML.html>>. Acesso em: maio 2011.

RYANN, E. K. **Field Guide to Learning Management Systems**, ASTD Learning Circuits, 2009.

RYU, H. Is mobile learning a necessary evil?: The goodness-of-fit of mobile learning situations. In: PERVASIVE LEARNING: DESIGN CHALLENGES AND REQUIREMENTS WORKSHOP, Toronto, 2007. **Proceedings of Pervasive Learning: Design Challenges and Requirements workshop held at PERVASIVE**. Toronto, Canada, 2007. p. 27-33.

SANTOS, J. L. O.; VERBERT, K.; GOVAERTS, S.; DUVAL, E. Addressing learner issues with StepUp!: an Evaluation. In: INTERNATIONAL CONFERENCE ON LEARNING ANALYTICS AND KNOWLEDGE (LAK), Leuven, 2013. **Proceedings of the international conference on learning analytics and knowledge**. ACM, Leuven, Belgium, 2013. p 143–152

SARGENT, R. G. VERIFICATION AND VALIDATION OF SIMULATION MODELS. IN: WINTER SIMULATION CONFERENCE, 2011. **Proceedings of the Winter Simulation Conference**. 2011.

SCHEUER, O.; ZINN, C. How did the e-learning session go? The student inspector. In: CONFERENCE ON ARTIFICIAL INTELLIGENCE IN EDUCATION, AMSTERDAM, 2007. **Proceedings of the Conference on Artificial Intelligence in Education: building technology rich learning contexts that work**. IOS Press, Amsterdam, 2007. p. 487–494

SCHAFFHAUSER, D. **Essentials of Prescriptive Analytics for Student Success**. Campus Technology, 2014. Disponível em: <<http://campustechnology.com/Articles/2014/06/25/12-Essentials-of-Prescriptive-Analytics-for-Student-Success.aspx?Page=1>> Acesso em: julho 2014.

SCHELLENS, T.; VALCKE, M. Fostering knowledge construction in university students through asynchronous discussion groups. **Computers & Education**, v. 46, n. 4, p. 349–370, 2006.

SCHILIT, B. N.; THEIMER, M. M. Disseminating active map information to mobile hosts. In: **IEEE Network**, v. 8, n. 5, p. 22-32, 1994.

SCHILIT, B.; ADAMS, N.; WANT, R. Context-Aware Computing Applications. In: WORKSHOP ON MOBILE COMPUTING SYSTEMS AND APPLICATIONS, 1994. **Proceedings of IEEE Workshop on Mobile Computing Systems and Applications**. 1994.

SCHMIDT, A.; WINTERHALTER, C. User Context Aware Delivery of E-Learning Material: Approach and Architecture. **Journal of Universal Computer Science**, v. 10, n. 1, p. 28-36, 2004.

SEABORNE, A. **From RDQL to SPARQL**. 2006. Disponível em: <<http://seaborne.blogspot.com/2006/04/from-rdql-to-sparql.html>>. Acesso em: agosto 2011.

SECOLSKY, C. On the Direct Measurement of Face Validity: A Comment on Nevo. In: **Journal of Educational Measurement**, v. 24, n. 1, Spring, p. 82-83, 1987.

SHUM, B. S.; FERGUSON, R. Social Learning Analytics. **Educational Technology & Society**, v. 15, n. 3, p. 3–26, 2012.

SIEMENS, G. **What are Learning Analytics?**. 2010. Disponível em: <<http://www.elearnspace.org/blog/2010/08/25/what-are-learning-analytics/>>. Acesso em: maio 2014.

SIEMENS, G. **Learning Analytics: the emergence of a discipline**. American Behavioral Scientist, 2013. Disponível em: <<http://abs.sagepub.com/content/early/2013/08/20/0002764213498851>>. Acesso em: 27 maio 2014.

SIEMENS, G., e LONG, P. **Penetrating the Fog: Analytics in Learning and Education**. EDUCAUSE Review, vol. 46, no. 5, 2011. Disponível em: <<https://net.educause.edu/ir/library/pdf/ERM1151.pdf>>. Acesso em: junho 2014.

SIEMENS, G.; BAKER, R. S. J. D. Learning Analytics and Educational Data Mining: Towards Communication and Collaboration. In: INTERNATIONAL CONFERENCE ON LEARNING ANALYTICS AND KNOWLEDGE, New York, 2012. **Proceedings of the 2nd International Conference on Learning Analytics and Knowledge**. New York, 2012, p. 252-254.

SMITH, V. C.; LANGE, A.; HUSTON, E. D. R. Predictive Modeling to Forecast Student Outcomes and Drive Effective Interventions in Online Community College Courses. **Journal of Asynchronous Learning Networks**, v. 16, n.3, p. 51-61, 2012.

SOHN, T. Context-aware computing support for the educationally disadvantaged. In: SYMPOSIUM ON HUMAN CENTRIC COMPUTING LANGUAGES AND ENVIRONMENTS (HCC). **Proceedings of the IEEE Symposium on Human Centric Computing Languages and Environments**, IEEE, 2003. p. 293-294.

STEVENSON, W. J. **Estatística aplicada à administração**. São Paulo: Harbra, 2001.

STRANG, T.; LINNHOF-POPIEN, C. A. Context Modeling Survey. In: WORKSHOP ON ADVANCED CONTEXT MODELLING, REASONING AND MANAGEMENT, UbiComp, Nottingham, 2004. **Proceedings of The Sixth International Conference on Ubiquitous Computing**. Nottingham, England, 2004.

STRIJBOS, J-W. Assessment of (Computer-Supported) Collaborative Learning. **IEEE Transactions on Learning Technologies**, v. 4, n. 1, 2011.

SYMONDS, K. W. (2003). **After the test: How schools are using data to close the achievement gap**. San Francisco: Bay Area School Reform Collaborative. Disponível em: < <http://racialequitytools.org/resourcefiles/symonds.pdf> >. Acesso em: abril 2014.

SYVÄNEN, A.; BEALE, R.; SHARPLES, M.; AHONEN, M.; LONSDALE, P. Supporting Pervasive Learning Environments: Adaptability and Context Awareness in Mobile Learning. In: INTERNATIONAL WORKSHOP ON WIRELESS AND MOBILE TECHNOLOGIES IN EDUCATION, Tokushima, 2005. **Proceedings of the 3rd IEEE International Workshop on Wireless and Mobile Technologies in Education**, Japan, 2005. p. 21-28.

THOMAS, S. Pervasive learning games: Explorations of hybrid educational gamescapes. **Simulation & Gaming**, v. 37 n. 1, p. 41-55, 2006.

VAN HARMELEN, M. Personal Learning Environments. In: INTERNATIONAL CONFERENCE ON ADVANCED LEARNING TECHNOLOGIES, 6., 2006. **Proceedings of International Conference on Advanced Learning Technologies**, IEEE, 2006.

VAN HARMELEN, H. Design trajectories: four experiments in PLE implementation. **Interactive Learning Environments**, v. 16, n. 1, p. 36 – 46, 2008.

VAN BARNEVELD, A., ARNOLD, K.E., e CAMPBELL, J.P. (2012). **Analytics in higher education: Establishing a common language. Educause Learning Initiative**. ELI Paper 1:2012, 1–11. Disponível em: < <http://net.educause.edu/ir/library/pdf/ELI3026.pdf> >. Acesso em: junho 2014.

VERBERT, K.; DRACHSLER, H.; MANOUSELIS, N.; WOLPERS, M.; VUORIKARI, R.; DUVAL, E. Data Set-Driven Research for Improving Recommender Systems for Learning. In: INTERNATIONAL CONFERENCE ON LEARNING ANALYTICS AND KNOWLEDGE (LAK), 1., 2011. **Proceedings of the First International Conference Learning Analytics and Knowledge**, 2011. p. 44-53.

VERBERT, K., MANOUSELIS, N., DRACHSLER, H., e DUVAL, E. Dataset-driven research to support learning and knowledge analytics. **Educational Technology and Society**, v. 15, n. 3, p. 133-148, 2012.

XIA, F.; ASABERE, N. Y.; AHMED, A. M.; LI, J.; KONG, X. Mobile Multimedia Recommendation in Smart Communities: A Survey. **IEEE Access Practical Innovations: Open Solutions**, v. 1, p. 606-624, 2013.

WANG, J.; ZENG, C.; HE, C.; HONG, L.; ZHOU, L.; WONG, R. K.; TIAN, J. Context-aware role mining for mobile service recommendation. In: ANNUAL ACM SYMPOSIUM ON APPLIED COMPUTING, 27., New York, 2012. **Proceedings of the 27th Annual ACM Symposium on Applied Computing.** New York, 2012. p.173-178.

WICKRAMARATHNE, T. L. **An Analytical Framework for Soft and Hard Data Fusion:** A Dempster-Shafer Belief Theoretic Approach. A dissertation of the University of Miami. Coral Gables, Florida, August 2012.

WILSON, B., SWITZER, S., PARRISH, P. Transformative learning experiences: How do we get students deeply engaged for lasting change? In: SELECTED RESEARCH AND DEVELOPMENT PRESENTATIONS, Washington., 2007. **Proceedings of selected research and development presentations.** Washington D. C: Association for Educational Communications and Technology, 2007.

WOLF, C. iWeaver: towards “learning style” – based e-Learning. In: AUSTRALASIAN COMPUTING EDUCATION CONFERENCE (ACE), 2003. **Proceedings of the Australasian Computing Education Conference - Conferences in Research and Practice in Information Technology.** v. 20, 2003.

WOLPERS, M.; NAJJAR, K.; VERBERT, K.; DUVAL, E. Tracking Actual Usage: the Attention Metadata Approach. **Educational Technology e Society**, v. 10, n. 3, p. 106-121, 2007.

YERVA, S. R.; JEUNG, H.; ABERER, K. Cloud based Social and Sensor Data Fusion. In: INTERNATIONAL CONFERENCE ON INFORMATION FUSION, 2012. **Proceedings of the International Conference on Information Fusion**, 2012.

YERVA, S. R.; SALTARIN, J.; JEUNG, H.; ABERER, K. Social and Sensor Data Fusion in the Cloud. In: INTERNATIONAL CONFERENCE ON MOBILE DATA MANAGEMENT, 13., 2012. **Proceedings of the 13th International Conference on Mobile Data Management**, IEEE, 2012.

YIN, C.; OGATA, H.; TABATA, Y.; YANO, Y. JAPELAS2: Supporting the Acquisition of Japanese Polite Expressions in Context-Aware Ubiquitous Learning, Mobile and Ubiquitous Technologies for Language Learning. **International Journal of Mobile Learning and Organization**, v. 4, n. 2, p.214-234, 2010.

YIN, R. K. **Estudo de Caso: planejamento e método.** 2. ed. São Paulo: Bookman, 2001.

YIN, R. K. **Case study research: design and methods**. 4. ed. California: Sage, 2009.

YU, Z.; ZHOU, X.; SHU, L. Towards a semantic infrastructure for context-aware e-learning. **Springer Science and Business Media**, 2009.

YU, Z.; ZHOU, X.; SHU, L. Towards a semantic infrastructure for context-aware e-learning. **Journal Multimedia Tools and Applications**, v. 47, n. 1, p. 71-96, 2010.

ZHAO, S.; WONGTHONGTHAN, P.; DILLON, T.; CHANG, E. Mapping relational data model to OWL ontology: Knowledge conceptualization in OWL. In: INTERNATIONAL CONFERENCE ON DIGITAL ECOSYSTEMS AND TECHNOLOGIES, 2., 2008. **Proceedings of 2nd IEEE International Conference on Digital Ecosystems and Technologies**, IEEE, 2008. p. 497-501.

Apêndice A

TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

Você está sendo convidado(a) para participar, como voluntário, em uma pesquisa. Após ser esclarecido(a) sobre as informações a seguir e de participar da apresentação e simulação do sistema a ser utilizado para a pesquisa, no caso de aceitar fazer parte do estudo, assine ao final deste documento, que está em duas vias. Uma delas é sua e a outra é da pesquisadora responsável.

Desde logo fica garantido o sigilo das informações. Em caso de recusa você não será penalizado(a) de forma alguma.

INFORMAÇÕES SOBRE A PESQUISA

Título do Projeto: MODELO E FRAMEWORK PARA O DESENVOLVIMENTO DE FERRAMENTAS ANALÍTICAS DE APOIO AO ENSINO, APRENDIZAGEM E GESTÃO EDUCACIONAL

Pesquisadora Responsável: Gislaine Cristina Micheloti Rosales

Telefone para contato: (16) 3332 6406 e (16) 98142 5920

O objetivo deste estudo é avaliar a percepção dos alunos sobre a utilização da ferramenta ViTracker, uma aplicação de Rastreamento e Visualização de Dados e Recomendações. Trata-se de um estudo exploratório, que para ser realizado é necessário que os participantes da pesquisa instalem em seu dispositivo pessoal, notebook ou PC, um conjunto de sensores lógicos capazes de capturar dados

referentes às interações do usuário com o ambiente digital, incluindo interações realizadas dentro e fora do browser.

Para coleta de dados, o estudante deverá seguir o protocolo de instalação do software, que encontra-se ao final deste documento. A princípio, os dados a serem coletados são os seguintes: cliques realizados em formulários; localização do usuário no Moodle; login no Moodle; recebimento e perda de foco (seleção de objetos); uso de barras de rolagem; redimensionamento da janela; pressionamento de teclas; materiais teóricos acessados no Moodle; atividades acessadas no Moodle; uso de portas USB; aplicações em execução no dispositivo do usuário; ferramentas de comunicação instantânea sendo usadas; uso da câmera; uso do microfone. No entanto, o aluno poderá instalar apenas os sensores cuja coleta de dados for de seu interesse.

Caso o aluno opte por instalar todos os sensores listados acima, poderá pausar a coleta de dados realizada por qualquer um dos sensores a qualquer momento durante a utilização do sistema. Para maiores esclarecimentos, será realizada uma apresentação do sistema ViTracker, onde também serão feitos esclarecimentos sobre a coleta de dados a partir do dispositivo do aluno.

Apenas após ler este documento e participar da apresentação para esclarecimentos que será feita pela pesquisadora, o aluno poderá decidir sobre assinar ou não este termo.

Gislaine Cristina Micheloti Rosales

CONSENTIMENTO

Eu, _____, _____,
abaixo assinado, concordo em participar do estudo
_____, como sujeito. Fui
devidamente informado e esclarecido pelo pesquisador
_____ sobre a pesquisa, os procedimentos nela
envolvidos, assim como os possíveis riscos e benefícios decorrentes de minha
participação. Foi-me garantido o sigilo das informações e que posso retirar meu
consentimento a qualquer momento, sem que isto leve à qualquer penalidade ou
interrupção de meu acompanhamento/ assistência/tratamento.

Local e data _____/_____/_____/_____

Nome: _____

Assinatura do sujeito ou responsável: _____

PROTOCOLO DE INSTALAÇÃO DO VITRACKER – MODO USUÁRIO

1) Instalação

- Sensores: Para a instalação dos sensores na máquina do usuário, basta copiar o arquivo "Sensores" para qualquer diretório e extraí-lo.

- Plugin: Para instalar o *plugin* no *Google Chrome*, abra-o e clique em menu > ferramentas > extensões. Ative o modo de desenvolvedor e arraste o arquivo "LMS2Plugin.crx" (que se encontra no diretório *Plugin*) para este menu. Para instruções mais detalhadas acesse: <http://www.tecmundo.com.br/tutorial/26055-google-chrome-como-instalar-extensoes-que-nao-estao-na-web-store.htm>

2) Execução

Para executar os sensores execute o arquivo "start.bat" dentro da pasta de instalação.

Configure corretamente o arquivo "lms2_user_config.xml" (também dentro da pasta de instalação) com o endereço do servidor que rodará o framework. Este endereço deve ser configurado no campo <jade_server_address>. Se for necessário, em caso de conflito, é possível alterar a porta do mini webserver Jetty, no campo <jetty_port> do arquivo de configuração.

Para interromper a execução, basta fechar a janela ou pressionar CTRL + C.

Apêndice B

QUESTIONÁRIO - ALUNOS

1ª ETAPA: QUESTÕES PARA SEREM RESPONDIDAS ANTES DO USO DO SISTEMA

1. Quais ferramentas do Moodle você usa para ajudar na organização de seus estudos?

- agenda/calendário
- cronograma
- bloco de anotações
- outras. Indique os nomes: _____
- não uso ferramentas do Moodle para ajudar na organização dos meus estudos

2. Quais aplicativos *online* você usa para te ajudar na organização dos estudos?

- Evernote
- Google Calender
- My Study Life
- Remember the Milk
- outros. Indique os nomes: _____
- não uso aplicativos online para ajudar na organização dos meus estudos

3. Imagine que o Moodle seja inteligente. Ele sabe o quê e como você deve estudar, o tempo de dedicação necessário para cada estudo e até mesmo prevê seu desempenho final na disciplina. Considere que o Moodle tenha todas as funcionalidades que você imaginar e responda:

Quais funcionalidades você gostaria que ele tivesse?

Exemplos: uma agenda inteligente que lembre seus deadlines por sms; um assistente que te ajuda a decidir o que estudar para a prova.

2ª ETAPA: QUESTÕES PARA SEREM RESPONDIDAS APÓS USO DO SISTEMA (AO FINAL DAS DISCIPLINAS)

1. Informações Pessoais (4 questões)

1.1 Idade _____

1.2 Gênero: _____

1.3 Qual seu nível de escolaridade?

Possuo curso superior completo Possuo curso superior incompleto

Ensino médio completo

Cursando 1º ano do curso Superior

Cursando 2º ano do curso Superior

Cursando 3º ano do curso Superior

1.4 Há quanto tempo você usa o Moodle?

Menos de 6 meses

Menos de 1 ano

Menos de 3 anos

Menos de 6 anos

Mais de 6 anos

2. Em relação à ajuda oferecida pelo sistema (8 questões)

2.1 Em que medida o sistema ajudou você a decidir seu ritmo de estudo (tempo de dedicação a cada material e atividade)?

	<input type="radio"/>					
Atrapalhou Muito	-2	-1	0	1	2	Ajudou Muito

2.2 Comente sua resposta:

2.3 Em que medida o sistema ajudou você a gerenciar seus estudos (decidir o que executar primeiro)?

	<input type="radio"/>					
Atrapalhou Muito	-2	-1	0	1	2	Ajudou muito

2.4 Comente sua resposta:

2.5 Em que medida o sistema ajudou você a monitorar seus estudos (identificar quais materiais já estudados e outros que ainda precisavam ser estudados, atividades já feitas e outras atividades pendentes em um período de tempo)?

	<input type="radio"/>					
Atrapalhou muito	-2	-1	0	1	2	Ajudou muito

2.6 Comente sua resposta:

2.7 Em que medida o sistema ajudou você a refletir sobre sua aprendizagem a fim de atingir suas próprias metas estabelecidas para sua aprendizagem?

	<input type="radio"/>					
Atrapalhou muito	-2	-1	0	1	2	Ajudou muito

2.8 Comente sua resposta:

3. Em relação às recomendações do sistema (7 questões)

3.1 Em que medida você mudou sua atitude/ação no ambiente depois de uma recomendação do sistema ou depois de visualizar informações sobre seu desempenho em relação aos demais colegas?

	<input type="radio"/>					
Não mudei	0	1	2	3	4	Mudei muito

3.2 Comente sua resposta:

3.3 Em que medida você seguiu as recomendações feitas pelo sistema?

	<input type="radio"/>					
Não segui	0	1	2	3	4	Segui Completamente

3.4 Comente sua resposta:

3.5 Em que medida as recomendações foram úteis para você (conforme exemplo da fig. 1)?

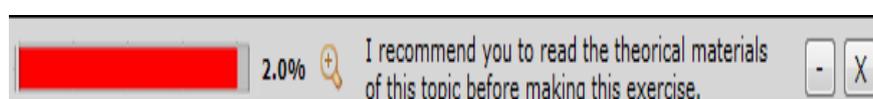


Fig. 1 – Recomendação em questionário

	<input type="radio"/>					
Extremamente Inúteis	-2	-1	0	1	2	Extremamente úteis

3.6 Comente sua resposta, indicando como as recomendações poderiam ser mais úteis (*ignore a pergunta se você respondeu extremamente úteis*)

3.7 Quais outras recomendações o sistema não trouxe e que seriam úteis para você ou para os demais usuários?

4. Em relação às informações disponibilizadas pelo sistema (6 questões)

4.1 Quão úteis foram as informações sumarizadas no gráfico de barra sobre seu progresso acadêmico (conforme fig. 2)?



Fig. 2 – Barra de progresso

	<input type="radio"/>					
Extremamente Inúteis	-2	-1	0	1	2	Extremamente úteis

4.2 O que pode ser alterado/acrescentado para que essa informação seja mais útil para você? (ignore a pergunta se você respondeu *Extremamente úteis*)

4.3 Quão úteis foram as informações de seu progresso acadêmico detalhado (conforme fig. 3)?

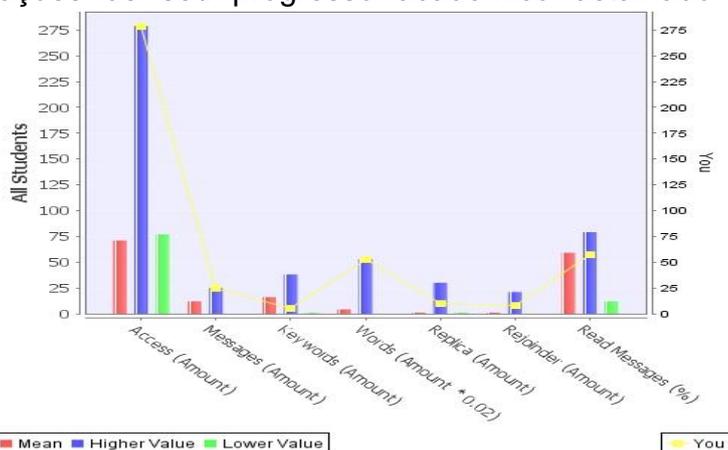


Fig. 3 – Gráfico com desempenho do aluno em cada indicador de fórum

Extremamente Inúteis ○ ○ ○ ○ ○ Extremamente úteis
 -2 -1 0 1 2

4.4 O que poderia ser alterado/acrescentado para que essa informação seja mais útil para você? (ignore a pergunta se você respondeu *Extremamente úteis*)

4.5 Quão úteis foram as informações sobre seu nível de atuação nas atividades (conforme exemplo da fig. 4)?

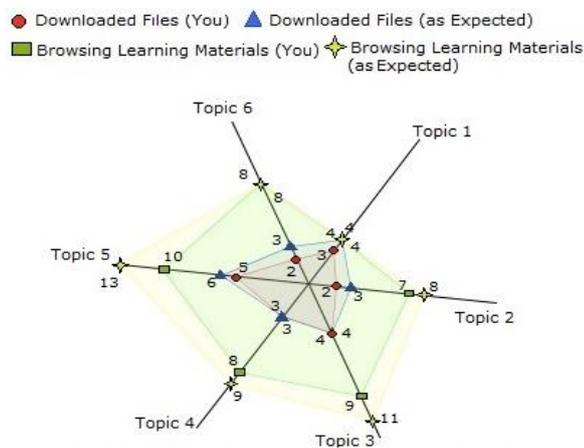


Fig. 4 – *Download* de materiais teóricos e Leitura de materiais teóricos online (realizados e esperado) em 6 unidades de aprendizagem diferentes

	<input type="radio"/>					
Extremamente Inúteis	-2	-1	0	1	2	Extremamente úteis

4.6 O que poderia ser alterado/acrescentado para que essa informação seja mais útil para você? *(ignore a pergunta se você respondeu Extremamente úteis)*

5 Em relação à facilidade de uso do sistema e motivação do estudante (6 questões)

5.1 Em que medida o sistema é fácil ou difícil de usar?

	<input type="radio"/>					
Extremamente difícil de usar	-2	-1	0	1	2	Extremamente fácil de usar

5.2 O que você sugere de alterações no sistema que poderiam facilitar seu uso ou o uso das informações apresentadas?

5.3 Quanto o sistema facilitou suas atividades?

	<input type="radio"/>					
Atrapalhou muito	-2	-1	0	1	2	Facilitou muito

5.4 Comente sua resposta:

5.5 Em que medida o sistema te motivou a estudar?

	<input type="radio"/>					
Não motivou	0	1	2	3	4	Motivou muito

5.6 Comente sua resposta:

Apêndice C

QUESTIONÁRIO - INSTRUTORES

1. Informações Pessoais (2 questões)

1.1 - Há quanto tempo você usa o *Moodle*?

- Menos de 6 meses Menos de 1 ano Menos de 3 anos Menos de 6 anos Mais de 6 anos

1.2 – Há quanto tempo você é instrutor (tutor ou professor) em cursos de EaD?

- Menos de 6 meses Menos de 1 ano Menos de 3 anos Menos de 6 anos Mais de 6 anos

2. Em relação ao seu gerenciamento pessoal, avaliação e supervisão dos alunos (8 questões)

2.1 Em que medida você aceitaria que o sistema apresentado o/a ajudasse a gerenciar suas atividades no ambiente virtual (decidir o que executar primeiro)?

- Não aceitaria -2 -1 0 1 2 Aceitaria completamente

2.2 Comente sua resposta:

2.3 Em que medida você aceitaria o sistema o/a ajudasse a acompanhar/supervisionar seus alunos?

	<input type="radio"/>					
Não	-2	-1	0	1	2	Aceitaria
aceitaria						completamente

2.4 Comente sua resposta:

2.5 Quais outras informações você gostaria de ter para o/a ajudar a supervisionar seus alunos e que o sistema não traz?

2.6 Em que medida você aceitaria que as informações trazidas pelo sistema o/a ajudassem a avaliar seus alunos?

	<input type="radio"/>					
Não	-2	-1	0	1	2	Aceitaria
aceitaria						completamente

2.7 Comente sua resposta:

2.8 Quais outras informações você gostaria de ter para o/a ajudar a avaliar seus alunos e que o sistema não traz?

3 Em relação ao apoio na tomada de decisões (6 questões)

3.1 Em que medida você aceitaria que o sistema o/a ajudasse a decidir sobre suas atividades docentes?

	<input type="radio"/>					
Não						Aceitaria
aceitaria	-2	-1	0	1	2	completamente

3.2 Comente sua resposta:

3.3 Em que medida você adaptaria sua estratégia de ensino para promover a aprendizagem e motivar os alunos a partir de informações que o sistema traz?

	<input type="radio"/>					
Não						Adaptaria
adaptaria	-2	-1	0	1	2	completamente

3.4 Comente sua resposta:

3.5 Você alteraria alguma atividade (mudança no enunciado, nos critérios de avaliação ou no valor/peso etc.) a partir de informações analíticas apresentadas pelo sistema?

Sim Não

3.6 Comente sua resposta:



Apêndice D

QUESTIONÁRIO - GESTORES

Q1 - Do ponto de vista de gestão acadêmica, o/a Sr.(a) acredita que o Modelo Conceitual proposto, e implementado pelo *framework*, pode apoiar a tomada de decisões no nível institucional?

Q2 - Sob o mesmo ponto de vista, o/a Sr.(a) considera que as visualizações oferecidas pelo sistema são adequadas e podem apoiar as decisões no nível institucional?

Q3 – Para cada um dos itens a seguir, indique se o/a Sr.(a) acredita que o Modelo Conceitual e as visualizações oferecidas podem apoiar a gestão acadêmica e a tomada de decisões no nível institucional?

Q3.1 - Observação de práticas pedagógicas;

Q3.2 - Monitoramento da retenção;

Q3.3 - Detecção de evasão;

Q3.4 - Monitoramento do grau de conclusão;

Q3.5 - Observação de estratégias de intervenção;

Q3.6 – Monitoramento de equipes de trabalho (por exemplo: pedagógica, *designers* instrucionais, tutores e professores);

Q3.7 - Reclassificação de pessoal;

Q3.8 - Identificação de necessidade de capacitação;

Q3.9 - (Re)Planejamento pedagógico;

Q3.10 - Avaliação de currículos, programas e da própria instituição;

Q3.11 - Replanejamento de políticas institucionais.

Q4 - Quais outros tipos de aplicações poderiam ser implementadas a partir do Modelo Conceitual proposto a fim de auxiliar a tomada de decisão no nível institucional?

Q5 - Das tecnologias atuais de seu conhecimento, o/a Sr.(a) acha que esse Modelo Conceitual avança no apoio à gestão educacional?

Q6 - O/A Sr.(a) estaria disposto que sua equipe de TI integrasse esse sistema ao seu SGA?

Apêndice E

QUESTIONÁRIO - DESENVOLVEDORES

Q1 - Dada sua experiência no desenvolvimento de ferramentas para apoio à educação *online*, o/a Sr(a) acredita que a implementação do *framework* abrange todas as etapas necessárias para facilitar a construção de aplicações educacionais analíticas, incluindo a coleta de dados, o tratamento e a análise sobre os mesmos?

Q2 - Do ponto de vista de Engenharia de *Software*, o/a Sr(a) acredita que o *framework* proposto implementa, satisfatoriamente, as etapas de coleta de dados, tratamento dos dados e análises previstas no Modelo Conceitual?

Q3 – O/A Sr(a) acredita que o *framework* proposto facilita a construção de ferramentas educacionais analíticas para tomada de decisão nos níveis de ensino, aprendizagem e gestão educacional?

Q4 – O/A Sr(a) considera que o *framework* oferece um caminho simples e unificado tanto para aquisição de dados a partir de diversas e variadas fontes, quanto para as análises dos dados coletados?

Q5 - As interfaces de programação disponíveis no *framework* facilitam a construção de aplicações, sem a necessidade de conhecer a fundo o código?

Apêndice F

QUESTIONÁRIO - ESPECIALISTAS

Q1 - Dada sua atuação como pesquisador(a) na área de educação a distância, especialmente no desenvolvimento de ferramentas de apoio à educação *online*, o/a Sr(a) acredita que o **Modelo Conceitual proposto é adequado para apoiar o desenvolvimento de aplicações educacionais analíticas?**

Q2 - O Modelo Conceitual proposto apresenta as seguintes etapas principais: coleta de dados, análises e decisão. Dada sua experiência com pesquisas na área de educação a distância *online*, o/a Sr(a) considera que essas etapas são suficientes e adequadas para apoiar a construção de ferramentas analíticas com objetivo de melhorar a tomada de decisões nos níveis ensino, aprendizagem e gestão educacional, atendendo portanto diferentes perfis de usuários de EaD, tais como estudantes, instrutores e gestores?

Q3 - Sobre a **validação do Modelo Conceitual** proposto, o/a Sr(a) acredita que as **avaliações realizadas** validam satisfatoriamente o Modelo?

Apêndice G

FIGURAS E TABELAS COMPLEMENTARES AO CAPÍTULO 6

FIGURAS

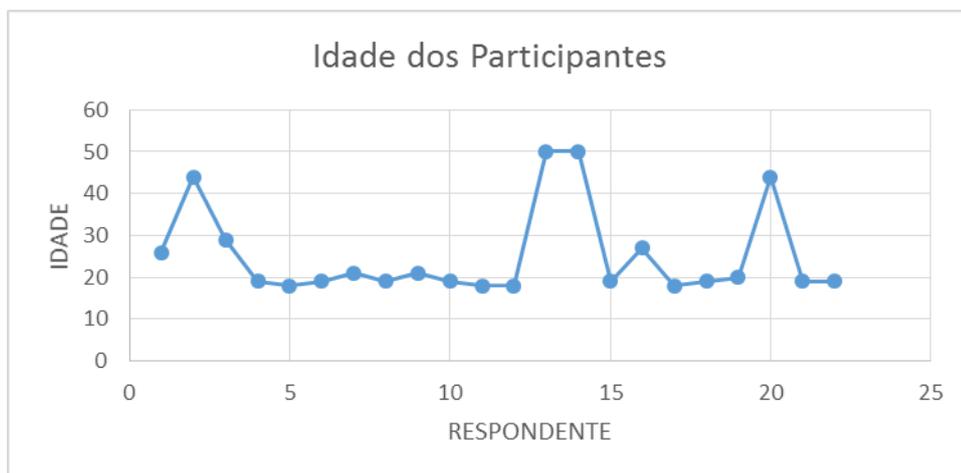


Figura G.1 – Idade dos alunos participantes. Fonte: elaborado pela autora.

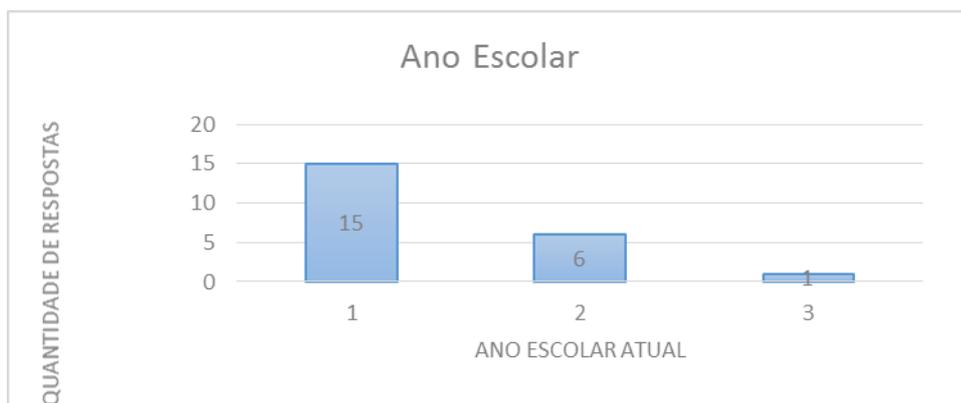


Figura G.2 – Ano escolar. Fonte: elaborado pela autora.

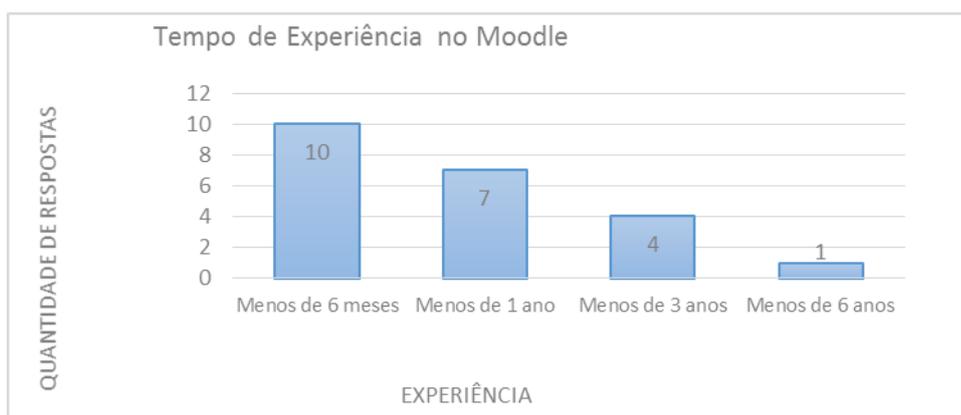


Figura G.3 – Tempo de Experiência. Fonte: elaborado pela autora.

TABELAS

Tabela G.1 – Caracterização da amostra dos estudantes.

Idade		
N	Válido	22
	Ausente	0
Mediana		19,00
Modo		19
Mínimo		18
Máximo		50

Tabela G.2 – Estatísticas: tópico 2 do instrumento de avaliação.

		Ajuda a Decidir o Tempo de Dedicção (2.1)	Autogerenciamento dos Estudos (2.3)	Automonitoramento (2.5)	Ajuda a Atingir Metas (2.7)
N	Válido	22	22	22	22
	Ausente	0	0	0	0
Mediana		2,000	2,000	2,000	2,000
Modo		2	2	2	2
Mínimo		1	1	1	1
Máximo		2	2	2	2

Tabela G.3 – Frequências: itens 3.1 e 3.3 do instrumento de avaliação.

		Mudou de Atitude/Ação (3.1)	Seguiu Recomendações (3.3)
N	Válido	22	22
	Ausente	0	0
Mediana		4,000	3,500
Modo		4	4
Mínimo		2	2
Máximo		4	4

Tabela G.4 – Frequências: item 3.5 do instrumento de avaliação.

Utilidade das Recomendações 3.5		
N	Válido	22
	Ausente	0
Mediana		2,000
Modo		2
Mínimo		1
Máximo		2

Tabela G.5 – Frequências: itens 4.1, 4.3 e 4.5 do instrumento de avaliação.

		Utilidade das informações Sumarizadas (4.1)	Utilidade das informações Detalhadas (4.3)	Utilidade das informações sobre Nível de Atuação (4.5)
N	Válido	22	22	22
	Ausente	0	0	0
Mediana		2,000	2,000	2,000
Modo		2	2	2
Mínimo		0	1	0
Máximo		2	2	2

Tabela G.6 – Frequências: itens 5.1 e 5.3 do instrumento de avaliação.

		Em que medida o sistema é Fácil ou Difícil de Usar (5.1)	Quanto o sistema Facilitou as Atividades (5.3)
N	Válido	22	22
	Ausente	0	0
Mediana		2,000	2,000
Modo		2	2
Mínimo		0	1
Máximo		2	2

Tabela G.7 – Frequências: item 5.5 do instrumento de avaliação.

		Motivação para Estudar (5.5)
N	Válido	22
	Ausente	0
Mediana		4,000
Modo		4
Mínimo		3
Máximo		4

Apêndice H

LISTA DE PUBLICAÇÕES

ROSALES, G. C. M.; DUARTE, F. V.; OTSUKA, J. L. ; ARAUJO, R. B. Educacional Data Collection Framework Based on Hybrid-Sensors-Network. In: IADIS WWW/INTERNET (ICWI), Fort Worth Texas, 2013. **Proceedings of IADIS WWW/INTERNET**, v. 12, 2013, p. 1-10.

ROSALES, G. C. M.; DUARTE, F. V.; ZANOTTO, M. A. C.; ARAUJO, R. B. Providing Self-Regulated Learning From Monitoring of Student Personal Workspace. IN: IADIS WWW/INTERNET (ICWI), Fort Worth Texas, 2013. **Proceedings of IADIS WWW/Internet**, . v. 12. 2013. p. 1-10.

ROSALES, G. C. M.; ARAUJO, R. B.; OTSUKA, J. L ; BEDER, D. M. Data Collection Framework to the Learning Monitoring using Logical Sensor Network. In: **International Conference on Advanced Learning Technologies (ICALT)**, Roma, 2012. **Proceedings of the 12th IEEE International Conference on Advanced Learning Technologies**, 2012.

STOCCO, J. A.; ROSALES, G. C. M.; ARAUJO, R. B.; OTSUKA, J. L.;. Uma Abordagem para Prover Avaliações Online Seguras. In: SIMPÓSIO BRASILEIRO DE INFORMÁTICA NA EDUCAÇÃO (SBIE), 2012, Rio de Janeiro. **Anais do XXIII Simpósio Brasileiro de Informática na Educação**, 2012.

CAMPOS, U.; ROSALES, G. C. M.; CAMARGO, W.; ARAUJO, R. B.; OTSUKA, J. L. Agente de Recomendações de Estudo e de Organização de Atividades Didáticas. In: SIMPÓSIO INTERNACIONAL DE EDUCAÇÃO A DISTÂNCIA, São Carlos, 2012. **Anais do I Simpósio Internacional de Educação a Distância**, 2012.

STOCCO, J. A.; OTSUKA, J. L.; ROSALES, G. C. M.; ARAUJO, R. B.; BEDER, D. M. Logical and physical sensors-based online assessment security support. In: INTERNATIONAL CONFERENCE ON ADVANCED LEARNING TECHNOLOGIES (ICALT), Roma, 2012. **Proceedings of the International Conference on Advanced Learning Technologies**, 2012.

ROSALES, G. C. M.; BEDER, D. M.; OTSUKA, J. L.; ARAUJO, R. B. Using Logical Sensor Network to Support Context-Aware Learning Assessment. WORKSHOP DO INCT-SEC - INSTITUTO NACIONAL DE CIÊNCIA E TECNOLOGIA EM SISTEMAS EMBARCADOS CRÍTICOS, Águas de Lindóia, 2011. **Anais do Workshop Interno do INCT-SEC**, 2011.

ROSALES, G. C. M.; ARAUJO, R. B.; OTSUKA, J. L. Using Logical Sensors Network to the Accurate Monitoring of the Learning Process in Distance Education Courses. In: **International Conference on Advanced Learning Technologies (ICALT)**, Athens Georgia, 2011. **Proceedings of the 11th IEEE International Conference on Advanced Learning Technologies**, 2011.

OTSUKA, J. L.; ABIB, S.; ROSALES, G. C. M.; ZANOTTO, M. A. C.. Reflexões sobre o Bacharelado em Sistemas de Informação na modalidade de Educação a Distância da UFSCar. In: WEI - WORKSHOP SOBRE EDUCAÇÃO EM COMPUTAÇÃO, Belo Horizonte, 2010. **Anais do XXX Congresso da SBC**, 2010.

ROSALES, G. C. M.; ARAUJO, R. B.; OTSUKA, J. L.; BEDER, D. M. Logical Sensor Networks to the Monitoring of Learning Events, Florianópolis. In: INTERNATIONAL INFORMATION AND TELECOMMUNICATION TECHNOLOGIES SYMPOSIUM. **Proceedings of 8th International Information and Telecommunication Technologies Symposium**, 2009.

ROSALES, G. C. M.; ARAUJO, R. B.; OTSUKA, J. L. ; BEDER, D. M. Aplicação de redes de sensores lógicos no monitoramento de ambientes de aprendizagem. In: SIMPÓSIO BRASILEIRO DE INFORMÁTICA NA EDUCAÇÃO, FLORIANÓPOLIS. **Anais do XX Simpósio Brasileiro de Informática na Educação**, 2009.
