

ERNANE ROSA MARTINS

Organizador



CIÊNCIA DA COMPUTAÇÃO

TECNOLOGIAS EMERGENTES EM COMPUTAÇÃO

VOLUME 2



editora científica

ERNANE ROSA MARTINS

Organizador

CIÊNCIA DA COMPUTAÇÃO

TECNOLOGIAS EMERGENTES EM COMPUTAÇÃO

VOLUME 2

1ª EDIÇÃO



editora científica

2021 - GUARUJÁ - SP

Copyright© 2021 por Editora Científica Digital

Copyright da Edição © 2021 Editora Científica Digital

Copyright do Texto © 2021 Os Autores



Todo o conteúdo deste livro está licenciado sob uma Licença de Atribuição Creative Commons. Atribuição 4.0 Internacional (CC BY 4.0).

**Dados Internacionais de Catalogação na Publicação (CIP)
(eDOC BRASIL, Belo Horizonte/MG)**

C569 Ciência da computação [livro eletrônico] : tecnologias emergentes em computação: volume 2 / Organizador Ernane Rosa Martins. – Guarujá, SP: Científica Digital, 2021.

Formato: PDF

Requisitos de sistema: Adobe Acrobat Reader

Modo de acesso: World Wide Web

Inclui bibliografia

ISBN 978-65-89826-02-6

DOI 10.37885/978-65-89826-02-6

1. Computação – Pesquisa – Brasil. 2. Tecnologias emergentes.
I. Martins, Ernane Rosa.

CDD 004

Elaborado por Maurício Amormino Júnior – CRB6/2422

Parecer e Revisão Por Pares

Os textos que compõem esta obra foram submetidos para avaliação do Conselho Editorial da Editora Científica Digital, bem como revisados por pares, sendo indicados para a publicação.

O conteúdo dos capítulos e seus dados e sua forma, correção e confiabilidade são de responsabilidade exclusiva dos autores. É permitido o download e compartilhamento desta obra desde que no formato Acesso Livre (Open Access) com os créditos atribuídos aos respectivos autores, mas sem a possibilidade de alteração de nenhuma forma ou utilização para fins comerciais.



editora científica

EDITORA CIENTÍFICA DIGITAL LTDA

Guarujá - São Paulo - Brasil

www.editoracientifica.org - contato@editoracientifica.org

CORPO EDITORIAL

Direção Editorial

Reinaldo Cardoso

João Batista Quintela

Editor Científico

Prof. Dr. Robson José de Oliveira

Assistentes Editoriais

Elielson Ramos Jr.

Erick Braga Freire

Bianca Moreira

Sandra Cardoso

Bibliotecário

Maurício Amormino Júnior - CRB6/2422

Jurídico

Dr. Alandelon Cardoso Lima - OAB/SP-307852



editora científica

CONSELHO EDITORIAL

MESTRES, MESTRAS, DOUTORES E DOUTORAS

Robson José de Oliveira

Universidade Federal do Piauí, Brasil

Carlos Alberto Martins Cordeiro

Universidade Federal do Pará, Brasil

Rogério de Melo Grillo

Universidade Estadual de Campinas, Brasil

Eloisa Rosotti Navarro

Universidade Federal de São Carlos, Brasil

Ernane Rosa Martins

Instituto Federal de Educação, Ciência e Tecnologia de Goiás, Brasil

Rossano Sartori Dal Molin

Universidade Federal do Rio Grande do Sul, Brasil

Carlos Alexandre Oelke

Universidade Federal do Pampa, Brasil

Domingos Bombo Damião

Universidade Agostinho Neto, Angola

Edilson Coelho Sampaio

Universidade da Amazônia, Brasil

Elson Ferreira Costa

Universidade do Estado do Pará, Brasil

Reinaldo Eduardo da Silva Sales

Instituto Federal do Pará, Brasil

Patrício Francisco da Silva

Universidade CEUMA, Brasil

Auristela Correa Castro

Universidade Federal do Pará, Brasil

Dalízia Amaral Cruz

Universidade Federal do Pará, Brasil

Susana Jorge Ferreira

Universidade de Évora, Portugal

Fabricio Gomes Gonçalves

Universidade Federal do Espírito Santo, Brasil

Erival Gonçalves Prata

Universidade Federal do Pará, Brasil

Gevair Campos

Faculdade CNEC Unai, Brasil

Flávio Aparecido de Almeida

Faculdade Unida de Vitória, Brasil

Mauro Vinicius Dutra Girão

Centro Universitário Ina, Brasil

Clóvis Luciano Giacomet

Universidade Federal do Amapá, Brasil

Giovanna Moraes

Universidade Federal de Uberlândia, Brasil

André Cutrim Carvalho

Universidade Federal do Pará, Brasil

Dennis Soares Leite

Universidade de São Paulo, Brasil

Silvani Verruck

Universidade Federal de Santa Catarina, Brasil

Osvaldo Contador Junior

Faculdade de Tecnologia de Jahu, Brasil

Claudia Maria Rinhel-Silva

Universidade Paulista, Brasil

Silvana Lima Vieira

Universidade do Estado da Bahia, Brasil

Cristina Berger Fadel

Universidade Estadual de Ponta Grossa, Brasil

Graciete Barros Silva

Universidade Estadual de Roraima, Brasil



editora científica

Carlos Roberto de Lima

Universidade Federal de Campina Grande, Brasil

Wesley Viana Evangelista

Universidade do Estado de Mato Grosso, Brasil

Cristiano Marins

Universidade Federal Fluminense, Brasil

Marcelo da Fonseca Ferreira da Silva

Escola Superior de Ciências da Santa Casa de Misericórdia de Vitória, Brasil

Daniel Luciano Gevehr

Faculdades Integradas de Taquara, Brasil

Silvio Almeida Junior

Universidade de Franca, Brasil

Juliana Campos Pinheiro

Universidade Federal do Rio Grande do Norte, Brasil

Raimundo Nonato Ferreira do Nascimento

Universidade Federal do Piauí, Brasil

Antônio Marcos Mota Miranda

Instituto Evandro Chagas, Brasil

Maria Cristina Zago

Centro Universitário UNIFAAT, Brasil

Samylla Maira Costa Siqueira

Universidade Federal da Bahia, Brasil

Gloria Maria de Franca

Universidade Federal do Rio Grande do Norte, Brasil

Carla da Silva Sousa

Instituto Federal de Educação, Ciência e Tecnologia Baiano, Brasil

Dennys Ramon de Melo Fernandes Almeida

Universidade Federal do Rio Grande do Norte, Brasil

Mário Celso Neves de Andrade

Universidade de São Paulo, Brasil

Julianno Pizzano Ayoub

Universidade Estadual do Centro-Oeste, Brasil

Ricardo Pereira Sepini

Universidade Federal de São João Del-Rei, Brasil

Maria do Carmo de Sousa

Universidade Federal de São Carlos, Brasil

Flávio Campos de Moraes

Universidade Federal de Pernambuco, Brasil

Jonatas Brito de Alencar Neto

Universidade Federal do Ceará, Brasil

Reginaldo da Silva Sales

Instituto Federal de Educação, Ciência e Tecnologia do Pará, Brasil

Iramirton Figuerêdo Moreira

Universidade Federal de Alagoas, Brasil

Moisés de Souza Mendonça

Instituto Federal de Educação, Ciência e Tecnologia do Pará, Brasil

Bianca Anacleto Araújo de Sousa

Universidade Federal Rural de Pernambuco, Brasil

Pedro Afonso Cortez

Universidade Metodista de São Paulo, Brasil

Bianca Cerqueira Martins

Universidade Federal do Acre, Brasil

Vitor Afonso Hoeflich

Universidade Federal do Paraná, Brasil

Francisco de Sousa Lima

Instituto Federal de Educação, Ciência e Tecnologia Baiano, Brasil

Sayonara Cotrim Sabioni

Instituto Federal de Educação, Ciência e Tecnologia Baiano, Brasil

Thais Ranielle Souza de Oliveira

Centro Universitário Euroamericano, Brasil

Cynthia Mafra Fonseca de Lima

Universidade Federal de Alagoas, Brasil

Marcos Reis Gonçalves

Centro Universitário Tiradentes, Brasil

Rosemary Laís Galati

Universidade Federal de Mato Grosso, Brasil

Maria Fernanda Soares Queiroz

Universidade Federal de Mato Grosso, Brasil



Letícia Cunha da Hungria

Universidade Federal Rural da Amazônia, Brasil

Dioniso de Souza Sampaio

Universidade Federal do Pará, Brasil

Leonardo Augusto Couto Finelli

Universidade Estadual de Montes Claros, Brasil

Danielly de Sousa Nóbrega

Instituto Federal de Educação, Ciência e Tecnologia do Acre, Brasil

Mauro Luiz Costa Campello

Universidade Paulista, Brasil

Livia Fernandes dos Santos

Instituto Federal de Educação, Ciência e Tecnologia do Acre, Brasil

Sonia Aparecida Cabral

Secretaria da Educação do Estado de São Paulo, Brasil

Camila de Moura Vogt

Universidade Federal do Pará, Brasil

José Martins Juliano Eustáquio

Universidade de Uberaba, Brasil

Walmir Fernandes Pereira

Miami University of Science and Technology, Estados Unidos da América

Liege Coutinho Goulart Dornellas

Universidade Presidente Antônio Carlos, Brasil

Ticiano Azevedo Bastos

Universidade Federal de Ouro Preto, Brasil

Jónata Ferreira De Moura

Universidade Federal do Maranhão, Brasil

Daniela Remião de Macedo

Faculdade de Belas Artes da Universidade de Lisboa, Portugal

Francisco Carlos Alberto Fonteles Holanda

Universidade Federal do Pará, Brasil

Bruna Almeida da Silva

Universidade do Estado do Pará, Brasil

Adriana Leite de Andrade

Universidade Católica de Petrópolis, Brasil

Clecia Simone Gonçalves Rosa Pacheco

Instituto Federal do Sertão Pernambucano, Brasil

Claudiomir da Silva Santos

Instituto Federal de Educação, Ciência e Tecnologia do Sul de Minas, Brasil

Fabício dos Santos Ritá

Instituto Federal de Educação, Ciência e Tecnologia do Sul de Minas, Brasil, Brasil

Ronei Aparecido Barbosa

Instituto Federal de Educação, Ciência e Tecnologia do Sul de Minas, Brasil

Julio Onésio Ferreira Melo

Universidade Federal de São João Del-Rei, Brasil

Juliano José Corbi

Universidade de São Paulo, Brasil

Alessandra de Souza Martins

Universidade Estadual de Ponta Grossa, Brasil

Francisco Sérgio Lopes Vasconcelos Filho

Universidade Federal do Cariri, Brasil

Thadeu Borges Souza Santos

Universidade do Estado da Bahia, Brasil

Francine Náthalie Ferraresi Rodriguess Queluz

Universidade São Francisco, Brasil

Maria Luzete Costa Cavalcante

Universidade Federal do Ceará, Brasil

Luciane Martins de Oliveira Matos

Faculdade do Ensino Superior de Linhares, Brasil

Rosenery Pimentel Nascimento

Universidade Federal do Espírito Santo, Brasil

Lívia Silveira Duarte Aquino

Universidade Federal do Cariri, Brasil

Irlane Maia de Oliveira

Universidade Federal do Amazonas, Brasil

Xaene Maria Fernandes Mendonça

Universidade Federal do Pará, Brasil



Thaís de Oliveira Carvalho Granado Santos

Universidade Federal do Pará, Brasil

Fábio Ferreira de Carvalho Junior

Fundação Getúlio Vargas, Brasil

Anderson Nunes Lopes

Universidade Luterana do Brasil, Brasil

Iara Margolis Ribeiro

Centro Universitário Boa Viagem, Brasil

Carlos Alberto da Silva

Universidade Federal do Ceará

Keila de Souza Silva

Universidade Estadual de Maringá, Brasil

Francisco das Chagas Alves do Nascimento

Universidade Federal do Pará, Brasil

Réia Sílvia Lemos da Costa e Silva Gomes

Universidade Federal do Pará, Brasil

Priscyla Lima de Andrade

Centro Universitário UniFBV, Brasil

Aleteia Hummes Thaines

Faculdades Integradas de Taquara, Brasil

Darlindo Ferreira de Lima

Universidade Federal de Pernambuco, Brasil

Sílvia Raquel Santos de Moraes

Universidade Federal do Vale do São Francisco, Brasil



APRESENTAÇÃO

A Ciência da Computação é uma área com amplas possibilidades de atuação, tais como: o desenvolvimento de softwares e de aplicativos mobile, o gerenciamento de informações, e a atuação acadêmica. A abrangência desta área exige dos seus profissionais conhecimentos diversos, tais como: pensamento criativo, novos idiomas, capacidade de comunicação, além da constante atualização dos conhecimentos. Os profissionais da área de Ciência da Computação estão atualmente sendo cada vez mais prestigiados, importantes, reconhecidos, valorizados e requisitados por empresas de todos os portes e setores. Necessitando de profissionais qualificados, que apresentem potencial para promover inovação, desenvolvimento e eficiência.

Dentro deste contexto, a presente obra aborda assuntos relevantes sobre a computação, tais como: programação de computadores, desenvolvimento de software, mineração de dados, usabilidade de software, aprendizagem de máquina, gerenciamento de redes, modelagem de processos, performance de aplicações, pensamento computacional, robótica, aplicações paralelas de alto desempenho, sistema especialista, software livre e inteligência artificial.

Nesse sentido, esta obra apresenta enorme potencial para contribuir com análises e discussões aprofundadas sobre assuntos relevantes da área da computação, podendo servir de referência para novas pesquisas e estudos nesta área. Agradecemos em especial aos autores dos capítulos que a compõem, e desejamos aos nossos leitores, inúmeras e boas reflexões sobre as temáticas abordadas.

Ernane Rosa Martins



editora científica

SUMÁRIO

CAPÍTULO 01

ANÁLISE DE COMPONENTES LATENTES DA APRENDIZAGEM DE PROGRAMAÇÃO PARA CLASSIFICAÇÃO DE PERFIS

Márcia Gonçalves de Oliveira; Nátaly A. Jiménez Monroy; Eliana Zandonade; Elias Silva de Oliveira

DOI: 10.37885/210304001 14

CAPÍTULO 02

ANÁLISE SOBRE A EVOLUÇÃO DA MINERAÇÃO DE DADOS EDUCACIONAIS SEGUNDO A PERSPECTIVA DOS PRINCIPAIS ESTUDOS DA ÁREA: UMA REVISÃO NARRATIVA

Vanessa Faria de Souza; Mayara Faria de Souza

DOI: 10.37885/210303626 27

CAPÍTULO 03

AValiação DA USABILIDADE DO SCRATCH UTILIZANDO O MÉTODO SYSTEM USABILITY SCALE (SUS)

Wendell Bento Geraldes; Ernane Rosa Martins; Ulisses Rodrigues Afonseca

DOI: 10.37885/210202975 47

CAPÍTULO 04

AValiação AUTOMÁTICA DE REDAÇÕES NA LÍNGUA PORTUGUESA BASEADA NA COLETA DE ATRIBUTOS E APRENDIZAGEM DE MÁQUINA

Silvério Sirotheau; Eloi Favero; João Carlos Alves dos Santos; Simone Negrão; Marco Lima do Nascimento

DOI: 10.37885/210303887 56

CAPÍTULO 05

CAPACITAÇÃO EM TEMPOS DE PANDEMIA: CONTRIBUIÇÕES DO MODELO LV DE AValiação EM POSSIBILITAR O GROWING FOCUS ON MEASURING LEARNING

Gilvandenys Leite Sales; Bento Silva; Universidade do Minho; José Alberto Lencastre; Universidade do Minho; João Batista da Silva

DOI: 10.37885/210303985 69

CAPÍTULO 06

GERENCIAMENTO DE REDES UTILIZANDO ZABBIX

Adriane Santos Rezende; Joseane Santos de Jesus; José Damião de Melo; Eduardo Henrique do Lago Silva; Mayka de Souza Lima

DOI: 10.37885/210303998 86

SUMÁRIO

CAPÍTULO 07

GOOGLEMÁTICA: A MATEMÁTICA POR TRÁS DO GOOGLE

Naiara Beber; Arthur Roberto Fronza; Luma Carina Picolli; Andresa Laurett da Silva

DOI: 10.37885/201202590 95

CAPÍTULO 08

IMPORTÂNCIA DA TECNOLOGIA DA INFORMAÇÃO: ANÁLISE DO ATENDIMENTO DE PACIENTES DE UMA MATERNIDADE

João Cardim Ferreira Lima; Luan Lyon Lopes Nascimento; Adriano Varella de Moraes

DOI: 10.37885/210102803 106

CAPÍTULO 09

MENSURAÇÃO DO CAPITAL SOCIAL ACUMULADO A PARTIR DE INTERAÇÕES SOCIAIS EM PÁGINAS INSTITUCIONAIS NO FACEBOOK

Kaíque Matheus R. Cunha; Alan Keller Gomes

DOI: 10.37885/210303621 122

CAPÍTULO 10

MODELAGEM DE PROCESSOS PARA REPRESENTAÇÃO E GESTÃO DA MEMÓRIA CORPORATIVA BASEADO EM CASOS

José Damião de Melo; Daniel Antonio de Jesus Melo; Bruno Barros de Souza; Claudia de Medeiros Lima; Vania de Jesus; Valdenice de Jesus Melo

DOI: 10.37885/210303881 139

CAPÍTULO 11

PERFORMANCE DE APLICAÇÕES WEB: UM ESTUDO DE CASO NA EMPRESA TUHL SOLUÇÕES

Daniel da Rosa Vasconcelos; Sidnei Renato Silveira; Antônio Rodrigo Delepiane de Vit; Cristiano Bertolini; Fábio José Parreira; Guilherme Bernardino da Cunha; Nara Martini Bigolin

DOI: 10.37885/210303676 152

CAPÍTULO 12

PROBLEMAS INVESTIGATIVOS NAS AULAS DE MATEMÁTICA DA ESCOLA BÁSICA COM O PENSAMENTO COMPUTACIONAL: UM RELATO DE EXPERIÊNCIA MOBILIZADOR À APRENDIZAGEM

Aline Silva de Bona; Natália Bernardo Nunes; Rafaela da Silva Bobsin; Anelise Lemke Kologeski

DOI: 10.37885/210303952 171

SUMÁRIO

CAPÍTULO 13

PYRALELO: UM SISTEMA PARA SUPORTE À PROGRAMAÇÃO VISUAL DE APLICAÇÕES PARALELAS DE ALTO DESEMPENHO

André Ricardo Dantas Bezerra Landim; Guilherme Álvaro Rodrigues Maia Esmeraldo

DOI: 10.37885/210303615..... 186

CAPÍTULO 14

ROBÓTICA EM TEMPOS DE PANDEMIA: UMA ABORDAGEM USANDO PROGRAMAÇÃO EM BLOCOS

Wilson Rogério Soares e Silva; Klévia Letícia Almeida; Natalle do Socorro da Costa Freitas

DOI: 10.37885/210203196..... 196

CAPÍTULO 15

SISTEMA ESPECIALISTA PARA O DOMÍNIO DO LICENCIAMENTO AMBIENTAL: ESTUDO DE CASO COM SHELL EXPERT SINTA

José Damião de Melo; Alvaro Luis Enrique Adriazola Uribe; Daniel Antonio de Jesus Melo; Mikaele Costa Mendonça; Ubirajara Rodrigues

Xavier; Valdenice de Jesus Melo; Patrick Henrique da Silva Brito; Evandro de Barros Costa

DOI: 10.37885/210303823..... 206

CAPÍTULO 16

SOFTWARE = [(IDEIA U MATEMÁTICA) W LÓGICA]

Shuichi Murakami

DOI: 10.37885/201202612..... 218

CAPÍTULO 17

SOFTWARE LIVRE NA GESTÃO DE TI: UM ESTUDO DE CASO APLICADO AO CAMPUS PARAÍSO DO TOCANTINS DO IFTO

Kálita Fernandes da Silva Fonteles; Gislaíne Pereira Sales

DOI: 10.37885/210303604..... 236

CAPÍTULO 18

UMA ABORDAGEM UBÍQUA PARA ACOMPANHAMENTO CLÍNICO DE PACIENTES

Alexandre Renato Rodrigues de Souza; João Ladislau Barbará Lopes; Fabrício Neitzke Ferreira; Leonardo Costa Reichow; Ana Marilza

Pernas Fleischmann; Adenauer Correa Yamin

DOI: 10.37885/210303856..... 248

SUMÁRIO

CAPÍTULO 19

UMA EXPERIÊNCIA EM ENSINO DE PROGRAMAÇÃO DE COMPUTADORES COM ALUNOS DA REDE PÚBLICA USANDO A LINGUAGEM *SCRATCH*

Wendell Bento Geraldes; Ernane Rosa Martins; Ulisses Rodrigues Afonseca

DOI: 10.37885/210202974.....263

CAPÍTULO 20

USO DE INTELIGÊNCIA ARTIFICIAL PARA AUXILIAR O COMBATE À OBESIDADE: ESTUDO DE CASO EM PYTHON

Edjair Aguiar Gomes Filho; Mateus Amorim Silva; Leonardo Corsino Campello; Ricardo Argenton Ramos; Brauliro Gonçalves Leal

DOI: 10.37885/201202574.....274

CAPÍTULO 21

UTILIZANDO SERVIÇOS EM NUVEM PARA REDUZIR OS CUSTOS DE DESENVOLVIMENTO E FORNECIMENTO DE SOFTWARE

Mário Santos Sousa; Karlos Kelvin Santos; Guilherme Álvaro Rodrigues Maia Esmeraldo

DOI: 10.37885/210303820283

SOBRE O ORGANIZADOR.....293

ÍNDICE REMISSIVO294

Análise de componentes latentes da aprendizagem de programação para classificação de perfis

I Márcia Gonçalves de **Oliveira**
UFES

I Nátaly A. Jiménez **Monroy**
UFES

I Eliana **Zandonade**
UFES

I Elias Silva de **Oliveira**
UFES

RESUMO

Este trabalho tem como objetivo reconhecer, a partir de programas desenvolvidos por alunos e de uma rede de variáveis representantes do domínio da programação, componentes latentes que caracterizem os desempenhos de uma turma em cada atividade de programação. A identificação das componentes latentes é realizada pela técnica de análise fatorial para representar perfis de aprendizagem na prática da programação e, por clustering, formam-se agrupamentos de perfis similares. De cada agrupamento formado, são selecionadas amostras de perfis pré-classificadas para gerar modelos que orientem a classificação dos demais perfis de cada agrupamento por níveis de aprendizagem. Os resultados indicam que a combinação das técnicas de análise fatorial e de clustering melhora a classificação de perfis.

Palavras-chave: Análise Fatorial, Clustering, Aprendizagem de Programação, Mapeamento de Perfis, Classificação de Perfis.

INTRODUÇÃO

A programação de computadores é o processo de escrever em linguagem formal instruções sequenciadas logicamente com o propósito de resolver um problema através de uma solução automatizada. Mas, por envolver a combinação de várias habilidades cognitivas, a prática da programação é complexa. Para se ter uma ideia da complexidade de programar, para desenvolver um programa, são necessárias as seguintes atividades cognitivas (PEA e KURLAND, 1984):

1. Compreender o problema
2. Planejar ou projetar uma solução
3. Escrever código do plano de solução
4. Compreender a escrita do programa
5. Depurar o programa

Para contemplar a aprendizagem de programação, hoje existem sofisticadas tecnologias para apresentar conteúdos de programação (POWERS, *et al.* 2006), para instruir (PILLAY, 2003), para prever comportamentos (MAVRIKIS, 2010), para dar instruções e exercícios personalizados de acordo com os perfis de alunos (MAZZA e DIMITROVA, 2007), para recomendar atividades (OLIVEIRA, *et al.* 2013) e até para avaliação automática de exercícios de programação (MOREIRA e FAVERO, 2009; NAUDE *et al.* 2010). Mas carecemos de tecnologias que de fato ofereçam uma prática assistida, isto é, planejada, monitorada, controlada e com *feedback* imediato, no domínio da aprendizagem de programação.

Para melhor assistir e regular a prática da programação, no entanto, é necessário conhecer continuamente estados de aprendizagem de alunos representando-os através de perfis que sejam mapeados em diferentes variáveis de avaliação que apontem suas habilidades e dificuldades de aprendizagem e os ajudem na remediação da aprendizagem (ANDERSON, 2000, MAZZA e DIMITROVA, 2007; OLIVEIRA, *et al.* 2013). Caminhando nessa direção, o trabalho de Oliveira e Oliveira (2014) chama a atenção para a representação de perfis por componentes de habilidades que representem desempenhos em variáveis de avaliação do domínio da aprendizagem de programação.

Com essa representação, Oliveira e Oliveira (2014) propõem um Núcleo de Avaliação Diagnóstica (NAD) que reúne as funções de mapeamento de perfis e de correção semiautomática de exercícios de programação. No NAD, o mapeamento de perfis é realizado a partir dos desempenhos dos alunos em componentes de habilidades representando o domínio de aprendizagem da programação (OLIVEIRA e OLIVEIRA, 2014). Dessa forma, representa-se

uma turma por uma matriz cognitiva $A_{m \times n}$, em que m é o número de alunos e n , o número de componentes de habilidades.

Os desempenhos mapeados em componentes de habilidades são obtidos por tarefas (listas de exercícios e provas), onde cada atividade requer como resposta do aluno um programa de computador desenvolvido em Linguagem C (OLIVEIRA e OLIVEIRA, 2014). No programa submetido, a frequência de ocorrência das palavras reservadas, dos operadores e dos símbolos bem como os indicadores de execução representa os valores das componentes de habilidades de um perfil de aluno.

Este trabalho estende o Núcleo de Avaliação Diagnóstica de Oliveira e Oliveira (2014) ao propor uma redução de dimensionalidade da matriz A através da técnica de análise fatorial. Essa nova matriz é representada pelas componentes latentes das relações entre as componentes de habilidades da matriz A . Através dessa nova representação, objetiva-se promover melhoramentos nos processos de classificação de perfis.

Os resultados de experimentação deste trabalho indicam que a combinação das técnicas de análise fatorial e de *clustering* melhoram a classificação de perfis. A contribuição desta proposta para o domínio da aprendizagem de programação é, portanto, oferecer um mecanismo de seleção de características e amostras de perfis que melhor representem a prática de programação e possibilitem a classificação de novos perfis conforme os níveis de aprendizagem dos estudantes.

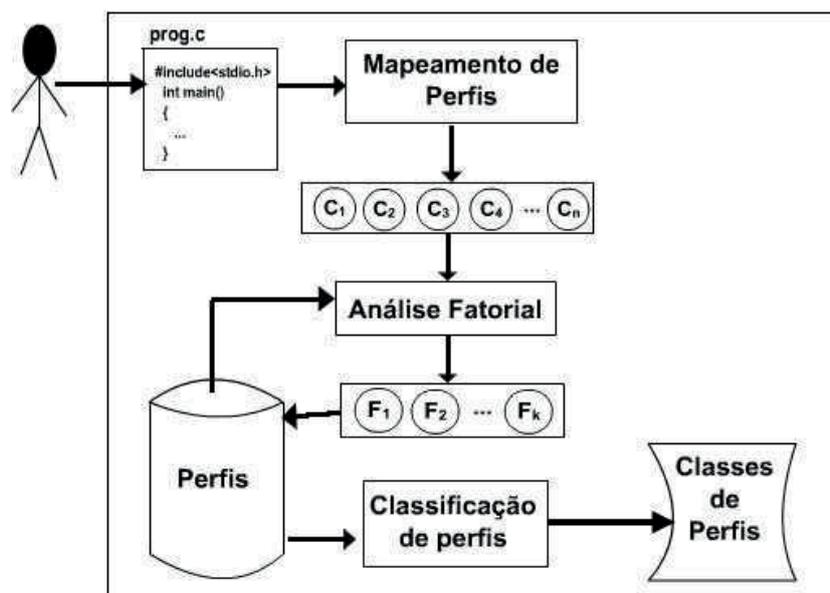
OBJETIVO

Reconhecer, a partir de programas desenvolvidos por alunos e de uma rede de variáveis representantes do domínio da programação, componentes latentes que caracterizem os desempenhos de uma turma em cada atividade de programação.

MÉTODOS

A estratégia de mapeamento e classificação de perfis deste trabalho é apresentada na Figura 1. O módulo de mapeamento de perfis recebe programas de computador em Linguagem C e os transforma em vetores cujas dimensões são chamadas de componentes de habilidades (C_i). Cada componente de habilidade é a frequência de ocorrência de palavras reservadas, símbolos, operadores e funções da Linguagem C e *tokens* indicadores de sucessos de compilação e execução de programas.

Figura 1. Mapeamento e classificação de perfis



O módulo de *Análise Fatorial*, conforme a Figura 1, redimensiona os perfis gerados pelo módulo de *Mapeamento de Perfis* extraindo as componentes latentes através das relações entre as componentes de habilidades C_i . Em seguida, os perfis, agora representados por fatores F_j que são os valores das componentes latentes, são submetidos ao algoritmo de *clustering Bisecting K-means* do módulo de *Classificação de Perfis* e são agrupados conforme as suas similaridades.

De cada *cluster*, são selecionadas 2/3 de amostras aleatórias de perfis para gerar modelos de regressão para prever escores indicadores de níveis de aprendizagem. Em seguida, os escores dos demais 1/3 dos perfis de cada *cluster* são preditos e uma classe de nível de aprendizagem é associada a cada um desses perfis.

O objetivo dessa metodologia é mostrar como a combinação das técnicas de análise fatorial e de *clustering* pode melhorar o processo de classificação de perfis.

ANÁLISE FATORIAL

A análise fatorial é uma técnica de análise multivariada que consiste na redução da dimensionalidade de um conjunto de dados, encontrando grupos homogêneos a partir de um grande número de variáveis. Esses grupos homogêneos, também chamados de fatores, são formados por variáveis muito correlacionadas entre si. O ideal é que os fatores sejam independentes uns dos outros. Portanto, a análise fatorial é uma técnica que busca explicar ao máximo a informação contida nos dados, usando a menor quantidade possível de dimensões.

De acordo com Härdle e Simar (2012), a análise é baseada em um modelo onde o vetor observado é representado por uma parte sistemática e uma parte de erro não observado. As componentes do vetor de erros são consideradas não-correlacionadas ou

independentes, enquanto a parte sistemática é considerada uma combinação linear de um número relativamente pequeno de fatores não observados.

Frequentemente as influências dos k fatores costumam ser divididas em comuns e específicas. Por exemplo, há fatores altamente informativos que são comuns para todas as p componentes de \mathbf{X} e fatores que são específicos apenas para algumas componentes. Nesse caso, a matriz \mathbf{X} pode se expressa como

$$\mathbf{X} = \mathbf{QF} + \mathbf{U} + \mu,$$

onde \mathbf{Q} é uma matriz ($p \times k$) de coeficientes (não-aleatórios) dos fatores comuns \mathbf{F} (de dimensão ($k \times 1$)), \mathbf{U} é uma matriz ($p \times 1$) dos fatores (aleatórios) específicos e μ é o vetor de médias de \mathbf{X} .

Assumindo que um modelo fatorial com k fatores foi encontrado razoável, isto é, a maioria das variações das p variáveis em \mathbf{X} foram explicadas por k fatores fixos, podemos interpretar os fatores F_l , $l = 1, \dots, k$ calculando suas correlações com as variáveis X_j , $j = 1, \dots, p$, para obter a matriz \mathbf{PXF} . Essa correlação é dada por

$$\mathbf{PXF} = \mathbf{D}^{-1/2} \mathbf{Q},$$

onde $\mathbf{D} = \text{diag} \{ \sigma_{X_1 X_1}, \dots, \sigma_{X_p X_p} \}$, sendo $\sigma_{X_j X_j}$ a variância de X_j . Usando essa informação podemos determinar quais das variáveis originais X_1, \dots, X_p influenciam nos fatores não observados F_1, \dots, F_k (ANDERSON, 1984; BISHOP, *et al.* 1975; MORRISON, 1990).

■ O CLUSTERING

O *clustering* é uma técnica que utiliza a abordagem de aprendizagem não-supervisionada para agrupamento de padrões em classes ou *clusters* considerando as características semelhantes desses padrões. O objetivo do *clustering* é formar grupos caracterizados por alta homogeneidade entre padrões de um mesmo grupo e heterogeneidade entre padrões de grupos distintos.

O *Bisecting K-means* é uma variação do algoritmo *K-means*. Ele começa com um simples *cluster* e continuamente seleciona um *cluster* para dividir em dois *sub-clusters* até alcançar o número K de *clusters* desejados (LOOKS, *et al.* 2007).

De acordo com Steinbach *et al.* (2000), o *Bisecting K-means* tem apresentado melhor performance do que o *K-means* em muitos casos devido ao fato de produzir *clusters* de tamanhos mais uniformes em vez de *clusters* de tamanhos variáveis.

METODOLOGIA DE EXPERIMENTAÇÃO

A metodologia experimental deste trabalho foi realizada conforme os seguintes passos:

1. Seleção de 100 amostras de programas em Linguagem C desenvolvidos por estudantes para uma atividade de programação
2. Mapeamento das amostras de programas em vetores reunidos em uma matriz A
3. Análise multivariada fatorial da matriz A
4. Clusterização da matriz A
5. Separação dos conjuntos de treino e de teste de cada *cluster*
6. Criação de modelos de regressão linear para cada cluster a partir do conjunto de treino pré-classificado
7. Predição de escores do conjunto de teste de cada *cluster*
8. Classificação das amostras do conjunto de teste de cada *cluster*
9. Avaliação de resultados

Os 100 programas reunidos em uma base chamada *Base-p* foram representados por uma matriz $A_{m \times n}$, onde m é o número de linhas representadas pelos 100 programas desenvolvidos por alunos e n , o número de colunas representando as 60 componentes de habilidades, que são variáveis cujos valores são a frequência de ocorrência de palavras reservadas, operadores, símbolos e funções da Linguagem C. As variáveis de execução indicando se um programa compilou e executou assumem valores 0 e 1, que significam, respectivamente, *falso* e *verdadeiro*. A Figura 2 mostra como os programas desenvolvidos por alunos foram vetorialmente representados na matriz A . Cada linha dessa matriz corresponde a um perfil de aluno mapeado a partir dos desempenhos desse aluno em cada componente de habilidade.

Figura 2. Representação vetorial de programas em Linguagem C

100	60
1	1 3 0 0 0 0 0 9 2 0 0 0 0 0 2
1	1 0 0 0 0 0 0 9 0 0 0 0 0 0 2
0	0 0 1 0 0 0 9 0 0 0 0 0 0 0 2
1	1 5 4 2 0 1 0 7 0 0 0 0 0 0 2
1	1 1 3 0 0 0 0 2 1 0 0 0 0 0 1
1	1 0 0 0 0 0 0 4 4 0 0 0 0 0 2
1	1 0 0 0 0 0 0 6 5 0 0 0 0 0 2

Após a geração da Matriz A , foi realizada a análise multivariada fatorial com rotação *Varimax* para agrupar as variáveis correlacionadas e criar possíveis fatores representantes

da atividade de programação aplicada em turmas de Engenharia e de Computação. O programa estatístico utilizado foi o SPSS, versão 18.0.

Das 60 variáveis iniciais, restaram para análise apenas 35, pois as demais possuíam valor zero. Dessa análise, consideraram-se doze fatores, sendo que o percentual de variabilidade explicado por eles foi de 82.09%. Adotou-se como critério de decisão do número de fatores autovalores maiores que 1. Os autovalores representam a variabilidade explicada por cada fator.

A Figura 3 apresenta as cargas fatoriais de quatro dos doze fatores (ou componentes latentes) do modelo gerado. Nota-se que as variáveis agrupadas em diferentes cores definem cada fator.

Na Tabela 1 são apresentados os doze fatores selecionados para caracterizar o exercício de programação resolvido por 100 alunos. As características principais são representadas por identificadores de palavras reservadas (*cbreak*, *ccase*, *cswitch*, *cfor*, *tchar*, *cif*, *cint*, *cmain*, *cscanf*, *celse*), operadores aritméticos (+, -, *, ++, -), operadores lógicos (*opOr*, *opNOT*), operadores relacionais (\geq , \leq , $==$) e indicadores de compilação e execução (*compila*, *executa*) da Linguagem C.

Figura 3. Cargas fatoriais

Nome do fator	1	2	3	4	5	6	7	8	9	10	11	12
opmais	.934	.005	-.008	-.053	.203	-.052	-.044	.027	.154	.126	.009	.020
opmult	.899	.026	-.149	-.040	.287	-.069	-.071	.027	.173	.136	-.029	.026
opmenos	.844	-.016	-.055	.007	.199	.110	-.048	.020	-.128	.087	.090	.152
endereco	.774	-.098	-.291	-.017	-.064	.015	.010	.121	.206	-.247	-.085	-.095
cimprimir	.678	-.010	-.159	-.041	.285	.006	.186	.197	.076	.061	.024	-.256
cbreak	.004	.974	.094	.037	.046	-.035	.038	.036	.009	-.030	.017	.025
ccase	-.012	.986	-.015	-.015	.039	-.012	.023	.009	.008	-.022	.060	-.006
cswitch	-.012	.986	-.015	-.015	.039	-.012	.023	.009	.008	-.022	.060	-.006
tvoid	-.199	.487	.072	-.065	.095	.390	-.016	.153	.009	.014	.026	-.480
opmenor	.254	-.082	.432	.171	.201	.077	-.015	-.047	-.011	.333	-.323	-.288
opdec	-.002	-.015	.809	.174	-.040	-.108	.133	.095	.038	-.079	.018	-.010
creturn	-.004	.227	.571	.304	.008	-.071	-.082	.573	.054	.019	.053	-.175
cfor	-.285	.028	.807	-.065	-.174	.183	-.129	-.049	.040	.037	-.017	-.001
opinc	-.293	.071	.851	-.046	-.079	.160	-.128	-.022	.035	-.022	-.110	.133
tchar	.046	-.016	.351	.817	.127	-.084	-.012	.044	.075	.035	-.229	.069
opor	.024	-.037	-.177	.853	-.095	.120	.030	-.041	-.015	.012	.128	-.122
opnot	-.051	-.022	.046	.920	-.089	.075	-.040	.020	.000	.033	.065	-.021
cwhile	-.148	.095	.274	.710	.215	-.041	.032	.074	.037	-.110	-.143	.236

Tabela 1. Identificação de componentes latentes

Fator	Características principais	Componente Latente
1	<i>opmais(+), opmult (*), opmenos (-)</i>	Operação Aritmética
2	<i>cbreak, ccase, cswitch</i>	Seleção
3	<i>opdec (-), cfor, opinc (++)</i>	Repetição Para
4	<i>opOR (), opNOT (!), tchar</i>	Expressão Lógica
5	<i>cif</i>	Condição se verdadeira
6	<i>opmaiorigual (≥)</i>	Comparação maior ou igual
7	<i>funciona, compila</i>	Execução
8	<i>cint, opmenorigual (≤)</i>	Comparação menor ou igual
9	<i>cmain</i>	Início de programa
10	<i>cscanf</i>	Entrada de dados
11	<i>celse</i>	Condição se falsa
12	<i>opigual (==)</i>	Comparação igual

Os fatores 1 a 12 da Tabela 1 foram nomeados conforme as maiores cargas fatoriais de cada grupo de componentes de habilidades. O *Fator 1* (na Figura 3), por exemplo, foi nomeado *Operação Aritmética* porque as componentes de maior carga fatorial que o formam são operadores aritméticos. Da mesma forma, o *Fator 2* foi nomeado *Seleção* porque as componentes *cbreak*, *ccase* e *cswitch*, que possuem maior carga fatorial, identificam, respectivamente, as palavras-chave *break*, *case* e *switch* da estrutura de seleção *switch* da Linguagem C.

Após a realização da análise fatorial, que reduziu o número de 35 componentes de habilidades de cada perfil a doze componentes latentes, a nova matriz reduzida e a matriz original foram submetidas ao algoritmo de *clustering Bisecting k-means*. Para a realização do *clustering*, foi utilizada a medida de similaridade *cosseño* e escolheu-se $k = 11$ para o número de *clusters* a serem formados.

A partir de cada *cluster* formado, foram selecionadas 2/3 de suas amostras para criar um modelo de regressão linear para estimar os escores dos 1/3 das amostras restantes de cada *cluster* (OLIVEIRA, 2013). Conforme os valores de escores obtidos, os níveis de aprendizagem são classificados de acordo com a escala da Tabela 2.

Tabela 2. Faixas de scores para classificação de níveis de aprendizagem

Escala de Classificação	
Scores	Classes
Acima de 4.0	A
De 2.5 a 4.0	B
De 1.5 a 2.5	C
Abaixo de 1.5	D

RESULTADOS

A Tabela 3 apresenta os resultados de predição de escores para classificação de alunos através da análise dos erros médio, mínimo e máximo. Para os testes foram utilizadas a *Base-p* e o *Cluster-p*, que é o maior *cluster* formado da *Base-p*. A *Base-p* foi dividida em 52 amostras de treino (*Tr*) e 48 amostras de teste (*Te*). Já o *Cluster-p* foi dividido em quinze amostras de treino e sete amostras de teste. Na Tabela 3 há também os resultados sem (*sAF*) e com a análise fatorial (*cAF*), respectivamente.

Tabela 3. Resultados de predição de escores

Erro	Base-p (52Tr - 48Te)		Cluster-p (15Tr - 7Te)	
	sAF	cAF	sAF	cAF
Erro Médio	0.7506	0.5964	1.4547	0.5615
Erro Mínimo	0.0096	0.0492	0.4300	0.0383
Erro Máximo	3.5000	3.5000	2.3000	0.9731

De acordo com a Tabela 3, evidencia-se que os escores de classificação são melhores preditos quando as estratégias de análise fatorial e de *clustering* são combinadas, conforme indicam os resultados do *Cluster-p* com análise fatorial (*cAF*). Por outro lado, a aplicação da análise fatorial não fez grande diferença nos resultados da *Base-p*.

A Tabela 4 apresenta os resultados de classificação de perfis de aprendizagem da *Base-p* e do *Cluster-p* sem e com a análise fatorial.

Tabela 4. Resultados de classificação

	Acertos	
	sAF	cAF
Base-p	57%	55%
Cluster-p	0	85.7%

Os resultados iniciais de experimentação da combinação das técnicas de análise fatorial e de *clustering* em amostras reais de exercícios de programação apontam, portanto, para êxitos na classificação de perfis. Isso porque selecionam-se melhor as amostras de perfis (por *clustering*) e as componentes (por análise fatorial) para gerar os modelos de regressão utilizados na predição dos *scores* de classificação de níveis de aprendizagem.

DISCUSSÃO

As técnicas de análise multivariada como a análise fatorial e o *clustering* foram aplicadas para avaliação da aprendizagem em diferentes domínios do conhecimento com o objetivo de identificar fatores que caracterizem um domínio e para reconhecimento de classes de perfis de aprendizagem.

O trabalho de (LAW, *et al.* 2010), por exemplo, por análise fatorial, apresenta um estudo preliminar que investiga os fatores-chave de motivação que afetam a aprendizagem entre estudantes de cursos de programação. Esses cursos são suportados por um sistema *online*, o *PASS (Programming Assignment Assessment System)*, que visa fornecer uma infraestrutura para facilitar a aprendizagem de programação. Este trabalho avança a proposta de Law *et. al* (2010) ao analisar as relações de componentes de habilidades geradas a partir informações de códigos-fontes escritos por estudantes para classificação de perfis de alunos conforme indicadores de habilidades e dificuldades.

No artigo de Blikstein (2011), é descrita uma técnica automatizada para avaliar, analisar e visualizar os estudantes que aprendem programação de computadores. Assim, registram-se centenas de códigos dos alunos e empregam-se diferentes técnicas quantitativas para extrair os comportamentos dos alunos e categorizá-los em termos de suas experiências em programação. Em relação a Blikstein (2011), temos o diferencial de categorizar alunos a partir das relações entre componentes de habilidades representadas em componentes latentes geradas por análise fatorial. Dessa forma, utilizando algoritmos de clustering, reunimos padrões de códigos-fontes de estudantes em classes de perfis, conforme as similaridades reconhecidas entre as componentes latentes representantes de cada perfil de estudante.

Em um trabalho mais recente, Oliveira (2013) mapeia perfis em componentes de habilidades e aplica a técnica de *clustering* para classificar esses perfis. Além disso, Oliveira (2013) mostra como algoritmos de *clustering* melhoram a seleção de características e amostras para composição de modelos de regressão utilizados na avaliação semiautomática de exercícios de programação.

Concluindo, assim como Mesic e Muratovic (2011) selecionam fatores que influenciam a dificuldade de itens de Física para identificação de preditores de modelos de regressão, avançamos os trabalhos de Oliveira (2013) e de Oliveira e Oliveira (2014) ao reduzir, por análise fatorial, o número de variáveis de perfis ao mapear as relações entre as componentes de habilidades reconhecidas em códigos-fontes em componentes latentes.

■ CONSIDERAÇÕES FINAIS

Este trabalho apresentou uma estratégia de combinação das técnicas de análise fatorial e de *clustering* para mapeamento e classificação de perfis. Os primeiros resultados indicaram que essa estratégia pode ser aplicada para melhor selecionar amostras de perfis e componentes latentes de forma a melhorar os resultados de classificação de perfis.

A proposta dessa estratégia pode ser estendida a outros domínios, desde que se definam as componentes de habilidades que representem a aprendizagem de um domínio e que essas componentes possam ser quantificadas.

O mapeamento de perfis foi realizado identificando componentes latentes que explicam as relações entre as componentes de habilidades. Como trabalhos futuros a partir deste, sugere-se que se desenvolvam mais pesquisas em relação à representação de perfis de forma que esses realmente reflitam um modelo de aprendiz e que, reduzindo a dimensionalidade, de fato caracterizem habilidades e dificuldades de aprendizagem de programação.

Espera-se, portanto, que os estudos realizados neste trabalho se consolidem em um passo inicial, mas relevante para a avaliação da aprendizagem de programação ao oferecer aos professores a possibilidade de compreenderem melhor a prática de programação de seus alunos e assim assisti-los individualmente.

■ REFERÊNCIAS

1. ANDERSON, John R. **Cognitive psychology and its implications**. Macmillan, 2005. Anderson, T. W. (1984). *An Introduction to Multivariate Statistical Analysis*. Wiley, New York, NY, second edition.
2. BISHOP, Yvonne M.; FIENBERG, Stephen E.; HOLLAND, Paul W. **Discrete multivariate analysis: theory and practice**. Springer Science & Business Media, 2007.
3. BLIKSTEIN, Paulo. Using learning analytics to assess students' behavior in open-ended programming tasks. In: **Proceedings of the 1st international conference on learning analytics and knowledge**. 2011. p. 110-116.
4. HÄRDLE, Wolfgang Karl; SIMAR, Léopold. Cluster analysis. In: **Applied Multivariate Statistical Analysis**. Springer, Berlin, Heidelberg, 2012. p. 331-349. Law, K. M. Y., Lee, V. C. S., and Yu, Y. T. (2010). Learning motivation in e-learning facilitated computer programming courses. *Comput. Educ.*, 55(1):218–228.
5. Looks, M., Levine, A., Covington, G., Loui, R., Lockwood, J., and Cho, Y. (2007). Streaming Hierarchical Clustering for Concept Mining. *Aerospace Conference, 2007 IEEE*, pages 1–12.
6. MAVRIKIS, Manolis. Machine–Learning Assessment of Students' Behavior within Interactive Learning Environments. **Handbook of Educational Data Mining**, p. 441, 2010.
7. MAZZA, Riccardo; DIMITROVA, Vania. CourseVis: A graphical student monitoring tool for supporting instructors in web-based distance courses. **International Journal of Human-Computer Studies**, v. 65, n. 2, p. 125-139, 2007.
8. MESIC, Vanes; MURATOVIC, Hasnija. Identifying predictors of physics item difficulty: A linear regression approach. **Physical Review Special Topics-Physics Education Research**, v. 7, n. 1, p. 010110, 2011.
9. MOREIRA, Mireille Pinheiro; FAVERO, Eloi Luiz. Um ambiente para ensino de programação com feedback automático de exercícios. In: **Workshop sobre Educação em Computação (WEI 2009)**. 2009.
10. MORRISON, D. F. *Multivariate Statistical Methods*, Ed. 3 McGraw. 1990.

11. NAUDÉ, Kevin A.; GREYLING, Jean H.; VOGTS, Dieter. Marking student programs using graph similarity. **Computers & Education**, v. 54, n. 2, p. 545-561, 2010.
12. OLIVEIRA, Márcia G.; OLIVEIRA, Elias. Avaliar para nivelar e formar: um sistema online de avaliação formativa para alunos de Biblioteconomia. In: **Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE)**. 2008. p. 166-175.
13. OLIVEIRA, Márcia Gonçalves; CIARELLI, Patrick Marques; OLIVEIRA, Elias. Recommendation of programming activities by multi-label classification for a formative assessment of students. **Expert Systems with Applications**, v. 40, n. 16, p. 6641-6651, 2013.
14. OLIVEIRA, M. G. **Núcleos de avaliações diagnóstica e formativa para regulação da aprendizagem de programação**. 2013. Tese de Doutorado. Tese de doutorado, Universidade Federal do Espírito Santo.
15. OLIVEIRA, Marcia; OLIVEIRA, Elias. Metodologia de Diagnóstico e Regulação de Componentes de Habilidades da Aprendizagem de Programação. In: **Anais do XXII Workshop sobre Educação em Computação**. SBC, 2014. p. 120-129.
16. PEA, Roy D.; KURLAND, D. Midian. On the cognitive effects of learning computer programming. **New ideas in psychology**, v. 2, n. 2, p. 137-168, 1984.
17. PILLAY, Nelishia. Developing intelligent programming tutors for novice programmers. **ACM SIGCSE Bulletin**, v. 35, n. 2, p. 78-82, 2003.
18. PIMENTEL, Edson Pinheiro; FRANÇA, Vilma Fernandes de; OMAR, Nizam. A caminho de um ambiente de avaliação e acompanhamento contínuo da aprendizagem em Programação de Computadores. In: **II Workshop de Educação em Computação e Informática do Estado de Minas Gerais (WEIMIG'2003)**. Poços de Caldas, MG, Brasil. 2003.
19. POWERS, Kris et al. Tools for teaching introductory programming: what works?. In: **Proceedings of the 37th SIGCSE technical symposium on Computer science education**. 2006. p. 560-561.
20. STEINBACH, Michael; KARYPIS, George; KUMAR, Vipin. A comparison of document clustering techniques. 2000.

Análise sobre a evolução da mineração de dados educacionais segundo a perspectiva dos principais estudos da área: uma revisão narrativa

I Vanessa Faria de **Souza**
IFRS

I Mayara Faria de **Souza**
UTFPR

RESUMO

Com o aumento da disponibilidade de dados, sobretudo no contexto educacional, surgiram áreas específicas para extração de informações relevantes, como a Mineração de Dados Educacionais (MDE), que integra inúmeras técnicas que dão suporte a captação, processamento e análises desses conjuntos de registros. Essa área tem evoluído constantemente nos últimos 20 anos desde os primeiros trabalhos publicados, e tem conquistado cada vez mais espaço, tanto entre pesquisadores da área da educação, como da ciência de dados. Nesse sentido, esse estudo de revisão narrativa aborda e apresenta os principais aspectos da MDE, seus objetivos e processo, e apresenta o contexto de evolução da área com base em estudos de revisão sistemática de literatura que sintetizaram boa parte dos estudos de MDE das últimas duas décadas.

Palavras-chave: Mineração de Dados Educacionais, Revisão Narrativa, Evolução.

■ INTRODUÇÃO

As organizações de diversos seguimentos têm se mostrado muito eficientes em capturar, organizar e armazenar bases dados de grande dimensão, conseguidos por meio operações diárias ou pesquisas, entretanto a chave para avançar é saber como utilizar adequadamente essa grande quantidade de dados para transformá-la em conhecimentos que possam ser utilizados em suas atividades, sejam elas comerciais ou científicas. Nesse sentido, o conceito de Mineração de Dados (*Data Mining*) está se tornando cada vez mais popular como um método para exploração de informações, que podem apontar conhecimentos novos, que possam orientar em decisões em situações de certeza restrita.

Para Aggarwal (2015) a Mineração de Dados (MD) é definida como:

A mineração de dados é o estudo de coleta, limpeza, processamento, análise e obtenção de informações e ideias úteis de dados. Existe uma grande variação em termos de domínios problemáticos, aplicativos, formulações e representações de dados encontradas em aplicativos reais. Portanto, “Mineração de dados” é um termo abrangente usado para descrever esses diferentes aspectos de processamento de dados (AGGARWAL, 2015).

Com a crescente adoção de MD, essa passou a ser empregada com sucesso também no contexto educacional, auxiliando em diversos cenários, e ficou conhecida como Mineração de Dados Educacionais (MDE). Nesse sentido este estudo tem como objetivo sistematizar estudos que contribuíram para a consolidação da MDE como uma área de pesquisa em potencial expansão, para isso são apresentados primeiramente sua definição, seus objetivos e processo de funcionamento e em seguida é detalhado o contexto de evolução da área.

■ DESENVOLVIMENTO

Nessa seção de desenvolvimento serão abordados os principais aspectos sobre a Mineração de Dados Educacionais, descritos com base nos principais autores da área. Para sintetizar o processo de evolução foram selecionados os estudos de revisão sistemática de literatura mais proeminentes e abrangentes disponíveis entre as bases de dados: IEEE Xplore, ACM Digital Library, ERIC, Scopus e ScienceDirect. Nesse sentido essa seção inicia abordando as definições, objetivos e processo da MDE, em seguido é apresentado o seu contexto de evolução e consolidação.

Definições e Objetivos da Mineração de Dados Educacionais

Nos últimos anos a educação tem se modificado, em decorrência do avanço tecnológico disponível que direcionou a uma instrumentação do setor educacional, tanto em

softwares voltados para o ensino, como na administração digital dos registros acadêmicos pelos gestores das instituições, bem como no uso da internet para a aprendizagem, em especial pela popularização do *e-learning*. Todos esses fatores impulsionaram um crescimento exponencial no volume de dados educacionais, e para se analisar uma grande quantidade de dados, é imprescindível contar com recursos computacionais, caso contrário a tarefa torna-se impraticável.

Dessa forma, as técnicas de mineração de dados estão ganhando cada vez mais importância no setor educacional, pois são uma forma de acompanhar, analisar e avaliar o processo de aprendizagem. Provavelmente, as técnicas de mineração de dados podem fornecer aos formuladores de políticas educacionais modelos para apoiar seus objetivos de aprimorar a eficiência e a qualidade do ensino e da aprendizagem. Além disso, o uso de diferentes técnicas de mineração de dados pode ser visto como base para uma mudança sistêmica, capaz de impactar de maneira positiva nas soluções de problemas específicos das Instituições de Ensino, por exemplo, viabilizando soluções que envolvem a personalização dos ambientes educacionais ou fornecendo suporte para o processo de tomada de decisão no ambiente educacional.

Nesse cenário, destaca-se a MDE que utiliza as técnicas da MD para extrair informações relevantes de conjuntos diversificados de dados educacionais. Segundo a Sociedade Internacional de Mineração de Dados Educacionais¹, a MDE pode ser definida da seguinte forma:

É uma disciplina emergente, preocupada com o desenvolvimento de métodos para explorar dados únicos e cada vez mais em larga escala, provenientes de contextos educacionais e usa esses métodos para entender melhor os alunos e as configurações em que aprendem (EDM, 2020).

Em outras palavras, a MD refere-se a um conjunto de técnicas computacionais para extrair informações de grandes massas de dados, e quando os dados analisados são provenientes de contextos educacionais, chama-se MDE (ROMERO; VENTURA, 2013). Igualmente, De Los Reyes *et al.* (2019) define MDE como uma área voltada ao desenvolvimento de métodos para explorar dados oriundos de ambientes educacionais e utilizá-los para compreender melhor os processos de ensino e aprendizagem. Nessa acepção, Baker, Isotani e Carvalho (2011) alegam que a MDE é definida como a área de pesquisa que tem como foco o desenvolvimento de técnicas para explorar conjuntos de dados coletados em ambientes educacionais. Conforme os autores, a natureza destes dados é mais diversa do que a observada nos dados tradicionalmente utilizados em tarefas de mineração, demandando adaptações e novas abordagens. Ao mesmo tempo, essa diversidade nos dados

1 <http://educationaldatamining.org/>

representa um potencial de implementação de recursos fundamentais para auxílio na melhoria da educação (BAKER; ISOTANI; CARVALHO, 2011; DE LOS REYES *et al.*, 2019; RIGO *et al.*, 2014).

Sendo assim, necessita-se de técnicas e ferramentas que auxiliem na tarefa de verificar, interpretar e relacionar esses dados, com o intuito de gerar conhecimento útil e relevante, o que, segundo De Los Reyes *et al.* (2019) já era um objetivo das técnicas de MD, empregadas para identificar padrões de comportamento e encontrar *insights* que provoquem melhorias em produtos e serviços. Romero e Ventura (2007) elencam questões que diferenciam a MDE da MD em outros domínios:

1. *Objetivos*: que podem se relacionar à pesquisa (a) aplicada, que busca responder questões práticas, por exemplo: como melhorar o processo de aprendizagem; e (b) pura, com a finalidade de por exemplo dar sentido às observações. Na maioria das vezes esses objetivos são difíceis de quantificar e exigem seu próprio conjunto especial de técnicas de medição.
2. *Dados*: em ambientes educacionais, existem muitos tipos diferentes de dados disponíveis para mineração. Esses dados são específicos da área educacional, portanto, possuem informações semânticas intrínsecas, relacionamentos com outros dados, e vários níveis de hierarquia significativa.
3. *Técnicas*: problemas educacionais têm algumas características especiais que exigem que a questão da mineração seja tratada de uma maneira diferente. Embora, a maioria das técnicas tradicionais de MD possam ser aplicadas diretamente, outras não podem e devem ser adaptadas ao problema educacional específico. Exemplo disso, é que em se tratando de cenários comuns de MD, as variáveis são em sua maioria numéricas, tratáveis diretamente por algoritmos de *Machine Learning*, enquanto que em ambientes educacionais a grande maioria é categórica, o que implica esforço em pré-processamento e transformações, para codificar essas variáveis em numéricas, para que então possam ser interpretadas pelos algoritmos.

O processo de Mineração de Dados Educacionais

Além de definir os diferenciais da MDE é importante explicar como é seu funcionamento, pois o processo de MDE não é trivial contendo várias fases. Nesse sentido, será explicado o processo de MD de acordo com Aggarwal (2015), depois será apresentada uma sequência de etapas considerada mais adequada, baseada em pesquisas (BAKER; ISOTANI; CARVALHO, 2011; DE LOS REYES *et al.*, 2019; RIGO *et al.*, 2014; ROMERO; VENTURA, 2013, 2020) específicas de MDE e também em experiências com aplicações.

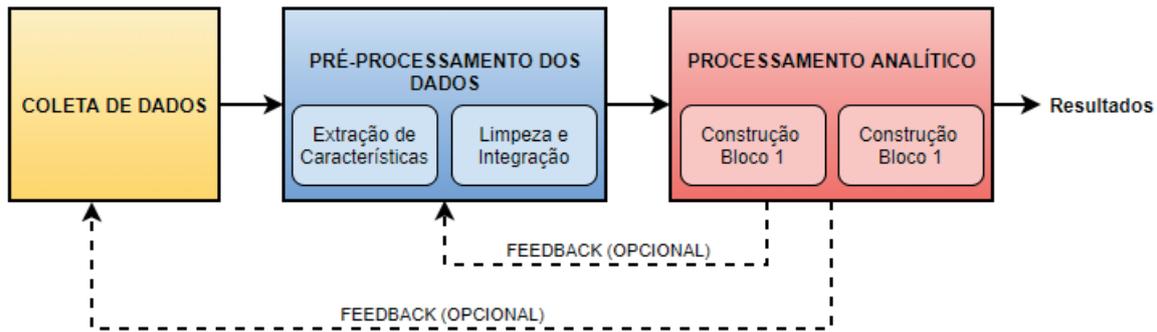
Aggarwal (2015) determina que o fluxo de trabalho de um processo típico de MD contém as seguintes fases:

1. Coleta de dados: A coleta de dados pode exigir o uso de hardware especializado, como uma rede de sensores, trabalho manual, como a coleta de pesquisas com usuários ou ferramentas de software como um mecanismo de rastreamento de documentos da Web para coletar documentos. Esta etapa é específica da plataforma e geralmente fora do domínio do analista de mineração de dados. Após a fase de coleta, os dados geralmente são armazenados em um banco de dados ou, em geral, um *data warehouse*²(tradução livre – armazém de dados) para processamento.
2. Extração de recursos e limpeza de dados (Pré-Processamento e Transformação): Quando os dados são coletados, eles geralmente não estão em um formato adequado para processamento. Para tornar os dados adequados para processamento, é essencial transformá-los em um formato que seja interpretável aos algoritmos de mineração, como multidimensionais, séries temporais ou formato semiestruturado. O formato multidimensional é o mais comum, no qual diferentes campos de dados correspondem às diferentes propriedades medidas que são chamadas de atributos. A fase de extração de recursos geralmente é realizada em paralelo com a limpeza de dados, onde partes ausentes e incorretas dos dados são estimadas ou corrigidas. Em muitos casos, os dados podem ser extraídos de várias fontes e precisam ser integrados em um formato unificado para processamento. O resultado final deste procedimento é um conjunto de dados estruturados, que pode ser efetivamente usado por um programa de computador. Depois da fase de extração de recursos, os dados podem ser armazenados novamente em um banco de dados para processamento.
3. Processamento analítico e algoritmos: A parte final do processo de mineração é projetar métodos analíticos eficazes a partir dos dados processados.

A sequência das etapas do processo de MD proposto por Aggarwal (2015) é apresentada na Figura 01.

² É um repositório central de informações que ficam disponíveis para serem analisadas e dão suporte ao processo de tomada de decisão.

Figura 1. Processo de MD proposto por Aggarwal (2015)



Fonte: Adaptado de Aggarwal (2015)

A partir do processo de MD apresentado, pesquisas em MDE e na experiência com aplicações, algumas das etapas foram ajustadas para condizer efetivamente com o que é executado, levando em conta o fator usuário (programador ou cientista de dados), nesse sentido o processo de MDE pode ser formado pelo seguinte conjunto de etapas:

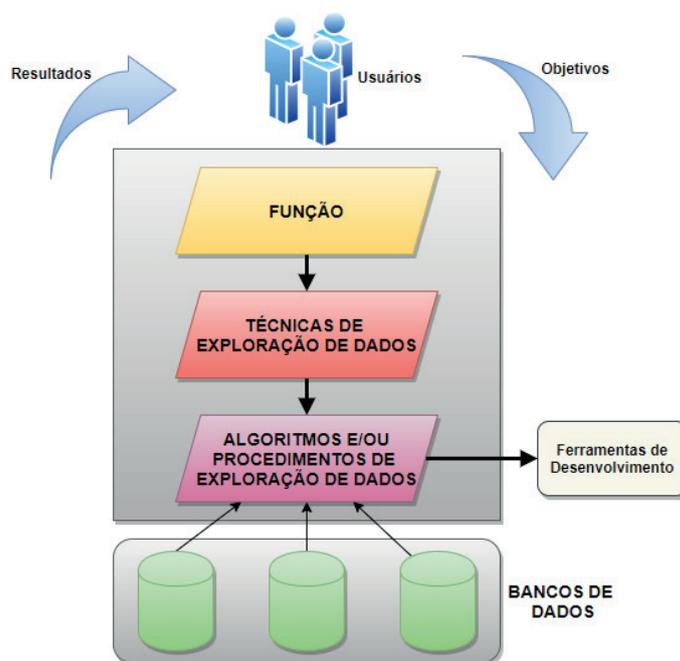
1. Definição da função da MDE: Será realizada a determinação do objetivo do processo MDE, para qual finalidade ela está sendo aplicada, como por exemplo: identificação de padrões, detecção de desvio, segmentação, sistemas de recomendação, análise de ligações e regras de associação, sumarização e visualização, mineração de textos, afinidade em grupos, descrição de grupos, para isso é preciso especificar que tipo de conhecimento pretende-se extrair dos dados.
2. Formatação dos dados que serão utilizados: Diversos tipos de armazenamentos de dados e de bancos de dados podem ser manipulados no processo de mineração, cabe ao usuário definir qual formato é o mais adequado para aplicação das técnicas de mineração selecionadas, outro ponto importante é que baseado no tipo de conjunto de dados disponível para análise é que definem-se os padrões, relacionamentos ou informações que se consegue minerar. Nessa etapas todas as incoerências na base devem ser corrigidas e caso for necessário podem ser acrescentados mais atributos que sejam resultantes da combinação de outros ou que possam ser deduzidos de outros, como a idade que pode ser calculada a partir da data de nascimento, ou o total de atividades realizadas que pode ser efetuada por meio de uma soma.
3. Definição das Técnicas de MDE: A definição das técnicas é um passo importante, pois elas devem ser específicas para o tratamento da função de MDE estabelecida, as técnicas mais utilizadas para MDE são: a Estatística Descritiva, a Aprendizagem de Máquina e mais recentemente tem sido empregada a Aprendizagem Profunda, cabe salientar que cada uma dessas técnicas possui algoritmos, rotinas e/ou pro-

cedimentos específicos para manipulação dos dados. A literatura (AGGARWAL, 2015; BAKER; ISOTANI; CARVALHO, 2011; DE LOS REYES *et al.*, 2019; RIGO *et al.*, 2014), em muitos casos, não deixa claro as diferenças entre funcionalidades, funções e técnicas. Por exemplo, um problema que requer identificar alunos desistentes ou concluintes, para isso deve-se pensar em técnicas que possibilitem identificar esses padrões de comportamento nos alunos nessas duas categorias, isso seria a função, o propósito da MDE o de classificar esses alunos, entretanto também é possível entender como a técnica que seria aplicada, algoritmos de Aprendizagem de Máquina para classificação, nesse caso, por isso as vezes fica confusa a interpretação.

4. Delineamento de como essas técnicas serão aplicadas: Nessa etapa são selecionadas as ferramentas que vão dar suporte ao desenvolvimento de sistemas capazes de processar os dados, e gerar os resultados esperados.

Na Figura 04 pode-se observar as relações entre as etapas enunciadas, tendo em vista explicar a interatividade da função da MDE com as técnicas a serem utilizadas.

Figura 2. Etapas do processo de MDE



Fonte: Adaptado Aggarwal (2015) e Rometo e Ventura (2020)

A MDE, levando em consideração seus objetivos, técnicas e processo de funcionamento, formam uma importante metodologia para apoio a qualquer cenário educacional. Por meio da MDE, talvez seja possível acompanhar e compreender o processo de aprendizagem, bem como outros fatores que a influenciam. Como por exemplo, identificar que tipo de abordagem

instrucional (e.g. aprendizagem individual ou colaborativa) proporciona mais benefícios ao aluno, observando variáveis que representem seu engajamento com o curso. Além disso, abre-se a possibilidade de verificar se o aluno está aprendendo ou confuso, identificar níveis de motivação, envolvimento nas atividades on-line, descoberta de elementos ou indicadores comportamentais de conclusão e sucesso em um curso, identificar padrões de interação, descobrir estratégias que contribuam para a permanência dos estudantes (PURSEL *et al.*, 2016), bem como detectar possíveis fraudes, ou trapanças no sistema de aprendizagem. Estes fatores podem ajudar a personalizar o ambiente e os métodos de ensino, para oferecer melhores condições de aprendizagem (BAKER; ISOTANI; CARVALHO, 2011).

Em suma, a MDE tem se desenvolvido e pode ser considerada como uma das formas mais promissoras para extração de informações de bases de dados educacionais e suas técnicas têm se tornado cada vez mais eficientes e eficazes, graças ao número crescente de dados disponíveis e dos avanços computacionais. Para entender melhor como se deu esse desenvolvimento e como a MDE tem sido empregada no decorrer do tempo, são descritos os aspectos de sua evolução na sequência.

A evolução da Mineração de Dados Educacionais

A disponibilidade de grandes bases de dados educacionais, fomentada pelas modernas plataformas e mídias educacionais, combinadas com avanços na computação, formam a composição ideal para o surgimento da MDE. Embora existam relatos sobre publicações a respeito deste tema desde 1995 (ROMERO; VENTURA, 2007) o primeiro *workshop* foi realizado em 2005, em Pittsburgh, Pensilvânia, tendo sido seguido por várias oficinas e, em 2008, ocorreu a 1.^a Conferência Internacional sobre MDE realizada em Montreal, Quebec. As conferências anuais sobre MDE impulsionaram o surgimento do *Journal of Educational Data Mining*, que publicou sua primeira edição em 2009, na sequência, o primeiro manual de MDE foi publicado em 2010 (ROMERO *et al.*, 2010). Posteriormente, em 2011 a Sociedade Internacional de Mineração de Dados Educacionais foi formada com o objetivo de promover pesquisa científica na área interdisciplinar da MDE, organizando as conferências e os periódicos. No campo das publicações, uma primeira revisão de literatura foi apresentada por Romero e Ventura (2007), seguido de um modelo teórico proposto por Baker e Yacef (2009), e uma revisão bem mais abrangente sobre MDE foi desenvolvida por Romero e Ventura (2010). Na sequência, outras publicações iniciaram um amplo movimento de pesquisas nesse âmbito, e tiveram grande notoriedade.

Neste sentido, muitas publicações sobre MDE surgiram nos últimos anos, e algumas delas se dedicaram a implementar revisões de literatura sobre essa área. Tais estudos auxiliam no entendimento de como a MDE tem sido aplicada em diversos contextos educacionais,

seus objetivos, as técnicas mais utilizadas, verificação de resultados alcançados e validação dos benefícios proporcionados, identificação de avanços e também desafios que têm sido relatados por pesquisadores da área.

À vista disso, chamaram a atenção algumas publicações, que trouxeram grandes contribuições para pesquisadores interessados em MDE, quais sejam: Shahiri, Husain e Rashid (2015); Sukhija, Jindal e Aggarwal (2016); Schwendimann *et al.* (2017); Aldowah, Al-Samarraie e Fauzy (2019); e Romero e Ventura (2020), e foram fundamentais para um aperfeiçoamento e compreensão da evolução da MDE no decorrer de sua consolidação como área de pesquisa e são sintetizadas na sequência.

A primeira revisão sobre MDE analisada foi desenvolvida por Shahiri, Husain e Rashid (2015), essa forneceu uma visão geral das técnicas de mineração de dados que eram usadas para prever o desempenho dos alunos, em publicações datadas entre 2002 e 2015. O estudo também se concentrou em como os algoritmos de previsão poderiam ser usados para identificar os atributos mais importantes dentre a diversidade de dados dos alunos. Nessa revisão, os autores seguiram duas questões de pesquisa para estruturar os resultados: 1) Quais são os atributos mais importantes empregados na previsão do desempenho dos alunos; e 2) Quais as técnicas/algoritmos de previsão mais eficientes. Quanto aos principais atributos, Shahiri, Husain e Rashid (2015) apontam que foram usados com frequência a média cumulativa de notas e a avaliação interna usada por 10 dos 30 artigos selecionados para a revisão. Os autores também chegaram à conclusão, que a *Machine Learning* era a técnica mais usada e quanto à eficácia dos algoritmos as Redes Neurais tiveram a maior precisão (98%) para previsão do desempenho dos alunos, seguida das Árvores de Decisão (91%), depois as Maquinas de Vetores de Suporte e KNN com a mesma eficácia (83%), por fim, o método menos preciso foi o Naive Bayes (76%). Os autores afirmaram ainda, que prever o desempenho dos alunos é muito útil para ajudar educadores e alunos a melhorar o processo de ensino e aprendizagem. No entanto, é importante ressaltar que os índices de eficácia são resultado da interação entre a complexidade da questão de pesquisa com a qualidade (e algumas vezes, a extensão da base) dos dados, não sendo uma avaliação a respeito dos métodos em si. Aponta-se como ponto negativo da revisão, que algumas das pesquisas citadas pelos autores não apontam uma diversidade de métricas de avaliação dos algoritmos – uma informação importante, que deveria ser divulgada como parte dos resultados

Na sequência foi analisada a revisão sistemática desenvolvida por Sukhija, Jindal e Aggarwal (2016) que descreveram a evolução da MDE, trazendo à tona os aspectos e resultados de vários estudos divididos em 3 gerações: 1.^a geração de 2001 a 2005; 2.^a geração de 2006 a 2010; e 3.^a geração, de 2011 até 2015. No período de 2001 a 2005, as pesquisas se basearam no uso da MDE como uma ferramenta para antecipar os padrões que ajudam

na avaliação de cursos on-line. Dessa forma, os registros de dados e registros de atividades dos alunos foram usados para analisar seu comportamento. Os autores evidenciaram, que o início da literatura relacionada à MDE, produziu pesquisas com inclinação para o ambiente de aprendizagem baseado na *Web*, devido especialmente à grande disponibilidade de dados em cursos on-line. No final deste período, as pesquisas estavam com foco no uso de algoritmos evolutivos para mineração de dados da internet.

Referente ao período de 2006 a 2010, a MDE evoluiu e os estudos passaram a buscar a aplicação de algoritmos mais eficientes. Os bancos de dados usados se tornaram provenientes de sistemas de ensino à distância, baseados na *Web* e vinculados a grandes Instituições de Ensino, bem como o tamanho desses bancos de dados aumentou. Além disso, houve uma inclinação dos pesquisadores para análises preditivas de dados, com relação a prever os problemas e identificar os alunos em potencial, com alta probabilidade, de apresentar um desempenho acadêmico ruim, nesse sentido, sistemas de apoio à decisão para equilibrar a demanda e a oferta educacional também foram desenvolvidos. Durante esse período, a implementação da classificação baseada em Árvores de Decisão e Redes Neurais se acentuaram no contexto educacional, com diferentes objetivos. Ademais, citam que a pesquisa na área foi direcionada a dados relacionados ao ensino superior, bem como técnicas como *Online Analytical Processing (OLAP)*³ em combinação com DELPHI⁴ foram muito utilizadas.

Finalmente, no que tange ao período de 2011 a 2015, a MDE evoluiu para incorporar técnicas melhores e mais eficientes, conseguindo integrar novas e mais eficientes regras de associações, ferramentas como WEKA⁵ começaram a ganhar popularidade, algumas pesquisas se voltaram a utilizar dados também do ensino médio e a comunidade de pesquisa trabalhou em função de uma aceitação comercial das técnicas de MD na educação. Os pesquisadores forneceram evidências para maior correlação entre diferentes construtos no sistema educacional, levando a uma maior anuência dos resultados obtidos entre estudantes e autoridades. Além disso, extensos estudos foram realizados na busca de encontrar uma solução para as altas taxas de abandono em diferentes contextos acadêmicos, nessa perspectiva também uma grande quantidade de estudos para estimar o desempenho dos estudantes, foram implementados. Sukhija, Jindal e Aggarwal (2016) ainda expuseram, que os bancos de dados usados neste período ficaram consideravelmente maiores do que os anteriores e essa crescente no volume dos dados foi acompanhada do desenvolvimento de novas técnicas para MDE, como por exemplo, a retomada dos estudos com *Deep Learning*.

3 OLAP é um conceito de interface com o usuário que proporciona a capacidade para manipular e analisar um grande volume de dados sob múltiplas perspectivas.

4 O Delphi é comumente conhecido como uma linguagem de programação. Mas, na verdade abrange um kit de desenvolvimento de software e ganhou propagação entre profissionais da área de desenvolvimento no segmento de aplicações de desktop, entretanto hoje é integrado e utilizado também para aplicações *Web* e *mobile*.

5 Weka é um Software livre do tipo *open source* para mineração de dados, desenvolvido em Java. Ao longo dos anos se consolidou como a ferramenta de data mining mais utilizada em ambientes acadêmicos.

Sukhija, Jindal e Aggarwal (2016) apontam cinco lacunas na área de MDE: (1) indisponibilidade de conjuntos de dados consistentes que sejam grandes o suficiente para refletir o sistema educacional e seu funcionamento; (2) necessidade de integração e versatilidade nos conjuntos de dados; (3) grande parte das técnicas de mineração foram aplicadas isoladamente e poucos trabalhos foram realizados utilizando técnicas híbridas; (4) havia falta de confiança das autoridades nos resultados da MDE e (5) necessidade de comparar métodos. Pode-se dizer que embora as descobertas dos autores fossem fortemente fundamentadas, o cenário se modificou bastante deste então.

No que se refere, à primeira lacuna ressalta-se que com a evolução dos cursos *e-learning*, bases com milhões de dados estão disponíveis, como exemplo, pode-se aludir ao trabalho desenvolvido por Northcutt, Ho e Chuang (2016), onde foi utilizada uma base de dados gerada a partir de uma plataforma de *Massive Open Online Course* (Curso Online Aberto e Massivo) com 1.893.092 de usuários, que produziram em média de 200 a 1500 interações com a plataforma, cada um, por curso realizado, portanto, a indisponibilidade de conjuntos de dados já não se configura mais como um problema. No que tange à segunda limitação pode-se dizer que em relação à integração das bases, ela se mantém, pois, não é possível integrar duas bases de forma simples, sem necessidade de um grande esforço de pré-processamento. Em relação à versatilidade dos dados – qualidade de não ser colinear, ou seja, dos dados não estarem relacionados – pode-se dizer que houve mudança, pois, vários tipos diferentes de dados são usados nos modelos.

A terceira lacuna, sobre uso de métodos híbridos, pode ser considerada a que mais não coincide com a realidade dos experimentos realizados na área de MDE atualmente, pois muitos pesquisadores têm empregado técnicas de MDE em conjunto com outras ferramentas de pesquisa como em Gallén e Caro (2017) que utilizaram algoritmos de Agrupamento e um Questionário respondido pelos alunos para analisar os motivos pelos quais uma pessoa se inscreve em um MOOC. Ademais, pode-se citar como exemplo o trabalho desenvolvido por Nen-Fu et al. (2018) que propõem um método para usar resultados de questionários e tipos de grupos predefinidos, por meio do algoritmo K-means, para classificar os alunos em MOOCs, com o intuito entender sua auto-organização nas fases iniciais de um curso, pois acreditam haver um ligação entre a motivação e comportamento de aprendizagem. Com relação, à quarta limitação supõe-se que com os grandes avanços tecnológicos disponíveis e a consolidação da Inteligência Artificial (IA), presente no cotidiano das pessoas, a aceitação da MDE como aporte para tomada de decisões no setor acadêmico tenha crescido.

Enfim, quanto à quinta lacuna, é possível destacar que devido aos avanços tecnológicos muitos pesquisadores tem se dedicado a comparar novas técnicas de MDE com outras mais consolidadas, para verificação da eficácia. Nesse sentido, Gao et al. (2019) propuseram um

novo modelo para analisar o perfil de aprendizagem e o engajamento de alunos em MOOCs e para validar sua precisão os autores o compararam com algoritmos amplamente utilizados – Regressão Linear e Máquinas de Vetores de Suporte. Além desse exemplo, aponta-se a pesquisa de Waheed *et al.* (2020), que tinha como objetivo prever o abandono em MOOCs, utilizando um rede neural de Aprendizado Profundo, que foi comparada aos algoritmos de Regressão Logística e Máquinas de Vetores de Suporte.

A outra revisão incluída neste estudo foi elaborada por Schwendimann *et al.* (2016), que incluía, além de MDE Análise de Aprendizado (AA) e painéis de aprendizado. Os autores cunharam o termo “painéis de aprendizado” e o definiram da seguinte forma: “uma única exibição que agrega diferentes indicadores sobre aluno(s), processo(s) de aprendizagem e/ou contexto(s) de aprendizagem em uma ou várias visualizações”. Nesse sentido, os autores afirmaram que os painéis de aprendizado estão se tornando populares devido ao aumento do uso de tecnologias educacionais, como Sistemas de Gerenciamento de Aprendizagem (termo em inglês: *Learning Management Systems* - LMS) para a execução de diversos tipos de cursos em EAD como os MOOCs, que na opinião deles constituem a base para o desenvolvimento das áreas de AA e MDE.

A revisão de Schwendimann *et al.* (2016) foi realizada em 6 bases de dados: ACM Digital Library, IEEE Xplore, SpringerLink, Science Direct, Wiley e Google Scholar, retornando 346 artigos no total, dos quais 55 artigos foram incluídos na análise final. A revisão distinguiu entre 2 tipos de contribuições: Artigos que contribuíram com uma proposta teórica ou referencial (3 artigos; 5%); Artigos que descreveram a implementação de um painel de aprendizado (39 artigos; 71%), além de 13 artigos (5%) que apresentaram uma combinação dessas duas. A revisão finalizou delineando questões em aberto e futuras linhas sobre como trabalhar na área de painéis de aprendizado, indicando ainda que há uma necessidade longitudinal de pesquisas e de captação de dados em ambientes virtuais de aprendizado, assim como estudos que comparem sistematicamente designs de painéis diferentes.

A princípio o foco na MDE, pela revisão relatada, parece ficar difuso, entretanto no decorrer da análise percebe-se sua relação quando Schwendimann *et al.* (2016) citam os painéis de aprendizagem como subsídios para o desenvolvimento das áreas de MDE e AA. Tais painéis se constituem como uma forma sistemática de organização dos dados disponíveis dos alunos, para a finalidade de investigação. Os autores salientam ainda que os painéis apresentados nas publicações analisadas utilizaram principalmente os relatórios de Logs das atividades dos estudantes como fonte de dados, e algumas dessas pesquisas usaram Interface de Programação de Aplicativos (termo em inglês: *Application Programming Interface* – APIs) externas, bem como atividades escritas dos estudantes, e inclusive bancos de dados institucionais. Nesse sentido, a revisão de Schwendimann *et al.* (2016) auxiliou

a entender melhor quais tipos de dados podem servir ao propósito de examinar, por meio de MDE, o comportamento dos alunos, resultante da navegação em uma plataforma de ensino e aprendizagem.

Na sequência, a revisão de Aldowah, Al-Samarraie e Fauzy (2019) teve como foco o tema Mineração de Dados Educacionais e Análise de Aprendizado para o século XXI no ensino superior. Os autores relataram que as revisões anteriores sobre MDE e AA forneceram informações substanciais sobre a base teórica desse campo em rápido crescimento, no entanto, tais estudos não consideraram a associação entre diferentes técnicas de MDE e AA na resolução de problemas educacionais específicos e não produziram uma classificação clara das dimensões em que essas técnicas poderiam ser aplicadas com sucesso no ensino superior.

Aldowah, Al-Samarraie e Fauzy (2019) sugerem que essa revisão poderia ser aplicada como um guia para futuros estudos sobre o uso de técnicas de MDE e AA com o propósito de resolver problemas específicos de ensino e aprendizagem. Os autores conduziram a revisão com o intuito de responder a duas questões: “Como usar MDE e AA para resolver desafios práticos na educação?” e “Quais técnicas de mineração são mais adequadas para esses problemas?”. Para responder a essas perguntas, buscaram artigos das bases de periódicos e conferências: Scopus, Web of Science, Google Scholar, ERIC, Science Direct, DBLP, ACM Digital Library, IEEEXplore e Springer, encontrando 491 artigos publicados de 2000 até 2017.

Aldowah, Al-Samarraie e Fauzy (2019) afirmaram que as técnicas podem ser agrupadas em 4 grandes dimensões: Análise de Aprendizagem Suportada por Computador (*Computer-Supported Learning Analytics – CSLA*); Análise Preditiva Suportada por Computador (*Computer-Supported Predictive Analytics – CSPA*); Análise Comportamental Suportada por Computador (*Computer-Supported Behavioral Analytics – CSBA*); e Análise de Visualização Suportada por Computador (*Computer-Supported Visualization Analytics – CSVA*). Os resultados trazidos pelos autores podem ser sintetizados da seguinte forma: (1) As pesquisas sobre CSLA (120 artigos) concentraram-se principalmente no uso de análise estatística de dados para executar tarefas analíticas sofisticadas, a fim de investigar os comportamentos de aprendizagem colaborativa e de busca de informações dos alunos no contexto de um curso; (2) Os estudos sobre CSPA (253 artigos) focaram, em sua grande maioria, no uso de funções preditivas ou variáveis contínuas para sugerir maneiras eficazes de melhorar o aprendizado e o desempenho dos alunos, bem como avaliar a adequação do aprendizado; (3) As publicações sobre a dimensão CSBA (80 artigos), em maior parte, criaram modelos de comportamento, ações e conhecimento; (4) Os estudos sobre CSVA (38 artigos) concentraram-se em métodos para explorar visualmente os dados – usando gráficos interativos por

exemplo – para destacar informações úteis e produzir decisões precisas sobre as informações novas descobertas nos dados.

Os autores relataram que mineração sequencial de padrões, mineração de texto, mineração de correlação, detecção de outlier e mineração de estimativa de densidade, não são comumente usadas devido à complexidade na obtenção dos atributos necessários para regular ou adaptar-se às necessidades dos dados educacionais. Além disso, descobriram que na CSPA há uma taxa mais alta de tarefas de classificação, devido essa ser aceita como uma técnica eficaz para prever padrões de interesse e formar modelos de aprendizagem, promovendo tarefas específicas nesse sentido. Pode-se afirmar que essa revisão forneceu informações substanciais sobre a base teórica, metodológica e objetivos dessas áreas em expansão.

Finalizando a explanação sobre as revisões focadas em MDE e temas correlatos, que deram suporte ao desenvolvimento desse estudo, destaca-se Romero e Ventura (2020), que efetuaram um estudo sobre a MDE e AA, atualizando revisões anteriores (ROMERO; VENTURA, 2007, 2010, 2013, 2017). Partindo dessa premissa, a publicação de 2020 forneceu informações sobre o estado da arte, revisando as publicações da área no sentido de elucidar: os principais marcos; o ciclo de descoberta de conhecimento; os ambientes educacionais mais utilizados; as ferramentas específicas desenvolvidas; os conjuntos de dados disponíveis gratuitamente; os métodos e técnicas mais empregados; os principais objetivos; e por fim, as tendências futuras nessa área de pesquisa. Devido à grande amplitude da revisão, são abordados os itens considerados mais relevantes: as principais alterações na MDE e na AA e suas conclusões.

No que diz respeito as principais mudanças, Romero e Ventura (2020) afirmaram que na de 2010 a 2020 a MDE, como área de pesquisa, evoluiu enormemente e uma ampla gama de expressões relacionadas surgiram no estado da arte, os autores citam como principais as apresentadas no Quadro 1. Quanto às conclusões, Romero e Ventura (2020), apresentaram apontamentos relacionados aos seguintes aspectos: (1) A importância e evolução da MDE e AA; (2) Uma avaliação se as tendências encontradas na publicação de Romero e Ventura (2013) haviam se concretizado; e (4) tendências em MDE e AA.

Quadro 1. Expressões Relacionadas a MDE

EXPRESSÕES	DEFINIÇÃO
<i>A Academic Analytics (AA) e a Institutional Analytics (IA)</i>	Se preocupam com a coleta, análise e visualização de atividades do programa acadêmico como: cursos de graduação, cursos EAD, avaliação de cursos, alocação de recursos e gerenciamento para gerar <i>insights</i> institucionais. Portanto, estão focadas no desafio político/econômico.
<i>Teaching Analytics (TA)</i>	Refere-se à análise das atividades de ensino, aprendizagem, dados de desempenho, bem como do design, desenvolvimento e avaliação dessas atividades, está focada no desafio educacional do ponto de vista dos instrutores.
<i>Data-Driven Education (DDE) e Data-Driven Decision-Making in Education (DDDM)</i>	Referem-se a coletar e analisar sistematicamente vários tipos de dados educacionais, para orientar uma série de decisões com o intuito de ajudar a melhorar o sucesso de alunos e escolas.
<i>Big Data in Education (BDE)</i>	Denota à aplicação de <i>Big Data</i> – conotação básica resumida em volume, variedade e valor – para dados do ambiente educacional.
<i>Educational Data Science (EDS)</i>	Edefinida como o uso de dados coletados de ambientes/configurações educacionais para resolução de problemas nesse contexto. A ciência de dados é um conceito para unificar estatísticas, análise de dados, <i>Machine Learning</i> , métodos e técnicas relacionados.

Fonte: Romero e Ventura (2020)

No que se refere à importância da MDE e a AA, os autores salientaram que essas são duas comunidades interdisciplinares de cientistas da computação, cientistas de aprendizagem, psicometristas e pesquisadores de diversas áreas, mas todos com o mesmo objetivo, o de melhorar o aprendizado a partir dos dados. Os autores sugeriram que a área cresceu rapidamente nas últimas duas décadas, com duas conferências anuais⁶, dois periódicos específicos⁷ e com o aumento do número de livros, artigos, pesquisas e resenhas. Além disso, há uma corrente para mudar as pesquisas restritas a laboratórios para o mercado em geral, assim empregando a MDE e AA em instituições educacionais e escolas de todo o mundo. Os autores declararam, que em 2020 toda pesquisa educacional com relativa importância para o cenário envolve análise e mineração de dados. Isso indica que essas áreas se tornarão em breve maduras e amplamente utilizadas não apenas pelos pesquisadores, mas também por instrutores, administradores educacionais e negócios relacionados, em todo o mundo.

Com relação à comparação entre as tendências levantadas em sua pesquisa anterior (ROMERO; VENTURA, 2013) os autores destacaram que em 2013 dois direcionamentos foram levantados: o primeiro dizia respeito as ferramentas para MDE e AA, o qual ainda não foi completamente alcançado; e o segundo se referia ao desenvolvimento de uma cultura baseada em dados, que continua sendo um desafio (ROMERO; VENTURA, 2020). Referente à primeira tendência, de acordo com Romero e Ventura (2013), as ferramentas para MDE e AA deveriam ser disponibilizadas gratuitamente, para que uma população mais ampla pudesse utilizá-las, o que aconteceu, pois, uma grande variedade de ferramentas de finalidade específica está disponível de forma gratuita. Porém os autores ressaltaram que ainda é necessário desenvolver ferramentas para uso geral, que possam ser aplicadas em

6 Conferência Internacional sobre Learning Analytics & Knowledge (LAK) – Conferência Internacional sobre Educational Data Mining.

7 Journal of Educational Data Mining – Journal of Learning Analytics

várias tarefas e resolver diferentes problemas educacionais com a mesma interface/ferramenta, além disso, é necessário melhorar a portabilidade dos modelos obtidos por essas ferramentas (ROMERO; VENTURA, 2020). A segunda tendência pressupõe que educadores e instituições deveriam desenvolver uma cultura baseada em dados, utilizando-se deles para tomar decisões e melhorar seus processos de ensino, aprendizagem e administrativos. Todavia, de acordo com a pesquisa de 2020, a maioria das instituições e profissionais de ensino continuam cientes dos benefícios proporcionados pela análise de dados em larga escala, mas não adotaram essa cultura de forma efetiva nos seus processos gerenciais.

Romero e Ventura (2020) ainda sistematizaram os principais desafios relatados acerca de MDE e AA, quais sejam: transferibilidade e generalização; eficácia e aplicabilidade e interpretabilidade. Transferibilidade e Generalização se referem à utilização de modelos comuns para vários contextos, o que ainda não ocorre na prática. Eficácia e Aplicabilidade dizem respeito à realização de ações de intervenção a partir das análises de dados, assim há uma grande diversidade de modelos desenvolvidos os quais não se tem informações se foram ou são aplicados na prática das Instituições de Ensino e se são efetivos na resolução de problemas dessas instituições. Interpretabilidade refere-se à capacidade dos usuários compreenderem os modelos gerados para as análises de dados, acarretando um mau aproveitamento de tais modelos.

Finalmente, os autores propuseram algumas ideias, que disseram ser “visionárias e pessoais” que, em suas opiniões, podem formar tendências e direcionamentos muito promissores para a áreas de MDE e AA, para a década de 2020 resumidos no Quadro 2.

Quadro 2. Tendências para MDE e AA

TENDÊNCIA	ANÁLISE
<i>Levar em consideração todos os dados pessoais dos alunos durante toda a vida</i>	Atualmente, as informações consideradas na MDE e AA baseiam-se principalmente na interação de alunos com um único ambiente educacional, mas em um futuro próximo, graças ao grande volume de dados e à Internet das Coisas (IoT), pesquisadores serão capazes de ter informações disponíveis para cada aluno desde o nascimento até o momento e em tempo real. Isso implica a integração não apenas dos dados tradicionais de desempenho coletados das Instituições de Ensino e ambientes educacionais que cada aluno utilizou, mas também as informações sobre o status de cada aluno sob diferentes pontos de vista, como médico, familiar, econômico, religioso, sexual, relacionamento emocional, psicológico e assim por diante. Todos esses dados podem ser coletados a partir de várias fontes e elas poderiam ser fundidas para serem usadas com o intuito de melhorar e personalizar o processo de aprendizagem de cada aluno em cada momento específico de sua vida, o que suscitaria um novo nível de precisão.
<i>Aplicação e integração da MDE e AA aos futuros ambientes educacionais tecnológicos</i>	Na última década, os grandes avanços em tecnologias inovadoras permitiram o desenvolvimento de novos sistemas educacionais a partir de dispositivos móveis, a onipresente à realidade virtual, ambientes de realidade aumentada, hologramas e nas próximas décadas, satos quânticos serão associados à aplicação da Inteligência Artificial. Nesse contexto, não é errado pensar que os instrutores poderiam ser substituídos por máquinas sem que os alunos percebessem a mudança graças a avanços atuais em robôs humanoides inteligentes, agentes de conversação ou interfaces de voz e assistentes em ambientes educacionais. Mas esses sistemas precisam de técnicas de MDE e AA para analisar terabytes de dados e gerar modelos de análise portáteis em tempo real, a fim de enfrentar os desafios educacionais específicos desses futuros ambientes virtuais de aprendizagem.
<i>Análise e Mineração de Dados coletados diretamente do cérebro dos alunos para uma melhor compreensão do aprendizado</i>	O cérebro é o fator-chave para realmente entender como os alunos aprendem. Os avanços promissores na neurociência humana e neurotecnologia generalizada (interfaces cérebro-computador) estão dando origem a oportunidades sem precedentes de obter, coletar, compartilhar e manipular qualquer tipo de informação coletada do cérebro humano. Num futuro próximo, esses dados íntimos, sobre o estado psicológico e a atividade neural do aluno poderão ser analisados e minerados em tempo real, graças aos futuros dispositivos de eletroencefalografia de alta qualidade. Esses dados cerebrais, juntamente com outros dados multimodais poderiam ser integrados e usado pelos pesquisadores de MDE e AA, a fim de alcançar uma compreensão completa do processo de aprendizado.

Fonte: Adaptado Romero e Ventura (2020)

■ CONSIDERAÇÕES FINAIS

Em conclusão, com a análise e interpretação dessas publicações, que sistematizaram boa parte do estado da arte na área, em destaque dos últimos 20 anos, foi possível compreender o início do processo de adoção de MDE, quais eram as principais técnicas, os dados utilizados para formação de bancos/bases e os resultados que foram alcançados, enfim, de entender o modo pelo qual MDE se estabeleceu como uma área consolidada de pesquisa.

Além disso, explorar como ocorreu o processo de sua evolução, o qual foi constatado, teve início em decorrência fundamentalmente devido a dois fatores: 1) A adoção de grandes bases de dados na educação, impulsionado sobretudo pelo surgimento de cursos *e-learning*, como os do tipo MOOCs; e 2) O avanço das tecnologias computacionais, que são indispensáveis para aplicação das técnicas de MDE. Tais avanços, proporcionaram melhorias nas técnicas/ferramentas de MD já existentes, simplificando assim as tarefas dos pesquisadores e aperfeiçoando os resultados obtidos. Dessa forma, tais técnicas/ferramentas puderam ser aplicadas e testadas ao grande volume de dados educacionais disponíveis, consumando assim a MDE como um importante conjunto de técnicas e estratégia de pesquisa sobre dados educacionais.

Por fim, essas publicações propiciaram ainda, visualizar o cenário futuro para os próximos estudos com análises de dados educacionais. O que despertou interesse em descobrir quais seriam as principais tendências temáticas de pesquisas com MDE a contexto específicos da educação, como: no presencial, no ensino remoto, em cursos de Educação à Distância ou cursos *e-learning* diversos. Todavia, essa é uma proposta para um trabalho futuro, no qual pretende-se selecionar um contexto específico e analisar quais os propósitos da MDE, identificar tendências e tópicos de estudos mais proeminente, assim como encontrar quais as técnicas e algoritmos de MDE mais utilizados, além das principais oportunidades de pesquisa e os desafios mais relatados, formando um estudo amplo que apoie pesquisadores da área.

■ AGRADECIMENTOS

Agradeço ao Instituto Federal de Educação, Ciência e Tecnologia do Rio Grande do Sul (IFRS) pela oportunidade de capacitação que possibilitou o desenvolvimento dessa pesquisa.

■ REFERÊNCIAS

1. AGGARWAL, Charu C. **Data Mining: The Textbook**. 1. ed. New York, USA: Springer, 2015. v. 1 *E-book*. Disponível em: <https://doi.org/10.1007/978-3-319-14142-8>
2. ALDOWAH, Hanan; AL-SAMARRAIE, Hosam; FAUZY, Wan Mohamad. Educational data mining and learning analytics for 21st century higher education: A review and synthesis. **Telematics and Informatics**, [S. l.], v. 37, p. 13–49, 2019. Disponível em: <https://doi.org/10.1016/j.tele.2019.01.007>
3. BAKER, Ryan; ISOTANI, Seiji; CARVALHO, Adriana. Mineração de Dados Educacionais: Oportunidades para o Brasil. **Revista Brasileira de Informática na Educação**, [S. l.], v. 19, n. 02, p. 3–13, 2011. Disponível em: <https://doi.org/10.5753/rbie.2011.19.02.03>
4. BAKER, Ryan S. J. D.; YACEF, Kalina. The State of Educational Data Mining in 2009 : A Review and Future Visions. **Journal of Educational Data Mining**, [S. l.], v. 1, n. 1, p. 3–17, 2009. Disponível em: <https://doi.org/10.5281/zenodo.3554657>
5. DE LOS REYES, Daniel A. Guimarães *et al.* Predição de sucesso acadêmico de estudantes: uma análise sobre a demanda por uma abordagem baseada em transfer learning. **Revista Brasileira de Informática na Educação**, [S. l.], v. 27, n. 1, p. 1–25, 2019. Disponível em: <https://doi.org/10.5753/rbie.2019.27.01.01>
6. EDM. **Educational Data Mining**. [s. l.], 2020. Disponível em: <http://educationaldatamining.org/>. Acesso em: 31 maio. 2020.
7. GALLEN, Rosa Cabedo; CARO, Edmundo Tovar. An exploratory analysis of why a person enrolls in a Massive Open Online Course within MOOCKnowledge data collection. *In*: 2017, Athens, Greece. **Global Engineering Education Conference, (EDUCON)**. Athens, Greece: IEEE, 2017. p. 1600–1605. Disponível em: <https://doi.org/10.1109/EDUCON.2017.7943062>
8. GAO, Lina *et al.* Modeling the effort and learning ability of students in MOOCs. **IEEE Access**, [S. l.], v. 7, p. 128035–128042, 2019. Disponível em: <https://doi.org/10.1109/ACCESS.2019.2937985>
9. NEN-FU, Huang *et al.* The Clustering Analysis System Based on Students' Motivation and Learning Behavior. *In*: 2018, **Proceedings of 2018 Learning With MOOCs, LWMOOCs 2018**. : IEEE, 2018. p. 117–119. Disponível em: <https://doi.org/10.1109/LWMOOCs.2018.8534611>
10. PURSEL, B. K. *et al.* Understanding MOOC students: Motivations and behaviours indicative of MOOC completion. **Journal of Computer Assisted Learning**, [S. l.], v. 32, n. 3, p. 202–217, 2016. Disponível em: <https://doi.org/10.1111/jcal.12131>
11. RIGO, Sandro José *et al.* Minerando Dados Educacionais com foco na evasão escolar: oportunidades, desafios e necessidades. **Revista Brasileira de Informática na Educação**, [S. l.], v. 22, n. 01, p. 168–177, 2014. Disponível em: <https://doi.org/10.5753/RBIE.2014.22.01.132>
12. ROMERO, C.; VENTURA, S. Educational data mining: A survey from 1995 to 2005. **Expert Systems with Applications**, [S. l.], v. 33, n. 1, p. 135–146, 2007. Disponível em: <https://doi.org/10.1016/j.eswa.2006.04.005>
13. ROMERO, Cristbal; VENTURA, Sebastin. Educational data mining: A review of the state of the art. **IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews**, [S. l.], v. 40, n. 6, p. 601–618, 2010. Disponível em: <https://doi.org/10.1109/TS-MCC.2010.2053532>

14. ROMERO, Cristóbal *et al.* **Handbook of Educational Data Mining**. 1. ed. Boca Raton, USA: CRC Press - Taylor & Francis, 2010. *E-book*. Disponível em: <https://doi.org/10.1201/b10274>
15. ROMERO, Cristobal; VENTURA, Sebastian. Data mining in education. **Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery**, [S. l.], v. 3, n. 1, p. 12–27, 2013. Disponível em: <https://doi.org/10.1002/widm.1075>
16. ROMERO, Cristobal; VENTURA, Sebastian. Educational data mining and learning analytics: An updated survey. **Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery**, [S. l.], v. 10, n. 3, p. 1–21, 2020. Disponível em: <https://doi.org/10.1002/widm.1355>
17. ROMERO, Cristóbal; VENTURA, Sebastián. Educational data science in massive open online courses. **Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery**, [S. l.], v. 7, n. 1, 2017. Disponível em: <https://doi.org/10.1002/widm.1187>
18. SCHWENDIMANN, Beat A. *et al.* Perceiving learning at a glance: A systematic literature review of learning dashboard research. **IEEE Transactions on Learning Technologies**, [S. l.], v. 10, n. 1, p. 30–41, 2017. Disponível em: <https://doi.org/10.1109/TLT.2016.2599522>
19. SHAHIRI, Amirah Mohamed; HUSAIN, Wahidah; RASHID, Nur'Aini Abdul. A Review on Predicting Student's Performance Using Data Mining Techniques. **Procedia Computer Science**, [S. l.], v. 72, p. 414–422, 2015. Disponível em: <https://doi.org/10.1016/j.procs.2015.12.157>
20. SUKHIJA, Karan; JINDAL, Manish; AGGARWAL, Naveen. The recent state of educational data mining: A survey and future visions. *In*: 2015, Amritsar, India. **3rd International Conference on MOOCs, Innovation and Technology in Education (MITE)**. Amritsar, India: IEEE, 2015. p. 354–359. Disponível em: <https://doi.org/10.1109/MITE.2015.7375344>
21. WAHEED, Hajra *et al.* Predicting academic performance of students from VLE big data using deep learning models. **Computers in Human Behavior**, [S. l.], v. 104, p. 1–13, 2020. Disponível em: <https://doi.org/10.1016/j.chb.2019.106189>

Avaliação da Usabilidade do Scratch utilizando o Método System Usability Scale (SUS)

I Wendell Bento **Geraldes**
IFG

I Ernane Rosa **Martins**
IFG

I Ulisses Rodrigues **Afonseca**
IFG

RESUMO

A linguagem visual de programação Scratch foi criada em 2007 no MIT, tornando-se largamente utilizada no mundo inteiro. Este artigo tem como objetivo avaliar a usabilidade da linguagem em sua versão online disponível na internet. Para esta finalidade foi utilizado o método System Usability Scale (SUS), criado por John Brooke em 1986, que é considerado um dos questionários mais confiáveis e válidos para medir a usabilidade percebida pelos usuários. A coleta de dados foi realizada através de um questionário online contendo 10 perguntas, com uma escala de 1 a 5, onde 1 significa Discordo totalmente e 5 Concordo totalmente. Os resultados obtidos demonstraram uma avaliação positiva dos usuários em relação a linguagem Scratch.

Palavras-chave: Scratch, Usabilidade.

■ INTRODUÇÃO

O *Scratch* é uma linguagem de programação visual desenvolvida por uma equipe do *Massachusetts Institute of Technology - MIT* e lançada em 2007, possui uma comunidade de usuários espalhada por todo o mundo e conta com mais de 42 milhões de projetos compartilhados. Atualmente estão registrados mais de 41 milhões de usuários, sendo que somente no Brasil são cerca de 506 mil [Scratch 2019]. A linguagem está disponível por meio de uma plataforma Web, podendo também ser instalada em computadores pessoais que não possuem acesso a internet. O *Scratch* permite a criação de jogos, animações e histórias interativas digitais. Estes projetos podem ser compartilhados e modificados por outros usuários da plataforma [Eloy et al. 2017].

O uso do *Scratch* pode possibilitar aos estudantes a criação de seus próprios conteúdos relacionados a sua área curricular, com inúmeras vantagens, tais como: motivação, diversão, comprometimento e entusiasmo, além de apresentar melhorias relacionadas ao pensamento computacional e as práticas computacionais [Sáez-López et al. 2016]. A linguagem possui uma interface simples e muito intuitiva, sendo baseada no uso de blocos que são peças usadas para criar os códigos. Os blocos conectam-se uns aos outros como peças de um quebra-cabeça, onde cada tipo de dado (evento, comando, valor reportado, valor booleano, ou fim do script) tem sua própria forma específica que se encaixam perfeitamente uns aos outros. Uma série de blocos conectados são chamados de *Scripts* [Scratch 2019]. Apesar de ser bastante conhecida, não existem muitas pesquisas sobre sua usabilidade.

Este trabalho tem como objetivo avaliar a usabilidade da linguagem *Scratch*, utilizando o método *System Usability Scale (SUS)*, que é uma métrica de avaliação da usabilidade de sistemas computacionais, constituído por um questionário padronizado, considerado o mais utilizado do mundo por sua praticidade e rapidez. Quando os pesquisadores e profissionais precisam de uma medida de usabilidade, estes certamente consideram seriamente o uso do SUS [Lewis 2018]. Este é considerado um dos questionários mais confiáveis e válidos para medir a usabilidade percebida pelos usuários, mesmo com a utilização de pequenas amostras [Revythi and Tselios 2019].

■ USABILIDADE

Segundo [Nielsen 2006], usabilidade é um atributo de qualidade que avalia a facilidade de uso de uma interface, sendo definida por cinco componentes:

1. Capacidade de aprendizagem: a facilidade de utilizar o sistema pela primeira vez;
2. Eficiência: rapidez para executar as tarefas;

3. Memorização: o processo de lembrar como utilizar o sistema, após um tempo sem utilizar;
4. Erros: ausência de erros apresentados pelo sistema;
5. Satisfação: design agradável.

[Shackel 1986] sugere quatro critérios para a avaliação das tarefas sob o ponto de vista da usabilidade:

1. Eficácia: avalia o desempenho de uma tarefa tendo como foco a análise da velocidade de execução e quantidade de erros;
2. Aprendizagem: avalia o desempenho desde a instalação do produto até o início do uso;
3. Flexibilidade: Avalia a capacidade de adaptação a novas tarefas além das já especificadas no sistema;
4. Atitude: Avalia o desempenho com relação ao conforto ou satisfação do usuário ao utilizar o sistema.

[Bastien and Scapin 1993] apresentaram oito critérios ergonômicos para construção de interfaces de sistemas:

1. Condução: refere-se aos meios disponíveis para aconselhar, orientar, informar, e conduzir o usuário na interação com o computador (mensagens, alarmes, rótulos, etc.).
2. Carga de Trabalho: está relacionada aos elementos da interface, devido ao importante papel na redução da carga cognitiva e perceptiva do usuário, possibilitando aumento da eficiência do diálogo.
3. Controle Explícito: refere-se ao controle que os usuários tem sobre o processamento de suas ações pelo sistema.
4. Adaptabilidade: capacidade do software em se comportar conforme as necessidades e preferências do usuário.
5. Gestão de erros: compreende os mecanismos capazes de prevenir e reduzir erros causados pelo sistema, além de rotinas para facilitar sua correção quando os mesmos ocorrerem.
6. Consistência: o modo que os aspectos de interface são mantidos em contextos semelhantes, e se diferenciam em contextos distintos.
7. Significância do código: relação semântica entre os códigos e os itens ou ações às que eles se referem, demonstra a clareza com que os códigos e denominações são apresentados para o usuário.

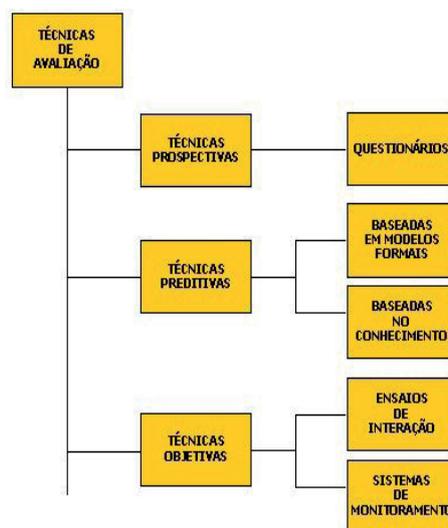
8. Compatibilidade: é a relação entre as características dos usuários e as tarefas, saídas e entradas do sistema.

■ TÉCNICAS DE AVALIAÇÃO DE USABILIDADE

Para [Cybis 2003] existem três tipos de técnicas de avaliação ergonômica (Figura 1):

1. Técnicas prospectivas, que buscam a opinião do usuário sobre a interação com o sistema;
2. Técnicas preditivas ou diagnósticas, que buscam prever os erros de projeto de interfaces sem a participação de usuários;
3. Técnicas objetivas ou empíricas, que buscam constatar os problemas a partir da observação do usuário interagindo com o sistema.

Figura 1. Técnicas de avaliação de usabilidade



Fonte: [Cybis 2003]

As técnicas prospectivas estão baseadas na aplicação de questionários/entrevistas com o usuário para avaliar sua satisfação ou insatisfação em relação ao sistema e sua operação. Ela mostra-se bastante pertinente na medida em que é, o usuário a pessoa que melhor conhece o software. Nada mais natural em buscar suas opiniões para orientar revisões de projeto. Este tipo de técnica pode ser empregado para aumentar a efetividade de avaliações analíticas, realizadas por especialistas que diagnosticam problemas de usabilidade. Apoiados pelas respostas do questionário de satisfação, estes podem centrar suas análises sobre os pontos problemáticos no sistema, apontados pelo usuário [Cybis 2003].

■ SYSTEM USABILITY SCALE (SUS)

Existem vários questionários disponíveis para profissionais de usabilidade que podem auxiliar a avaliação da usabilidade de um produto ou serviço. Muitos destes questionários são usados para avaliar tipos específicos de *interfaces*, enquanto outros podem ser usados para avaliar uma faixa mais ampla de tipos de *interface*. A escala *System Usability Scale* (SUS) é uma das que podem ser usadas para avaliar a usabilidade de uma variedade de produtos ou serviços. As características do SUS que tornam seu uso atrativo, são:

- Composto de apenas dez afirmações: é relativamente rápido e fácil para os participantes do estudo concluírem e para os administradores pontuarem;
- Não proprietário: rentável para uso e pode ser pontuado muito rapidamente, imediatamente após a conclusão;
- Independente de tecnologia: pode ser usado por um amplo grupo de profissionais de usabilidade para avaliar praticamente qualquer tipo de interface de usuário; Pontuação única: variando de 0 a 100, sendo relativamente fácil de entender por uma ampla gama de pessoas de outras disciplinas que trabalham em equipes de projeto [Bangor et al. 2009].

No estudo realizado por [Bangor et al. 2009] os resultados de 2.324 pesquisas do SUS em 206 testes de usabilidade coletados ao longo de um período de dez anos, verificou-se que o SUS era altamente confiável ($\alpha = 0,91$) e útil em uma ampla gama de tipos de *interface* [Bangor et al. 2009].

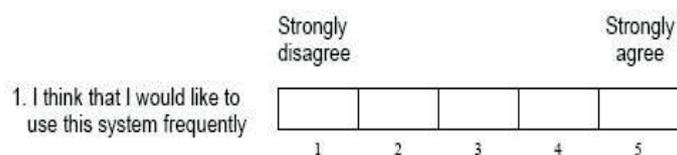
O conceito de aplicar um adjetivo (bom, ruim, excelente, etc.) à usabilidade do produto era atraente porque é familiar para a maioria das pessoas que trabalham em equipes de *design*, independentemente de sua disciplina. Ter um ponto de referência familiar e fácil de entender que pode ser facilmente entendido por engenheiros e gerentes de projeto facilita a comunicação dos resultados do teste. [Bangor et al. 2009] relataram os resultados de um estudo piloto que procurou mapear adjetivos descritivos (por exemplo, bom, péssimo, etc.) ao alcance das pontuações do SUS [Bangor et al. 2009].

■ METODOLOGIA

Neste estudo foi realizada uma pesquisa exploratória através de um formulário *on-line* com objetivo de conhecer as opiniões dos usuários acerca da usabilidade da linguagem de programação visual *Scratch*. Para coleta de dados foi utilizado um questionário SUS - *System Usability Scale* composto por 10 itens, com 5 opções de respostas. O respondente do

questionário SUS assinala sua resposta numa escala Likert que varia de 1 (Discordo totalmente) a 5 (Concordo totalmente) observado na Figura 2 [Boucinha and Tarouco 2013].

Figura 2. Exemplo de questão do SUS



Fonte: [Boucinha and Tarouco 2013]

O questionário SUS foi construído em língua inglesa, porém optou-se por utilizar uma versão traduzida para o português conforme a Tabela 1, por ser mais adequada ao público alvo da pesquisa [Teixeira 2015]. A análise dos resultados do SUS foi a soma da contribuição individual de cada item. Para os itens ímpares deve-se subtrair 1 à resposta do usuário, ao passo que para os itens pares deve-se subtrair 5 a resposta do usuário. Depois de obter o score de cada item, soma-se os scores e multiplica-se o resultado por 2,5. O resultado obtido será um índice de satisfação do utilizador que varia de 0 a 100 [Boucinha and Tarouco 2013]. A pesquisa pode ser classificada como descritiva. Para [Gil 2010] uma pesquisa descritiva tem como objetivo a descrição de características de uma determinada população, utilizando técnicas padronizadas de coleta de dados, como questionários e observação sistemática [Gil 2010].

Tabela 1. Questionário System Usability Scale traduzido para o português.

Item	Item correspondente em português
1	Acho que gostaria de utilizar este produto com frequência
2	Eu acho o sistema desnecessariamente complexo
3	Eu achei o sistema fácil de usar.
4	Eu acho que precisaria de ajuda de uma pessoa com conhecimentos técnicos para usar o sistema.
5	Eu acho que as várias funções do sistema estão muito bem integradas.
6	Eu acho que o sistema apresenta muita inconsistência.
7	Eu imagino que as pessoas aprenderão como usar esse sistema rapidamente.
8	Eu achei o sistema atrapalhado de usar.
9	Eu me senti confiante ao usar o sistema.
10	Eu precisei aprender várias coisas novas antes de conseguir usar o sistema.

■ RESULTADOS

A coleta de dados foi realizada entre os dias 13/05/2019 e 11/06/2019, sendo obtidas 83 respostas ao questionário SUS. Esta quantidade de respondentes é 21% do tamanho esperado da amostra, que era de 384 respostas.

A avaliação da usabilidade realizada pelos usuários da linguagem *Scratch* por meio do questionário SUS revelou que a média geral do *SUS-Score* foi de 75,27. Este indicador mostra que a linguagem possui uma boa usabilidade [Teixeira 2015].

■ CONSIDERAÇÕES FINAIS

Este trabalho teve como objetivo avaliar a usabilidade da linguagem *Scratch*, utilizando o método *System Usability Scale (SUS)*. O resultado encontrado com a aplicação do questionário SUS na avaliação da usabilidade da linguagem *Scratch* na percepção do usuário, teve como classificação o status de bom, por ter alcançado a pontuação de 75,27 [Bangor et al. 2009]. Permitindo inferir que a linguagem *Scratch* atende as necessidades de seus usuários quanto a usabilidade.

Embora a aplicação do questionário SUS tenha sido amplamente divulgada nos meios de comunicação digitais, a amostra coletada pode ser considerada pequena diante do número de usuários brasileiros da linguagem *Scratch*. Sendo fundamental a realização de novos estudos para a ampliação da amostra e confirmação do resultado encontrado, permitindo uma melhor análise e discussão deste tema. Podendo ainda ser utilizados outros questionários gratuitos na avaliação, tais como: *After Scenario Questionnaire (ASQ)* da IBM e *Usefulness, Satisfaction and Ease of Use (USE)*.

■ REFERÊNCIAS

1. Bangor, A., Kortum, P., and Miller, J. (2009). Determining what individual sus scores mean: Adding an adjective rating scale. *JUS - The Journal of Usability Studies*, 4(3):113–123.
2. Bastien, J. C. and Scapin, D. L. (1993). Human factors criteria, principles, and recommendations for hci: methodological and standardisation issues. *INRIA*.
3. Boucinha, R. M. and Tarouco, L. M. R. (2013). Avaliação de ambiente virtual de aprendizagem com o uso do sus - system usability scale. *Revista Renote*, 11(3):1–10. Novas Tecnologias na Educação.
4. Cybis, W. (2003). Engenharia de usabilidade: Uma abordagem ergonômica. Laboratório de utilizabilidade de informática (labiutil). Disponível em: <http://www.labiutil.inf.ufsc.br/hiperdocumento/conteudo.html>. Acesso em 11 de agosto de 2019.
5. Eloy, A., Lopes, R., and Angelo, I. (2017). Uso do scratch no brasil com objetivos educacionais: uma revisão sistemática. *Revista Renote*, 15(1):1–10. Novas Tecnologias na Educação.
6. Gil, A. C. (2010). *Como elaborar projetos de pesquisa*. Atlas, 5 edition. Lewis, J. (2018). The system usability scale: Past, present, and future. *International Journals of Human – Computer Interaction*, 34(7):577–590.

7. Nielsen, J. (2006). Usability 101: Introduction to usability. Disponível em: [http:// www.ingenieriasimple.com/usabilidad/IntroToUsability.pdf](http://www.ingenieriasimple.com/usabilidad/IntroToUsability.pdf). Acesso em 11 de junho de 2019.
8. Revythi, A. and Tselios, N. (2019). Extension of technology acceptance model by using system usability scale to assess behavioral intention to use e-learning. *Education and Information Technologies*, 24(4):2341–2355.
9. Scratch (2019). Estatísticas. Disponível em: [<https://scratch.mit.edu/statistics/>](https://scratch.mit.edu/statistics/). Acesso em 11 de junho de 2019.
10. Shackel, B. (1986). Ergonomics in design for usability. In: HCI 86 Conference on People and Computer. New York: Cambridge University Press.
11. Sáez-López, J., Román-González, M., and Vázquez-Cano, E. (2016). Visual programming languages integrated across the curriculum in elementary schools: A two years case study using scratch in five schools. *Computer & Education*, 97:129–141.
12. Teixeira, F. (2015). O que é o sus (system usability scale) e como usá-lo em seu site. Disponível em: <https://brasil.uxdesign.cc/o-que-%C3%A9-o-sus-system-usability-scale-e-como-us%C3%A1-lo-em-seu-site-6d63224481c8>>. Acesso em 11 de junho de 2019.

Avaliação automática de redações na língua portuguesa baseada na coleta de atributos e aprendizagem de máquina

| Silvério **Sirotheau**
UFPA

| Eloi **Favero**
UFPA

| João Carlos Alves dos **Santos**
UFPA

| Simone **Negrão**
UFPA

| Marco Lima do **Nascimento**
UFPA

|

RESUMO

Ambientes virtuais demandam métodos de avaliação automática para questões discursivas. Na literatura encontramos métodos promissores para textos na língua inglesa, porém, para o português os estudos são apenas preliminares. Esta pesquisa foca numa abordagem de avaliação automática de redações em português, baseada na coleta de atributos e em métodos de aprendizagem de máquina. Nos experimentos utilizou-se 1000 redações de um concurso público. Na coleta de atributos explorou-se quatro dimensões: Léxica, Sintática, Conteúdo e Coerência. Como resultado foram obtidos índices Kappa quadrado (KQ) de 0.68 do sistema contra humanos versus um KQ de 0.56 de humano contra humano.

Palavras-chave: Avaliação Automática, Coleta de Atributos, Aprendizagem de Máquina.

■ INTRODUÇÃO

A avaliação é uma tarefa central no processo educativo, visto que é uma maneira de qualificar o conhecimento dos alunos em relação aos conceitos ensinados (Mohler e Mihalcea, 2009; Rodrigues e Araújo, 2012). O resultado da avaliação é também utilizado como uma ferramenta de *feedback* para orientação do aluno. Tradicionalmente a avaliação é feita pelo professor de forma manual, mas com o avanço tecnológico várias tarefas podem ser automatizadas (Bull e Mckenna, 2001; Amália *et al.*, 2019). Sistemas com essa finalidade são desenvolvidos pela pesquisa em avaliação automática de textos (AAT), relacionada à subárea Processamento de Linguagem Natural (PLN).

Em muitas plataformas de ensino, questões discursivas (ou ensaios) servem ao propósito de verificação e aferição da aprendizagem do aluno, em particular à capacidade de escrita e estruturação do discurso argumentativo (Page, 1966; Yang, 2012; Lee, 2014; Zupanc e Bosnic, 2015). Neste contexto, o desenvolvimento de abordagens que permitem automatizar a correção ganha relevância, por apresentar vantagens: possibilita *feedback* imediato; permite múltiplas submissões em tempo e lugar determinados pelo usuário; tem eficiência estável; e baixo custo.

Pesquisas em AAT iniciou-se na década de 60 (Page, 1966; Hearst, 2000; Noorbehbahani e Kardan, 2011), porém, somente na década de 90, com o uso de técnicas de PLN, houve um avanço considerável. Hoje, em termos de definir o escore de um ensaio, os sistemas já alcançam boa acurácia em relação aos avaliadores humanos (Shermis e Hammer, 2012). Porém, os sistemas deixam muito a desejar quanto ao *feedback* inteligente e orientador para o aluno, similar ao oferecido por um avaliador humano (Attali e Burstein, 2006; Haley *et al.*, 2007).

Em provas discursivas avaliadas por dois especialistas humanos de forma independente, os avaliadores podem divergir nos seus escores. Por exemplo, considerando um único texto avaliado por dois especialistas humanos, numa escala de 0 a 10, se os seus escores divergem em 2 pontos, então dizemos que o percentual de concordância ou acurácia Humano *versus* Humano (HxH) é de 80%. Assim, num mesmo texto avaliado por um sistema e por um especialista humano, podemos medir a acurácia Sistema *versus* Humano (SxH). Um sistema possui bom desempenho quando a acurácia SxH é próxima ou superior à acurácia HxH. Para a literatura inglesa, vários estudos relatam boa aproximação e até superação de acurácia SxH contra HxH (Shermis e Hammer 2012; Srivastava *et al.*, 2017).

Neste trabalho propomos uma abordagem para avaliar redações em português centrada na coleta de atributos. Serão considerados atributos relacionados a 4 dimensões: Léxica, Sintática, do Conteúdo e da Coerência. Recentemente surgiram também abordagens de coleta automática de atributos (mais de conteúdo) tais como *Convolutional Neural Networks* (Dong, Zhang e Yang 2017; Dasgupta *et al.*, 2018). Estas abordagens têm alcançado e até

superado as acurácias do escore final das abordagens de coleta de atributo manuais, porém elas têm restrições para dar *feedback* para o estudante. A metodologia que usa coleta de atributos permite enriquecer o escore com *feedback* orientador para o aluno, por exemplo, mostrando os erros gramaticais e/ou informando em qual grupo de atributos o texto teve deficiência (Zupanc e Bosnic, 2017). A meta é desenvolver uma acurácia SxH próxima da HxH, assim por exemplo, num evento real um dos avaliadores poderá ser substituído pelo computador.

Nesta investigação busca-se esclarecer algumas questões de pesquisas (QP) em torno do uso dos atributos nas dimensões: Léxica, Sintática, Conteúdo e Coerência. Qual a contribuição de cada uma das quatro dimensões (QP1-2-3-4) na acurácia final? Explora-se também o cruzamento entre as dimensões: (QP5) Qual a acurácia da combinação das dimensões duas a duas: léxica x sintática, léxica x conteúdo, léxica x coerência, sintática x conteúdo, sintática x coerência e conteúdo x coerência? Além disso, como a dimensão Conteúdo tem a maior influência explora-se conteúdo + léxica + coerência, conteúdo + sintática + coerência e conteúdo + léxica + sintática. (QP6) Com base nas quatro dimensões, a acurácia final do método SxH alcança a acurácia dos avaliadores humanos HxH?

Além desta introdução este artigo está organizado da seguinte forma: A seção 2 apresenta os trabalhos relacionados. A seção 3 apresenta o corpus de estudo. A seção 4 apresenta a abordagem proposta. A seção 5 apresenta os resultados e discussão. Finalmente, a seção 6 apresenta a conclusão.

■ TRABALHOS RELACIONADOS

Quanto às abordagens centradas na coleta de atributos, temos sistemas com algumas dezenas até com várias centenas: 60 atributos (*Intelligent Essay Assessor*); 90 atributos (Bookette); 150 atributos (*Semantic Automated Grader for Essays*); até mais de 400 atributos (IntelliMetric). Fonseca *et al.*, 2018, utilizou 681 atributos num trabalho com textos em língua Portuguesa.

O *Intelligent Essay Assessor* (IEA) é um sistema de avaliação automática de ensaios. Ele usa Análise Semântica Latente (LSA) e PLN para coletar 60 atributos (Foltz, Laham e Landauer, 1999). Compreende um modelo de aprendizagem mecânica que induz a semelhança semântica de palavras baseada em grandes corpora de texto relevante para o domínio do tema e concentra-se mais em conteúdo do que em qualidade de escrita; possui também um módulo de *feedback*. Ele alcança uma acurácia KQ de 0.73 (Shermis e Hamner, 2012).

O Bookette (Rich *et al.*, 2013) usa técnicas de PLN para extrair 90 atributos que descrevem a qualidade do texto. O sistema usa redes neurais para gerar escore e *feedback*. O sistema produz um *feedback* baseado no desempenho dos atributos, um *feedback* holístico

e também comentários sobre as convenções de gramática e ortografia. O sistema alcança uma acurácia KQ de 0.70 (Shermis e Hamner, 2012).

Zupanc e Bosnic (2017) estenderam o sistema SAGE (*Semantic Automated Grader for Essays*) com atributos de coerência, baseados em janelas sequenciais (1/4 do texto, com passo de 10 palavras) de um ensaio, visualizado num espaço semântico. As medidas de coerência são o resultado das mudanças entre as janelas para estimar a coerência do texto. O sistema estendido com os novos atributos de coerências alcança uma acurácia KQ de 0.93.

O Intellimetric (Elliot, 2003) é um sistema que analisa elementos do discurso para formar um sentido semântico baseado em duas categorias: a) estrutura de atributos sintático - estruturais e mecânicos e b) conteúdo - discurso/retórica e atributos de conteúdo/conceito. O sistema usa múltiplas abordagens de aprendizagem de máquina (incluindo análise linear, análise bayesiana e LSA) para combinar os diferentes modelos em um único desfecho final (Rudner, Garcia e Welch, 2006). O sistema alcança uma acurácia KQ de 0.76 (Shermis e Hamner, 2012).

Para a língua portuguesa, Amorim e Veloso (2017) desenvolveram um experimento em AAT usando um conjunto de dados disponível na *web*, com 1840 redações (média de 300 palavras, do tipo ENEM) sobre 96 temas diferentes. Estas redações foram avaliadas em relação a 5 aspectos: Linguagem formal, foco no tema, organização do discurso, organização do argumento, solução proposta. A pesquisa se baseou em menos de 30 atributos, em classes como: sofisticação léxica, gramática e mecânica (erros léxicos e sintáticos e pontuação). Eles alcançaram uma acurácia KQ 0.42.

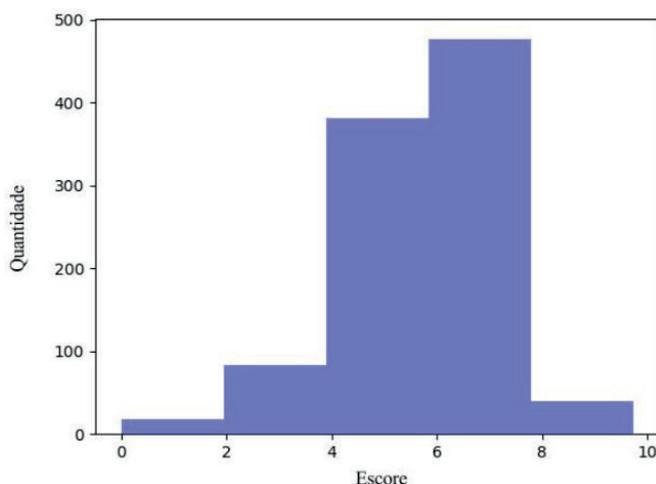
Também para a língua portuguesa, Fonseca *et al.*, (2018) alcançaram uma acurácia KQ de 0.75 no score geral em provas tipo ENEM, avaliadas nos mesmos 5 aspectos. Eles utilizaram uma abordagem de *Deep Neural Network* (DNN) e também baseada em um grande número de atributos (681) coletados manualmente. Os atributos foram coletados nas categorias: POS *tags*; POS *n-grams*; tokens *n-grams*; expressões específicas (agentes sociais, conectivos, etc.) e métricas de contagens. A abordagem baseada nos atributos foi vencedora na avaliação dos quatro primeiros aspectos. E a abordagem DNN foi vencedora para avaliar o quinto aspecto, a solução proposta.

Shermis e Hammer (2012) relataram que para dois humanos os escores variaram nas taxas de concordância de 0.61 a 0.85 medido por KQ; e nas pontuações das máquinas contra os humanos a concordância variou de 0.60 a 0.84 em KQ. Na língua portuguesa os estudos ainda são incipientes, portanto, este trabalho é mais uma contribuição para a maturidade desta área de pesquisa na nossa língua.

■ CORPUS DE ESTUDO

O corpus da pesquisa foi composto por uma amostra de 1.000 redações de um concurso público do edital nº 26/2016 da Universidade Federal do Oeste do Pará. Estas redações passaram por um processo de digitalização manual onde não foi feito nenhum tipo de correção ortográfica e nem alterações nos aspectos gramaticais do texto original. Todas as redações foram previamente avaliadas por dois avaliadores humanos, recebendo uma pontuação inteira entre 0 e 10, com passo de 0.25; sendo que cada avaliador não conhece a pontuação do outro. Foram feitas verificações de discrepâncias: se as duas pontuações divergem por mais de um ponto, então um terceiro avaliador atribui uma pontuação para resolver a discrepância. A Figura 1 apresenta, num histograma, as classes de notas do corpus atribuídas pelos avaliadores humanos.

Figure 1. Histograma dos escores do corpus de redações, dos valores atribuídos pelos avaliadores humanos.



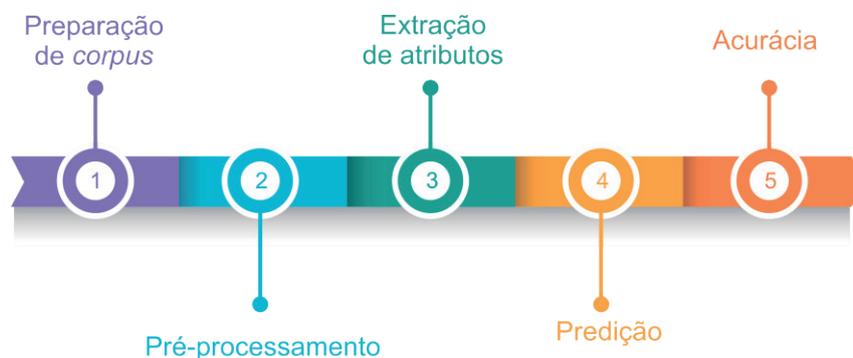
O escore final é média da avaliação de 3 competências: (i) tema está relacionada com compreender o tema e não fugir do que é proposto; (ii) coerência está relacionada com selecionar, relacionar, organizar e interpretar informações, fatos, opiniões e argumentos em defesa de um ponto de vista e conhecimento dos mecanismos linguísticos necessários para a construção da argumentação e; (iii) regras está relacionada com domínio da escrita formal da língua portuguesa.

■ ABORDAGEM PROPOSTA

Na literatura existe uma tendência de uso de uma arquitetura de *pipeline* similar à da Figura 2 para sistema de AAT (Burrows, Gurevych e Stein, 2015). Ela contém 5 etapas: preparação de corpus, (2) pré-processamento, (3) coleta de atributos, (4) modelo de predição e (5) avaliação. A saída de uma etapa é entrada para a próxima. A arquitetura *pipeline* é bem

comum em vários campos da pesquisa e PLN, por exemplo, na extração informação, extração de relações e preenchimento de *templates* (Wachsmuth, Stein e Engels, 2011). Em cada uma dessas etapas os pesquisadores utilizam diferentes métodos e técnicas para no final gerar o escore de cada resposta.

Figure 2. Arquitetura em *pipeline* para avaliação de textos composta de 5 etapas.



Na etapa de **preparação do corpus** as respostas das questões abertas são digitadas e organizadas numa coleção. Cada resposta da base tem pelo menos 2 escores de avaliadores humanos, permitindo a medida da acurácia HxH, para depois ser comparada com a acurácia SxH.

No **pré-processamento** busca-se uma representação “normalizada” do documento, deixando apenas a informação relevante para o processo de avaliação. As respostas foram separadas em sentenças e em seguidas vetorizadas em *tokens*. Após isso, três técnicas de pré-processamento foram aplicadas, com o uso biblioteca *Natural Language Toolkit* (NLTK): (1) Remoção de Caracteres Especiais, pontuação, acentuação e conversão para minúsculas (RCE); (2) Remoção de *stopword* (RSW) e; (3) Remoção de sufixos (*stemmer*) (RSU). As técnicas foram combinadas da seguinte forma: a) sem pré-processamento (-RCE, -RSW, -RSU); b) com remoção de caracteres especiais (+RCE, -RSW, -RSU); c) com remoção de caracteres especiais e *stopword* (+RCE, +RSW, -RSU) e; d) com remoção de caracteres especiais, *stopword* e aplicação de *stemmer* (+RCE, +RSW, +RSU). Em seguida, ainda nesta etapa, os *tokens* foram classificados morfologicamente com o etiquetador o Aelius (Alencar, 2010).

Na etapa de **extração de atributos**, procura-se abranger as principais classes de atributos citadas na literatura recente para a língua inglesa (Zupanc e Bosnic, 2017; Palma e Atkinson, 2018, Vajjala, 2018). Foram extraídos mais de 140 atributos agrupados em 4 dimensões (Tabela 1):

- **Léxica** descreve a coleta de atributos em um aspecto individual das palavras. Esta dimensão tem três principais categorias: (i) estatística de superfície, coleta estatística baseado em contagem de palavras; (ii) diversidade, coleta medidas que repre-

sentam o quanto é diverso o vocabulário utilizado; (iii) legibilidade, mede o grau de facilidade da leitura do texto.

- **Sintática** descreve atributos que retratam o aspecto individual de cada sentença, compreende duas categorias: (1) número de cada *PoS tag* (*part-of-speech tagging*), como por exemplo, número de nomes (*noun*) e verbos (*verb*) (2) Erros, léxicos e sintáticos, conta o número de erros de sentenças mal formuladas, por exemplo, erros de concordância e pontuação.
- **Conteúdo** descreve atributos relacionadas com as medidas de similaridade entre as respostas dos alunos e a resposta de referência.
- **Coerência** descreve atributos relacionados à coerência do discurso de cada resposta (Zupanc e Bosnic, 2017).

Table 1. Exemplo de alguns atributos utilizados no experimento.

Atributos	
Léxica	<i>nº de caracteres, nº de diferentes palavras, nº de palavras, nº de palavras curtas, nº de palavras longas, nº médio de palavras, nº de stopword, nº de sentenças, comprimento de palavra mais frequente, nº de sílaba (...)</i>
Sintática	<i>nº de cada etiquetas (SR = ser, HV = haver, ET = estar, TR = ter, VB = verbo, Gênero, Número, Substantivo, Nome próprio, Pronomes, Preposição...), nº de diferentes etiquetas, nº de erros ortográficos (...)</i>
Conteúdo	<i>foram aplicadas as medidas (Cosseno e Distância Euclidiana) contra vetores de respostas, incluindo também as variações no tipo de pré-processamento (SSW, CST, CSW)</i>
Coerência	<i>foram utilizados quatro modelos combinando técnicas de pré-processamento e as duas medidas (Cosseno e Distância Euclidiana), geraram 3(a, b, c) x 2(min/máx. med.) x 2(cos, dist) x 3(pre).</i>

Na etapa de **predição**, utilizou-se o algoritmo de aprendizado de máquina supervisionado *Random Forest* que aceita na entrada de centenas de atributos em tarefas de regressão e/ou predição. Para este tipo de problema, onde temos mais de 100 atributos, o algoritmo tem duas vantagens: tem bom desempenho e retorna uma classificação da importância dos melhores atributos (Fernández-Delgado *et al.*, 2014). Para validação dos resultados utilizamos a abordagem *Cross-validation*, particionando o conjunto de dados em 5 *folds*; a acurácia coletada é a média dos 5 testes.

Na etapa de **acurácia**, seleciona-se as melhores combinações das etapas anteriores (por exemplo, diferentes combinações de pré-processamento) buscando maximizar o resultado final. Para medir a acurácia usou-se Kappa Quadrático - KQ (Fleiss e Cohen, 1973), que mede o grau de concordância entre duas classes com uma certa flexibilidade em relação à concordância exata. O KQ mede também a concordância parcial: se devia predizer 6, mas se resultou em 5, não é totalmente errado. Essa métrica geralmente varia de 0 (pouca concordância entre avaliadores) a 1 (concordância completa entre avaliadores). A escala de interpretação do KQ é: i) $< 0.00 \rightarrow$ “Pobre”, ii) $0.00 - 0.20 \rightarrow$ “Fraco”, iii) $0.21 - 0.40 \rightarrow$

“Razoável”, iv) 0.41 - 0.60 → “Moderado”, v) 0.61 - 0.80 → “Substancial” e vi) 0.81 - 1.00 → “Quase perfeito”.

O KQ é calculado criando-se uma matriz de acordo com a equações 1 e 2. Cada célula da matriz, i.e., $O_{i,j}$ corresponde a uma resposta que pontuam i do avaliador humano e j do sistema. $W_{i,j}$ contém os pesos calculados conforme a Equação 1 e a matriz. $E_{i,j}$ contém as pontuações esperadas dos avaliadores humanos, obtidas pela multiplicação dos vetores de histograma das duas pontuações. Os subscritos em matriz $O_{i,j}$ correspondem ao número de ensaios que pontuam i do avaliador humano e j do sistema. No final do processo KQ é calculado de acordo com a equação 2.

$$W_{i,j} = \frac{(i-j)^2}{(N-1)^2} \quad (1)$$

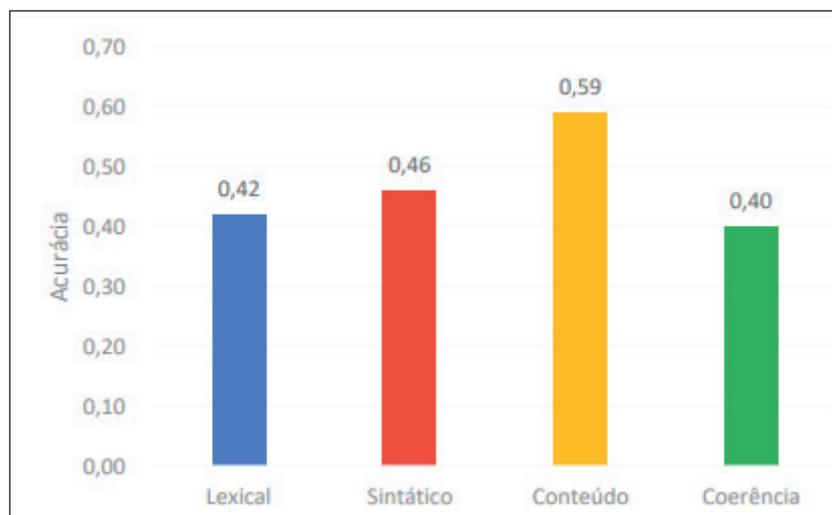
$$\kappa = 1 - \frac{\sum_{i,j} W_{i,j} O_{i,j}}{\sum_{i,j} W_{i,j} E_{i,j}} \quad (2)$$

RESULTADOS E DISCUSSÃO

Realizou-se experimentos no corpus com 1.000 redações. Foram trabalhados com mais de 140 atributos, a grande maioria adaptados da língua inglesa. Em parte, pretendia-se verificar se os atributos para língua inglesa são adequados para o português. Aplicou-se a abordagem com os dados do corpus com o objetivo de maximizar o valor SxH buscando uma aproximação com HxH.

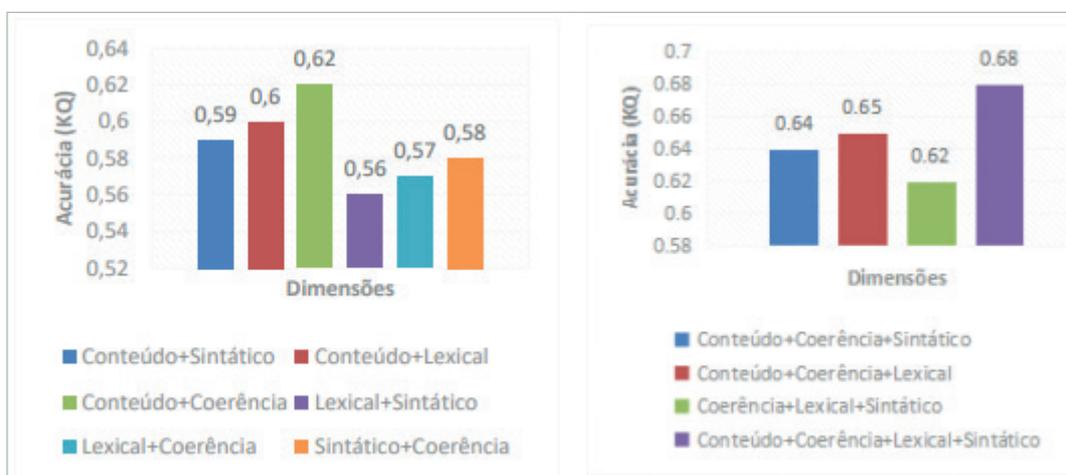
Em relação a questão de pesquisa QP1-2-3-4: Qual a contribuição de cada uma das quatro dimensões? Esta comparação foi feita considerando-se somente os atributos de uma dimensão para predizer o escore dos humanos. A Figura 3 mostra a contribuição das dimensões (léxica, sintática, conteúdo e coerência) apresentando os valores de KQ (0.42, 0.46, 0.59 e 0.40), respectivamente, que são valores apenas moderados na interpretação do KQ. No entanto, o valor 0.59 para conteúdo se aproxima do 0.60 que já está na escala de interpretação “substancial”. A melhor contribuição é do conteúdo.

Figure 3. A contribuição de cada uma das dimensões (Léxica, Sintática, Conteúdo e Coerência) na acurácia final (Métrica: KQ).



Respondendo à questão de pesquisa (QP5), que busca explorar o cruzamento entre as dimensões. Qual a acurácia da combinação das dimensões duas a duas: Léxica x Sintática, ...? Com duas dimensões, a menor acurácia vem da combinação da dimensão léxica+sintática com KQ 0.56; já o melhor desempenho é com a combinação das dimensões de conteúdo+coerência com KQ 0.62 (Figura 4).

Figure 4. Explorando a combinação das dimensões 2 a 2 e 3 a 3 na contribuição da acurácia.



Além disso, como a dimensão Conteúdo tem a maior influência explora-se Conteúdo + Léxica + Coerência, Conteúdo + Sintática + Coerência e Conteúdo + Léxica + Sintática (Figura 4). Com três dimensões a pior combinação vem de conteúdo + léxica + sintática com 0.62 e a melhor vem de conteúdo + coerência + léxica, com KQ 0.65. Combinando as quatro dimensões responde-se à questão de pesquisa (QP6): O método de avaliação alcança a acurácia dos avaliadores humanos? A combinação das quatro dimensões resultou uma acurácia final de KQ 0.68 contra a acurácia HxH com KQ 0.56. Portanto, o sistema supera a acurácia humana.

■ CONCLUSÃO

O objetivo deste trabalho foi desenvolver um método de avaliação automática de respostas discursivas para redação, coletando-se atributos em 4 dimensões. Foram classificadas numa espécie de taxonomia mais de 140 atributos. A maior parte delas veio de trabalhos relacionados da língua Inglesa as quais foram ajustadas para o Português.

Para realização dos experimentos utilizou-se uma arquitetura *pipeline* linear de 5 etapas. Utilizamos a técnica *Random Forest*, que permite a manipulação de um grande número de atributos além de retornar a relevância de cada atributo na etapa de classificação. Como resultado obtivemos um kappa quadrático (KQ) 0.68 SxH contra 0.56 HxH. Um resultado KQ de 0.68 é considerado “substancial”. Este resultado superou a acurácia medida entre os avaliadores humanos. Com relação aos trabalhos prévios da língua portuguesa, esta acurácia supera o do trabalho de Amorim e Veloso (2017) com KQ 0.42 mas ainda é inferior ao valor de acurácia encontrado por Fonseca *et. al.*, 2018, com KQ 0.75. Este resultado mostra que esta tecnologia está alcançando um estado de maturidade para ser utilizada em aplicações práticas, por exemplo, em ambientes virtuais de ensino. Como trabalho futuro pode-se estudar diversas frentes como, aprofundar a importância de cada atributo associado ao score e/ou as competências e aplicar técnicas de *deep learning* na abordagem.

■ REFERÊNCIAS

1. Alencar, L. F. de. (2010) Aelius: uma ferramenta para anotação automática de corpora usando NLTK. IX Encontro de Linguística de Corpus. Porto Alegre, PUCRS.
2. Amália, A. et al. (2019) Automated Bahasa Indonesia essay evaluation with latent semantic analysis. In: Journal of Physics: Conference Series. IOP Publishing.
3. Amorim, E. e Veloso, A. (2017). multi-aspect analysis of automatic essay scoring for Brazilian Portuguese. In Proceedings of the Student Research Workshop at the 15th Conference of the European Chapter of the Association for Computational Linguistics (pp. 94-102).
4. Attali, Y. e Burstein, J. (2006). Automated essay scoring with e-rater® V. 2. The Journal of Technology, Learning and Assessment, 4(3).
5. Bull, J. e Mckenna, C. (2001) A Blueprint for Computer Assisted Assessment. Taylor & Francis Editora.
6. Burrows, S, Gurevych, I. e Stein, B. (2015) The eras and trends of automatic short answer grading. International Journal of Artificial Intelligence in Education.
7. Dasgupta, I., Guo, D., Stuhlmüller, A., Gershman, S. J., e Goodman, N. D. (2018). Evaluating compositionality in sentence embeddings. arXiv preprint arXiv:1802.04302.

8. Dong, F., Zhang, Y. e Yang, J.(2017). Attention-based recurrent convolutional neural network for automatic essay scoring. In Proceedings of the 21st Conference on Computational Natural Language Learning (CoNLL 2017)(pp. 153-162)
9. Elliot, S. (2003). IntelliMetric: From here to validity. Automated essay scoring: A cross-disciplinary perspective, 71-86.
10. Fernández-Delgado, M., Cernadas, E., Barro, S. e Amorim, D. (2014). Do we need hundreds of classifiers to solve real world classification problems?. The Journal of Machine Learning Research, 15(1), 3133-3181.
11. Fleiss, J. L., e Cohen, J. (1973). The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. Educational and psychological measurement,33(3),613-619.
12. Foltz, P. W., Laham, D. e Landauer, T. K.(1999). The intelligent essay assessor: Applications to educational technology. Interactive Multimedia Electronic Journal of Computer-Enhanced Learning, 1(2), 939-944.
13. Fonseca, E., Medeiros, I., Kamikawachi, D., e Bokan, A. (2018). Automatically grading brazilian student essays. In International Conference on Computational Processing of the Portuguese Language (pp. 170-179). Springer, Cham.
14. Haley, D. T. et al. (2007) Seeing the whole picture: evaluating automated assessment systems. ITALICS.
15. Hearst, M. A. (2000) The debate on automated essay grading. IEE Intelligeng Systems archive.
16. Lee, I. (2014). Teachers' reflection on implementation of innovative feedback approaches in EFL writing. English Teaching, 69(1), 23-40.
17. Mohler, M. e Mihalcea, R. (2009) Text-to-text semantic similarity for automatic short answer grading. EACL'09 - Proceedings of the 12th Conference of the European Chapter of the Association for Computational Linguistics.
18. Noorbehbahani, F. e Kardan, A. A. (2011) The automatic assessment of free text answers using a modified bleu algorithm. Computer & Education.
19. Page, E. B. (1966) The imminence of grading essay by computer. The Phi Delta Kappan.
20. Palma, D. e Atkinson, J. (2018) Coherence-Based Automatic Essay Assessment. IEEE Intelligent Systems, v. 33, n. 5, p. 26-36.
21. Rich, C. S.; Schneider, M. C. e D'brot, J. M. (2013) Applications of automated essay evaluation in West Virginia. In: Handbook of Automated Essay Evaluation. Routledge. p. 121-145.
22. Rodrigues, F. e Araújo, L. (2012) Automatic Assessment of Short Free Text Answers.In: CSE-DU (2). p. 50-57.
23. Rudner, L. M., Garcia, V. e Welch, C. (2006). An evaluation of IntelliMetric™ essay scoring system. The Journal of Technology, Learning and Assessment, 4(4)
24. Shermis, M. D., e Hamner, B. (2012). Contrasting state-of-the-art automated scoring of essays: Analysis. In Annual national council on measurement in education meeting (pp. 14-16).

25. Srivastava, N.; Hinton, G.; Krizhevsky, A.; Sutskever, I.; Salakhutdinov, R. (2014) Dropout: a simple way to prevent neural networks from overfitting. *The Journal of Machine Learning Research*, 15(1):1929-1958.
26. Vajjala, S. (2018) "Automated assessment of non-native learner essays: Investigating the role of linguistic features". *International Journal of Artificial Intelligence in Education*, v. 28, n. 1, p. 79-105.
27. Wachsmuth, H., Stein, B., e Engels, G. (2011). Constructing efficient information extraction pipelines. In *Proceedings of the 20th ACM international conference on Information and knowledge management* (pp. 2237-2240). ACM.
28. Yang, W. (2012). A study of students' perceptions and attitudes towards genre-based ESP writing instruction. *The Asian ESP Journal*, 8(3), 50-73.
29. Zupanc, K. e Bosnic, Z. (2015) Automated essay evaluation augmented with semantic coherence measures. *IEEE International Conference on Data Mining (ICDM)*.
30. Zupanc, K. e Bosnic, Z. (2017). Automated essay evaluation with semantic analysis. *Know.-Based Syst.*, 120(C):118 – 132. DOI: 10.1016/j.knosys.2017.01.006

Capacitação em tempos de pandemia: contribuições do modelo LV de avaliação em possibilitar o Growing Focus on Measuring Learning

| Gilvandenys Leite **Sales**
IFCE

| Bento **Silva**

| Universidade do **Minho**
UMINHO

| José Alberto **Lencastre**

| Universidade do **Minho**
UMINHO

| João Batista da **Silva**
UECE

RESUMO

Face a demanda urgente por capacitação diante da pandemia do novo Coronavírus, que impactou os mais diversos setores, inclusive o educacional, e para atender ao que está estabelecido no plano emergencial de isolamento, que prediz ser necessário promover aulas remotas, criou-se o curso Jornada de Metodologias Ativas (JMA). Considerando essa carência, esta pesquisa tem como objetivo investigar as contribuições do Modelo Learning Vectors (Modelo LV), uma ferramenta de avaliação formativa e Não-linear, para potencializar a implementação da tendência educacional “Growing Focus on Measuring Learning” ao longo da formação inicial e continuada de professores durante a aplicação do Curso JMA. Ao fazer uso de tecnologias emergentes e essenciais a esta nova práxis pedagógica ancorado no AVA Moodle e no Modelo LV, com o intuito de difundir o uso de metodologias ativas e oferecer orientações e sugestões de como implementá-las em uma sala de aula remota, espera-se que seus atores possam reinventar a sala de aula e fazer o “Novo Normal” pedagógico alinhado às tendências educacionais contemporâneas.

Palavras-chave: Capacitação de Professores, Avaliação, Modelo LV, Tendências Educacionais Contemporâneas.

■ INTRODUÇÃO

Grandes são os desafios para se implementar Tecnologias Digitais de Informação e Comunicação (TDIC): falta de equipamentos nas escolas; iniquidade de acesso à internet para os alunos mais carentes; inexistência de políticas públicas para democratizar a compra de equipamentos (notebooks, smartphones, computadores de mesa, entre outros) pela maior parte dos alunos e professores, pois tais equipamentos ainda são muito caros aqui no Brasil; falta de formação adequada de professores; falta de políticas públicas para possibilitar a replicação em larga escala de experiências exitosas; entre outras.

Não obstante, nos últimos anos foi percebido que, indubitavelmente, as TDIC são ferramentas essenciais para o processo educativo contemporâneo. A pandemia de 2020, causada pelo novo corona vírus (COVID-19), contribuiu para evidenciar e corroborar o que já havia sido preconizado pelas pesquisas científicas (VALENTE, 1999; MAIA; BARRETO, 2012; FREEMAN et al. 2017; LINHARES et al., 2017; BECKER et al. 2018) a respeito da importância/necessidade de implementar TDIC na educação, como forma de potencializar e contribuir para o ensino e aprendizagem.

Contudo, vale destacar que não basta apenas comprar equipamentos e colocá-los num laboratório de informática, é preciso também que o uso de tecnologias esteja associado a metodologia de ensino e de aprendizagem adequadas (CAVALCANTE; SALES; SILVA, 2018). Mas como fazer isso? Pensar em estratégias que associem tecnologias a metodologias de ensino adequadas é um desafio para todos os níveis educacionais, do infantil ao superior, e em vários países do mundo.

Nesse sentido, para superar este desafio, é imprescindível investir na formação adequada de professores (MAIA; BARRETO, 2012). Vale destacar ainda que, a falta de formação adequada de professores para usar TDIC com fins educacionais não é um problema local, apenas do Brasil ou da América Latina, mas outros países do mundo com os Estados Unidos da América e/ou continentes, como a Europa (JOHNSON et al., 2015).

Para Ribeirinha e Silva (2020) “As TDIC permitem que professores e alunos experienciem essas novas metodologias emergentes, como estamos a constatar neste momento de pandemia, em que, isolados (em casa), estão juntos no online a prosseguir com as aprendizagens”.

Com o propósito único de aplicar metodologias ativas e oferecer orientações e sugestões de como implementá-las em uma sala de aula utilizando TDIC como suporte, desenvolveu-se e aplicou-se um curso destinado, principalmente, à reinvenção da sala de aula, nestes dias que estão se intitulando de o “Novo Normal” devido à pandemia de Coronavírus.

O curso Jornada de Metodologias Ativas visou capacitar, aperfeiçoar e atualizar profissionais de educação na área de metodologias ativas com auxílio de ferramenta digitais

para melhorar o engajamento dos alunos em sala de aula remota e futuramente presencial ou híbrida.

A sala de aula e sua ressignificação, cujos limites pertencem agora ao mundo virtual, foi o foco do curso, onde se discutiu o que são metodologias ativas e como aplicá-las, como proporcionar maior engajamento dos alunos na sala de aula remota e como instrumentar professores para utilizar ferramentas digitais.

Buscou-se fazer uso do Ambiente Virtual de Aprendizagem (AVA) Moodle implementado com o Plugin de avaliação processual Learning Vectors (Modelo LV), o qual substituiu o bloco de Notas Moodle pelo bloco “Notas LV”, o que permitiu fazer uso da métrica qualitativa Fator β como elemento para gamificar e promover ranking dos cursistas.

Freeman et al. (2017) destacam que, além da importância de implementar de TDIC na educação, é preciso também saber o impacto das TDIC como inovação para produzir experiências autênticas de aprendizado.

De acordo com Moreira (2011), apesar de não haver, necessariamente, uma relação causa e efeito entre o ensino e a aprendizagem, já que a aprendizagem é uma atividade idiossincrática, o ensino tem por fim a aprendizagem, mas como saber se houve a aprendizagem?

Para isso é preciso avaliar, e esta não é uma tarefa tão trivial quanto parece, principalmente, na realidade atual do ensino-remoto emergencial, a qual desafiou as instituições de todos os níveis de ensino a desenvolver estratégias avaliativas aplicáveis ao mundo digital.

O problema central da avaliação no ensino remoto é porque, este tipo de ensino demanda o rompimento do modelo avaliativo tradicionalista que emprega, predominantemente, provas objetivas, unidisciplinares e terminais (LIMA et al, 2020).

Long e Siemens (2001, p. 38) destacam que a avaliação, em vez de ser uma atividade que normalmente é aplicada no final de um curso, deveria “ser realizada em tempo real à medida que os alunos demonstram domínio de conceitos ou ideias importantes”.

Diante dessas dificuldades, em especial na avaliação, o NMC Horizon Report, uma comunidade internacional de especialistas em tecnologias educacionais, apontaram algumas tendências educacionais contemporâneas para acelerar a taxa em que a tecnologia está sendo adotada no ensino fundamental e médio, dentre elas a tendência “Growing Focus on Measuring Learning” (Foco Crescente na Medição do Aprendizado). Growing Focus on Measuring Learning é uma tendência educacional contemporânea que “descreve a exploração e avaliação de uma ampla variedade de ferramentas de avaliação usadas para medir a prontidão acadêmica, o progresso escolar, a aquisição de habilidades e o desempenho dos alunos” (FREEMAN et al. 2017, p. 16).

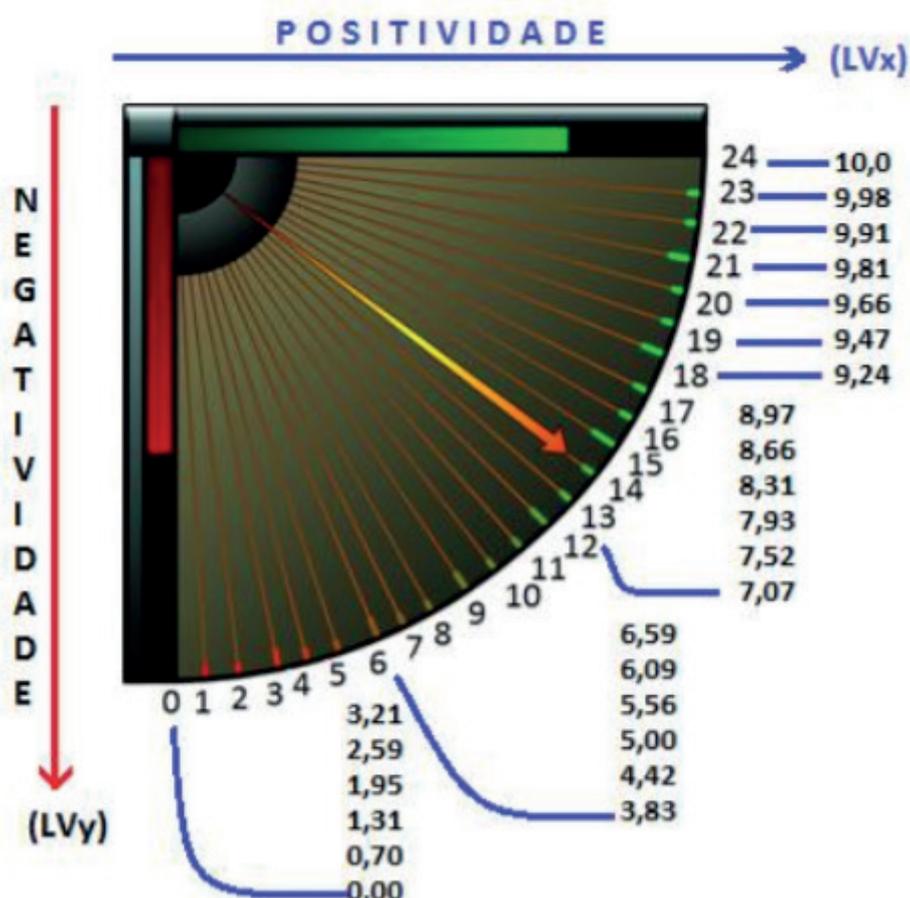
Para tanto, é necessário a disponibilidade de informações confiáveis do desempenho dos alunos em tempo real para que professores e gestores possam planejar suas atividades e tomar decisões, em especial com relação aos alunos que estão em risco (LONG; SIEMENS, 2011).

Todavia, algumas das maiores dificuldades para se implementar algumas tendências apontadas pelo o NMC Horizon Report, são os recursos que muitas vezes não tem ou são muito caros. Considerando essa carência, essa pesquisa tem como objetivo investigar as contribuições do Modelo LV, uma ferramenta de avaliação formativa e Não-linear, para potencializar a implementação da tendência educacional “Growing Focus on Measuring Learning” ao longo da formação inicial e continuada de professores durante a aplicação do Curso Jornada de Metodologias Ativas.

■ MÉTODO

O modelo de avaliação Learning Vectors (Modelo LV) é fundamentada na interação e na avaliação formativa e diagnóstica e visa otimizar a sobrecarga de trabalho no acompanhamento de alunos em Ambientes Virtuais de Aprendizagem (AVA). Em sua especificação o modelo LV faz uso de vetores e de equações trigonométricas (SALES, 2010; SALES; BARROSO; SOARES, 2012) representados no Vetor-aprendizagem (Figura 1).

Figura 1. Representação do Vetor-Aprendizagem do Modelo LV para consulta de notas.



Fonte: Sales et al. (2018)

A projeção horizontal do vetor (LVx) é resultante das notas nas atividades, a qual expressa em parte a positividade de desempenho do aluno. A projeção vertical representa a negatividade do desempenho (LVy), seus 25 possíveis valores de notas associadas aos ângulos também se encontram representadas no Vetor-Aprendizagem (Figura 1).

O processo de avaliação no Ambiente Virtual de Aprendizagem (AVA) é mediada por uma escala de menções qualitativas: Muito Bom, Bom, Regular, Fraco, Não Satisfatório (Tabela 1), categorizadas e associadas a uma escala iconográfica na forma de animação, os LV Gifs (SALES et al., 2018).

Tabela 1. LV Gifs e escala de menções qualitativas e valores do Fator β expressos em mili LV.

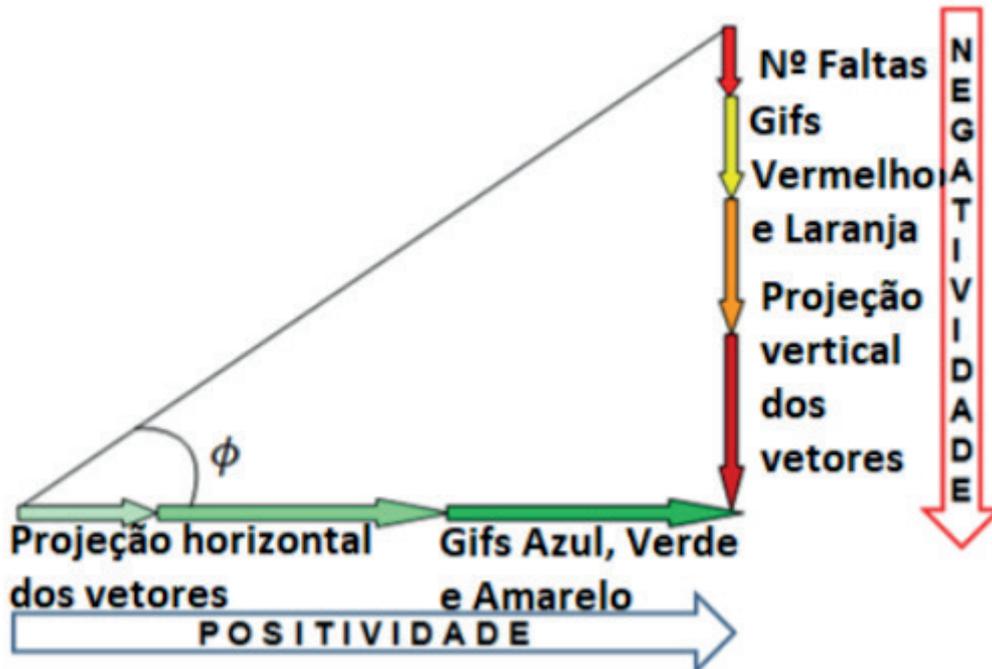
Menções Qualitativas LV Gifs	Não Satisfatório 	Baixo 	Médio 	Alto 	Muito Alto 
Fator β em mLV	$0 \leq \beta < 300$	$300 \leq \beta < 900$	$900 \leq \beta < 2620$	$2620 \leq \beta < 3780$	$\beta \geq 3780$

Fonte: adaptado de Sales (2010)

O Modelo LV contém uma métrica qualitativa não-linear para suporte ao acompanhamento da aprendizagem do aluno, o Fator β , cuja unidade de medida é o mili LV (mLV), seu limite superior, Fator $\beta \geq 3,78$, ou 378 mLV, significa o valor limítrofe a partir do qual se espera encontrar os alunos de “Muito Alto” desempenho, ou seja, alunos com pleno domínio das atividades apresentadas.

Esta métrica vale-se de seu caráter diagnóstico e está fundamentado na dimensão psicológica Positividade(P)/Negatividade(N) (Taxa P/N) do modelo de avaliação de desempenho denominado Meta Learning (Modelo ML) (LOSADA, 1999; LOSADA; HEAPHY, 2004; FREDRICKSON; LOSADA, 2005). As dimensões Positividade/Negatividade estão ilustradas na Figura 2.

Figura 2. Fator β : Positividade X Negatividade



Fonte: adaptado de Sales (2010)

Matematicamente é dado por:

$$\text{Fator } \beta = \frac{\text{Positividade (P)}}{\text{Negatividade (N)}}$$

Em que a Positividade (P) é o somatório das projeções horizontais das notas nas Atividades, acrescido do somatório do Número de Gifs de cor Azul, Verde e Amarela, respectivamente ponderados e associados às menções qualitativas “Muito Bom” (Gif Azul - peso 3), “Bom” (Gif Verde - peso 2) ou “Regular” (Gif Amarelo - peso 1).

A Negatividade (N) é o somatório das projeções verticais dos vetores correlacionados às notas de cada atividade, acrescido do Número Total de Faltas, mais o somatório dos Gifs de cor Laranja e Vermelha, associados às menções qualitativas “Fraco” (Gif Laranja - peso 1) ou “Não Satisfatório” (Gif Vermelho - peso 2).

Como resultado desta concepção de modelo de avaliação para o AVA Moodle, disponibiliza-se o Plugin Modelo LV, desenvolvido em PHP para ser compartilhado livremente no endereço <http://bit.ly/2ANXA1b> .

Quanto à abordagem, esta pesquisa é classificada como qualitativa. Com relação aos procedimentos técnicos foi realizado uma pesquisa exploratória, a qual é caracterizada proporcionar uma visão geral de um determinado fato ou fenômeno. De acordo com Gil (2008,

p 27), “este tipo de pesquisa é realizado especialmente quando o tema escolhido é pouco explorado e torna-se difícil sobre ele formular hipóteses precisas e operacionalizáveis”.

Nesse caso, o fenômeno investigado foi a avaliação formativa, utilizando um Indicador Qualitativo Não-Linear, de um curso Jornada de Metodologias Ativas, o qual foi ofertado com carga horária de 80 horas. Essa iniciativa é resultado dos programas de extensão do Instituto Federal do Ceará (IFCE). A equipe organizadora foi composta por sete tutores entre mestrandos e doutorandos e mais três administradores no suporte técnico do AVA Moodle customizado sob a denominação de Help Class Online.

O público-alvo é composto por alunos de graduação (licenciatura, pedagogia, bacharelado), professores em exercício, ou qualquer pessoa atuante na área educacional, como técnico administrativo educacional, coordenador e diretor educacional. Desta forma, participaram do curso 309 pessoas. A escolha do público-alvo se justifica porque de acordo com Freeman et al. (2017, p. 16), “os professores têm sido pioneiros na implementação de novas metodologias e ferramentas que avaliam a aprendizagem dos alunos de forma inovadora”. Todavia, acredita-se que todo quadro escolar (professores, coordenadores, gestores, entre outros) precisa estar envolvido para um melhor desempenho dos alunos.

O curso foi composto por 12 tópicos a serem trabalhados na modalidade remota emergencial com encontros síncronos e assíncronos. Síncronos nos encontros online em Live-aulas via a plataforma Stream Yard, pelo fato de permitir visualização direta no Youtube de um grande número de participantes simultaneamente, e assíncronos nas diversas atividades dispostas no ambiente virtual Moodle que será utilizado para gerenciamento acadêmico do cursista.

Os tópicos que moldaram o curso foram: À Guisa de Reflexões (abordagens pedagógicas sobre metodologias ativas); Mapa Conceitual e Mentimeter; Google Science Journal, Phet e Phyphox; Gamificação, Kahoot e Quizizz; Ensino Híbrido, Google Sala de Aula e salas de videoconferência; Instrução por Pares, Plickers e Socrative; Vamos Dominar o Moodle Cloud; Sala de Aula Invertida e Ensino Sob Medida; Vídeoaula e Edpuzzle; Aprendizagem Baseada em Projetos e Trello; Pensamento Computacional e Arduino.

Os dados foram coletados do AVA Help Class Online. A avaliação de aprendizagem no curso foi do tipo processual. Para isso foram realizados três tipos de atividades, Fóruns Pergunta e Respostas, Portfólios e Quizzes, sendo um total de 18 Quizzes do tipo: múltipla escolha, associação de colunas, verdadeiro ou falso e preenchimento de lacunas, e mais, atividades de exploração de simulações e plataformas de elaboração de quizzes. O controle de frequência foi automatizado no ambiente virtual facilitado pelo Plugin de Avaliação Learning Vectors, o qual auxiliou no cômputo de faltas às atividades.

Com relação aos aspectos éticos, essa pesquisa considerou as diretrizes da Resolução a Resolução N° 510, de 7 de abril de 2016 do Conselho Nacional de Saúde (CNS n° 510/2016), a qual dispõe sobre as normas aplicáveis a pesquisas em Ciências Humanas e Sociais. Sendo assim, de acordo com a CNS n° 510/2016, não precisam ser registradas no sistema CEP/CONEP as pesquisas que utilizem informações de acesso público, nos termos da Lei n° 12.527, de 18 de novembro de 2011, a que utilizem informações de domínio público e as pesquisas que utilizem bancos de dados, cujas informações são agregadas, sem possibilidade de identificação individual.

Apresentados os aspectos metodológicos da pesquisa, na próxima seção serão apresentados e discutidos os resultados.

■ RESULTADOS E DISCUSSÃO

O curso Jornada de Metodologias Ativas visou capacitar, aperfeiçoar e atualizar profissionais de educação na área de metodologias ativas com auxílio de ferramenta digitais para melhorar o engajamento dos alunos em sala na aula remota, presencial ou híbrida. Como suporte foi utilizado o Ambiente Virtual Moodle Help Class Online. Nesse ambiente foi implementado com o Plugin de avaliação processual Learning Vectors (LV), o qual substituiu o bloco de Notas Moodle pelo bloco “Notas LV”, o que permitiu fazer uso da métrica qualitativa Fator β como elemento quali-quantitativo para mensurar a avaliação.

Ao longo do curso os cursistas tiveram a oportunidade de consultar seu rendimento a qualquer momento, expresso por um valor de nota correspondente à média, limitada de zero a dez, com valor mínimo de 6,0 para sua aprovação, como também podia consultar seu desempenho representado pelo Fator β , cujos valores extrapolam estes limites por ser uma métrica não linear e, cuja unidade de medida é o mili LV.

Um exemplo do mapa de notas pode ser visualizado na Tabela 2. Este mapa estava sempre disponível, fazendo da avaliação um processo contínuo, formativo e com feedback imediato para alunos, professores e instituição.

Tabela 2. Exemplo do mapa de notas LV (Fator β Médio da Turma: 0.73)

Nome	Portfólios	Quizzes	Média	NTF* (hA / %)	Fator β (mili LV)	LV Gif	Situação
Aluno 1	1	8.65	9.7	0 / 0%	673		C
Aluno 2	1	8.85	9.9	0 / 0%	962		C
Aluno 3	1	7.62	8.6	4 / 5%	220		C
Aluno 4	1	8.65	9.7	0 / 0%	529		C
Aluno 5	0	6.16	6.2	4 / 5%	82		C
Aluno 6	0	0.48	0.5	72 / 90%	2		C
Aluno 7	1	8.08	9.1	0 / 0%	29		C
Aluno 8	0	0	0	72 / 90%	0		C

Fonte: Adaptado de helpclassonline.com.br

Com o auxílio dessa ferramenta foi possível que, tanto os alunos/cursistas quanto os professores pudessem acompanhar seu desempenho em tempo real. Conforme Freeman et al. (2017, p. 16), este tipo de avaliação diagnóstica, contínua e com feedback imediato apresenta-se como uma alternativa promissora à avaliação sumária, validando, assim, “a aprendizagem dos alunos por meio de portfólios digitais e avaliações autênticas que têm alunos realizando uma tarefa ou projeto em um ambiente real”.

Desta forma, o aluno, sempre que quiser, tem a oportunidade de consultar na coluna Portfólio (TABELA 1) suas notas das atividades enviadas (Atividades Tipo 2), e consequentemente visualizar seu desempenho instantâneo e em tempo real (Figura 3).

Figura 3. Imagem do Learning Vectors (LV) visualizada pelo aluno ao seu desempenho no AVA



Fonte: AVA Help Class Online

De acordo com Long e Siemens (2011, p. 40), é de extrema importância que os alunos recebam em tempo real, o que pode ser motivador e encorajador, as “informações sobre seu desempenho em relação aos seus pares ou sobre seu progresso em relação aos seus objetivos pessoais”.

Contudo não basta apenas isso, é preciso também que essas informações sejam disponibilizadas de forma acessível e oportuna (Freeman et al. 2017). Para isso, os resultados das atividades/avaliações foram convertidos em notas mensuráveis, de maneira que pudesse fornecer um panorama de sua situação. Ao consultar a coluna Quizzes (Tabela 2) o aluno confere suas notas nos questionários (TABELA 3). E, caso necessite, tem a oportunidade de tomar decisão e mudar de ação para se recuperar a tempo.

Tabela 3. Exemplo de notas de um determinado aluno nas Atividade de Quizzes

Atividades Tipo 1	Nota
Quiz 1 - Tópico 1 – Reflexões	9,4
Quiz 1 - Tópico 2 - Mapa Conceitual	9,8
Quiz 2 - Tópico 2 - Mentimeter e Brainstorming	10,0
Quiz 1 - Tópico 6 – Instrução por pares	10,0
Quiz 1 - Tópico 5 - Ensino Híbrido	9,3
Quiz 2 - Tópico 5 - Google Sala de Aula	10,0
Quiz 3 - Tópico 5 - Salas de videoconferência	10,0
Quiz 1 - Tópico 4 - Gamificação	10,0
Quiz 1 - Tópico 9 - Videoaula	10,0
Quiz 1 - Tópico 3 - Google Science Journal	9,8
Quiz 2 - Tópico 3 - Phet	10,0
Quiz 3 - Tópico 3 - Phypox	9,8
Quiz 1 - Tópico 7 – Moodle Cloud	10,0
Quiz 1 - Tópico 8 - Sala de Aula Invertida e Ensino sob Medida	10,0
Quiz 2 - Tópico 10 - Trello	10,0
Quiz 1 - Tópico 10 - Aprendizagem Baseada em Projetos	10,0
Quiz 1 - Tópico 11 - Pensamento Computacional	9,5
Quiz 2 - Tópico 11 - Arduino	9,3

Fonte: Próprio autor

O cálculo do Fator β , automatizado pelo Plugin LV, pode ser visualizado na simulação exposta na Tabela 4, tomou-se como exemplo o Aluno 2:

Tabela 4. Simulação de nota

Aluno(a): Aluno 2 / JMA 2020 / Carga Horária - 80 hA					
Atividades	Peso	Peso	Positividade	Média=	Negatividade
Tipo 1 (Quizzes)	(S)	Gobal (G)	LVx (P)	(S.G.P)	LVy (N)
T. 1 – Reflexões	6%	90%	9,4	0,508	3,41
T. 5 – Ens. Híbrido	6%		9,33	0,504	3,6
T. 4 – Gamificação	6%		10	0,54	0
T. 3 – GoogleScien	6%		9,75	0,527	2,22
T. 11 – PensComp	5%		9,5	0,428	3,12
T. 10 – ProjectBL	5%		10	0,45	0
T. 6 – Inst. Pares	5%		10	0,45	0
T. 5 – Videoc	6%		10	0,54	0
T. 3 – Phypox	5%		9,83	0,442	1,84
T. 5 – GoogleClass	6%		10	0,54	0
T. 3 – Phet	6%		10	0,54	0
T. 2 – Mapa Conceitual	6%		9,75	0,527	2,22
T. 7 – MoodleCloud	5%		10	0,45	0
T. 8 – SalaInvertida	5%		10	0,45	0
T. 9 – Videoaula	5%		10	0,45	0
T. 10 – Trello	6%		10	0,54	0
T. 11 – Arduino	5%		9,33	0,42	3,6
T. 2 – Menti/Brains	6%	10	0,54	0	
Soma parcial	100%		176,89	8,844	20,012
Atividades					
Tipo 2 (Portfólio)					
Kahoot e Quizizz	50%	10%	10	0,5	0
Moodle Cloud	50%		10	0,5	0
Soma parcial	100%		20	1	0
Soma Total		100%	196,89	9,844	20,01
Análise Quantitativa			Análise Qualitativa		
(Rendimento/Média)			(Desempenho/Fator β)		
Número Total de Faltas (NTF)	0		P - POSITIVIDADE (LVx+Gifs)	196,89+6=202,89	
Faltas Percentual (NTF%)	0		N – NEGATIVIDADE (LVy+ NTF+Gifs'+1)	20,01+0+0+1=21,01	
Nota QUIZZES	8,844		Fator $\beta = P/N$	9,66 ou 9660 mLV	
Nota PORTFÓLIO	1		Situação Final	AM	
Média Final	9,844		Conceito	A	

Fonte: Próprio autor

Na composição do Fator β faz-se o computo do número de LV Gifs atribuídos às atividades (Tabela 5). No caso, o Aluno 2, recebeu LV Gifs Azul, um em cada atividade de Portfólio, por isso na Tabela 5 foi acrescido o valor seis na positividade e zero na negatividade.

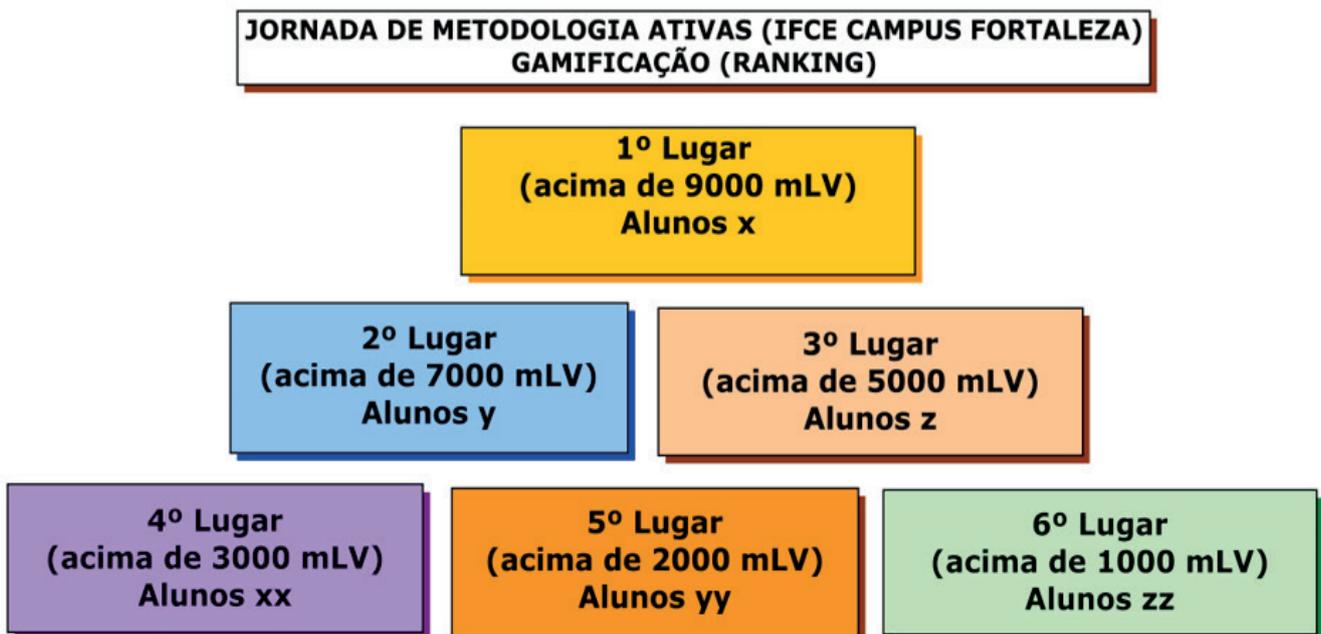
Tabela 5. Número de LV Gifs atribuídos às atividades

Atividades	Número de LV Gifs				
	Gifs de POSITIVIDADE			Gifs' de NEGATIVIDADE	
	Azul x(3)	Verde x(2)	Amarelo x(1)	Laranja x(1)	Vermelho x(2)
Portfólio 1	1				
Portfólio 2	1				
Total	6	0	0	0	0

Fonte: Próprio autor

Ao longo do curso procurou-se usar elementos de games na tentativa de gamificar o AVA, assim foi apresentado no mural do curso um podium com o ranking dos melhores alunos tomando por base os valores do Fator β (FIGURA 4). Para Barradas & Lencastre (2017) “A competitividade positiva é uma forma de elementos individuais competirem entre si para melhorar a sua posição num grupo, de uma forma cooperativa em que há respeito mútuo e interações que não prejudicam os outros membros do grupo.”

Figura 4. Podium - Fator β possibilitando o Ranking dos cursistas

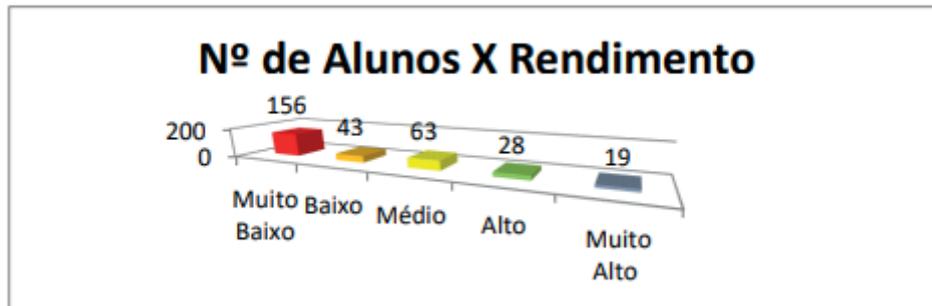


Fonte: Próprio autor

Tomando por base a nota mínima 6,0 para aprovação por média/rendimento e Fator β no intervalo de 90 miliLV a 2620 miliLV, o que corresponde a Médio Desempenho do cursista, ficam estabelecidas as condições sine qua non para receber a certificação da instituição.

Os resultados quanto ao desempenho, fundamento no Fator β e valores da Tabela 1, podem ser visualizados no Gráfico 1.

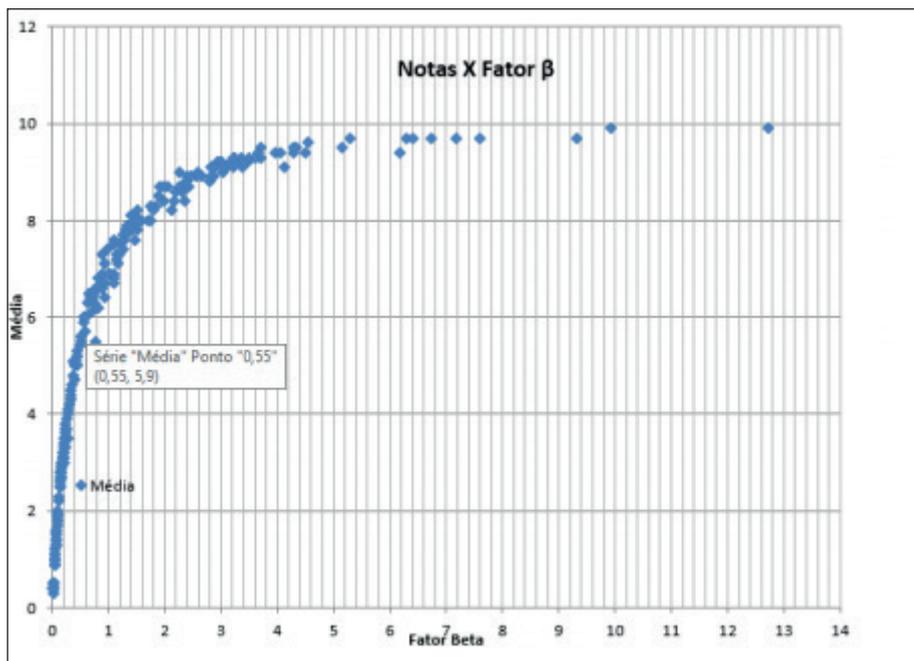
Gráfico 1. Classificação dos cursistas quanto ao Fator β



Fonte: Próprio autor

Uma análise deste gráfico revela que atingiram os critérios de aprovação no prazo regulamentar do curso em termos de desempenho, os cursistas classificados com rendimento Médio, Alto e Muito Alto, um total de 110 alunos. Aos demais foi estendido o prazo de conclusão, para que em seu ritmo e tempo, pudessem fazer as atividades. Esta flexibilização foi importante, por oportunizar aprendizagens, afinal o mais importante é atingir um grande número de profissionais da educação com esta capacitação.

Gráfico 2. Comparação entre a Média e o Fator β



Fonte: Próprio autor

Observa-se no Gráfico 2, que ao critério mínimo de média 6,0, associa-se o Fator β de 550 miliLV, que está no intervalo do critério de Médio Desempenho (900 a 2620 miliLV). O Gráfico 2 possibilita ainda fazer uma comparação entre alunos que obtiveram uma mesma média, mas que têm Fatores β diferentes, em destaque no gráfico alunos que tiraram média/Fator β : 9,7/6290 miliLV; 9,7/6410 miliLV; 9,7/6730 miliLV; 9,7/7170 miliLV; 9,7/7590

miliLV. Todos destaque tem Média 9,7. Ao analisar os valores do Fator β , pode-se concluir que o melhor aluno é quem tem esse valor maior de Fator β (7590 miliLV). Esta é mais uma funcionalidade desta métrica não linear, diferenciar alunos com médias iguais.

■ CONCLUSÃO

Um dos problemas do sistema de avaliação tradicionais é que eles só são realizados no final de uma etapa. Este fato deixa lacunas escancaradas de ações e oportunidade de intervenções com alunos que estão em risco de abandonar ou desistir do curso. Desta forma o Modelo LV oferece feedback imediato para que, tanto os alunos quanto professores e instituições possam tomar decisões antecipadas para evitar um possível insucesso acadêmico do aluno.

Os resultados observados no Modelo LV, explorados nessa pesquisa, coaduna com os pressupostos de Becker et al. (2018), de que deve haver mudança para uma aprendizagem que possa ser medida de forma clara, transparente e instantânea para apoiar os alunos em suas experiências de aprendizagem. Desta forma, acredita-se que os Learning Vectors (LV) mostraram-se uma ferramenta que se propõe a contribuir para promover a tendência educacional Growing Focus on Measuring Learning.

Acerca do curso Jornada de Metodologias Ativas IFCE 2020 que visou, principalmente capacitar professores para as demandas do século XXI, onde as incertezas presentes transformam-se em oportunidades de reflexão e momento de capacitação e apropriação de novas ferramentas, crê-se que sua maior contribuição, representada por esta capacitação remota, foi transformar professores para atuar numa sala de aula em que recursos tecnológicos são imprescindíveis e, inadiáveis são as razões para fazer da sala de aula realmente um espaço prazeroso de aprendizagem e de interações com o uso de TDIC.

A inserção do elemento de game, ranking, proporcionada pelo Fator β do Modelo LV de avaliação processual, serviu para motivar os cursistas e gerar empoderamento, além de que o uso do Plugin LV otimizou a carga de trabalho do professor no gerenciamento do curso, como o cômputo automatizado de faltas.

Espera-se que professores em exercício, ou que estejam em formação, possam incorporar esse novo fazer didático em suas novas salas de aula que se descortina em meio ao caos da pandemia que nos assola.

■ REFERÊNCIAS

1. BARRADAS, R.; LENCASTRE, J. A. Gamification e game-based learning: estratégias eficazes para promover a competitividade positiva nos processos de ensino e de aprendizagem. *Revista Investigar em Educação*, 6 (2), 11-37. 2017. Disponível em: <http://pages.ie.uminho.pt/inved/index.php/ie/article/view/118> Acessado em: 20.mar.2021
2. BECKER, S. A.; BROWN, M.; DAHLSTROM, E.; DAVIS, A.; DE PAUL, K.; DIAZ, V.; POMERANTZ, J. NMC Horizon Report. Higher Education Edition. Louisville, CO: EDUCAUSE. 2018. Disponível em: <https://library.educause.edu/~media/files/library/2018/8/2018horizonreport.pdf>
3. CAVALCANTE, A. A.; SALES, G. L.; SILVA, J. B. Digital technologies in Physics education: an experience report using the Kahoot as a tool of evaluation. *Research, Society and Development*, 7 (11), 01-17. 2018. Disponível em: <https://rsdjournal.org/index.php/rsd/article/view/456>
4. FREDRICKSON, B. L.; LOSADA, M. Positive Affect and the Complex Dynamics of Human Flourishing. *American Psychologist*, 60 (7), 678 – 686. 2005. DOI: <https://doi.org/10.1037/0003-066X.60.7.678>
5. FREEMAN, A.; ADAMS BECKER, S.; CUMMINS, M.; DAVIS, A; HALL GIESINGER, C. NMC/CoSN Horizon Report. K–12 Edition. Austin, Texas: The New Media Consortium, 2017.
6. LIMA, M. M. et al. Estratégias avaliativas no ensino remoto: avaliação terminal ou contínua? *Anais do 39º Seminário de atualização de práticas docentes*, 2020. Disponível em: <http://anais.unievangelica.edu.br/index.php/praticasdocentes/article/view/5758>
7. LONG, P.; SIEMENS, G. Penetrating the fog: Analytics in learning and education. *EDUCAUSE review*, 46 (5). 2015. Disponível em: <https://er.educause.edu/articles/2011/9/penetrating-the-fog-analytics-in-learning-andeducation>
8. LOSADA, M. The complex dynamics of high performance teams. *Mathematical and Computer Modelling*, v. 30, n. 9, 179-192. 1999. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0895717799001892>
9. LOSADA, M.; HEAPHY, E. The role of positivity and connectivity in the performance of business teams: A nonlinear dynamics model. *American Behavioral Scientist*, vol. 47, nº 6, 740-765. February. 2004. Disponível em: <https://journals.sagepub.com/doi/abs/10.1177/0002764203260208>
10. MAIA, D. L.; BARRETO, M. C. Tecnologias digitais na educação: uma análise das políticas públicas brasileiras. *Educação, Formação & Tecnologias*, 5 (1). 2012. Disponível em: <http://eft.educom.pt/index.php/eft/article/view/213> MOREIRA, M. A. Metodologias de pesquisa em ensino. São Paulo: Livraria da Física, 2011.
11. RIBEIRINHA, T.; SILVA, B. Avaliando a eficácia da componente online da “sala de aula invertida”: um estudo de investigação-ação. *Revista e-Curriculum*, 18(2), 568- 589, 2020. Disponível em: <https://revistas.pucsp.br/curriculum/article/view/47997>
12. SALES, G. L. Learning Vectors (LV): um modelo de avaliação da aprendizagem em EaD online aplicando métricas não-lineares. Tese Doutorado. Departamento de Engenharia de Informática. Universidade Federal do Ceará. 2010. 239f. Disponível em: <https://goo.gl/2ULNJU>

13. SALES, G. L.; ALBUQUERQUE, M. C. N.; LEITE, E. A. M.; PAZ, R. P. Dez Anos do Modelo Learning Vectors na Avaliação Formativa no AVA Moodle: de Emoticons a Gifs Animados In: Tecnologias da Educação: passado, presente, futuro. Coleção História da Educação. 1 ed. Fortaleza: UFC, v.1, p. 63-78, 2018.
14. SALES, G. L.; BARROSO, G. C.; SOARES, J. M. Learning Vectors (LV): Um Modelo de Avaliação Processual com Mensuração Não-Linear da Aprendizagem em EaD online. Revista Brasileira de Informática na Educação, 2012. Disponível em: <https://www.brie.org/pub/index.php/rbie/article/view/1368>

Gerenciamento de redes utilizando ZABBIX

| Adriane Santos **Rezende**

| Joseane Santos de **Jesus**

| José Damião de **Melo**

| Eduardo Henrique do Lago **Silva**

| Mayka de Souza **Lima**

RESUMO

Objetivo: Demonstrar como ocorre o gerenciamento de rede utilizando o software livre Zabbix, para o monitoramento e gerenciamento de redes de computadores de modo a ampliar sua disponibilidade ampliando o leque de opções do administrador de redes com o uso de uma ferramenta de código aberto, de forma que seja capaz de monitorar equipamentos remotos, analisando os dados de modo a garantir o funcionamento e operação dentro dos limites especificados além de controlar o sistema fazendo ajustes de acordo com as modificações ocorridas detectando comportamentos anormais. **Método:** Ao longo da pesquisa foi criado um ambiente virtual multiplataforma, para análise do desempenho do Zabbix, visando melhorias no tráfego de informações e a diminuição de problemas e falhas na rede. **Resultados:** Foi possível detectar as falhas na rede e solucionar o problema remotamente; analisar e controlar o desempenho, mesmo em ambiente virtual multiplataforma, atuando ativamente via monitoramento utilizando a interface web do Zabbix, com gráficos e mapas detalhados e em tempo real, informando o estado atual da rede. **Conclusão:** O Zabbix apresentou um número significativo de funcionalidades e possibilidades de uso, podendo gerenciar e monitorar uma rede de computadores de forma eficaz, com o intuito de auxiliar o gerente de redes a torná-la disponível a maior parte do tempo, comprovadamente demonstrado em nossa pesquisa via estudo de caso em ambiente virtualizado multiplataforma.

Palavras-chave: Gestão, Redes, Open Source, Zabbix.

■ INTRODUÇÃO

O gerenciamento de redes de computadores surgiu inicialmente, como meio de compartilhar dispositivos e periféricos. Entretanto, com a rápida evolução da tecnologia de redes computacionais e sua inserção em toda a sociedade, passaram a fazer parte do cotidiano das pessoas como uma ferramenta que oferece recursos e serviços, com destaque para a principal que é a internet.

Com o avanço da tecnologia, o mundo da interconexão de sistemas passou a conviver com a grande heterogeneidade de padrões, sistemas operacionais e equipamentos, tornando-se cada vez mais necessário o gerenciamento de redes computacionais.

As redes de computadores são compostas por vários dispositivos que se comunicam e compartilham recursos e para gerenciar esses sistemas e as próprias redes, é necessário utilizar ferramentas de gerenciamento automatizadas. Este gerenciamento é a coordenação de recursos materiais ou lógicos fisicamente distribuídos na rede.

Seu objetivo é fornecer ferramentas para que um administrador de redes seja capaz de monitorar equipamentos remotos, analisando os dados de modo a garantir o funcionamento e operação dentro dos limites especificados além de controlar o sistema fazendo ajustes de acordo com as modificações ocorridas detectando comportamentos anormais.

Dentre todos os protocolos de rede que fazem parte do TCP-IP, no quesito administração se destaca o SNMP - Simple Network Management Protocol (Protocolo Simples de Gerenciamento de Rede) que é um protocolo da camada de aplicação criado para transportar informações de gerência de rede entre os dispositivos gerenciados e os sistemas de gestão de redes, ele possibilita que administradores de rede gerenciem o desempenho da uma rede monitorando interfaces, processadores, memórias de equipamentos como roteadores, switches, dispositivos wireless e servidores. Com isso, os administradores de redes conseguem visualizar o status atual da rede, manter um histórico de atividades, bem como receber avisos de forma imediata para ajudar na resolução de problemas. (FACHINNI, 2015)

Já do ponto de vista de aplicações de gerenciamento, o sistema Zabbix vem se destacando como uma solução para empresas de pequeno, médio e grande porte. Seu objetivo é monitorar o tráfego e disponibilidade de serviços em uma rede. O Zabbix monitora diversos hosts por agente próprio, por SNMP, *json*, *icmp*, dentre outros (BONOMO, 2006).

Além disso, o Zabbix disponibiliza visualização em gráficos bem detalhados, mapas descritivos, faz autenticação segura de usuários, possui permissões flexíveis de usuários, sua interface é baseada em web, fornece ao administrador a capacidade de realizar a execução de comandos remotamente, exibir notificações ao administrador da rede via e-mail, indica escalabilidade dos equipamentos e hosts, dentre outras vantagens (BONOMO, 2006; BARBOSA, LEITE e SANTANA, 2014).

Para muitas empresas, manter a rede sempre ativa não é uma tarefa fácil, pois requer do administrador de rede o planejamento e levantamento das necessidades e objetivos da empresa. A disponibilidade de um sistema que permita atuar de forma centralizada e com maior assertividade é essencial para o sucesso da atividade de gerenciamento de redes e serviços.

Assim, visando uma melhor solução para os problemas na rede, é importante que exista uma ferramenta que auxilie o administrador de rede na resolução de problemas fazendo o gerenciamento e monitoramento.

A esta demanda podemos utilizar como resposta o Zabbix, que será apresentado melhor no decorrer deste artigo, além de uma comparação qualitativa do mesmo com outras ferramentas opensource para o mesmo fim, finalizando com uma discussão acerca de um caso de uso utilizando máquinas virtuais.

■ MÉTODO

O gerenciamento de rede é uma forma que o administrador de rede tem de auxiliá-lo a trabalhar, de modo a garantir a máxima eficiência e transparência da rede para seus usuários. Assim, é importante que o administrador utilize algum software que gerencie e monitore a rede. Para isto, existem vários softwares (ferramentas) livres que auxiliam o administrador no gerenciamento e monitoramento da rede.

Ao longo da pesquisa, foram identificadas a existência de várias ferramentas livres que auxiliam o administrador de rede, as principais ferramentas livres utilizadas são: Zabbix, Nagios e Cacti. Dentre elas, neste projeto escolhemos o Zabbix, como já foi descrito na fundamentação teórica, por ser a mais utilizada e por ter maior facilidade de implementação suporte. A seguir, descrevemos o conceito e funcionalidades do **Nagios** e **Cacti** e logo após um comparativo dentre elas, validando a escolha da ferramenta.

Nagios: é uma aplicação de monitoração de redes licenciado pelo sistema GPL, e tem com o objetivo monitorar equipamentos como: hosts, switches, roteadores, aplicações, carga do processador, capacidade em disco, etc. Ele notifica o administrador da rede quando serviços e/ou equipamentos estiverem indisponíveis. As notificações podem ser através de e-mails, SMS, ou qualquer outro meio definido pelo usuário através de plugin.

Os pontos negativos apresentados pelo Nagios são: relatórios escassos, a complexidade para ao adicionar um novo dispositivo na rede, sendo necessário editar vários arquivos de configuração onde estão contidas diversas informações (BARBOSA, LEITE e SANTANA, 2014).

Cacti: é uma ferramenta front-end (interfaces que permitem o usuário interagir como programa) ao RRDtool (ferramenta Round Robin Database ou sistema de base de dados

round Robin criado) , onde a mesma reporta gráficos, informando o estado dos elementos monitorados na rede através do SNMP. Essas informações são armazenadas em um banco de dados Mysql, e podem ser consultadas via web (BARBOSA, LEITE e SANTANA, 2014, 2014).

Na Tabela 1, apresentamos um comparativo, em algumas funcionalidades (ações), entre as ferramentas livres citadas anteriormente.

Tabela 1. Comparação das ferramentas livres Nagios, Cacti e Zabbix.

Ação	Nagios	Cacti	Zabbix
SLA Reports	Através de plugin	Não	Sim
Auto Discovery	Através de plugin	Através de plugin	Sim
Agente	Sim	Não	Sim
Permite Scripts Externos	Sim	Sim	Sim
Linguagem que foi escrito	Perl	PHP	C e PHP
SNMP	Através de plugin	Sim	Sim
Gatilhos / Alertas	Sim	Sim	Sim
Front-end Web	Controle Parcial	Controle Completo	Controle Completo
Método de Armazenamento de Dados	Mysql, MSSQL	RRDTool, Mysql, PostgreSQL em desenvolvimento	Oracle, Mysql, PostgreSQL, e SQLite.
Licenciamento	GPL	GPL	GPL
Geração de Gráficos / Mapas	Sim / Sim	Sim / Através de plugin	Sim / Sim
Eventos	Sim	Através de plugin	Sim

Fonte: Autores

Como podemos observar na tabela 1, o Zabbix é a ferramenta que possui um maior número de ações e vantagens, comparando-a com as outras. Após a validação da ferramenta, elaboramos uma simulação de uma rede através da criação de máquinas virtuais.

Para a criação das máquinas virtuais foi utilizada a ferramenta de virtualização Virtual Box. O Virtual Box é uma solução de virtualização de software gratuita, que permite a instalação de sistemas e aplicativos dentro de uma máquina física. O Virtual Box foi escolhido por ser uma ferramenta simples e leve, que é fornecida pela Oracle.

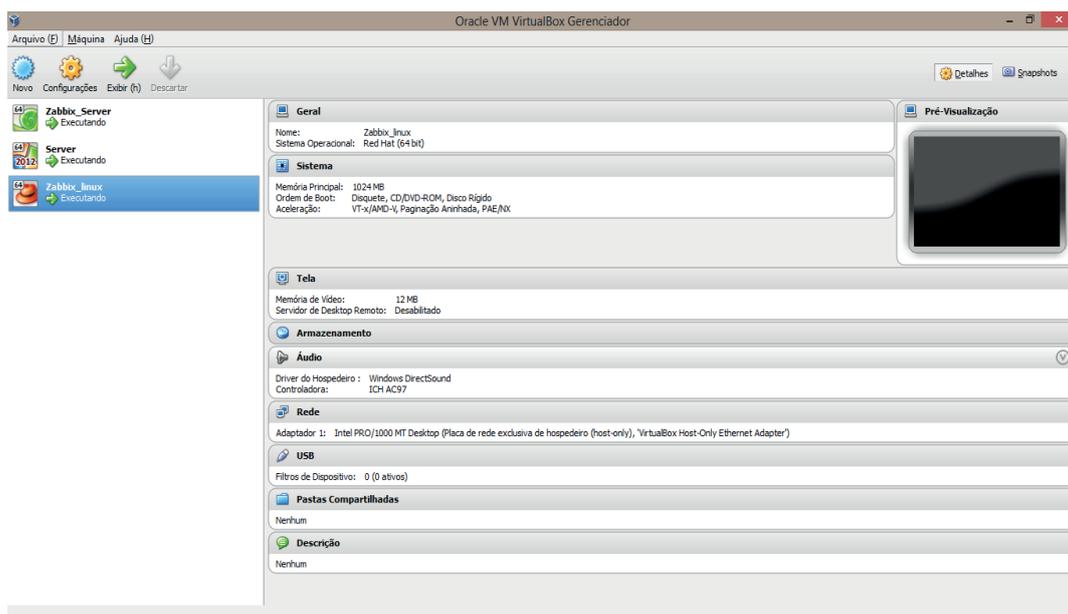
Na Figura 1 apresentamos as máquinas virtuais construídas no ambiente Virtual Box, que foram utilizadas na simulação da rede virtual, criados no ambiente multiplataforma utilizado para o estudo de caso de validação do Zabbix.

Assim, foram instaladas no Virtual Box três máquinas virtuais, duas com o sistema operacional Linux e uma com Windows. Como o gerente do Zabbix é configurado apenas no sistema operacional Linux, uma das máquinas virtuais Linux (openSUSE) foi utilizada para ser o gerente.

Além disso, o agente do Zabbix funciona em qualquer plataforma, sendo assim um agente foi instalado na máquina virtual com sistema operacional Windows (Windows 2012) e o outro com o sistema operacional Linux (Red Hat 64bits), como mostra a Figura 1.

Assim, foram efetuadas as instalações dos aplicativos agentes e gerente a serem monitorados. No caso do gerente, efetuamos o download do arquivo OVF em <http://www.zabbix.com/download.php>. para ser importado através do Virtual Box, sendo que a instalação foi feita no sistema operacional openSUSE.

Figura 1. Interface do Virtual Box.



Fonte: Autores

Para a instalação dos agentes em ambientes Linux e Windows, foi necessário também o download dos agentes em <http://www.zabbix.com/download.php>. Após isto, foram instalados os agentes nas máquinas. Na tela de monitoramento do Zabbix, é possível configurar e adicionar os dispositivos que serão monitorados, sendo assim, após a instalação dos agentes, eles foram adicionados na interface de monitoramento do Zabbix.

■ RESULTADOS E DISCUSSÃO

Podemos considerar que o Zabbix utiliza o modelo de gerenciamento de rede da ISO, sendo eles: é possível detectar as falhas na rede (através do gerenciamento de falhas) e solucionar o problema remotamente; analisa e controla o desempenho da rede (gerenciamento de desempenho); permite ao administrador de rede saiba quais dispositivos fazem parte (gerenciamento de configuração); controle de acesso de usuários e dispositivos aos recursos da rede (gerenciamento de contabilização); controle de acesso aos recursos de rede de acordo com alguma política definida (gerenciamento de segurança).

Dentre outras coisas, conseguimos visualizar o monitoramento da rede através da interface web disponibilizada pelo Zabbix. Além disso, observamos que o Zabbix disponibiliza gráficos e mapas detalhados e em tempo real, informando o estado atual da rede. Podemos visualizar, através de gráficos, o monitoramento do espaço da memória, número de processos rodando, carga do processador, swap, entre outros. O Zabbix também pode ser configurado com um slide show para que as telas fiquem girando a todo o momento.

Na Figura 2, mostra-se a interface via web do status do Zabbix. Nela podemos visualizar se o gerente Zabbix está rodando, a quantidade de hosts, a quantidade de itens, quantidade de briggers, o número de usuários online e o desempenho do servidor.

Figura 2. Interface do Zabbix

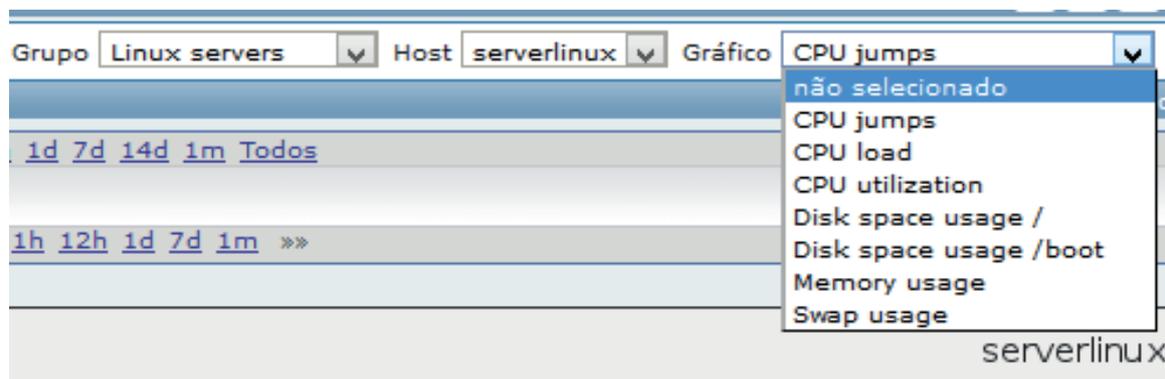


Fonte: Autores

Como foi dito anteriormente, o Zabbix disponibiliza gráficos informando a carga do processador, número de processos rodando, número de processador, espaço de memória, swap, entre outros.

A Figura 3, mostra estas opções que o agente Linux do Zabbix disponibiliza para a visualização em gráficos.

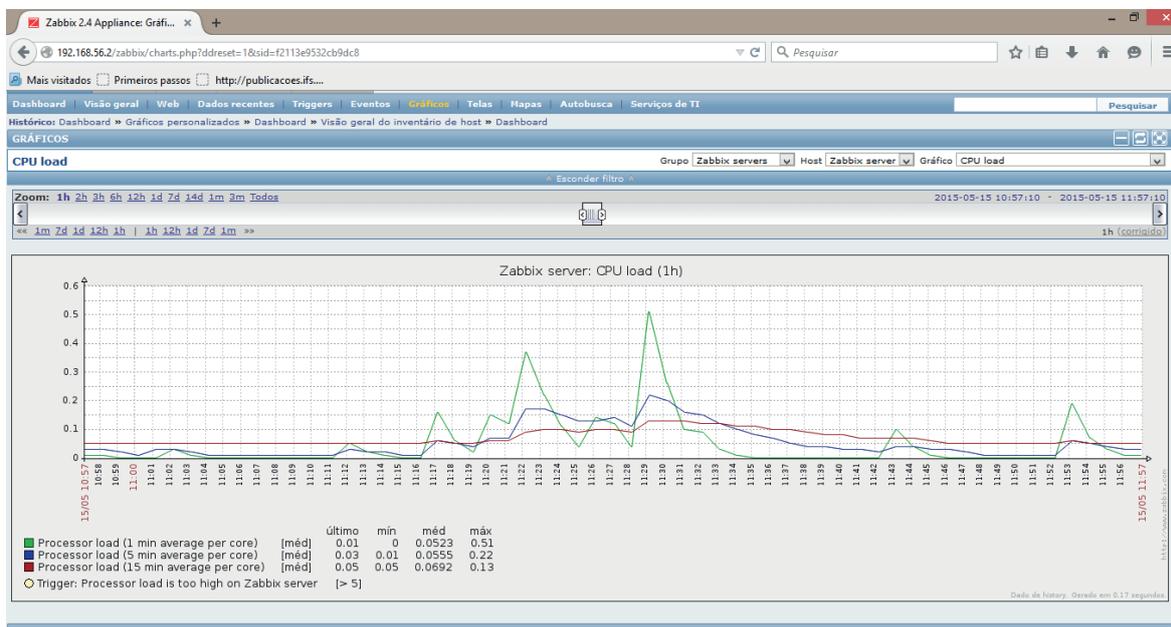
Figura 2. Opções para visualização em gráficos do agente Linux



Fonte: Autores

Na Figura 4, mostra o gráfico disponibilizado pelo Zabbix, da carga da CPU do servidor Linux.

Figura 2. Gráfico da carga da CPU



Fonte: Autores

CONCLUSÃO

Como o objetivo do projeto foi demonstrar como ocorre o gerenciamento de rede utilizando o software livre Zabbix, que monitora e gerencia a rede de modo a torná-la mais funcional, percebemos a partir das avaliações que o Zabbix é uma ferramenta que possui várias opções de configuração e ajustes que ampliam o leque de opções do administrador de redes.

Concluimos ainda que o Zabbix é a ferramenta, dentre as analisadas, que possui um maior número de funcionalidades e possibilidades de uso, podendo gerenciar e monitorar uma rede de computadores de forma eficaz com o intuito de auxiliar o gerente de redes a torna-la disponível a maior parte do tempo.

Ao longo da pesquisa surgiram novas possibilidades que não foram desenvolvidas, pois o tempo não seria suficiente para sua condução, uma das principais possibilidades é o aprofundamento da ferramenta Zabbix, preferencialmente adotando-a em um cenário de uso real, inclusive com a coleta da visão do administrador.

Uma segunda possibilidade é seu uso para fins que vão além do gerenciamento, se colocando como elemento de auditoria ou mesmo como banco de dados de comportamento organizacional para projetos futuros.

■ REFERÊNCIAS

1. BARBOSA, Daniel da Conceição, LEITE, Elson Paixão Silva, SANTANA, Marcélio de Oliveira. **Criação e Gerenciamento de Uma Rede Virtual de Computadores Simulando um Ambiente Real Analisando os Benefícios de Um Ambiente Monitorado**. 2014. Projeto Integrador V (Graduado em Gestão em tecnologia da Informação) – Faculdade de Administração e Negócios de Sergipe – FANESE, Aracaju, 2014.
2. BLACK, Tomas Lovis. **Comparação de Ferramentas de Gerenciamento de Redes**. 2008. Trabalho de Conclusão de Curso (Especialista) – Universidade Federal do Rio Grande do Sul, Porto Alegre, 2008.
3. BONOMO, Esley. **Gerenciamento e Monitoração de Redes de Computadores Utilizando-se Zabbix**. 2006. Monografia (Pós Graduação *Lato Sensu* em Administração de Redes Linux) – Universidade Federal de Lavras, Lavras, 2006.
4. CANDIDO, Wamilson Luiz. **Gerenciamento de Redes**. 2012. Trabalho de Conclusão de Curso (Técnico em Manutenção e Suporte em Informática) – Instituto Federal do Paraná, Paranaguá, 2011.
5. COMER, D. “Redes de Computadores e Internet”. Editora Person, Porto Alegre, 2ªed.,2001.
6. FACHINI, Thiago. **Implementação da Ferramenta Zabbix para Monitoramento Reativo**. (Tec. Rede de Computação) Universidade Luterana do Brasil – ULBRA, Canoas. Disponível na internet: <http://www.ulbra.inf.br/joomla/images/documentos/TCCs/2010_2/redes-thiago%20fachini.pdf>. Acesso em 20 de fevereiro de 2015.
7. FAGUNDES, Ana Carolina. **Guia Básico Para Elaboração de Referências Bibliográficas**-Segundo a ABNT. Departamento de Comunicação Institucional – Unifest. São Paulo, 2014.
8. KUROSE, J. F., ROSS, K. W., **Redes de Computadores e a Internet - Uma nova abordagem**, 5a Edição, São Paulo: Addison Wesley, 2010.
9. RAMOS, Bezaleel. **O que é Zabbix – Como funciona**, 2014. Disponível na Internet: <<http://blog.infolink.com.br/zabbix-como-funciona/>>. Acesso em 21 de fevereiro de 2015.
10. SANTOS, Cinthia Cardoso. **Gerenciamento de Redes com a Utilização de Software Livre**. (Curso de Sistemas de Informação) - Instituto de Estudos Superiores da Amazônia – IESAM, Belém. Disponível na internet:< <http://www3.iesam-pa.edu.br/ojs/index.php/sistemas/article/viewFile/442/374>>. Acesso em 10 de fevereiro de 2015.
11. Silva, Cassiano Ricardo G. Silva, Everton G. CASTRO, Ewerton C. SILVA, Marcus Vinícius G. **Gerenciamento de ativos**. São José dos Campos, 2011. Disponível na internet: < http://zabbixbrasil.org/files/TCC_Zabbix-FINAL.pdf >. Acesso em: 27 de março de 2015.
12. WHITE, Curt M: **Redes de Computadores e Comunicação de Dados**, 1ª edição. 2012.
13. ZABBIX SIA. **Zabbix Website**. Disponível na internet: <<http://www.zabbix.com>>. Acesso em 15 de março de 2015.

Googlemática: a matemática por trás do Google

| Naiara **Beber**
IMA

| Arthur Roberto **Fronza**
IMA

| Luma Carina **Picoli**
IMA

| Andresa Laurett da **Silva**
IMA

RESUMO

Objetivo: No mundo contemporâneo, as facilidades advindas dos avanços tecnológicos permitem que inúmeras informações estejam acessíveis na internet. Para acessá-las, comumente utilizam-se mecanismos de pesquisa, como o Google. A questão levantada foi: como o Google seleciona as páginas a serem exibidas para o usuário? Assim, o presente trabalho objetivou entender sua forma de execução e seus algoritmos de busca.

Métodos: Partindo desse pressuposto, foram realizadas pesquisas em sites, artigos e dissertações, além de duas pesquisas online sobre os métodos de compra empregados na região.

Resultados: Assim, descobriu-se que os métodos utilizados pelo PageRank do Google são baseados em milhões de dados, entre eles a quantidade de acessos e a relevância da página.

Conclusão: Com isso, percebeu-se que a tecnologia é muito utilizada, possui um funcionamento complexo, o que garante que os resultados sejam relevantes e próximos das necessidades de cada usuário.

Palavras-chave: PageRank, Mecanismo de Busca, Google.

■ INTRODUÇÃO

Fruto da tentativa de comunicação em meio ao caos da Guerra Fria, a Internet foi uma descoberta, inicialmente com fins militares, que mais tarde se popularizou e mostrou que veio para ficar. Após os anos de 1990, encaminhava-se para um dos meios de comunicação em massa, o que incitava a criação de sites de busca que facilitassem a sua utilização, bem como seus mecanismos.

Nesse contexto evolutivo da Internet, é inegável sua contribuição no dia a dia de uma sociedade modernizada e tecnológica. Os sistemas de busca, por sua vez, são complexos sistemas capazes de desempenhar funções comunicativas, de entretenimento, informativas e muito mais. Sua contribuição mais notável é a forma como atendem às necessidades de uma sociedade afoita por praticidade e rapidez. Com a multiplicidade de tarefas por eles realizadas, torna-se relevante compreender seu funcionamento, a título de curiosidade sobre a ferramenta diretamente influenciadora das relações comerciais e pessoais.

Visando tal reconhecimento dos mecanismos, analisou-se as preferências dos moradores na região em dois momentos: (a) na execução principal do presente projeto, no ano de 2017; e (b) no decorrer da pandemia de COVID-19, no ano de 2020, relacionadas ao comércio online e à influência dos anúncios em suas compras. Acredita-se que o comparativo de dois momentos divergentes social e economicamente elucidará as mudanças comportamentais e a relevância da proposta aqui apresentada.

Um dos buscadores de referência utilizados na atualidade é o Google, criado na década de 1990, por Larry Page e Sergey Brin. O nome “Google” vem de um trocadilho feito com o termo matemático “Googol”: número representado pelo dígito um seguido de cem dígitos zeros e representa o propósito do navegador de organizar um volume muito grande de dados na internet.

Além disso, buscou-se compreender o funcionamento das disposições dos resultados de uma pesquisa dentro do Google através do sistema PageRank e refletir sobre a relevância do algoritmo para sociedade atual.

■ OBJETIVOS

Compreender os mecanismos empregados pelo Google, o principal mecanismo de busca da atualidade, e como ele seleciona a ordem das páginas que irão aparecer quando realizada uma pesquisa. Investigar a importância deste mecanismo para a sociedade, em especial para os e-commerces, e como ele afeta os hábitos de consumo da população.

■ MÉTODOS

O trabalho foi desenvolvido por um grupo de alunos a partir da discussão sobre tecnologia, onde pensou-se em analisar como o Google classifica as páginas de uma busca.

Através de pesquisas em artigos e dissertações sobre a história da internet e os sistemas de buscas, analisou-se o algoritmo por trás do Search Engine Optimization Off Page (SEO Off Page), sistema utilizado pelo Google para ordenar as páginas em uma pesquisa levando em consideração a relevância que elas possuem, focando na ferramenta do PageRank. Com algumas simulações de páginas indicando para outras páginas, utilizou-se dos conceitos de probabilidade e álgebra. Conseqüentemente, conceitos de matrizes, álgebra linear, determinante, cadeia de Markov (OLIVEIRA, et. al, 2016) e regra do escalonamento para definir o PageRank de uma determinada página, foram necessárias.

Conforme a discussão evoluiu, o potencial que o Google pode exercer sobre as pessoas com seus anúncios ficou mais evidente. Realizaram-se, então, duas pesquisas online, a primeira, em 2017, no ano de produção inicial do trabalho, e a segunda, em 2020, para atualizar sua relevância. O objetivo das pesquisas foi verificar as influências desses anúncios em compras online e conhecer o comportamento das pessoas da região em relação às compras em lojas virtuais e físicas. Para tanto utilizou-se das medidas de tendência central para melhor representar a problemática.

■ RESULTADOS

Nos dias modernos, pesquisar é sinônimo de buscar na internet. Mas, como vimos, este é um conceito recente. No tempo que antecedeu a invenção desse imenso banco de dados, o acesso ao conhecimento era mais difícil e menos efetivo: as enciclopédias eram grandes compilados de informações subdivididas em tópicos por, geralmente, ordem alfabética. Isso resultava em pesquisas demoradas e resultados diversos, que não obedeciam a uma ordem de importância. O surgimento da internet, mesmo representando um excelente avanço, todavia, exigia uma forma eficaz de arranjar os produtos de uma busca.

Para isso, desenvolveu-se milhares de buscadores que objetivam organizar tudo aquilo que está disponível on-line. Entretanto, ao se pensar em busca, logo pensa-se em Google. Através de pesquisas, descobriu-se que o diferencial do Google é o algoritmo por trás dele, chamado Search Engine Optimization, uma ferramenta de pesquisa que facilita a busca do internauta ao obter mecanismos que filtrem as informações de maneira rápida e eficaz. Esse mecanismo utiliza-se de vários algoritmos, sendo um deles o PageRank, uma fórmula matemática desenvolvida pelos fundadores do Google, Larry Page e Sergey Brin, que classifica as páginas atribuindo um valor a cada uma com base nos links que esta mesma página

recebe. Ou seja, é a probabilidade de uma pessoa chegar a essa página, clicando em hiperligações, de modo que a soma dos PageRanks de todas as páginas da Web será sempre 1.

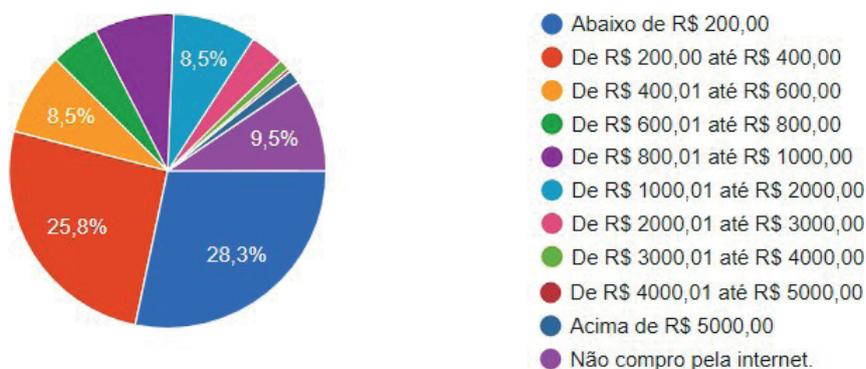
Entendendo o funcionamento do mecanismo mais comum e de seu algoritmo, mostrou-se necessário compreender hábitos dos usuários desses mecanismos, bem como a influência destes nas práticas de e-commerce. Segundo Alfredo Soares (2016)

É importante que o lojista utilize das técnicas de SEO para filtrar o termo que é procurado pelo cliente no Google. Exemplo: empresas do segmento esportivo acabam utilizando a palavra “tênis” em seus anúncios, mas isso dificulta ao cliente achar resultado menos preciso e tornando uma busca irrelevante para a loja virtual.

Para tanto, organizaram-se duas pesquisas para elencar dados essenciais no desenvolvimento desta problemática. A primeira pesquisa foi produzida e divulgada no ano de 2017, quando se deu início a esse estudo. A segunda, no ano de 2020, período de atualização do presente artigo. A conduta de realizar essa coleta em dois momentos visa corroborar a pertinência da proposta.

No ano de 2017, participaram da pesquisa 283 pessoas, sendo 70% do sexo feminino. Dentre o total de pesquisados, 57,2% dizem já ter sofrido influências de anúncios de um site de busca. Foi possível ainda, determinar que a idade mediana foi de 24,6 anos aproximadamente. A maior parte do grupo pesquisado costuma comprar mensalmente, sendo que 34,3% geralmente conhecem e compram o produto em loja virtual. Ao analisar o preço dos itens comprados, chegou-se ao valor médio de R\$588,89, porém, quanto ao Gráfico 2, foi constatado que o valor mais frequente é abaixo de R\$200,00, o que faz levar em consideração a mediana como mais adequada para representar a problemática, concluindo que o valor da compra ficou em torno de R\$526,03, aproximadamente.

Gráfico 2. Valor do produto que costuma comprar em lojas virtuais – Rio do Sul – 14/06/2017 a 25/06/2017.



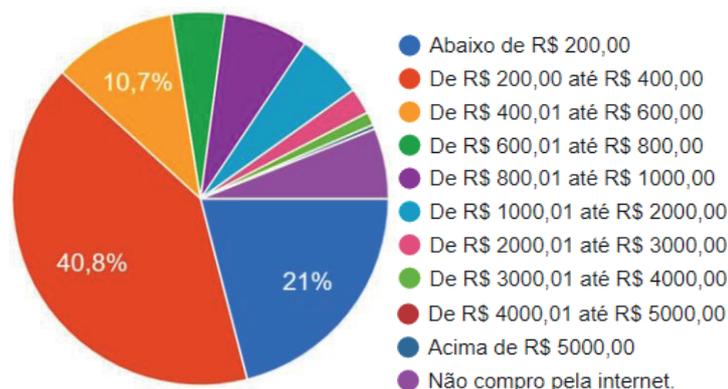
Os produtos mais comprados são da categoria de vestuário e acessórios, seguidos por livros, telefonia e informática.

No ano de 2020, foram consultados 262 consumidores, sendo 29,4% homens. As faixas etárias de maior destaque foram de 18 a 23 anos, com 29,8%; e acima de 38 anos, com 30,2%. As três categorias de maior procura foram livros (50,8%), seguido de eletrodomésticos (42,7%) e vestuário e acessórios (40,5%).

Na análise dos resultados, percebeu-se que 64,5% dos votantes já foram influenciados por anúncios na hora de suas compras e 51,9% já compraram algo não planejado por conta da publicidade. A frequência de consumo mais evidente foi mensal de produtos conhecidos e comprados na loja virtual, 44,3%.

Com relação aos valores, a mediana identifica valores de compra em torno de R\$342,06, aproximadamente, sendo o valor mais frequente de R\$200,00 a R\$400,00. Segundo os valores encontrados na segunda etapa da pesquisa, a moda e a mediana se encontram na mesma faixa de valor, identificando uma melhor distribuição de valores. Com isso, reforça-se a importância do conhecimento de diversos parâmetros estatísticos para a compreensão mais coesa dos dados analisados.

Gráfico 3. Valor do produto que costuma comprar em lojas virtuais – Rio do Sul – 01/12/2020 a 09/12/2020.



■ DISCUSSÃO

Tendo em vista os resultados obtidos nas pesquisas realizadas anteriormente, percebe-se a importância dos anúncios nos hábitos de compras da população. Posto isso, torna-se relevante compreender, na prática, exemplificando matematicamente como o Google seleciona os resultados mais relevantes utilizando-se dos recursos disponibilizados, como o PageRank.

A internet, atualmente, é composta por trilhões de páginas, na qual, ao efetuarmos uma pesquisa procuramos por informações de forma rápida e precisa. As ferramentas que facilitam tais pesquisas na internet foram os motores de busca, dos quais um deles se destaca: o Google. (BATTI, 2015).

O PageRank utiliza um sistema de processo estocástico, ou seja, um processo que descreve procedimentos de um sistema em um determinado período de tempo (Oliveira,

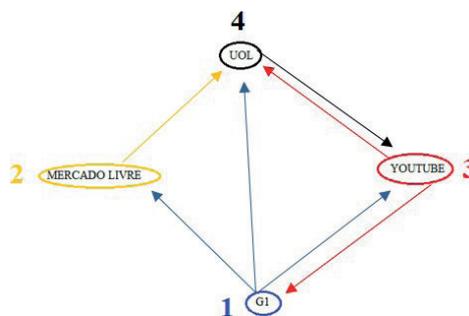
2016). No presente caso, o processo estocástico é markoviano, pois é estacionário e o seu comportamento futuro depende do estado presente, independentemente dos estados visitados no passado (ALVES, 1997, apud OLIVEIRA, 2016).

Uma cadeia de Markov em tempo é um processo estocástico em que a variável t representa intervalo de tempo contável, onde que $\{X(t) | t = 0, 1, 2, 3, \dots\}$, obedecendo a propriedade markoviana (OLIVEIRA, 2016). As cadeias podem ser representadas por meio de uma matriz de transição P que contém as probabilidades de transição entre estados em determinado momento (OLIVEIRA, 2016).

A cadeia markoviana utilizada por esse mecanismo de busca é denominada não ergódica ou redutível. Isso quer dizer que existem elementos da cadeia onde, em número determinado de passos, não é possível retornar a qualquer outro elemento. Essa característica denota a existência de sumidouros, como podemos ver no diagrama a seguir.

Considere um conjunto com quatro páginas, G1, Mercado Livre, Youtube e UOL, com seus links conforme o esquema da Figura 1.

Figura 1. Simulação de links entre páginas da internet



Para identificar a classificação – ordem de importância – dessas páginas, é necessário verificar os links que cada uma recebe e a importância que cada link fornece. Em outras palavras, esses links podem ser interpretados como voto de confiança que outros sites dão para o seu conteúdo. Segundo Batti (2015):

O Google pontua a importância de uma página em função dos links que se conectam a ela. Um link da página A para a página B indica que a página A “votou” pela importância da página B, isto é, se um site X cita (possui um link para) outro site Y, esse site Y é, mesmo que minimamente, importante para o site X.

Veja, que a página G1 (Figura 1) recebe link da página Youtube, e esta, fornece link a duas páginas, a G1 e UOL, com isso G1 recebe metade da importância da página Youtube. Chamando G1 de X_1 e Youtube de X_3 , pode-se descrever a situação matematicamente:

$$X_1 = \frac{X_3}{2} \quad (1)$$

De forma análoga, pode-se representar a importância das outras páginas, Mercado Livre

$$X_2 = \frac{X_1}{3} \quad (2)$$

$$X_3 = \frac{X_1}{3} + X_4 \quad (3)$$

$$X_4 = \frac{X_1}{3} + X_2 + \frac{X_3}{2} \quad (4)$$

As quatro equações descritas acima, formam um sistema de equações, que podem ser representados matricialmente por:

$$\left[\begin{array}{cccc|c} 1 & 0 & -1/2 & 0 & 0 \\ -1/3 & 1 & 0 & 0 & 0 \\ -1/3 & 0 & 1 & -1 & 0 \\ -1/3 & -1 & -1/2 & 1 & 0 \end{array} \right]$$

A matriz representa um sistema linear homogêneo, visto que a soma dos elementos de cada coluna é zero. Com isso, o sistema possui infinitas soluções, sendo a trivial (0,0,0,0) uma delas. Por sua vez, essa solução não é interessante, já que a solução diferente de zero permite ordenar as páginas de acordo com sua pontuação, da maior para a menor (Batti, 2016).

Então, utilizando-se da regra de escalonamento, é possível chegar a solução $S = \left(\frac{3}{5}\alpha, \frac{1}{5}\alpha, \frac{6}{5}\alpha, \alpha\right)$, onde α é o valor da variável X_4 . Ao atribuir $\alpha = 5$, obtém-se como solução $S = (3,1,6,5)$ que define a página 3 como a de maior importância, seguida das páginas 4, 1 e 2. Isto quer dizer que a sequência de páginas do Google seriam Youtube, UOL, G1 e Mercado Livre, nessa ordem.

É possível resolver a mesma situação com o critério de que a soma dos PageRanks deve ser sempre 1, matematicamente poderia ser descrito por:

$$X_1 + X_2 + X_3 + X_4 = 1$$

Através de substituições, partindo das quatro equações já citadas, chega-se às importâncias das páginas, onde $X_1 = 0,20$, $X_2 = 0,06$, $X_3 = 0,40$ e $X_4 = 0,33$. Independentemente

do método escolhido para a resolução, as páginas continuarão com o mesmo grau de importância definido por um dos métodos apresentados.

Do esquema apresentado na Figura 1, pode-se compreender, de forma prática, como a probabilidade está presente no conceito do PageRank. Quando o internauta está na página do G1, ele pode através dos links, acessar Mercado Livre ou Youtube ou UOL, ou seja, a probabilidade do leitor acessar o Youtube através do G1 é de 1/3, enquanto a probabilidade de acessar o G1 através do Youtube é de 50%. A genialidade do algoritmo está no fato de não considerar apenas a quantidade de links que a página recebe, mas sim, aliado à qualidade desses links quando considerada a proporção entre eles.

O ranqueamento de páginas se mostra eficiente quando, ao pesquisar na internet, nota-se que logo nas primeiras indicações é possível encontrar as respostas pelas quais procurava-se.

Nesse sentido, o algoritmo não só beneficia a população de forma geral agregando mais conhecimento, como também contribui positivamente para empresas que possuem lojas virtuais. Ou seja, o e-commerce apresenta grandes vantagens ao utilizar-se desse conhecimento, pois, ao planejar e executar ações adequadas ao sistema, permite que ao fazer uma “[...] busca no Google seu anúncio seja direcionado ao público-alvo, otimizando o processo de busca dos seus clientes, facilitando dessa forma a decisão de compra, direcionando-as para a sua loja e tornando sua empresa visível na internet” (SOARES, 2016).

Quanto à execução de duas pesquisas em momentos econômicos e sociais tão diferentes, segundo o indicador macroeconômico SpendingPulse, da empresa norte-americana Mastercard identificou que nos meses de agosto, setembro e outubro, o e-commerce brasileiro teve uma alta de 84,5% se comparado ao mesmo período de 2019. Aliando esse dado ao resultado da execução da segunda pesquisa deste projeto, afere-se que, com relação ao ano de 2017, o valor mediano de compra diminuiu, resultado de diversos fatores, entre eles, o enfrentamento da pandemia de COVID-19 que alterou o cenário nacional para um momento de crise econômica. Em contrapartida, considerando as medidas de prevenção da doença, houve queda na porcentagem de pessoas que não utilizam a internet para compras, uma vez que o e-commerce cumpre com as determinações de isolamento e distanciamento social, conferindo maior proteção ao usuário.

■ CONCLUSÕES

Com os estudos e discussões levantadas acerca do assunto, pode-se concluir, primeiramente, que a Internet está presente constantemente no contexto social, e muito há de se entender na sua complexidade e seu funcionamento. O Google, como buscador mais utilizado, recebe destaque nesse ínterim. E, portanto, justifica a relevância desse estudo.

Segundo Oliveira, “o que o algoritmo do PageRank faz é, em sua essência, replicar esses cálculos [da cadeia de Markov] milhares de vezes em um universo muito maior de sites para trazer ao usuário uma maior comodidade e melhor qualidade na busca pela informação.”

Para que as páginas tenham maior visibilidade, é utilizado o complexo sistema PageRank, que analisa e compara milhões de dados, onde é levado em conta a frequência que determinada página é citada nas demais, e a relevância das mesmas. Como o Google trabalha com milhares delas, seus cálculos exigem rapidez e precisão. Mais uma vez, a matemática age diretamente no âmbito social.

Suas contribuições vão desde informações corriqueiras à demanda das grandes empresas, que utilizam desse recurso para obter mais visibilidade e, conseqüentemente, maiores lucros. Esta situação ficou evidente com as pesquisas realizadas, que indicaram que as pessoas sofrem influências dos anúncios nos sites de busca e que uma pequena parte do grupo pesquisado não costuma comprar pela internet. Sendo assim, não é condição suficiente que os envolvidos na empresa conheçam o algoritmo, mas que os mesmos utilizem de outros recursos que auxiliem na elevação do PageRank de sua página, além, é claro, de conhecer quem é seu público-alvo.

Por fim, nota-se a contribuição para uma pesquisa prática e que atenda às necessidades de quem a fez. O PageRank é apenas uma das diversas modalidades que o mundo atual, moderno e tecnológico tem desenvolvido para trazer benefícios à vida dos usuários.

Assim funciona a matemática por trás do Google.

■ AGRADECIMENTOS

Agradecemos a Deus, pelos dons a nós concedidos que nos proporcionaram o necessário no desenvolvimento deste projeto visando expandir o conhecimento que tivemos a oportunidade de explorar.

Agradecemos à nossa professora orientadora, pela dedicação, afincamento e zelo em empregar seu tempo e atenção a nós e ao projeto, pela solicitude, pelo amor e maestria com que lecionou esses e tantos outros temas.

Agradecemos ao Instituto Maria Auxiliadora pelo incentivo, por proporcionar sempre ótimas oportunidades aos discentes, e, essencialmente, pelo comprometimento com o ensino de qualidade.

Agradecemos aos colegas e amigos pelo tempo investido em colaborar com a pesquisa, logo, com o artigo.

Um agradecimento especial às nossas famílias, sem elas nada teria sido possível. Obrigado por nos conceder a oportunidade de compor a Instituição, por compreender a distância e o tempo investido nesse estudo, por nunca se queixarem da nossa ausência, uma

vez que sempre compreenderam o percurso como a nossa busca pela excelência do nosso projeto e do nosso conhecimento.

■ REFERÊNCIAS

1. BATTI, João Carlos Bez. **Um pouco de matemática por trás do algoritmo PageRank do Google**. 2015. 63f. Dissertação (Mestrado em Matemática) – Universidade Federal de Santa Catarina, Florianópolis, 2015.
2. OLIVEIRA, M. V. et al. PageRank: O funcionamento da ferramenta de busca do Google. **Caderno de graduação**, v.3, n.3, p.73-84, out. 2016.
3. SOARES, Alfredo. **Qual a importância do Google Adwords para o sucesso do seu e-commerce?**. 2016. Disponível em: <<https://www.ecommercebrasil.com.br/artigos/google-adwords-e-commerce>>. Acesso em: 18 jul. 2017.
4. VENDAS no e-commerce brasileiro disparam 87% em outubro. **Época negócios**, 2020. Disponível em: <<https://epocanegocios.globo.com/Economia/noticia/2020/11/epoca-negocios-vendas-no-e-commerce-brasileiro-disparam-87-em-outubro.html>>. Acesso em: 15 dez. 2020.

Importância da tecnologia da informação: análise do atendimento de pacientes de uma maternidade

| João Cardim Ferreira **Lima**

| Luan Lyon Lopes **Nascimento**

| Adriano Varella de **Morais**

RESUMO

A tecnologia da informação é fundamental para o gerenciamento e melhoria de processos pois possibilita maior facilidade na organização e agrupamento de dados. À medida que a quantidade de informações e sua complexidade aumenta, a necessidade da utilização da TI torna-se mais evidente. A pesquisa trata-se de um Estudo de caso da Maternidade Escola Januário Cicco, localizada no estado do Rio Grande do Norte. O objetivo consistiu-se em analisar como a Tecnologia da Informação promove o suporte ao processo de atendimento às pacientes da maternidade. Inicialmente foram expostos alguns temas encontrados na literatura acerca do tema de tecnologia da informação, gerenciamento de processos e indicadores de desempenho. Para acompanhamento *in loco* das atividades da Maternidade, foram realizadas visitas técnicas. Posteriormente elaborou-se o mapeamento e a análise do processo de atendimento. Por fim foram criadas propostas de soluções viáveis para a melhoria no processo de atendimento às pacientes e aumentar o nível de qualidade do serviço: desenvolvimento de um sistema de informação único para integrar todas as informações geradas e a criação de indicadores de desempenho para orientar o gerenciamento do hospital.

Palavras-chave: Ciência da Informação, Gestão em Saúde, Redes de Comunicação de Computadores.

■ INTRODUÇÃO

O novo contexto empresarial exige que sejam utilizadas técnicas modernas para armazenamento, processamento e gestão da informação, capazes de serem utilizadas como diferencial. A informação tornou-se, cada vez mais, um fator primordial de diferenciação, competitividade e sucesso para as empresas. É imprescindível que sejam desenvolvidos fluxos de informações confiáveis. Um fluxo de informações bem estruturado, coerente e adaptado às necessidades da empresa mostra-se fundamental para que a informação trafegue até o destinatário final, auxiliando as atividades das empresas, com o intuito de contribuir para o apoio às tomadas de decisões por parte dos gestores, possibilitando a disseminação do conhecimento e a aprendizagem organizacional.

A pesquisa foi realizada na Maternidade Escola Januário Cicco (MEJC), idealizada pelo médico norte-rio-grandense nascido em São José do Mipibu que deu nome à maternidade. Hoje, a Maternidade Escola é hospital de referência terciária do SUS e funciona não só como hospital, mas como um campo de ensino e aplicação prática para as profissões não apenas da área da saúde, como também de outras áreas, cumprindo um meritório trabalho de ensino, pesquisa e atenção à população da cidade de Natal e de seus municípios vizinhos. Suas bases de pesquisa principais são: Gestação de alto risco, Saúde da mulher e Cirurgia ginecológica.

O objetivo da pesquisa consistiu em analisar o processo de atendimento às pacientes da maternidade e compreender como a tecnologia da informação dá suporte a esse processo atendimento. Deve-se evidenciar que o trabalho foi demandado pela diretoria da instituição por acreditar que o processo de atendimento possui ineficiências, além de impactar na redução da disponibilidade de espaço físico para os leitos e dificuldades na disponibilização de documentos, sendo a TI – Tecnologia da Informação uma possível solução para otimização do processo de atendimento da Maternidade. Os *softwares* utilizados pela maternidade para o processamento das informações durante o atendimento serão citados e comentados no trabalho.

■ REFERENCIAL TEÓRICO

Tecnologia da Informação

A TI é um conjunto dos componentes tecnológicos individuais, normalmente organizados em sistemas de informação baseados em computador que potencializa a otimização da sistematização da informação e seus processos produtivos em um mercado competitivo,

representando um sistema estratégico com foco no cliente e no serviço, motivando esforços de melhoria e reengenharia de negócio (TURBAN *et al.* , 2005).

Alguns autores defendem que a tecnologia da informação se restringe apenas à aspectos técnicos, no entanto, outros autores, tais como Porter (1986) e Druker (2000) adotam um conceito mais abrangente de Tecnologia da Informação, incluindo nesse conceito os sistemas de informações e a utilização de *softwares e hardwares*, banco de dados, redes de telecomunicações e de recursos visuais para o fornecimento de informações confiáveis.

Importância da Tecnologia da Informação

A TI está redefinindo os fundamentos dos negócios. Atendimento ao cliente, operações, estratégias de produto, de *marketing*, de distribuição e até mesmo a gestão do conhecimento dependem demasiadamente, ou às vezes até totalmente, dos SI- Sistemas de Informação. A TI e seus custos passaram a fazer parte integrante do dia-a-dia das empresas. Para atender a essa complexidade das necessidades empresariais, não se pode desconsiderar a TI e seus recursos disponíveis, sendo muito difícil elaborar SI essenciais da empresa sem envolver esta moderna tecnologia. No entanto, muitas empresas ainda acreditam que o simples ato de informatizá-las, espalhando computadores e impressoras pelas unidades departamentais, ligando-os em rede e instalando sistemas aplicativos, possam organizar as mesmas (BALLONI, 2006).

A informação e seus respectivos sistemas desempenham funções fundamentais e estratégicas nas organizações. Em sua totalidade, a informação apresenta-se, sob a ótica da vantagem competitiva, como um recurso estratégico chave. Em resumo, o desafio gerencial central é como utilizar a TI para projetar e realizar a gestão de empresas de forma ética, efetiva e competitiva. Os SI são vitais para a gestão, organização e operação das empresas, exercendo impactos na estrutura organizacional, influenciando a cultura, as filosofias, as políticas, os processos e os seus modelos de gestão. Portanto a gestão da informação e o conhecimento dos SI são essenciais para se criar empresas competitivas, gerenciar globalmente as corporações e prover os clientes com produtos e serviços de valor (BALLONI, 2006).

Os conceitos de Planejamento Estratégico, Gestão em Sistemas e Tecnologias de Informação e os respectivos recursos precisam estar disseminados dentro da empresa e multiplicados entre os recursos humanos componentes da mesma. A informação e seus respectivos sistemas desempenham funções fundamentais e estratégicas nas organizações, e apresenta-se como recurso estratégico chave sob a ótica da vantagem competitiva. Como a Globalização dos mercados está, cada vez mais, forçando as empresas a buscarem por vantagens competitivas. Nesta era das “empresas digitais” torna-se estratégico que essas

empresas considerem seus clientes, fornecedores e mesmo competidores atuando como parceiros de negócios (BALLONI,2006).

Gestão de Processos

DAVENPORT (2004) define processo como uma ordenação específica das atividades de trabalho no tempo e no espaço, com um começo, um fim, *inputs e outputs* claramente identificados. Todo produto ou serviço oferecido por determinada organização é, necessariamente, resultado de um conjunto de atividades ou processos. Esses processos são oriundos dos diferentes níveis organizacionais, desde o operacional ao estratégico. Logo, não há um produto ou serviço ofertado sem a existência de um processo (WESKE, 2007 *apud* SOUSA; MEDEIROS, 2008).

SOUSA e MEDEIROS (2008) defendem a existência de alguns fatores que influenciam a adoção da gestão de processos por parte das empresas, esses fatores são: O rápido crescimento, aquisições e fusões e maior agilidade para aproveitar oportunidades de mercado, além disso, fatores como maior controle dos processos por parte da administração, redução de custos e necessidade de uma cultura de melhoria contínua também mostram-se importantes fatores para a adoção de uma administração focada no gerenciamento de processos.

Para gerenciar determinado processo é imprescindível que se tenha um conhecimento prático acerca do funcionamento desse processo. Esse conhecimento prático possibilita a identificação de possíveis oportunidades de melhoria. Algumas ferramentas são utilizadas para o gerenciamento de processos, destacando-se o uso de Fluxogramas. A ferramenta ajuda a quantificar alguns parâmetros, conhecer a fundo os processos e embasar possíveis decisões a partir de dados levantados durante as observações. (Fonte = nós mesmos)

Fluxogramas

Fluxogramas são formas de representar, por símbolos gráficos, a sequência dos passos de um trabalho para facilitar sua análise. Um fluxograma é um recurso visual utilizado pelos gerentes de produção para analisar sistemas produtivos, buscando identificar oportunidades de melhorar a eficiência dos processos (PEINADO; GRAEML, 2007).

O fluxograma destina-se à descrição de processos. Um processo é uma determinada combinação de equipamentos, pessoas, métodos, ferramentas e matéria-prima, que geram um produto ou serviço com determinadas características. (LINS,1993)

A técnica de representação gráfica fluxograma se utiliza de símbolos previamente convencionados, permitindo a descrição clara e precisa do fluxo ou sequência de um processo, bem como sua análise e redesenho (OLIVEIRA, 2009).

Os aspectos principais de um fluxograma, são os seguintes: (i) Padronizar a representação dos métodos e os procedimentos administrativos; (ii) Maior rapidez na descrição dos métodos administrativos; (iii) Facilitar a leitura e o entendimento; (iv) Facilitar a localização e a identificação dos aspectos mais importantes; (v) Maior flexibilidade; (vi) Melhor grau de análise (OLIVEIRA,2002).

Indicadores de Desempenho

Os indicadores são essenciais ao planejamento e ao controle dos processos das organizações. São essenciais ao planejamento pois possibilitam o estabelecimento de metas quantificadas e o seu desdobramento na organização. E são essenciais ao controle porque os resultados apresentados através dos indicadores são fundamentais para a análise crítica do desempenho da organização, para a tomada de decisões e para o replanejamento (TAKASHINA; FLORES, 1996).

Takashina e Flores (1996) consideram que os indicadores contribuem os seguintes aspectos:

- a. os indicadores estão intimamente ligados ao conceito da qualidade centrada no cliente. Eles devem ser gerados a partir das necessidades e expectativas dos clientes, traduzidas através das características da qualidade do produto ou serviço, sejam elas tangíveis ou não;
- b. os indicadores possibilitam o desdobramento das metas do negócio, na estrutura organizacional, assegurando que as melhorias obtidas em cada unidade contribuirão para os propósitos globais da organização;
- c. os indicadores devem estar sempre associados às áreas do negócio cujos desempenhos causam maior impacto no sucesso da organização. Assim, eles darão suporte à análise crítica dos resultados do negócio, às tomadas de decisão e ao replanejamento;
- d. os indicadores viabilizam a busca da melhoria contínua da qualidade dos produtos e serviços e da produtividade da organização, aumentando a satisfação dos seus clientes, sua competitividade e, conseqüentemente, sua participação no mercado. Neste sentido, o uso dos indicadores encoraja tanto melhorias incrementais quanto revolucionárias. A comparação com referenciais de excelência, particularmente, pode dar uma contribuição para a identificação de possibilidades mais amplas de melhorias.

■ MÉTODO DE PESQUISA

O estudo com abordagem qualitativa consiste em pesquisa com objetivo exploratório de natureza descritiva e o tema foi selecionado a partir de uma demanda inicial da organização analisada, que apresentou dificuldade no gerenciamento de processos que envolvem entrada e armazenamento de informações no atendimento das pacientes, segundo os próprios colaboradores.

Nas visitas técnicas foram entrevistados 05 servidores, todos com contato diário com processos que envolvem a entrada e armazenamento de dados e informações. Dos servidores entrevistados, 02 foram as Atendentes, que se comunicam diretamente com as pacientes, 02 Coordenadores e o Diretor Geral da MEJC – Maternidade Escola Januário Cicco.

Para instrumento de coleta de dados, utilizou-se entrevista focalizada, semiestruturada ou semi-dirigida, que é livre, no entanto possui foco nos processos da Maternidade que possuem entrada e armazenamento de dados e informações referentes ao atendimento das pacientes. Além das entrevistas foi realizada análise documental através da disponibilização de arquivos e documentos importantes pelo Diretor Geral. A pesquisa bibliográfica teve papel fundamental para o esclarecimento acerca de todos os conceitos envolvidos e compreensão geral do campo de estudo. (não alterei pois não teve fonte, nós que escrevemos)

O processo foi mapeado através da elaboração de um Fluxograma para uma melhor compreensão dos processos. Para isso foram mapeados todos os processos relacionados ao atendimento às pacientes. Para a construção do fluxograma e análise dos resultados utilizou-se a ferramenta Bizagi, que proporcionou uma visão mais detalhada do funcionamento do processo de atendimento dentro da Maternidade Escola Januario Cicco. Após o mapeamento, voltou-se para a literatura para busca de métodos para maior eficiência no gerenciamento destes processos e encontra-se o conceito de Business Process Management (BPM), que integra os conhecimentos de processos e utiliza tecnologia de TI para entrada e armazenamento de dados, otimizando a utilização destas informações, assim como busca e segurança.

Com o mapeamento dos processos e conhecimento acerca dos conceitos encontrados na literatura, foi possível a elaboração de propostas para melhoria dos processos dentro da empresa, que poderão refletir em ganho de eficiência, além de muitos outros benefícios que serão expostos ao decorrer do trabalho.

Figura 1. Esquemática das etapas de metodologia do estudo realizado.



Fonte: Elaboração própria

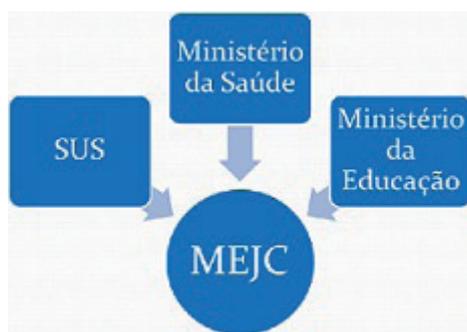
■ RESULTADOS

Idealizada pelo médico norte-rio-grandense nascido em São José do Mipibu que deu nome à maternidade, a Maternidade Escola Januário Cicco (MEJC) foi fundada em 19 de março de 1928, porém apenas cerca de 23 anos após a fundação, no início da década de 40, estava pronta para funcionar. Com o passar do tempo, foram aumentando o número de serviços. Instalou-se a Cátedra de Obstetrícia, integrou-se o Departamento de Toco-Ginecologia, estimulou-se a formação das bases de pesquisa, servindo como pólo de atração de conhecimento na área da medicina e ciência.

Em 2015, a Maternidade Escola é hospital de referência terciária do SUS e funciona como um campo de ensino e aplicação prática, cumprindo um meritório trabalho de ensino, pesquisa e atenção à população da cidade de Natal e de seus municípios vizinhos. Suas bases de pesquisa principais são em Gestaç o de alto risco, Sa de da mulher e Cirurgia ginecol gica.

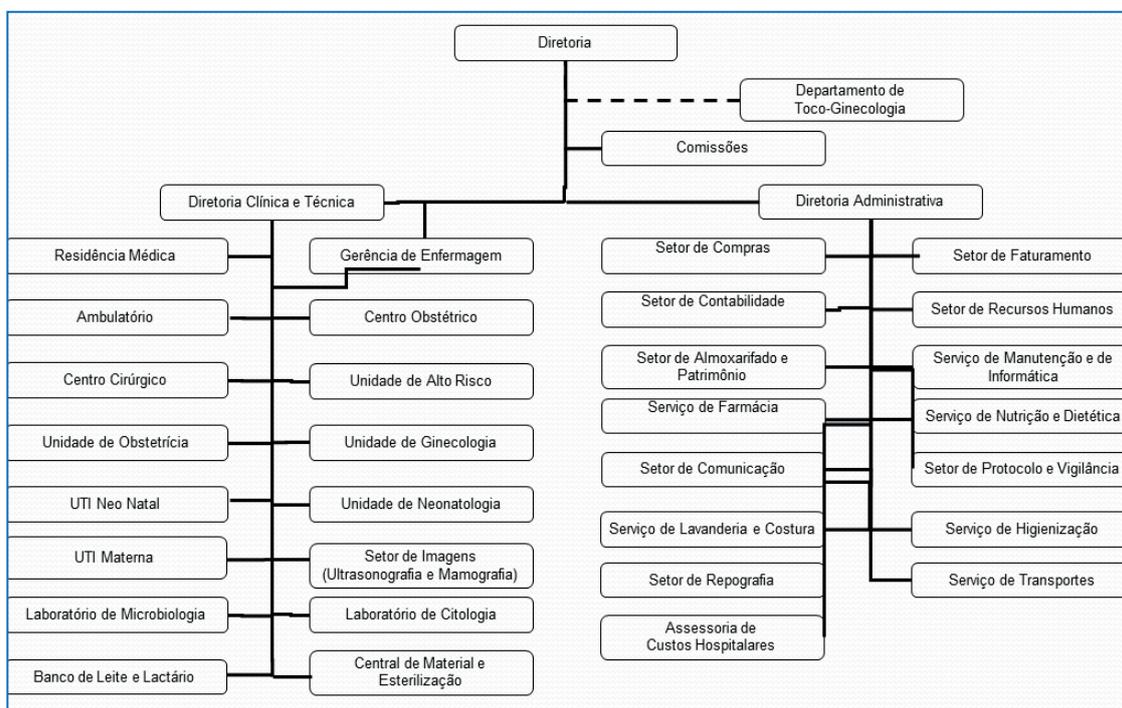
A MEJC   p blica e se submete   outros  rg os federais: Ao SUS, ao minist rio da sa de, e, ao minist rio da educa o, uma vez que se constitui uma maternidade escola, conforme a Figura 2.

Figura 2. Contexto no qual a maternidade encontra-se inserida



Fonte: Elaboração própria

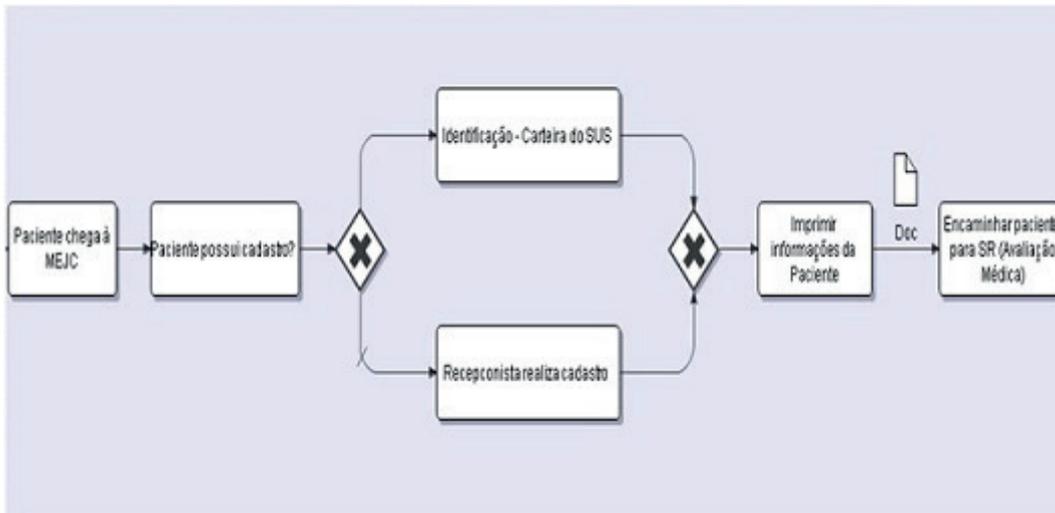
Figura 3. Organograma



Fonte: Elaboração própria

No estudo de caso é feita uma análise dos processos de atendimento do hospital que geram documentos. Também foram analisado os Sistemas de Informação utilizados e o armazenamento de documentos, dados e informações geradas nos processos de atendimento. Todo o processo de Atendimento à paciente segue descrito na Figura 4.

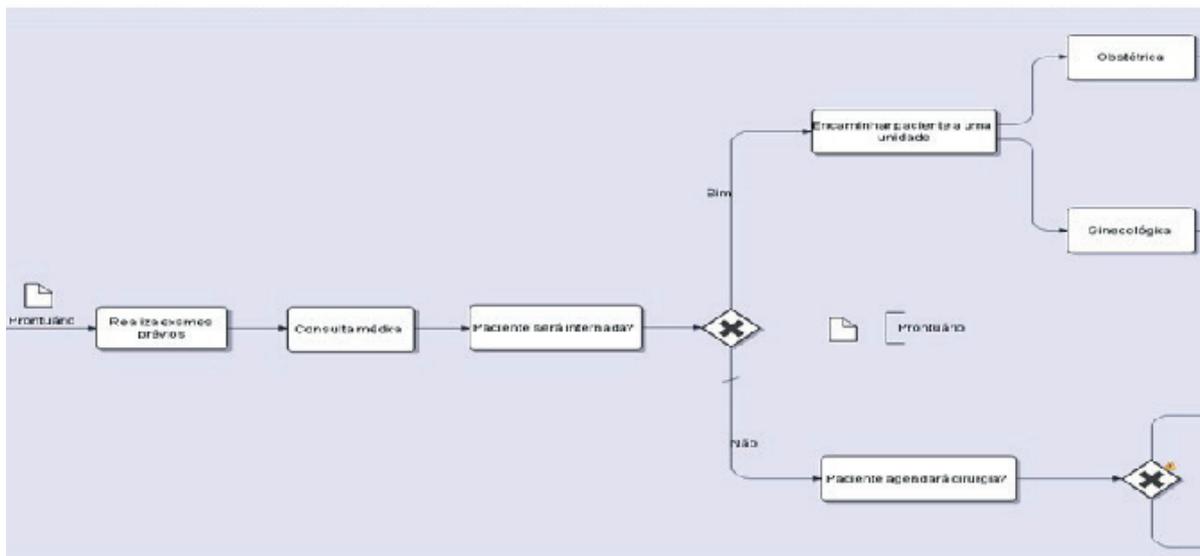
Figura 4. Fluxograma



Fonte: Elaboração própria

A Paciente chega à Maternidade e é verificado se ela possui cadastro, se não possuir a recepcionista realiza o cadastro, as informações do cadastro geram um documento que é encaminhado junto à paciente para a Avaliação Médica, de acordo com a Figura 5.

Figura 5. Fluxograma

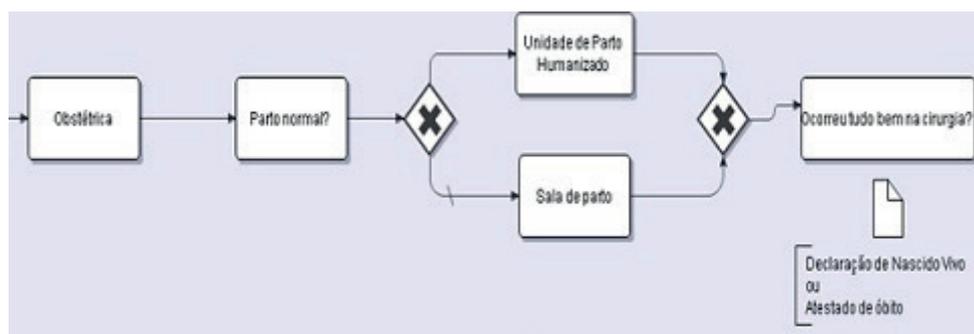


Fonte: Elaboração própria

Na Avaliação Médica é gerado um prontuário, que acompanha a paciente, que realiza exames prévios, e consulta médica, e é decidido se a paciente será internada e encaminhada a uma unidade Obstétrica ou Ginecológica, caso a paciente não seja internada, ela poderá agendar a cirurgia para outra data na secretaria ou ser encaminhada para outro hospital.

Caso a paciente seja encaminhada a uma unidade Obstétrica, o procedimento a ser seguido será segue descrito na Figura 6.

Figura 6. Fluxograma

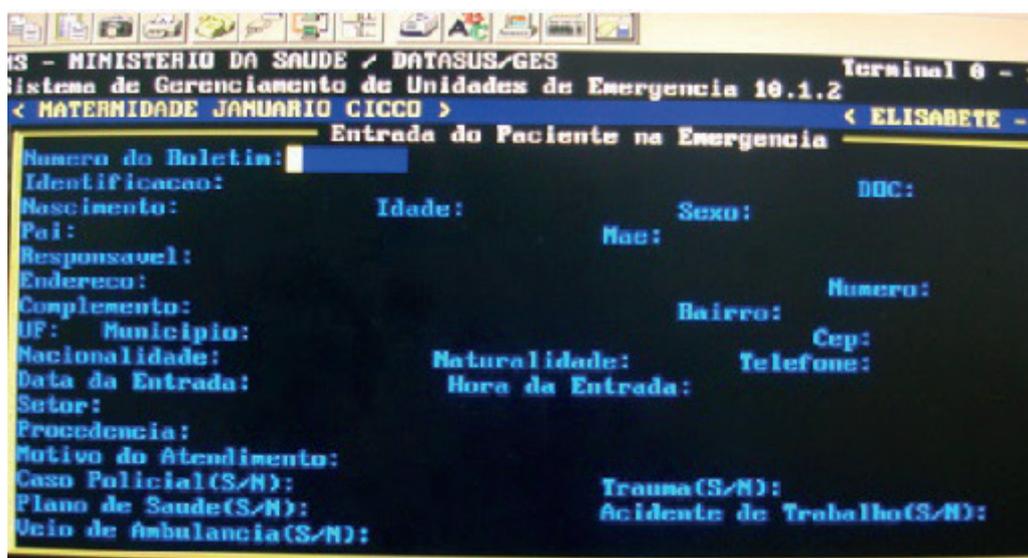


Fonte: Elaboração própria

Se for parto normal, ela vai para unidade de parto humanizado, senão vai para sala de parto. Depois é gerado um documento de declaração de Nascido Vivo ou Atestado de óbito. Caso ocorra tudo bem na cirurgia, a paciente será encaminhada até a unidade Ginecológica, onde serão feitos os procedimentos e ela é encaminhada ao leito, caso ocorra algum problema na cirurgia, a paciente será encaminhada à UTI.

O cadastro das pacientes fica armazenado no HOSPUB, Sistema integrado do SUS, que fornece soluções de gerenciamento e controle social em unidades hospitalares. A Figura 7, representa a imagem da interface do HOSPUB.

Figura 7. Tela inicial do HOSPUB para cadastro de pacientes

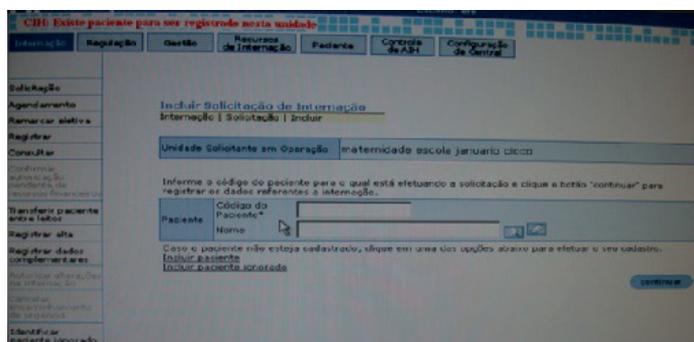


Fonte: Elaboração própria

As principais queixas apresentadas pelos colaboradores que utilizam o HOSPUB é que apesar dele ser integrado com o SUS, ele não é integrado com outros setores da Maternidade, causando dificuldade de encontrar e compartilhar documentos dentro da MEJC.

Já o SISREG, (sistema *on-line*, interligado ao DataSUS) está vinculado ao gerenciamento e operação das centrais de regulação, as capacitações dos RHs das Secretarias de Saúde e estabelecimentos agregados. Representado na Figura 8.

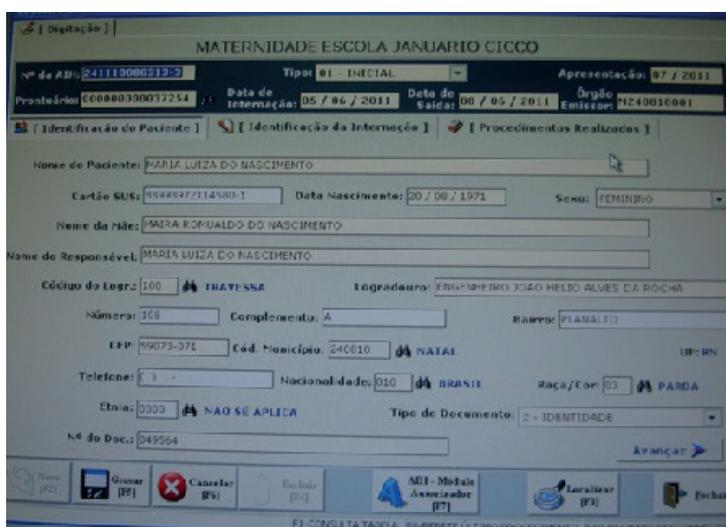
Figura 8. Tela do SISREG



Fonte: Elaboração própria

Na MEJC, também utilizam o SISAIH (Sistema de Informações Ambulatoriais do SUS) que possui abrangência Municipal e Estadual e facilita a digitação e remessa de dados através da emissão de relatórios gerenciais que possibilitam auditoria nas internações indevidas antes do efetivo pagamento. Segue a interface do seu sistema na Figura 9.

Figura 9. Interface do SISAIH



Fonte: Elaboração própria

A informação surge ao início do atendimento à paciente, utilizando-se de sistemas informatizados do SUS para o armazenamento de informações pessoais da paciente e de folhas de prontuário para registrar informações clínicas, que seguem o mesmo fluxo que as pacientes internadas (acompanhando a paciente por qualquer unidade que ela seja encaminhada dentro da maternidade). Ao fim, quando a paciente recebe alta, o prontuário é direcionado para o setor de contas para registro em um sistema *on-line* do SUS, para fins financeiros.

O primeiro momento é a chegada da paciente à Urgência da maternidade, o atendimento é direcionado a pacientes de alto risco (problemas ginecológicos ou obstétricos). O sistema para registro de pacientes é o HOSPUB (sistema do SUS possui a função de gerenciamento e controle social em unidades hospitalares). No sistema ficam armazenadas informações pessoais das pacientes, bem como o momento de entrada e saída da paciente.

O HOSPUB não é *on-line* e também não é integrado a nenhum outro sistema. Em muitos casos, as pacientes que ficam internadas não retornam mais à recepção da Urgência e a informação não chega para registro no sistema. Assim, estas pacientes ficam registradas como se ainda estivessem presentes ocupando os leitos da maternidade (o que mostra a ineficiência do sistema utilizado).

O SISREG auxilia os gestores na operacionalização dos Complexos Regulatórios da Assistência à Saúde, ou seja, é um sistema *on-line* de regulação que possibilita o gerenciamento de informações sobre a rede de atendimento, desde os procedimentos básicos, até as internações hospitalares. Esse sistema fundamenta-se na necessidade do controle das ações executadas em casa Unidade Hospitalar (regulação, avaliação e auditoria). Os recursos físicos e financeiros sendo gerenciados de forma transparente possibilitam a população um melhor acesso aos serviços do Sistema Único de Saúde.

A paciente geralmente passa pela SR, (onde são realizados procedimentos básicos como análise de pressão, temperatura corporal, teste rápido de HIV, dentre outros). Na sequência, a paciente é atendida por um médico em um dos consultórios do ambulatório, nesse momento é gerado o prontuário da paciente (documento de papel que contém todas as informações técnicas necessárias para o atendimento e serão anexados outros documentos relacionados aos procedimentos realizados pela paciente).

Caso a paciente esteja em situação confortável, busca-se a realização do parto normal. Em caso de uma complicação maior, a mulher é encaminhada para UTI Materna. Ao nascer, a criança recebe a declaração de nascido vivo e em caso de necessitar de algum procedimento, é aberto um prontuário do recém-nascido que o acompanhará em sua permanência na maternidade.

O Setor de Contas recolhe os prontuários nas enfermarias e UTI das pacientes que já receberam alta e dos recém nascidos que também já deixaram o hospital. Esse setor é responsável por analisar todos os procedimentos realizados por cada paciente, para lançar em um terceiro sistema do SUS essas informações. A documentação referente à paciente só pode ser integrada ao sistema quando for anexado o documento dos procedimentos realizados em assistência ao recém-nascido, que algumas vezes permanece na maternidade mesmo após a saída da mãe.

As atividades realizadas pelo setor de contas basicamente são recolher os prontuários, organizá-los, contabilizar o número de procedimentos e quais foram, codificar essa informação, pois cada procedimento possui um código a ser referido no sistema. Para assim, passar essas informações ao SISAIH, sistema *on-line* do SUS.

O SISAIH realiza a entrada de dados da Autorização de Internação Hospitalar (AIH). Este sistema foi desenvolvido para contribuir com o aprimoramento das áreas de Controle, Avaliação, Regulação e Auditoria das Secretarias Municipais e Estaduais de Saúde, auxiliando no trabalho dos colaboradores responsáveis pela alimentação dos sistemas de informação que, estão lotados nos setores de contas ou faturamento dos estabelecimentos de saúde que atendem pelo Sistema Único de Saúde em todo o país.

■ DISCUSSÃO

Os documentos que são gerados durante todo o processo são registrados em um livro com um número de protocolo, nome da paciente e a data de saída da maternidade. Toda a documentação da paciente é armazenada no prédio do Arquivo (setor responsável por estocar e controlar toda a documentação de todas as pacientes).

Uma ficha é utilizada de base para referência à documentação da paciente. Caso uma paciente deseja saber quando foi a última vez que esteve internada, deverá ser consultado o Arquivo.

Os colaboradores desse setor sabem onde está cada documento e isso é comprovado pela auditoria periódica. Entretanto, a falta de informatização gera uma situação mais complexa para realizar um trabalho da melhor forma.

O fator mais importante para suprir esse sistema rústico é o espaço físico requerido. O prédio do Arquivo armazena documentos desde o início da maternidade e são diversas prateleiras para que isso possa ocorrer. Um móvel próprio para arquivamento de pastas foi adquirido para melhor organizar essas informações, porém, sua capacidade de armazenamento é limitada.

■ CONCLUSÕES

A análise da importância da tecnologia da informação no atendimento às pacientes da MEJC foi realizada segundo as características relacionadas ao fluxo de informação e aos sistemas de informação utilizados pela maternidade durante o processo de atendimento às pacientes. Essa análise possibilitou a identificação de oportunidades de melhorias. Estas proposições, resultantes do estudo aqui apresentado, são baseadas nos fundamentos teóricos e nas perspectivas do hospital.

Pode-se perceber que, apesar do investimento feito em TI por parte da empresa, a tecnologia da informação na maternidade ainda não possui maturidade suficiente e adequada ao tamanho e à importância da MEJC para o estado. Não há um sistema que englobe todas as necessidades de informação por parte da empresa, ou seja, os sistemas de informação utilizados atualmente apresentam algumas restrições, além de não serem integrados, o que dificulta o acesso às informações por parte de alguns setores da maternidade.

Para sanar esse problema propõe-se a criação de um sistema único capaz de abranger todos os setores da empresa, o que facilitaria a comunicação e o acesso à informação. A integração desses sistemas traria benefícios tangíveis e intangíveis para a organização. Tangíveis: Diversos benefícios referentes à redução de desperdício de tempo (que é muito importante no ramo de serviços), em relação ao espaço físico (recurso mais caro da instituição, podendo ser transformado em novos leitos, aumentando a capacidade de atendimento da unidade), além da redução do quadro de funcionários do hospital. Intangíveis: Pode-se citar a melhora ao acesso à informação, o aperfeiçoamento dos processos e a padronização das atividades.

Propõe-se também a criação de indicadores de desempenho para avaliação das melhorias supracitadas, medindo a eficiência do recurso tempo, da capacidade de atendimento, do monitoramento dos operadores do sistema, e por fim da qualidade do serviço ofertado e os custos envolvidos. Estes indicadores deverão ser gerenciados para que os níveis se mantenham dentro de uma faixa desejável e caso algum saia da faixa deve-se buscar o problema solucionando-o o mais rápido possível.

Todos estes benefícios irão refletir potencialmente em todos os setores da organização, principalmente nos resultados financeiros do hospital, onde diversos custos serão reduzidos, proporcionando uma elevação significativa na qualidade de serviço prestado à população do estado do Rio Grande do Norte, podendo inclusive tornar-se modelo que poderá ser utilizado para outras Maternidades Públicas do País.

Além de investir na rede de telecomunicação, mostra-se necessário também investimento nos outros componentes da TI. A aquisição e renovação dos recursos tecnológicos, como impressoras e computadores, serão necessárias para dar apoio aos processos e manuseio dos dados. Desta forma é proposto adquirir computadores e instalar os softwares necessários. A capacitação dos colaboradores e treinamentos operacionais também é fundamental para o sucesso da T.I. na empresa.

As limitações da pesquisa estão relacionadas a análise apenas qualitativa dos desperdícios e ineficiências geradas pela ausência de um SI integrado em todo o hospital. Uma sugestão de trabalho futuro é o desenvolvimento de um *software* que integre todo hospital

e facilite o gerenciamento da informação e que inclusive possa ser replicado para outros hospitais e órgãos públicos que apresentem essa mesma necessidade.

■ REFERÊNCIAS

1. Turban E, Rainer RK, Potter RE. Administração de tecnologia da informação: Teoria e prática. Elsevier, 2005.
2. Porter ME. Estratégia competitiva: técnicas para análise de indústrias e da concorrência. Campus, 1986.
3. Drucker P. O futuro já chegou. Revista Exame, ano 34, n.6, edição 710, p.112-126, 2000.
4. Balloni AJ. Porque gesiti? Porque gestão em sistemas e tecnologias de informação? 2006.
5. Davenport TH. A Natureza da Reengenharia de Processos. 1994.
6. Sousa ME, Medeiros, JV. Afinal o que é Business Process Management (BPM)? Um novo conceito para um novo contexto. Revista Eletrônica de Sistemas de Informação, 2009.
7. Peinado J, Graeml AR. Administração da produção: operações industriais e de serviços. Curitiba: UnicenP, 2007.
8. Lins BFE. Ferramentas básicas da qualidade. Brasília, 1993.
9. Oliveira JW. Sistema de Informação. 2009
10. Oliveira DPR. Sistemas. Organização & Métodos: O&M - uma abordagem gerencial. Atlas, 2002.
11. Takashina N T, Flores MCX. Indicadores da qualidade e do desempenho: como estabelecer metas e medir resultados. Rio de Janeiro. 1996.

Mensuração do capital social acumulado a partir de interações sociais em páginas institucionais no facebook

| Kaíque Matheus R. **Cunha**

| Alan Keller **Gomes**

RESUMO

Funcionalidades que permitem a interação entre usuários dentro das Redes Sociais Online configuram práticas de sociabilidade em rede, ou seja, ações como Publicar, Comentar, Curtir e Compartilhar têm um significado especial quando estão sob o foco da Sociologia Digital. Com o apoio da teoria de Pierre Bourdieu, neste trabalho são identificados e capturados dados relacionados às práticas de sociabilidade em rede dentro de Páginas Institucionais no Facebook. Esses dados são utilizados no aprendizado da sequência de ações. A partir da frequência de ocorrências das sequências aprendidas, um indicador do Capital Social acumulado é apresentado. A partir desse indicador e do número de interações sociais, o volume de Capital Social é computado.

Palavras-chave: Redes Sociais Online, Capital Social, Interações Sociais, Páginas Institucionais, Pierre Bourdieu, Facebook.

■ INTRODUÇÃO

Em uma sociedade cada vez mais conectada e mediada por tecnologias interativas, as Redes Sociais Online (RSO) têm promovido mudanças sociais profundas, a tal ponto, que essas redes digitais se fundem e estendem as redes sociais da vida real. As RSO são facilitadores tecnológicos que contribuem para o fortalecimento de laços sociais da vida real e para o estabelecimento de novos laços (SANTOS; CYPRIANO, 2014). Além disso, elas incorporam em suas funcionalidades o caráter lúdico e simétrico que define a sociabilidade nas redes sociais do mundo real (SANTOS; CYPRIANO, 2011).

Funcionalidades como Publicar, Curtir, Compartilhar, Comentar, têm um significado especial se vistas sociologicamente: elas apontam para práticas de sociabilidade em rede, que são constituídas de formas de comunicação on-line, que levam à formação, à manutenção e à dinâmica de relações entre os agentes (i.e., páginas institucionais) e sua rede de conexões (i.e., seguidores) no campo (i.e., Rede Social Online) (STERNE, 2003) (SANTOS; CYPRIANO, 2014). As RSO oferecem novas formas de acumulação desse tipo de capital, bem como o acesso a recursos que não estão facilmente disponíveis no mundo real para manutenção das relações sociais (RECUERO, 2012).

Para Bourdieu, o capital acumulado a partir das conexões sociais que os agentes estabelecem em um determinado campo é denominado de Capital Social (BOURDIEU, 1986). Instituições que estão presentes nas RSO estão cada vez mais interessadas em saber de que maneira o Capital Social acumulado nesses meios digitais pode refletir no acúmulo desse capital no mundo real.

O Capital Social é acumulado a partir das práticas de sociabilidade em rede, que se cristalizam por meio das trocas realizadas entre os agentes presentes no campo. Nesta pesquisa, agentes são Páginas Institucionais (fanpages) e o campo é a rede Facebook. O volume do capital social acumulado por um agente depende da sua capacidade de mobilizar sua rede de conexões, na qual são estimuladas trocas e o reconhecimento mútuo entre integrantes dessa rede de conexões (BOURDIEU, 1986).

Outros autores estudaram o Capital Social (COLEMAN, 1988) (PUTNAM, 2001) (LIN et al., 2001) (LAW; CHANG, 2008), além de Bourdieu. Apesar de indicadores do acúmulo desse tipo de capital terem sido estudados, principalmente, a partir da estrutura da rede de conexões de cada agente, não há consenso em como o Capital Social pode ser medido. No intuito de contribuir para a superação dessa lacuna, o presente trabalho apresenta uma abordagem para a construção de um indicador e para o cômputo do Capital Social acumulado em Páginas Institucionais dentro do Facebook, a partir do número de interações sociais capturadas de cada página.

A coleta de dados é direcionada por uma estratégia que se apoia no arcabouço teórico de Bourdieu. Esses dados coletados são processados e utilizados na extração automática e na mensuração de interações sociais no formato de sequência de ações. As frequências de ocorrência das sequências são associadas com a capacidade de mobilização dos seguidores conectados a cada Página Institucional (agente).

Caracterizando as relações entre páginas e seguidores dentro de um sistema de pesos, que leva em conta o número de ações executadas por seguidores dessas páginas em cada interação social, é construído um indicador do Capital Social acumulado. Em seguida, esse indicador juntamente com o número de interações sociais capturadas, são utilizados para computar o volume de Capital Social acumulado em cada Página Institucional. O objetivo geral deste trabalho é mensurar o Capital Social acumulado em Páginas Institucionais no Facebook, levando-se em conta o número de interações sociais. A principal contribuição desta pesquisa é a abordagem empregada tanto para a construção de um indicador de acúmulo de Capital Social quanto na utilização desse indicador, juntamente com o número de interações sociais capturadas, para computar o volume de Capital Social acumulado.

Os resultados obtidos estão alinhados com as teorias de Bourdieu no que se refere 1) à representação do sistema de trocas entre agentes e sua rede de conexões, 2) à identificação da capacidade (incapacidade) de mobilização da rede de conexões de uma agente, e, 3) à capacidade (incapacidade) de mobilização de um agente que independe do tamanho da sua rede de conexões, 4) uma grande rede de contatos não garante um grande volume de Capital Social acumulado, ou seja, um volume maior de Capital Social acumulado depende da capacidade (efetiva) de mobilização da rede de conexões do agente, e 5) o Capital Social depende de um esforço incessante de sociabilidade, entretanto, um grande número de interações sociais não provoca necessariamente maior acúmulo de Capital Social, pois depende da efetiva mobilização da rede de conexões.

É importante observar que os dados coletados se referem a informações públicas e não pertencem ao perfil individual de qualquer seguidor de qualquer uma das Páginas Institucionais estudadas. As informações coletadas são mantidas sem identificação do perfil do usuário que a produziu. Ressalta-se que são explorados padrões de comportamento coletivos e impessoais, que podem ser extraídos automaticamente a partir de dados relacionados com as práticas de sociabilidade em rede.

Este trabalho está organizado da seguinte forma: Seção 2 detalha a metodologia empregada na pesquisa; Seção 3 apresenta e discute os resultados obtidos; Seção 4 elenca os trabalhos relacionados; e Seção 5 apresenta as considerações finais.

■ METODOLOGIA

Recorte da pesquisa

Dentro de uma RSO, a ação de Publicar inicia uma interação social. Essa ação também é chamada de estímulo social (SKINNER, 1953). Na Psicologia Social (MECHNER, 2008), cada publicação seguida da execução, ou não, de ações, pode ser compreendida como uma contingência comportamental, ou também, uma interação social (GOMES; PIMENTEL, 2011) (GOMES; PIMENTEL, 2014). Na Sociologia Digital (IGNATOW; ROBINSON, 2017), as práticas de sociabilidade dos agentes (i.e., página institucional) e sua rede de conexões (i.e., seguidores) no campo (i.e., RSO) são cristalizadas a partir da execução de funcionalidades intrínsecas à rede. Na rede Facebook, logo após uma publicação, as ações de Comentar, Curtir e Compartilhar ficam disponíveis e podem ser executadas pelos usuários conectados ao provedor do estímulo social (GOMES; PIMENTEL, 2014).

Dentro da rede Facebook, a partir da ação de Publicar ficam disponíveis para a execução as ações de Comentar, Curtir e Compartilhar. Além do compartilhamento para fora, a rede Facebook permite o compartilhamento de uma publicação para dentro da rede. A ação de compartilhar analisada nesta pesquisa restringe-se apenas ao compartilhamento para dentro do Facebook (compartilhamento interno). Para analisar as práticas de sociabilidade em rede, foi selecionada uma instituição pública de ensino profissional com 13 câmpus. Cada câmpus tem autonomia para gerir sua fanpage seguindo diretrizes estabelecidas institucionalmente, e ainda, pelo governo federal¹. O público conectado nessas páginas são, em sua imensa maioria, estudantes, professores e demais servidores. Cada página é sempre o único provedor de estímulos sociais e o público pode interagir somente executando as ações de curtir, comentar e compartilhar.

Identificação dos dados a serem capturados

O arcabouço teórico de Bourdieu (BOURDIEU, 1981) (BOURDIEU, 1986) (BOURDIEU, 1991), que fundamenta as pesquisas na Sociologia Digital (IGNATOW; ROBINSON, 2017), foi utilizado no direcionamento dos esforços na etapa de identificação dos dados a serem capturados. As funcionalidades de interação e comunicação que apontam para as práticas de sociabilidade em rede, foram identificadas, para esta pesquisa, como sendo as ações de Publicar, Curtir, Comentar e Compartilhar.

1 www.governodigital.gov.br/transformacao/compras/orientacoes/orientacoes-redes-sociais-1

Coleta, transformação e carregamento

Na etapa de coleta, transformação e carregamento dos dados identificados anteriormente, foi empregado o método ETL (*Extract/Transform/Load*) (KIMBALL; ROSS, 2002). O acesso aos dados é feito utilizando-se a API Facebook² durante o mês de janeiro de 2018. Foram coletados dados de todas as publicações (interações sociais) desde a data de criação de cada página. Atualmente, a API Facebook bloqueia o acesso aos dados de fanpages. Para contornar esse problema, até a liberação do acesso, é utilizado *Selenium*³ na coleta.

Scripts na linguagem Python⁴ foram implementados para o processamento do método ETL e os dados foram armazenados em um banco de dados MySQL⁵. Foi construída uma base de dados para cada página dentro do MySQL. Os dados capturados foram transformados para a etapa de extração de sequências de ações. Esses dados transformados foram carregados em uma outra base de dados que contém valores não-numéricos. A manipulação exclusiva de valores não-numéricos é uma restrição imposta pelo algoritmo de extração de sequências.

Extração e mensuração de sequências

Uma implementação do algoritmo GSP (*Generalized Sequential Pattern*) (AGRAWAL; SRIKANT, 1995) foi elaborada para o aprendizado automático das sequências das ações. Originalmente, o GSP não manipula valores numéricos e o número de sequências extraídas depende de um limiar (frequência mínima) arbitrado pelo usuário. Na implementação aqui realizada, suporte mínimo considerado para extração de sequências candidatas foi $\alpha \geq 0$, portanto, todas sequências que possam constar na base de dados foram extraídas. São geradas sequências de tamanho 1 até as de tamanho igual ao número de colunas da base, sem valores de atributos repetidos. Por fim, o número de ocorrências (frequência) de cada sequência é calculado.

Validação

A validação dos resultados é feita a partir do confronto dos valores das frequências de ocorrência das sequências aprendidas com contagens retornadas por consultas elaboradas em Linguagem de Consulta Estruturada - SQL⁶. Cada sequência extraída via GSP de tama-

2 <https://developers.facebook.com/>.

3 <https://www.seleniumhq.org/>

4 <https://www.python.org/>.

5 <https://www.mysql.com/>.

6 <https://www.oracle.com/br/database/technologies/appdev/sql.htm>

nho 3 é mapeada em uma consulta SQL. Na Listagem 1 é apresentado o mapeamento da sequência !K!L!S na sua respectiva consulta SQL.

Listagem 1. consulta em SQL

```
!K!L!S  
SELECT COUNT(*) FROM 'posts' WHERE status_type IS NOT NULL AND num_comments = 0  
AND num_likes = 0 AND num_shares = 0
```

Nessas consultas, são realizadas contagens iguais a 0 ou diferentes de 0 de cada ação. Considerando que as ações analisadas são Comentar, Curtir e Compartilhar, dessa forma, são feitas contagens do número de comentários, do número de curtidas e do número de compartilhamentos iguais a 0 ou diferentes de 0.

Havendo coincidência entre a frequência de ocorrência de cada sequência, obtidas a partir do aprendizado via GSP, e a contagem retornada pelas consultas SQL, esses resultados são considerados válidos. Em suma, os resultados são válidos se houver coincidência de valores entre o que foi aprendido via GSP com o que foi consultado via SQL.

Cômputo do Capital Social

Com o suporte teórico de Bourdieu (BOURDIEU, 1986), a capacidade de mobilização de cada página foi analisada a partir das frequências de ocorrência de cada sequência. Essa capacidade de mobilização se dá a partir de trocas realizadas entre agente e seguidores da página. Nesta pesquisa, foi considerado que as trocas se cristalizam dentro do Facebook a partir da execução ou não das ações Curtir, Comentar e Compartilhar, portanto, as trocas estão retratadas nas sequências aprendidas. A efetividade da capacidade de mobilização, por sua vez, é retratada pela execução de uma ou mais ações, sendo assim, as sequências que contém ações executadas no seu corpo estão associadas a efetividade da capacidade de mobilização do agente Página Institucional, como estudado em trabalho anterior pelos autores desta pesquisa (CUNHA; GOMES, 2019).

Caracterizando as relações entre cada página e seus seguidores dentro de um sistema de pesos, que leva em conta o número de ações executadas pelos seguidores dessas páginas em cada interação social, é construído um indicador do Capital Social acumulado, a partir das sequências aprendidas que estão ligadas à efetiva capacidade de mobilização. O número de ações executadas em cada interação social (sequência aprendida) é utilizado como peso para caracterizar a força da ligação de cada página com seus seguidores (GRANOVETTER, 1977). Em seguida, esse indicador juntamente com o número de

interações sociais capturadas de cada página institucional, são utilizados para computar o volume de Capital Social acumulado.

■ RESULTADOS E DISCUSSÃO

Identificação dos dados e aplicação do método ETL

Como resultado das fases de identificação e coleta dos dados foram construídas 13 bases de dados, uma para cada página institucional (fanpage). Nessas bases foram armazenadas informações como nome da página, identificador da página, identificador da postagem, conteúdo da postagem, tipo de mídia publicada na mensagem, data e hora da publicação, número de comentários, número de curtidas, número de compartilhamentos (internos).

Listagem 2. Exemplo de Base de Dados não-numérica

K, L, S K, L, IS K, IL, S K, IL, IS !K, IL, IS

Na transformação dos dados, valores numéricos foram transformados em valores não-numéricos. O número de comentários maior que 0 é representado por K e valores iguais a 0 por !K. O número de curtidas maior que 0 é representado por L e valores iguais a 0 por !L. O número de compartilhamentos maior que 0 é representado por S e valores iguais a 0 por !S.

Os dados transformados foram carregados em outra base com valores não- numéricos, ou seja, para cada base de dados coletados da RSO, foi criada uma (nova) base de dados não-numéricos. A partir de cada base de dados não-numéricos foi realizado o aprendizado de sequência de ações. Um exemplo de como os dados estão organizados nas bases não-numéricas é apresentado na Listagem 2.

Tabela 1. Mensuração da Incapacidade e Capacidade de Mobilização

1	2	3	4	5	6	7	8	9	10	11	12	13
ID	!K!L!S	total	I. Mob.	C. Mob.	# seg.							
TRD	8	292	0	1	108	304	0	325	1038	0,77%	99,23%	4967
CAT	11	143	0	2	32	143	0	129	460	2,39%	97,61%	2921
IPO	19	125	1	3	24	205	0	145	552	3,44%	96,56%	4086
POI	0	46	0	0	2	75	0	23	146	0,00%	100,00%	764
URT	54	314	1	10	76	394	0	219	1068	5,06%	94,94%	9638
CRI	12	126	0	0	15	131	0	72	356	3,37%	96,63%	1586
CER	71	471	2	20	116	410	1	312	1403	5,06%	94,94%	2905

1	2	3	4	5	6	7	8	9	10	11	12	13
HID	4	100	0	0	25	82	0	44	255	1,57%	98,43%	1982
CPB	23	427	2	3	68	349	0	190	1062	2,17%	97,83%	3679
PSS	5	178	1	2	35	91	0	57	369	1,36%	98,64%	2836
RVD	7	43	0	0	6	18	0	7	81	8,64%	91,36%	10293
IPA	11	3	0	9	1	24	1	2	51	21,57%	78,43%	542
MOR	546	210	0	0	52	331	0	217	1356	40,27%	59,73%	8819

Mensuração das sequências e da capacidade (incapacidade) de mobilização

Exemplos das sequências de ações extraídas automaticamente de cada base de dados são apresentadas Listagem 3. Para o cômputo do Capital Social, foram consideradas apenas as sequências de ações de tamanho 3, pois essas sequências trazem no seu corpo uma associação entre a execução e/ou não das ações Comentar, Curtir e Compartilhar.

Listagem 3. Exemplo de Sequências Candidatas

t1 :< K >, < L >, < S >, < !K >, < !L >, < !S >
t2 :< K, L >, < K, S >, < L, S >, < K, !L >, < K, !S >, < L, !S >, < !K, L >, < !K, S >, < !L, S >, < !K, !L >, < !K, !S >, < !L, !S >
t3 :< K, L, S >, < K, L, !S >, < K, !L, S >, < K, !L, !S >, < !K, L, S >, < !K, L, !S >, < !K, !L, S >, < !K, !L, !S >.

Na Tabela 1 são apresentados uma identificação da Página Institucional de cada Câmpus (coluna 1), o número de ocorrência de cada sequência de ações de tamanho 3 em valores absolutos (colunas 2 a 9), a totalização dos valores de ocorrência dessas sequências (coluna 10), o indicador da incapacidade de mobilização (coluna 11), o indicador da capacidade de mobilização (complementar a incapacidade de mobilização - coluna 12), número de seguidores de cada página (tamanho da rede de conexões - coluna 13).

É necessário ressaltar que somatório do número de ocorrências de todas as sequências de tamanho 3, para um determinado Câmpus, corresponde ao número de linhas da base de dados desse Câmpus, ou seja, na mensuração das sequências de tamanho 3, foram escaneadas todas as ocorrências de sequências de ações armazenadas na base de dados que contém o registro das interações sociais ocorridas dentro da página do Câmpus.

Dessa forma, para cada Câmpus (página institucional), o somatório do número de ocorrência das sequências !K!L!S, !KL!S, K!L!S, !K!LS, LK!S, !KLS, K!LS, KLS correspondem ao número de ocorrências das interações sociais armazenadas em cada base de dados.

Validação dos Resultados

Na Tabela 2 é apresentado o resultado da validação dos valores obtidos para a Página Institucional Câmpus TRD. Na segunda linha estão os números de ocorrência obtidos via GSP, e na terceira linha, os resultados das contagens feitas via consulta SQL. Para cada página foi construída uma tabela de validação tal como a Tabela 2. Os resultados foram considerados válidos, pois o algoritmo GSP aprendeu a partir de todos os dados constantes em cada base (nenhum dado ficou de fora).

Tabela 2. Exemplo de Validação de Resultados.

Câmpus TRD	!K!L!S	!KL!S	K!L!S	!K!L!S	LK!S	!K!L!S	K!L!S	K!L!S	total
GSP	8	292	0	1	108	304	0	325	1038
SQL	8	292	0	1	108	304	0	325	1038
≠	0	0	0	0	0	0	0	0	0

A validação dos resultados apresentados na Tabela 1 foi realizada a partir do confronto das frequências de ocorrências, obtidas a partir do aprendizado via GSP, com contagens obtidas a partir de consultas em SQL feitas em cada uma das 13 bases de dados. Os valores da Tabela 3 foram obtidos a partir da Tabela 1, portanto, validados os resultados da Tabela 1, os resultados apresentados na Tabela 3 também são considerados válidos.

Cômputo do Capital Social

Na Tabela 3 são apresentados uma identificação da página institucional de cada Câmpus (coluna 1) e o percentual de sequências com 0 ações executadas (colunas 2), correspondente ao valor percentual da coluna 2 da Tabela 1 (sequência !K!L!S). Na coluna 3, é apresentado o percentual de sequências com 1 ação executada, que corresponde ao percentual calculado a partir da soma das frequências de ocorrência das sequências !KL!S, K!L!S e !K!L!S.

Na coluna 4 da Tabela 3, é apresentado o percentual de sequências com 2 ações executadas, que corresponde ao percentual calculado a partir da soma das frequências de ocorrência das sequências LK!S, !K!L!S e K!L!S. Na coluna 5, é apresentado o percentual de sequências com 3 ações executadas, que corresponde ao percentual calculado a partir da frequência de ocorrência da sequência K!L!S (coluna 9 da Tabela 1).

Seguindo na Tabela 3, na coluna 6 é apresentado o indicador da capacidade de mobilização (mesmos valores da coluna 12 da Tabela 1). Na coluna 7 é apresentado o indicador do Capital Social acumulado. Na coluna 8 é apresentado o número de interações sociais capturadas de cada página (mesmos valores da coluna 10 da Tabela 1). Na coluna 9 é apresentado o volume de Capital Social acumulado a partir do número de interações sociais

capturadas de cada página. Na coluna 10 é apresentado o número de seguidores de cada página (tamanho da rede de conexões), valores da coluna 13 da Tabela 1.

Tabela 3. Indicador e Mensuração do Capital Social Acumulado.

1	2	3	4	5	6	7	8	9	10
ID	0 ação	1 ação	2 ações	3 ações	C. Mob.	Ind. CS	Int. Soc	CS acum.	# seg.
TRD	0,77%	28,23%	39,69%	31,31%	99,23%	33,59%	1038	348,66667	4967
CAT	2,39%	31,52%	38,04%	28,04%	97,61%	31,96%	460	147	2921
IPO	3,44%	23,37%	41,49%	26,27%	91,12%	30,86%	552	170,33333	4086
POI	0,00%	31,51%	52,74%	15,75%	100,00%	30,71%	146	44,833333	764
URT	5,06%	30,43%	44,01%	20,51%	94,94%	29,99%	1068	320,33333	9638
CRI	3,37%	35,39%	41,01%	20,22%	96,63%	29,68%	356	105,66667	1586
CER	5,06%	35,14%	37,56%	22,24%	94,94%	29,50%	1403	413,83333	2905
CPB	2,17%	40,68%	39,27%	117,89%	97,83%	28,81%	1062	306	3679
PSS	1,36%	49,05%	34,15%	15,45%	98,64%	27,28%	369	100,66667	2836
RVD	8,64%	53,09%	29,63%	8,64%	91,36%	23,05%	81	18,666667	10293
IPA	21,57%	23,53%	50,98%	3,92%	78,43%	22,88%	51	11,666667	542
MOR	40,27%	15,49%	28,24%	16,00%	59,73%	20,00%	1356	271,16667	8819

Seguindo na Tabela 3, na coluna 6 é apresentado o indicador da capacidade de mobilização (mesmos valores da coluna 12 da Tabela 1). Na coluna 7 é apresentado o indicador do Capital Social acumulado. Na coluna 8 é apresentado o número de interações sociais capturadas de cada página (mesmos valores da coluna 10 da Tabela 1). Na coluna 9 é apresentado o volume de Capital Social acumulado a partir do número de interações sociais capturadas de cada página. Na coluna 10 é apresentado o número de seguidores de cada página (tamanho da rede de conexões), valores da coluna 13 da Tabela 1.

O indicador do Capital Social apresentado na coluna 7 foi obtido a partir da média ponderada das colunas 3, 4 e 5, de modo que o percentual das sequências com 1 ação executada recebeu peso 1, o percentual das sequências com 2 ações executadas recebeu peso 2 e o percentual das sequências com 3 ações executadas recebeu peso 3. Esses pesos são atribuídos para refletir a força das ligações entre seguidores e agente página institucional. Sendo assim, as ligações mais fracas têm peso 1 e as ligações mais fortes tem peso 3. O Capital Social acumulado apresentado na coluna 9 foi obtido pela multiplicação do indicador do Capital Social (coluna 7) pelo número de interações sociais (coluna 8).

Discussão: mensuração das sequências e capacidade de mobilização

Nesta pesquisa, as práticas de sociabilidade são caracterizadas pela execução ou da não execução de ações Comentar, Curtir, Compartilhar, logo após uma publicação. Essas práticas são representadas no formato de sequências de ações de tamanho 3, cujas frequências de ocorrência são apresentadas na Tabela 2, colunas de 2 a 9.

Segundo Bourdieu, a mobilização de uma rede de contatos se dá a partir de trocas (BOURDIEU, 1986). As trocas, dentro de uma RSO, são realizadas a partir de funcionalidades que viabilizam a execução de ações como Publicar, Comentar, Curtir, Compartilhar (SANTOS; CYPRIANO, 2014). Agentes com capacidade de mobilização são capazes de promover a realização de trocas, a partir de uma publicação, nas quais seus seguidores executam ações subsequentes. Em contraposição, a incapacidade de mobilização está ligada com a não execução das ações subsequentes a uma publicação, ou seja, o agente publica conteúdo que não mobiliza seus seguidores.

Observando-se a Tabela 1, a incapacidade de mobilização pode ser compreendida a partir da sequência !K!L!S, que expressa que o agente publicou um conteúdo e que 0 ações foram executadas em seguida. Na coluna 11, se destacam os valores 21, 57% e 40, 27%, respectivamente, para as páginas institucionais dos Câmpus IPA e MOR, indicando que essas fanpages tem as mais altas incapacidades de mobilização (portanto, a mais baixa capacidade de mobilização) dentre os 13 agentes analisados. Observando-se o número de interações sociais, o Câmpus IPA tem o mais baixo número de interações sociais dentre os agentes estudados. Já o Câmpus MOR tem o segundo maior número de interações sociais. Esses resultados mostram que não basta prover um número grande de interações sociais para garantir uma alta capacidade de mobilização.

Além disso, Bourdieu destaca que a capacidade de mobilização de um agente depende do tamanho de sua rede de conexões (SANTOS; CYPRIANO, 2014). Continuando na Tabela 1, o número de seguidores de cada página é apresentado na coluna 13. Analisando o número de seguidores do Câmpus TRD, com maior capacidade de mobilização (menor incapacidade) e o Câmpus MOR com menor capacidade de mobilização, é possível observar que o Câmpus MOR tem quase o dobro de seguidores do Câmpus TRD. Portanto, os resultados apresentados na Tabela 1 estão alinhados com as teorias de Bourdieu no que se refere 1) à representação do sistema de trocas entre agentes e sua rede de conexões, 2) à identificação da capacidade (incapacidade) de mobilização da rede de conexões de uma agente, e, 3) à capacidade (incapacidade) de mobilização de um agente que independe do tamanho da sua rede de conexões.

Discussão: Cômputo do Capital Social

Na perspectiva de Bourdieu, o volume de Capital Social de um agente depende da rede de conexões que ele pode efetivamente mobilizar, portanto, a capacidade de mobilização é crucial para determinar o volume de Capital Social (BOURDIEU, 1986).

Na Tabela 3, os valores da capacidade de mobilização estão apresentados nas colunas 3 a 5. Considerando as ações Curtir, Comentar e Compartilhar, dependendo da força da ligação entre agente e seguidores, o agente é capaz de estimular trocas a partir da execução de 1 ação até 3 ações. Na presente pesquisa, esse número de ações executadas é utilizado como peso para caracterizar a força da ligação (GRANOVETTER, 1977).

É admitido, portanto, que o Capital Social acumulado a partir de ligações fortes tenham peso maior no cômputo do volume de Capital Social do que o Capital Social acumulado por ligações fracas. Como mencionado anteriormente, o indicador do Capital Social apresentado na coluna 7 foi obtido a partir da média ponderada das colunas 3, 4 e 5 da Tabela 3. Ao se propor um indicador do acúmulo de Capital Social a partir dessa média, o indicador promove um desbalanceamento em favor das ligações mais fortes entre o agente e seus seguidores.

Observando-se a capacidade de mobilização, os percentuais das ações executadas e o indicador de acúmulo de Capital Social (colunas de 3 até 7 na Tabela 3), o Câmpus POI tem 100% de capacidade de mobilização e o Câmpus TRD tem 99, 23%. Quando são observados os percentuais de ações executadas, o percentual da execução de 3 ações (ligação mais forte) é maior para o Câmpus TRD, com isso, o indicador do acúmulo de Capital Social do Câmpus TRD (33, 59%) é maior que o do Câmpus POI (30, 71%).

Bourdieu destaca que a manutenção e o crescimento do Capital Social dependem de um esforço incessante de sociabilidade, trabalho contínuo de trocas no qual o reconhecimento é afirmado e reafirmado. No caso das RSO, esse esforço é percebido pelo número de publicações que a página institucional se dispõe a fazer. Cada publicação dá origem a uma interação social, portanto, a reprodução do Capital Social depende do esforço constante de publicar conteúdo que provoca a efetiva mobilização da rede de conexões.

O Capital Social acumulado, apresentado na coluna 9 na Tabela 3, foi obtido pela multiplicação do indicador do Capital Social (coluna 7) pelo número de interações sociais (coluna 8). Destaca-se a página institucional do Câmpus CER com maior número de interações sociais e maior acúmulo de Capital Social.

No entanto, um maior número de interações sociais não provoca necessariamente maior acúmulo de Capital Social, pois depende da efetiva mobilização da rede de conexões. Observando-se os valores das colunas 7, 8 e 9 da Tabela 3 para o Câmpus MOR, é possível notar que, mesmo com o segundo maior número de interações sociais, essa página

institucional tem o menor valor para o indicador de Capital Social acumulado, o que afeta o volume de Capital Social acumulado.

Por fim, analisando os valores das colunas 7, 8 e 9 com a coluna 10 na Tabela 3, é possível perceber que uma rede de conexões ampla não garante necessariamente um grande acúmulo de Capital Social. O Câmpus RVD tem o maior número de seguidores, no entanto, o terceiro menor indicador do Capital Social e o segundo menor número de interações sociais. Sob a perspectiva de Bourdieu, o Capital Social acumulado nessa página pode aumentar, se o agente RVD investir esforços em provocar mais interações sociais que mobilizem efetivamente sua rede de contatos.

Sendo assim, os resultados apresentados na Tabela 3 estão alinhados com as teorias de Bourdieu no que se refere 4) uma grande rede de contatos não garante um grande volume de Capital Social acumulado, ou seja, um volume maior de Capital Social acumulado depende da capacidade (efetiva) de mobilização da rede de conexões do agente, e 5) o Capital Social depende de um esforço incessante de sociabilidade, entretanto, um grande número de interações sociais não provoca necessariamente maior acúmulo de Capital Social, pois depende da efetiva mobilização da rede de conexões.

Os resultados aqui apresentados limitam-se a realizar o cômputo do Capital Social acumulado em páginas de instituições de ensino dentro Facebook. Essas páginas têm a característica peculiar de que o único agente provedor de estímulos para as interações sociais é a instituição. Espera-se que abordagem aqui apresentada possa ser generalizada para qualquer agente, que seja único provedor de estímulos sociais, dentro de qualquer RSO. Em trabalhos futuros, o estudo aqui apresentado será ampliado para abarcar, por exemplo, o cômputo de Capital Social acumulado por “influenciadores digitais”.

■ TRABALHOS RELACIONADOS

Os estudos sobre a inferência do Capital Social de agentes presentes nas Redes Sociais Online têm evoluído gradativamente, de abordagens que analisam da estrutura da rede como um grafo, para abordagens que analisam a atividade dos usuários (práticas de sociabilidade em rede) dentro das RSO.

Na perspectiva de análise da estrutura da RSO para inferir o Capital Social, em (VENKATANATHAN et al. 2012) é apresentado um estudo sobre a relação entre a estrutura da rede Facebook, o Capital Social e como essa relação é moderada por traços de personalidade dos agentes. Em (LIU et al. 2014) é demonstrado como as informações estruturais implícitas das RSO podem fornecer informações importantes sobre a personalidade e o Capital Social dos usuários conectados nessas redes.

Em alinhamento com as conclusões obtidas nesses trabalhos, a pesquisa aqui apresentada destaca que agentes com maior volume de Capital Social acumulado são capazes de mobilizar uma rede de conexões mais ampla. Além disso, avança-se na compreensão de que o acúmulo do Capital Social depende das trocas, seguidas de reconhecimento mútuo, que são cristalizadas pelas funcionalidades disponíveis dentro da RSO.

Algoritmos de aprendizado de máquina voltados para tarefas preditivas, que empregam como técnica de aprendizado a regressão linear, são utilizados no estudo (SCHAMS et al. 2018) para investigar a hipótese de que o Capital Social acumulado por usuários do Twitter e do Facebook, pode ser extraído de propriedades de rede e da atividade de cada usuário dentro da rede. Neste trabalho, os usuários são ranqueados a partir de uma pontuação que é inferida automaticamente pelos algoritmos de classificação.

Avançando a partir da análise do trabalho de Schams et al., nesta pesquisa são objeto de análise os comportamentos coletivos dos usuários que fazem parte da rede de conexões de uma Página Institucional (agente). O Capital Social acumulado por um agente é computado a partir da utilização de um algoritmo de aprendizado de máquina voltado para tarefas descritivas, que emprega como técnica de aprendizado a associação, para extrair sequências de ações relacionadas às práticas de sociabilidade em rede dos seguidores de uma página. O Capital Social acumulado de cada agente é calculado a partir da mensuração das sequências de ações executadas por seguidores das páginas.

Em relação a esses trabalhos, a presente pesquisa traz, como contribuições, a extração automática de sequências de ações que caracterizam as práticas de sociabilidade em rede dentro das RSO, e ainda, o cômputo do Capital Social acumulado dentro de Páginas Institucionais no Facebook, a partir da mensuração das sequências de ações. Deve ser destacado que, na presente pesquisa, o volume de Capital Social acumulado é computado a partir das interações sociais capturadas (atividade dos seguidores) de cada página.

■ CONSIDERAÇÕES FINAIS

Neste trabalho foi apresentada uma abordagem para computar o Capital Social acumulado em Páginas Institucionais no Facebook, levando-se em conta o número de interações sociais que ocorrem dentro dessas páginas. O arcabouço teórico de Bourdieu foi utilizado no direcionamento da coleta de dados relativos às práticas de sociabilidade em rede, e ainda, no estabelecimento de um indicador do Capital Social acumulado e do volume de Capital Social acumulado.

Os resultados aqui apresentados estão alinhados com as teorias de Bourdieu no que se refere 1) à representação do sistema de trocas entre agentes e sua rede de conexões, 2) à identificação da capacidade (incapacidade) de mobilização da rede de conexões de uma

agente, e, 3) à capacidade (incapacidade) de mobilização de um agente que independe do tamanho da sua rede de conexões, 4) uma grande rede de contatos não garante um grande volume de Capital Social acumulado, ou seja, um volume maior de Capital Social acumulado depende da capacidade (efetiva) de mobilização da rede de conexões do agente, e 5) o Capital Social depende de um esforço incessante de sociabilidade, entretanto, um grande número de interações sociais não provoca necessariamente maior acúmulo de Capital Social, pois depende da efetiva mobilização da rede de conexões.

O escopo da pesquisa aqui apresentada está limitado a análise das ações executadas dentro de Páginas Institucionais por seguidores dessas páginas. O campo considerado limita-se a rede social Facebook. Apesar disso, deseja-se que a abordagem empregada para computar o Capital Social apresentada nesta pesquisa, possa ser generalizada para ser aplicada a qualquer página ou perfil institucional presente em qualquer Rede Social Online.

■ REFERÊNCIAS

1. Agrawal, R. e Srikant, R. (1995). Mining sequential patterns. In Proceedings of the International Conference on Data Engineering (ICDE '95), pages 3–14. IEEE Computer Society.
2. Bourdieu, P. (1981). Men and machines. In *Advances in Social Theory and Methodology*, pages 304–317. Routledge & Kegan Paul, Boston.
3. Bourdieu, P. (1986). The forms of capital. In *Handbook of Theory and Research for the Sociology of Education*, pages 241–258. Greenwood Press, New York.
4. Bourdieu, P. (1991). *Language and Symbolic Power*. Social Theory. John B. Thompson Ed., Polity Press.
5. Coleman, J. S. (1988). Social capital in the creation of human capital. *American Journal of Sociology*, 94:S95–S120.
6. Cunha, K. M. R. e Gomes, A. K. (2019). Measurement of mobilization capacity via actions sequence learning within institutional Facebook fan pages. In Proceedings of the 25th Brazilian Symposium on Multimedia and the Web, WebMedia 2019, pages 317–320. ACM.
7. Cunha, K. M. R. e Gomes, A. K. (2020). Mensuração do Capital Social Acumulado a partir de Interações Sociais em Páginas Institucionais no Facebook. *In: Anais do IX Brazilian Workshop on Social Network Analysis and Mining (BraSNAM), 2020, Cuiabá*. Porto Alegre: Sociedade Brasileira de Computação, 2020 . p. 85-96. ISSN 2595-6094.
8. Gomes, A. K. e Pimentel, M. d. G. C. (2011). A technique for human-readable representation and evaluation of media-based social interactions in social networks. In Proceedings of the 17th Brazilian Symposium on Multimedia and the Web, WebMedia 2011, pages 119–126. SBC.
9. Gomes, A. K. e Pimentel, M. d. G. C. (2014). Evaluation of Media-Based Social Interactions: Linking Collective Actions to Media Types, Applications, and Devices in Social Networks, pages 75–95. Springer Vienna, Vienna.

10. Granovetter, M. S. (1977). The strength of weak ties. In *Social networks*, pages 347–367. Elsevier.
11. Ignatow, G. e Robinson, L. (2017). Pierre Bourdieu: Theorizing the Digital. *Information, Communication & Society*, 20(7):950–966.
12. Kimball, R. e Ross, M. (2002). *The Data Warehouse Toolkit: The Complete Guide to Dimensional Modeling*. John Wiley & Sons, Inc., New York, NY, USA, 2nd edition.
13. Law, S. P.-M. e Chang, M. K. (2008). Fostering knowledge exchange in online communities: a social capital building approach. *International Conference on Information Systems, ICIS 2008 Proceedings*, page 173.
14. Lin, N., Cook, K. S., e Burt, R. S. (2001). *Social capital: Theory and research*. Transaction Publishers.
15. Liu, Y., Venkatanathan, J., Goncalves, J., Karapanos, E., e Kostakos, V. (2014). Modeling what friendship patterns on Facebook reveal about personality and social capital. *ACM Transactions on Computer-Human Interaction*, 21(3):17:1–17:20.
16. Mechner, F. (2008). Behavioral contingency analysis. *Behavioral Processes*, 78(2):124–144.
17. Putnam, R. (2001). Social capital: Measurement and consequences. *Canadian Journal of Policy Research*, 2(1):41–51.
18. Recuero, R. (2012). O capital social em rede: Como as redes sociais na internet estão gerando novas formas de capital social. *Contemporânea*, 10(3):597–617.
19. Santos, F. C. e Cypriano, C. P. (2011). Blogs e wikis: duas formas de colaboração em redes sociais. *Revista Ciência em Movimento - Rede Metodista*, 13:07 – 19.
20. Santos, F. C. e Cypriano, C. P. (2014). Redes sociais, redes de sociabilidade. *Revista Brasileira de Ciências Sociais*, 29:63–78.
21. Schams, S., Hauffa, J., e Groh, G. (2018). Analyzing a user’s contributive social capital based on activities in online social networks and media. In *Companion Proceedings of the Web Conference 2018, WWW ’18*, pages 1457–1464.
22. Skinner, B. F. (1953). *Science and Human Behavior*. New York: Macmillan. Sterne, J. (2003). Bourdieu, Technique and Technology. *Cultural Studies*, 17:367–389.
23. Venkatanathan, J., Karapanos, E., e Gonçalves, J. (2012). Network, Personality and Social Capital. In *Proceedings of the Annual Web Science Conference, WebSci’12*, pages 326–329. ACM.

Modelagem de processos para representação e gestão da memória corporativa baseado em casos

- | José Damião de **Melo**
- | Daniel Antonio de Jesus **Melo**
- | Bruno Barros de **Souza**
- | Claudia de Medeiros **Lima**
- | Vania de **Jesus**
- | Valdenice de Jesus **Melo**

RESUMO

Objetivo: Modelagem e construção de um protótipo de sistema apoiado em tecnologia da informação e comunicação, que permita a construção, gerenciamento, manutenção e acesso organizado à memória organizacional. **Metodologia:** Para sua obtenção, foram associadas técnicas de inteligência artificial pela metodologia de Raciocínio Baseado em Casos, gerando um sistema baseado em conhecimento, tendo como elemento de representação gráfico do processo em nível gerencial a *Business Process Modelling Notation* – BPMN. **Resultados:** A prototipagem possibilitou a realização de testes em uma base de casos, representando satisfatoriamente a memória organizacional, em uma interface gráfica, dos campos que definem o escopo do problema, e em seguida a recuperação de outros casos na base de dados em ordem de similaridade com os dados preenchidos, englobando dessa forma o total do ciclo de RBC permitindo o crescimento da capacidade da memória organizacional. **Conclusão:** Espera-se que os esforços processuais e cognitivos, possam ser elencados na memória organizacional, permitindo seu uso institucional, que os processos decisórios possam ser executados com maior grau de segurança e controle, e também que a memória organizacional esteja preservada, íntegra e disponível para todos na organização, favorecendo celeridade, criatividade e valorização de saberes, experiências e fomento à produção de conhecimento, contribuindo no desenvolvimento profissional individual e coletivamente.

Palavras-chave: Conhecimento, Decisão, Memória Organizacional, RBC.

■ INTRODUÇÃO

A atualidade nos confronta com uma composição social, econômica, política e tecnológica, que comumente tem sido chamada de Sociedade do Conhecimento (DRUCKER, 1993), cujo contexto tem se caracterizado pela alteração nas relações entre trabalho e capital, apresentando uma ruptura com os modelos derivados das revoluções industriais, transferindo a importância dos bens de capital e dos meios de produção, para um modelo em que os ativos intangíveis, formadores da base de conhecimento organizacional vêm se tornando os principais fatores de produção em diversos segmentos (CASTELLS, 2005).

A transição entre a importância antes direcionada aos bens de capital e agora ao conhecimento se mostra significativamente mais evidente com a separação dos meios e ferramentas de produção da mão-de-obra assalariada qualificada, pois o trabalhador do conhecimento agora é possuidor e detentor de ambos: Os meios de produção estão em seu poder, fazem parte dele e permitem que se desloquem juntamente com seu detentor (WERTHEIN, 2000).

Esta situação inverte, de certa maneira, o princípio capitalista da mais valia, possibilitando um maior equilíbrio na disputa de posições presentes no pacto de trabalho, entre o prestador e o tomador, compreendidos como aquele que possui o conhecimento e aquele que detém o capital: o primeiro, o contratado; o segundo, o contratante.

Para as empresas pequenas e médias, a possibilidade de negociação com relação a este novo elemento de pressão, qual seja a dependência do dono do capital financeiro ao capital intelectual de posse do empregado, pode significar perdas importantes, caso fique indisponível, seja pelo desligamento, falecimento ou mesmo negação de ação por parte do detentor.

Por outro lado, ao tempo em que se consolida como fator de produção, o conhecimento e a possibilidade de sua rápida distribuição pelo uso crescente de novas tecnologias da informação e comunicação, se articulam para criar uma nova estrutura de interação social, diminuindo virtualmente as distâncias geográficas e permitindo a criação de uma sociedade com o potencial de interagir em nível global, o que amplia a capacidade de operação das médias e pequenas empresas (CASTELLS, 2005).

Diante dessa realidade complexa, este artigo propõe uma discussão e proposição de modelagem para formular um processo criação de memória organizacional, utilizando sistemas baseados em conhecimento para múltiplos domínios, permitindo maior interação entre sujeito e objeto, entre ser humano e técnica.

Consolidando a relevância da temática desta pesquisa, é latente a necessidade da preservação da memória corporativa, seja pela manutenção do histórico organizacional, seja pela possibilidade de ampliar ação efetiva de novos projetos e processos, favorecendo a minimização de dependência cognitiva das pequenas e médias empresas para com os

detentores do conhecimento- os novos profissionais do conhecimento, sistemas automatizados e as redes de comunicação de dados que regem os negócios atualmente. Assim como, aponta-se a relevância da apropriação do potencial de novas técnicas e tecnologias para melhor disseminação e criação individual e coletiva de saberes.

Destacam-se como coadjuvantes à temática desta pesquisa ações que envolvem o mapeamento de processos, gestão por objetivos e melhorias sistêmicas, não sendo estas iniciativas as únicas existentes, mas que neste contexto se situam como exemplo de meios possíveis para a obtenção de mudança de comportamento e práticas que podem ser adotadas com o objetivo de aumentar a eficiência. (BATISTA et al, 2005; BATISTA, 2012; FONSECA e FRESNEDA, 2009, FRESNEDA et al, 2008; WIIG, 2000; ANGELONI, 2008)

Portanto, o conhecimento, representando na e pela memória organizacional, por sua própria natureza, está intrinsecamente vinculado à condição humana, incluindo neste contexto as relações sociais, ambientais, econômicas, tecnológicas e políticas que são derivadas da existência e do convívio social.

Estas relações situam os objetivos desta pesquisa enquanto um processo não trivial, demandando uma delimitação abstrata de pertencimento dos domínios em que o problema de pesquisa está situado e de esforços teóricos e metodológicos para a obtenção de resultados satisfatórios. Os quais, por conseguinte, estarão fundamentados na importância da relação e interação entre ser humano, técnica, organização e conhecimento, apontando para dois vieses: a importância da modelagem, proposta nesta pesquisa, para o sujeito e para a instituição.

Compreendendo a interferência das técnicas sobre as formas de conhecimento e sobre a relação do homem com esse conhecimento, concebe-se a utilização das tecnologias da informação e comunicação de maneira dialógica, a qual contraria a mera reprodução de saberes e informações. Seguindo a premissa de Levy (1993), propõe-se uma reflexão sobre os processos do conhecimento e do conhecer a partir das transformações sociotécnicas.

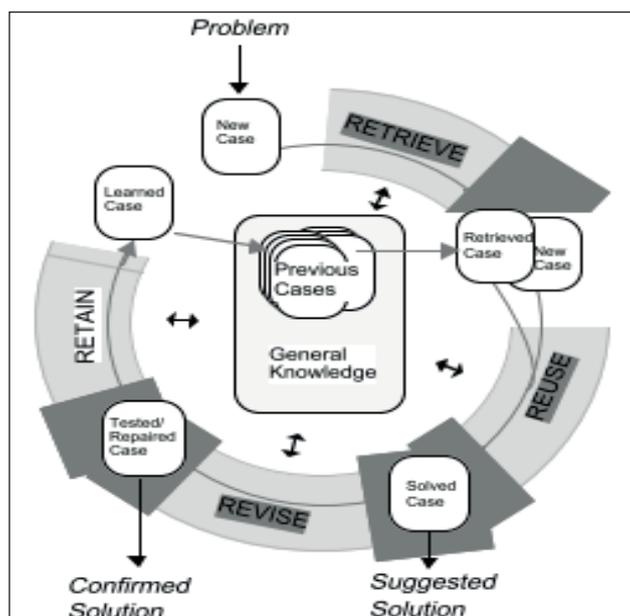
O autor defende a existência de uma inter-relação entre sujeito e objeto, entre homens e máquinas, como condição para o desenvolvimento da comunicação e do pensamento. Nessa compreensão, a técnica é considerada como dimensão de análise, resultado das experiências humanas, enquanto os homens e seus interesses são os verdadeiros agentes efetivos.

Para o ser humano, as técnicas podem auxiliar no desenvolvimento cognitivo individual e coletivo, ou seja, na imaginação e na aprendizagem estimuladas por experiências compartilhadas, e aperfeiçoadas para tomada de decisões. Por exemplo, um programa de informática que trabalha com estudos de caso, demonstra situações modelos que favorecem a reflexão e a criação de novas soluções para questões específicas, sem necessariamente partir do zero. Antes, aproveitando o compartilhamento de saberes e experiências

que poderão dar vazão a novos conhecimentos, de acordo com a realidade circunscrita em tempo e espaço específicos.

Pode-se considerar, portanto, a utilização e apropriação da técnica de estudos de casos para emancipação do sujeito, visto que permite maior complexidade nas análises e construção coletiva de conhecimento, aumentando as possibilidades de transformá-lo em sujeito ativo e criativo, valorizando o potencial individual e coletivo dentro das instituições

Figura 1. Ciclo Case Based Reasoning



Fonte: Aamodt e Plaza, 1994.

A gestão do conhecimento e, em particular, a construção de soluções baseadas em conhecimento e os esforços para a criação de memória organizacional, tem se firmado como estratégia de ação e campo de pesquisa, tratando da problemática de como o conhecimento pode vir a ser criado, armazenado e disseminado nas organizações, em que a Memória Corporativa coloca-se junto a outras práticas de gestão de maneira sistêmica e estratégica para possibilitar uma maior efetividade dos negócios. (DAVENPORT e PRUSAK, 2003; BATISTA, 2012; WATSON, 2003; DRUCKER, 1993).

Neste sentido, buscamos contribuir com o apoio à decisão, através da construção de uma metodologia com suporte de recursos de tecnologia da informação e comunicação, para a representação e processamento do conhecimento pela modelagem e proposição de uma estratégia de criação e manutenção da Memória Corporativa ou Organizacional, considerada parte inalienável do processo de Gestão do Conhecimento, baseado em casos e utilizando paradigmas de sistemas baseados em conhecimento, adaptando os postulados encontrados em Aamodt e Plaza (1994), Watson (2013) e Melo (2012) para esse fim.

■ MÉTODO

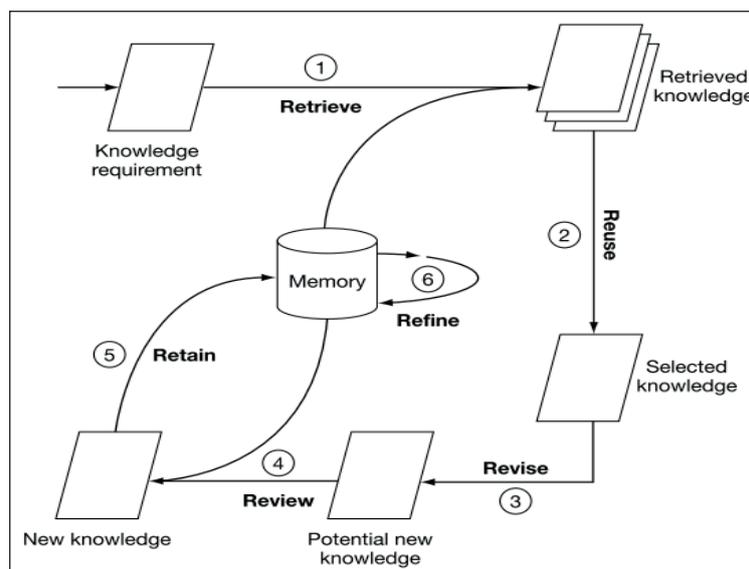
A pesquisa foi estruturada em três etapas distintas, porém, interligadas e interdependentes entre si: inicialmente, ocorreu uma etapa de levantamento bibliográfico e documental; em seguida, foi desenvolvido o processo de criação, acesso e atualização da memória corporativa e finalmente o processo de gestão do conhecimento modelado foi validado empiricamente com o desenvolvimento de um protótipo suportado por ferramentas de tecnologia da informação e comunicação, na forma de simulação.

A estratégia de condução da pesquisa englobou os seguintes marcos de regulação ou pontos-chave de controle:

1. Identificação de fontes bibliográficas e documentais;
2. Modelagem do processo de gestão da memória organizacional e representação em BPMN;
3. Validação empírica do processo modelado pelo desenvolvimento de um protótipo com base em Raciocínio Baseado em Casos.

O ciclo da Gestão do Conhecimento proposto por Watson (2003), diz respeito à aquisição, análise, preservação e uso dos ativos de conhecimento de uma organização de maneira controlada. O autor faz uma relação conceitual entre o ciclo de Gestão do Conhecimento e o ciclo da metodologia de Raciocínio Baseado em Casos.

Figura 2. Ciclo RBC ampliado para um Ciclo de Knowledge Management.



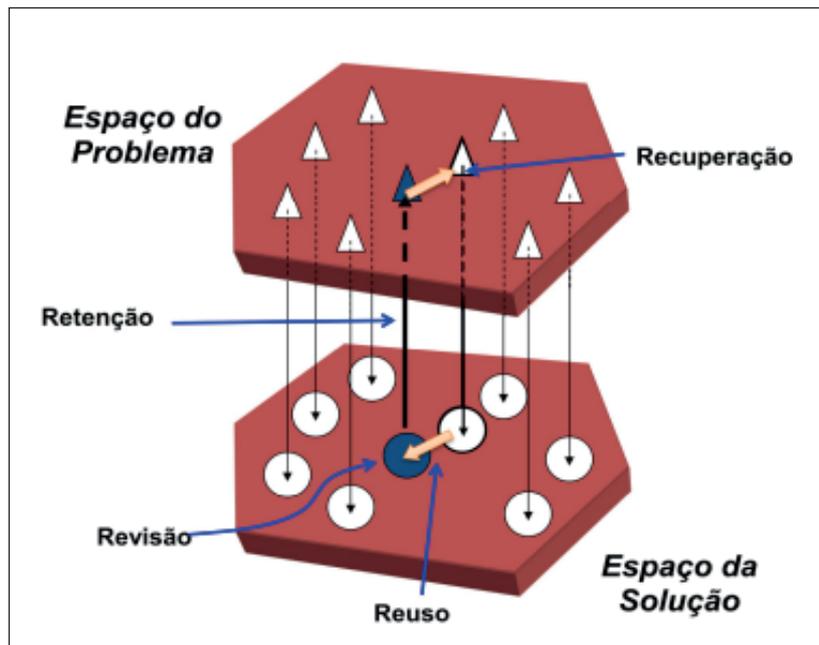
Fonte: Watson, 2003.

- (1) Recuperar o conhecimento que combine com os requerimentos de conhecimento do problema;
- (2) Reutilizar uma seleção do conhecimento recuperado;
- (3) Revisar ou adaptar o conhecimento de forma que ele possa ser usado para solucionar o problema;
- (4) Rever o conhecimento novo para definir se ele deve ser retido;
- (5) Reter o conhecimento, caso o processo de revisão tenha sucesso;
- (6) Refinar o conhecimento na memória de casos ou base de casos, caso seja necessário.

Watson também propôs mapear diretamente os seis R's, numa ampliação do ciclo de RBC original para executar as atividades necessárias à composição de um ciclo de Gestão do Conhecimento, pois os processos de Recuperação, Reuso e Revisão, dão suporte a aquisição do conhecimento; os processos ou atividades de Rever e Refinar permitem a análise do conhecimento e por sua vez a previsão de criação de Memória, suporta a preservação do conhecimento e finalmente a Recuperação, Reuso e Revisão, permitem o uso do conhecimento. (WATSON, 2003; AAMODT e PLAZA, 1994).

Através da determinação de utilidade dos casos recuperados, determina-se a usabilidade dos casos recuperados para a solução dos problemas. Nesse estudo, prevemos a possibilidade de existência de múltiplos atributos para a análise de similaridade, onde cada um desses atributos será parte da representação do caso. Cabe esclarecer, que um Caso pode ser entendido como a relação entre um problema e sua solução, conforme ilustrado na Figura 3, que nos traz claramente como ocorre o ciclo de RBC, com a definição dos espaços de problema e da solução.

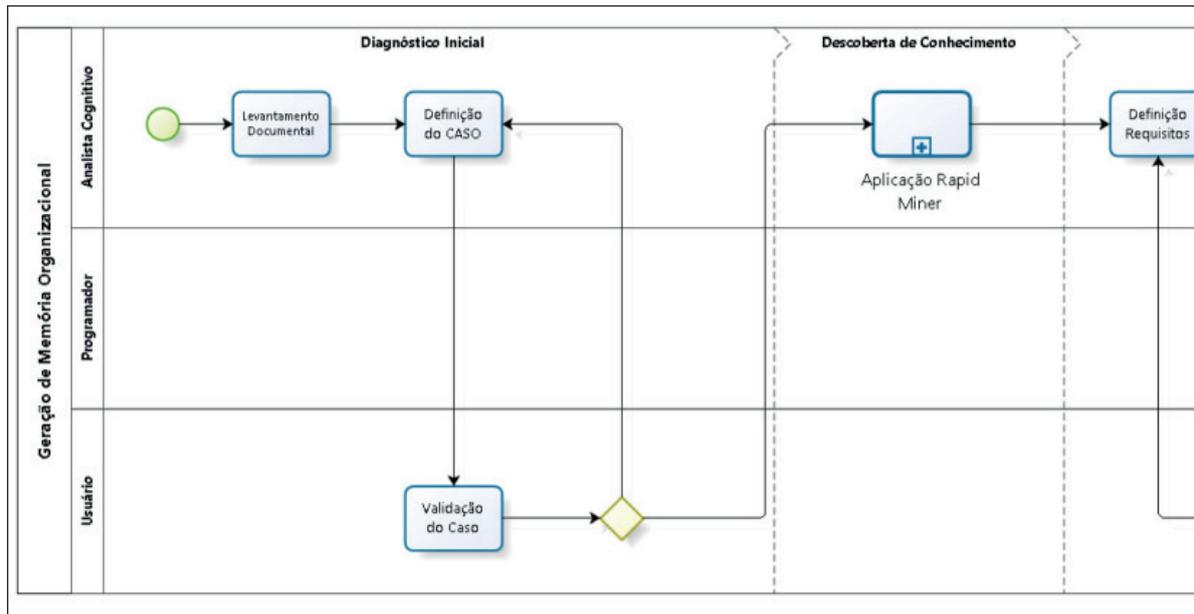
Figura 3. Espaço de Problema e Espaço de Solução



Fonte: Melo, 2012.

A Modelagem de Processos de Negócio e Notação (BPMN) é um padrão desenvolvido pelo OMG – Object Management Group, com o objetivo primário de ser facilmente compreendida pelos integrantes da organização, especialistas ou não, tornando-os capazes de representarem semânticas de processos complexos, de forma gráfica e intuitiva, conforme evidente na Figura 4.

Figura 4. Notação BPMN para o processo de Memória Organizacional

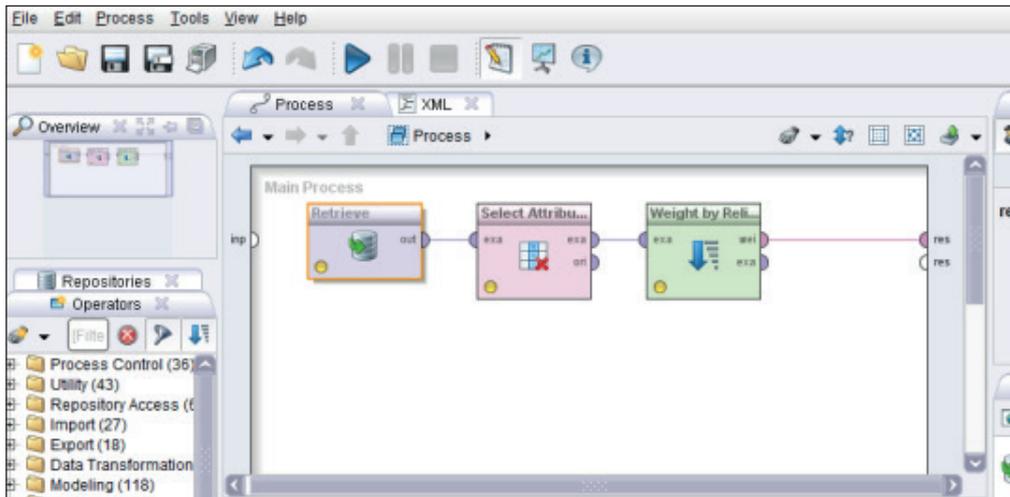


Fonte: Autores

Estão previstos em nosso modelo de construção de Memória Organizacional, três agentes para atuarem no processo: O usuário, representando o domínio da aplicação; o analista cognitivo, que tem a missão de efetuar os levantamentos necessários para a definição dos casos que comporão a memória, efetuando os levantamentos iniciais necessários, executar os processos de descoberta de conhecimento, para a definição dos requisitos necessários para o sistema de informações que efetuará a automação do processo; e o programador, agente especializado que fará a implementação do protótipo, em um primeiro momento, que logo após a validação, será construído o software definitivo.

O RapidMiner é uma ferramenta open-source de data mining que integra em um ambiente recursos de aprendizagem de máquina, mineração de dados, mineração de texto, análise preditiva e análise de negócios, para definir qual o peso no cálculo da similaridade entre os casos, base a comparação de todos os atributos com relação a um atributo alvo, definido como mais importante para o domínio, através da execução do algoritmo *Weight By Relief*. (MELO, 2012; WANGEHEIN e WANGENHEIN, 2006;).

Figura 5. Detalhe da execução do processo WeighByRelief no Rapid Miner

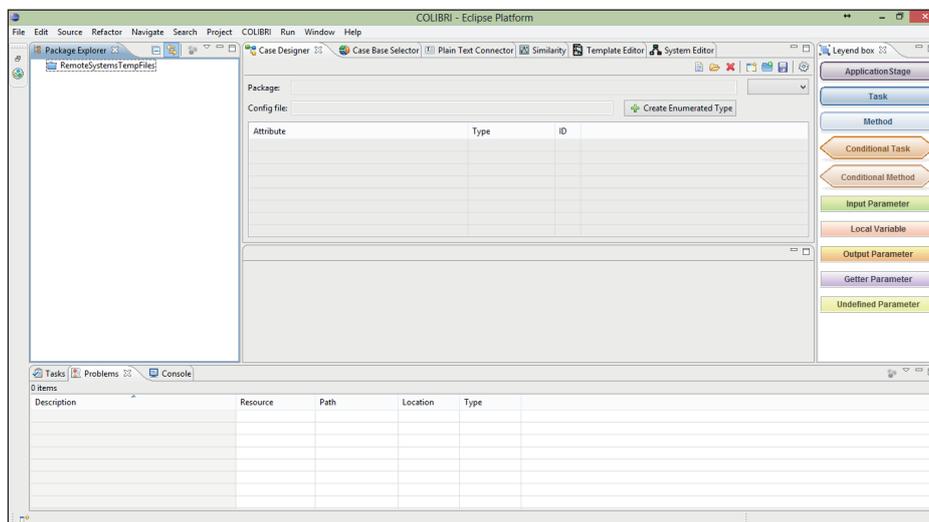


Fonte: Autores

Como passo final do processo de criação de memória organizacional, coloca-se a definição dos elementos e estratégias para a construção do sistema de tecnologia da informação, da forma mais rápida e menos complicada possível.

Foram analisados no andamento da pesquisa três opções: utilização do MyCBR em sua versão como extensão ao Protegé; desenvolvimento direto com a utilização do framework JColibri e a utilização do COLIBRI Studio. (GAIA, 2015. MELO, 2012. MyCBR, 2015).

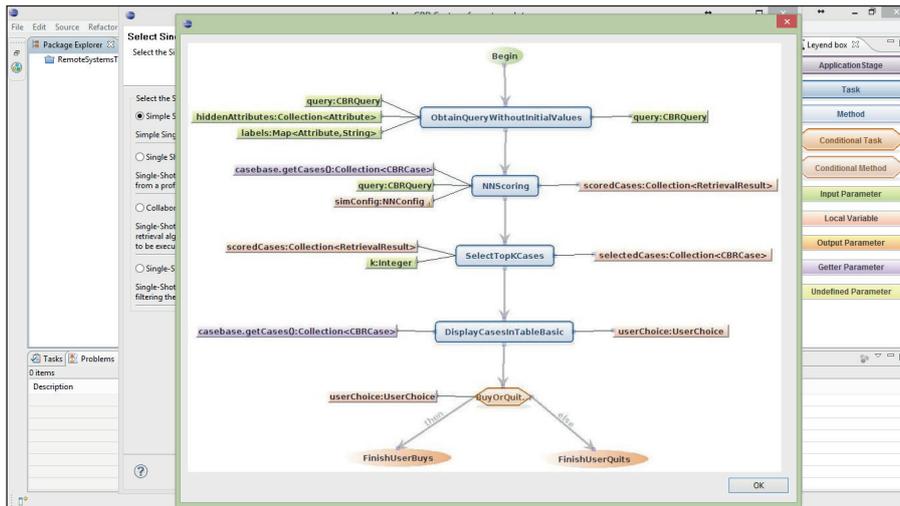
Figura 6. IDE do COLIBRI Studio



Fonte: Autores

Após avaliação das ferramentas, elencamos como elemento de construção de nosso protótipo o COLIBRI Studio, desenvolvido pelo Grupo GAIA - Group for Artificial Intelligence Applications.

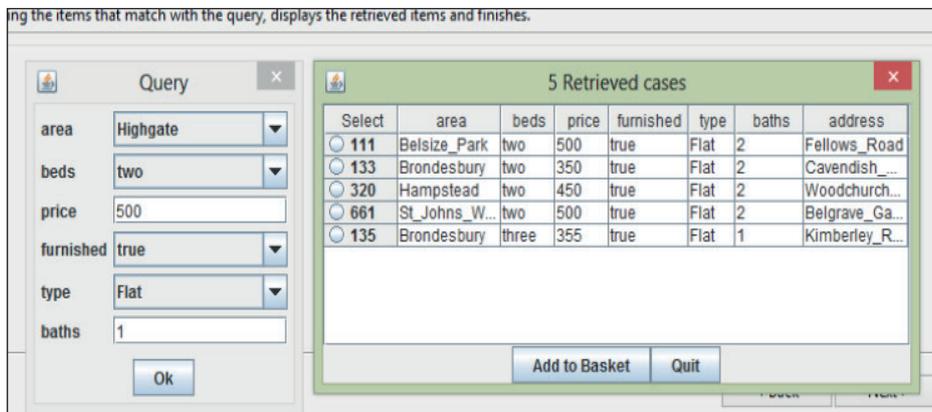
Figura 7. COLIBRI Studio – Template Single Shot, com KNN



Fonte: Autores

Esta solução permite a criação de uma solução de RBC da forma mais automatizada possível, inclusive gerando trechos de código em linguagem JAVA e permitindo o rápido desenvolvimento de aplicações, com o mínimo de conhecimento de programação (GAIA, 2105).

Figura 8. Prototipação do Sistema – Modelos



Fonte: Autores

Além disso, o COLIBRI Studio oferece templates, que funcionam como modelos para a construção de sistemas que utilizarão em seu cerne Raciocínio Baseado em Casos, guiando os desenvolvedores e analistas, permitindo uma rápida prototipação para seus projetos.

Na definição de nosso protótipo para a validação do processo de criação da memória, foi utilizado um determinado *template* ou modelo, para a verificação da viabilidade do processo modelado. Após a análise e definição de atributos que formam a definição de um caso, a etapa posterior tratou do nível de relevância dos atributos da base de casos da Memória Organizacional, elencando quais atributos são mais importantes para se determinar o nível de similaridade entre casos diferentes.

■ RESULTADOS E DISCUSSÃO

O objetivo da construção de modelos, a exemplo do nosso, é representar um processo visando sua compreensão e que por mais detalhado que seja, não será uma representação integral e completa do processo físico, antes, elencará os atributos principais que permitam a representação e análise do mesmo, em nossa proposição, o principal objetivo é permitir o entendimento por parte de todos os agentes envolvidos, em torno do conhecimento e da criação da memória organizacional. (PAVANI et al, 2011 apud MELO, 2012)

Portanto, adotar a notação BPMN apresenta diversas vantagens com relação a ser utilizada como notação para a construção de modelos de processo: é um padrão internacional aceito pela comunidade técnica e acadêmica; permite que o processo seja modelado de forma unificada e padronizada; diminui significativamente a lacuna de compreensão entre o desenho dos processos de negócio e sua implementação.

Além das questões técnicas que facilitam e dinamizam os processos organizacionais, adota-se a perspectiva interativa colaborativa, sobre as quais se fundamenta a ideia do compartilhamento de conhecimentos individuais e coletivos e o aprimoramento das técnicas e tecnologias da informação e comunicação.

A prototipagem tornou possível ainda realização de testes em uma base de casos, representando satisfatoriamente a memória organizacional, em uma interface gráfica, dos campos que definem o escopo do problema, e em seguida a recuperação de outros casos na base de dados em ordem de similaridade com os dados preenchidos, englobando dessa forma o total do ciclo de RBC, obtendo-se assim a memória organizacional e também sua estrutura de acesso e manutenção.

Cabe destacar que sistemas baseados em conhecimento tem a capacidade de aprendizado presente, especificamente quanto a RBC, a retenção dos casos e a ampliação da base representam esse incremento de conhecimento, pois o caso novo ao ser incluído será armazenado como uma solução válida e alvo de consulta em ciclos posteriores, mesmo que a adaptação seja nula, permitindo o crescimento da capacidade da memória organizacional.

■ CONCLUSÃO

Discussões importantes e propostas metodológicas para a adoção de práticas de gestão do conhecimento têm ocorrido com certo grau de continuidade, conforme é possível perceber nas referências, quase sempre fomentadas pela necessidade de aumento de efetividade da execução dos processos para o atendimento das demandas da sociedade.

Nesse contexto apresentamos aqui uma proposta de representação e gestão da memória corporativa ou organizacional, elemento essencial a iniciativas que visem gestão

do conhecimento, buscando sua generalização, utilizando-se do paradigma conceitual do raciocínio baseado em casos

Ao situarmos nossa pesquisa em campo tão amplo, trabalhos futuros e complementares são perceptíveis, em especial aqueles que permitam a validação do modelo proposto em uma empresa ou entidade real; a ampliação do objeto da pesquisa para a inclusão de agentes automatizados para a captura de informações e os canais preferenciais de aquisição e difusão do conhecimento possam se efetivar.

Espera-se que os esforços processuais e cognitivos, capturados dos agentes que atuam nos processos, possam ser elencados na memória organizacional, permitindo ainda seu uso institucional, contribuindo para que os processos decisórios possam ser executados com maior grau de segurança e controle e também que o conhecimento necessário para estas tarefas possa ser mantido, compartilhado e a memória organizacional esteja preservada, íntegra e disponível para todos na organização. Favorecendo, inclusive, celeridade, criatividade e valorização de saberes, experiências e fomento à produção de conhecimento, fato que contribui no desenvolvimento da autoestima profissional individual e coletiva.

■ REFERÊNCIAS

1. AAMODT A., PLAZA E. Case-Based Reasoning: Foundational Issues, Methodological Variations, and System Approaches. **AI Communications**. IOS Press, Vol. 7: 1, pp. 39-59. 1994
2. ANGELONI, Maria Terezinha. Org. **Gestão do conhecimento no Brasil: casos, experiências e práticas de empresas públicas**. Rio de Janeiro, Qualitmark: 2008
3. BATISTA, Fábio Ferreira. **Modelo de gestão do conhecimento para a administração pública brasileira: como implementar a gestão do conhecimento para produzir resultados em benefício do cidadão**. Brasília: Ipea, 2012.
4. BEPPLER, F. D. **Emprego de RBC para Recuperação Inteligente de Informações**. Dissertação (Pós-Graduação em Engenharia da Produção) - Universidade Federal de Santa Catarina - UFSC, Florianópolis, 2002.
5. BRANDALISE, Loreni Teresinha. **Modelo de medição de percepção e comportamento: uma revisão**, 2005. Disponível em < <http://www.lgti.ufsc.br/brandalise.pdf>>. Acesso em 13 de junho de 2014
6. CASTELLS, Manuel. **A sociedade em rede**. 8ª Edição. São Paulo: Paz e Terra, 2005.
7. DAVENPORT, Thomas H.; PRUSAK, Lawrence. **Conhecimento empresarial: como as organizações gerenciam o seu capital intelectual**. Rio de Janeiro: Elsevier, 2003, 14ª Edição.
8. DRUCKER, Peter. **A sociedade pós-capitalista**. São Paulo: Pioneira, 1993.
9. HAIR, Joseph F. ET al. **Fundamentos de métodos de pesquisa em administração**. Porto Alegre, Bookman, Reimpressão, 2006.

10. KASTER, Daniel dos Santos. **Combinando bancos de dados e raciocínio baseado em casas para apoio a decisão em planejamento ambiental**. Dissertação de Mestrado Programa de Pós-graduação em Ciência da Computação, Instituto de Computação, Universidade de Campinas – UNICAMP. São Paulo, 2006.
11. KOLODNER, Janet L. An Introduction to Case-Based Reasoning. **Artificial Intelligence Review** 6, 3--34, 1992.
12. LEVY, P. **As tecnologias da inteligência: O futuro do pensamento na era da informática**. Trad. Carlos Irineu da Costa. São Paulo: ED 34, 2000.
13. MELO, J. D. **Gestão do Conhecimento e Apoio a Decisão na Gestão Ambiental: C@BR@L – Case Based Reasoning aplicado ao Licenciamento Ambiental**. Dissertação de Mestrado Programa de Pós-graduação em Modelagem Computacional de Conhecimento, Instituto de Computação, Maceió: Universidade Federal de Alagoas, 2012.
14. TAKEUCHI, Hirotaka, NONAKA, Ikujiro. **Gestão do Conhecimento**. Porto Alegre; Bookman, 2008.
15. WANGENHEIM, Cristiane Gresse von. WANGENHEIM, Aldo von. **Raciocínio Baseado em Casos**. São Paulo, Manole:2003
16. WATSON, Ian Duncan. **Applying knowledge management: techniques for building corporate memories**. Morgan Kaufmann, San Francisco. 2003.
17. WERTHEIN, Jorge. **A sociedade da informação e seus desafios**. In Ci. Inf., Brasília, v. 29, n. 2, p. 71-77, maio/ago. 2000, Disponível em < <http://www.scielo.br/pdf/ci/v29n2/a09v29n2.pdf>>. Acesso em 13 de junho de 2014.
18. WIIG, Karl M. **Application of Knowledge Management in Public Administration**, Paper, Taiwan: 2000. Disponível em < <http://productivity.tavanir.org.ir/kmgroupp/maghale/pe3.pdf>>. Acesso em: 13 jul. 2014.
19. YIN, Robert K. **Estudo de caso: planejamento e métodos**. Tradução Ana Thorrel, 4. ed. Porto Alegre: Bookman, 2010.

Performance de aplicações web: um estudo de caso na Empresa Tuhl Soluções

| Daniel da Rosa **Vasconcelos**
UFSM

| Sidnei Renato **Silveira**
UFSM

| Antônio Rodrigo Delepiane de **Vit**
UFSM

| Cristiano **Bertolini**
UFSM

| Fábio José **Parreira**
UFSM

| Guilherme Bernardino da **Cunha**
UFSM

| Nara Martini **Bigolin**
UFSM

RESUMO

Este artigo apresenta o desenvolvimento de um estudo de caso sobre técnicas e ferramentas a serem aplicadas para melhorar o desempenho de aplicações *web*, desenvolvidas na empresa Tuhl Soluções. O principal resultado consistiu em otimizar os recursos computacionais no lado do servidor, permitindo um ganho de performance em aplicações *web*, impactando diretamente na qualidade do serviço prestado pela empresa aos seus clientes.

Palavras-chave: Desempenho de Aplicações web, Otimização de Recursos Computacionais, Estudo de Caso.

■ INTRODUÇÃO

Os avanços tecnológicos trouxeram, também, alguns problemas relacionados à entrega de serviços na Internet como, por exemplo, a baixa performance de aplicações *web*, já que estas necessitam de maiores recursos de processamento dos dispositivos que acessam estas aplicações. Surgem, então, uma série de fatores ligados aos aspectos de performance, como a necessidade de otimização de aplicações, tanto no lado do servidor, como no lado do cliente. As empresas que trabalham com este tipo de tecnologia, devem estar sempre atentas ao mercado digital, visto que acontecem diversas mudanças que impactam diretamente na prestação destes serviços. No ano de 2009 a gigante dos serviços de busca, *Google*, fez um experimento com um grupo restrito de usuários, aumentando a latência do acesso ao serviço em *300ms*, passando de *100ms* para *400ms*. Os resultados do experimento mostraram que, aumentar a latência do serviço de busca, reduziu o número de pesquisas diárias (GOMES, 2012).

Neste contexto, a motivação para o desenvolvimento deste trabalho surgiu a partir do trabalho desenvolvido na empresa Tuhl Soluções, onde um dos autores atua como especialista na área de infraestrutura. Esta área é responsável pelos servidores de rede existentes na nuvem. Em trabalhos já desenvolvidos para clientes da empresa, notou-se a necessidade de um aprimoramento na entrega de serviços relacionados a aplicações *web*, envolvendo a melhoria da performance das mesmas.

A empresa Tuhl Soluções possui mais de 5 anos de atividades no mercado de soluções digitais. Seu foco é desenvolver sistemas *web*, *sites* e aplicativos móveis para as plataformas *iOS* e *Android*. A empresa atende clientes dos mais variados segmentos, desde grandes indústrias, como também empresas de pequeno porte. Inicialmente, a Tuhl Soluções era focada na entrega de um serviço em específico, um sistema bastante robusto para executar diversos tipos de pesquisas, seja de mercado, satisfação, pesquisa de opinião, eleitoral, entre outros. Conforme a empresa foi crescendo, notou-se a necessidade de ampliar a gama de serviços, visto que a empresa era especialista na área de desenvolvimento de *software*, podendo beneficiar com seus trabalhos demais empresas da região.

Neste sentido, o principal objetivo deste trabalho foi o de estudar técnicas e ferramentas que pudessem ser utilizadas para melhorar a performance de aplicações *web* desenvolvidas na empresa Tuhl Soluções.

■ FUNDAMENTAÇÃO TEÓRICA

Performance de Aplicações Web

A *World Wide Web* (WWW) ou simplesmente *web*, foi projetada, inicialmente, para possibilitar a troca de informações entre cientistas, e neste contexto não eram necessárias grandes preocupações com a questão de desempenho. Com o avanço das décadas e a inclusão de novos tipos de usuários, aliado a mudanças de conteúdos manipulados (tais como os conteúdos e multimídia), tornou-se indispensável um aperfeiçoamento para suprir e escalar as aplicações *web*, de forma que se obtenha performance no acesso, visto que empresas estão investindo um valor alto para obter melhores resultados (DANTAS, 2003).

Essa preocupação com o desempenho de serviços relacionados à entrega de conteúdo na *web* é um assunto importante, visto que o atraso na entrega da informação ao usuário que está no outro lado da tela pode gerar grandes perdas, impactando muitas vezes diretamente na questão financeira.

Imaginemos um cenário onde uma empresa possui um site para comercialização de produtos (*e-Commerce*) e, cada vez que um usuário busca o produto desejado dentro desse ambiente, tenha que aguardar um período de 10 segundos. Neste cenário, imaginemos o impacto que estes 10 segundos podem fazer no resultado financeiro dessa empresa ao final de cada dia? A gigante *Amazon*, do comércio eletrônico, mostrou que o intervalo de 100ms de atraso no carregamento de suas páginas resultava em uma baixa no seu faturamento equivalente a 1% (LINDEN, 2008). Já a *Microsoft* concluiu que a diferença de 2000 ms (2 segundos) referente à latência do seu site, causava uma redução de 4,3% em seu faturamento (SCHURMAN, 2009).

Ao diminuir o tempo de carregamento em 2,2 segundos da *landing page* do *Firefox*, a *Mozilla* obteve um aumento de 15% em seus números de *downloads*, expressando um ganho superior a 60 milhões de cópias anuais (CUTLER, 2010).

Estudos desenvolvidos em grandes corporações mundiais mostram o impacto expressivo que os aspectos ligados à performance podem causar em uma empresa. Além dessas empresas globais, é interessante analisar este mesmo impacto em serviços prestados por empresas nacionais (ROCHA, 2013).

Segundo dados do ano de 2012, a cada minuto o portal de notícias Terra recebia aproximadamente 535 mil autenticações, enviava 32 mil e recebia 17 mil *e-mails*, fazia o carregamento de 150KB em sua capa e 540 KB em *banners*, recebia 3.6 milhões de *hits* e consumia 2.7 TBits de rede. Se fosse possível uma redução de apenas 1 KB, no intervalo de 1 minuto haveria uma economia de 3.4 GB, em 1 hora uma economia de 180 GB e em um dia, um valor total de 4.2 TB (GOMES, 2012).

No site Peixe Urbano existiam mais de 20 milhões de usuários ativos, gerando mais de 22 milhões de acessos e mais de 75 milhões de *page views* por mês. Após algumas otimizações, a página principal do *site* teve uma redução em seu tamanho de 3 MB para 267 KB, reduzindo o tempo de carregamento de 7.5 segundos para 4.3 segundos. Essa redução gerou uma economia diária de R\$ 16.000,00 em custos com servidores, além de aumentar a receita em mais de R\$ 1.000.000,00 por dia (KEPPELEN, 2012).

Técnicas que podem ser utilizadas para obter uma melhor performance

Em se tratando de otimização e ganho de performance, pode-se avaliar todo um contexto de estrutura de um serviço que fornece ou hospeda uma aplicação *web*. As técnicas de aumento de performance podem ser aplicadas tanto no lado do servidor, como no lado do cliente (onde acontece o processamento de dados no próprio *browser* do usuário).

A arquitetura cliente-servidor é utilizada neste contexto de aplicações *web* onde o usuário, quando quer acessar um *site*, por exemplo, deve solicitar informações ao servidor por meio do seu navegador e, posteriormente, o servidor responde enviando os dados necessários para realizar o acesso. Essas informações recebidas serão interpretadas pelo *browser* que, então, fará a renderização do *site* que logo é exibido na tela do usuário (LOPES, 2016).

Para obter um ganho de performance em aplicações *web* existem inúmeras técnicas a serem utilizadas. Nas próximas subseções serão apresentadas algumas delas, especificamente empregadas no lado do servidor, que é o foco deste trabalho.

Rede de Distribuição de Conteúdo

O conceito de *Content Delivery Network* (CDN) é bastante empregado em situações onde o objetivo é ter um ganho no intervalo de tempo do acesso a um conteúdo na *web*. Essa técnica permite que os dados de uma determinada aplicação *web* estejam replicados em diversos locais, espalhando estes geograficamente para que, assim, um usuário, ao solicitar o conteúdo por meio de uma requisição via internet, tenha a resposta da fonte (servidor) mais próxima. Uma rede de distribuição de conteúdo é uma coleção de servidores de Internet distribuídos por múltiplas localidades, que distribuem conteúdo para usuários de forma mais eficiente (SOUDERS, 2007).

O recurso de distribuição de conteúdo na rede é indicado para conteúdos estáticos, como imagens, vídeos e até mesmo arquivos *JavaScript* e CSS (*Cascade Style Sheets*). Como desvantagem, pode ser citado o fato de que, em alguns casos, pode ocorrer um atraso no carregamento de algum conteúdo, em virtude de existirem outras requisições de acesso em outros sites hospedados neste mesmo servidor (LOPES, 2016).

Otimização de imagens

É muito comum que aplicações *web* possuam uma série de imagens distribuídas em suas páginas, como por exemplo, um *site*. Em se tratando de performance, o tamanho físico dessas imagens pode afetar diretamente o desempenho de uma aplicação *web* (LOPES, 2016).

Imaginemos um portal de notícias, onde toda notícia possui uma imagem de capa e, além disso, existam galerias de fotos dentro de cada uma destas notícias. Neste cenário, cada vez que um usuário solicitar o acesso ao conteúdo de uma notícia por meio do seu navegador, terá que buscar essas imagens no servidor (por meio de um segmento de rede) para que, então, consiga visualizar as mesmas em seu *browser*. Neste processo, existe uma taxa considerável de *download*, pois geralmente uma imagem não otimizada tem o tamanho aproximado de 600KB. Logo, se a página tiver 15 fotos anexadas, o usuário terá que fazer o *download* de 9MB (apenas de arquivos relacionados à mídia). Há casos em que as imagens possuem tamanhos equivalentes a 3MB ou mais, tudo depende do tratamento que essa imagem recebeu antes de ser enviada ao servidor, e também o dispositivo que gerou a mesma.

Para otimizar o carregamento dessas páginas e obter um ganho de performance, existem algumas técnicas relacionadas à compressão de imagens, permitindo que as mesmas mantenham um padrão de qualidade visual ao olho humano. Um exemplo bem comum, que pode ser feito antes de enviar a imagem para o servidor, é alterar a resolução desta usando programas como o *Adobe Photoshop* ou o *Gimp*, entre outros relacionados ao tratamento de imagens. Se uma página na *web* exibe as imagens em um quadro de 800 px (*pixels*) por 600 px, não é necessário enviar para o servidor uma imagem com tamanho superior a isso. Neste simples procedimento, já se consegue um ganho considerável em relação ao tamanho do arquivo.

Outra técnica para otimizar imagens e arquivos de mídia é o formato *WebP*, criado pela empresa *Google*. Esse formato, lançado em 2010, começou a ganhar visibilidade e vem sendo adotado por vários desenvolvedores e empresas (NILSON, 2014). O *WebP* é um formato de imagem que estabelece os mesmos padrões de arquivos como PNG (*Portable Network Graphics*) e JPEG (*Joint Photographic Experts Group*) com tamanhos inferiores. Este formato possui o recurso de transparência nativo em arquivos PNG, possuindo um tamanho 26% menor que o PNG. Se comparado ao JPG, o *WebP* pode obter uma taxa de 34% de redução em seu tamanho. O *WebP* é suportado nativamente nos navegadores *Google Chrome* e *Opera*, entre outras ferramentas e bibliotecas de *software* (GOOGLE, 2016).

Uma desvantagem dessa técnica é o fato de aumentar o volume físico de dados no servidor, de modo que o custo operacional aumenta gradativamente. O motivo disso se dá pelo fato de que nem todos os navegadores (como exemplo, o *Safari*) suportam o recurso *WebP*. Logo, torna-se necessário disponibilizar imagens em outros formatos para esses

navegadores em específico. Exemplos de navegadores que ainda não suportam *WebP* são o *Mozilla Firefox* e o *Safari*.

Podemos analisar o comportamento de uso do formato *WebP* em diferentes navegadores acessando, por exemplo, o *site* do Jornal Folha do Noroeste, disponível em <<https://www.folhadonoroste.com.br>>, onde só serão carregadas imagens *WebP* se o *browser* tiver suporte a essa tecnologia, caso contrário, serão carregadas imagens no formato JPEG.

Compressão GZIP no Webserver

O GZIP é um compressor genérico que pode reduzir 30% ou mais o tamanho de arquivos em uma aplicação *web*, como arquivos CSS, HTML (*HyperText Markup Language*) e *Javascript*, entre outros. Em um servidor que possui esse recurso ativado, sempre que um usuário solicita o conteúdo da aplicação *web*, o servidor antes de servir faz uma compressão dos arquivos específicos e, então, envia novamente ao navegador do usuário que, por sua vez, faz a descompressão e carrega os arquivos originais no *browser* (LOURENÇO, 2017).

Basicamente, o GZIP encontra cadeias semelhantes em um arquivo de texto e substitui essas *strings* de forma temporária, resultando em um arquivo menor. Essa técnica de compressão traz grandes benefícios em uma aplicação *web*, pois arquivos como HTML e CSS possuem bastante texto repetido, além de espaços. Como resultado dessa compressão é possível obter uma redução de 60% a 70% no tamanho de páginas *web* (HERTEL, 2017).

Ativar o GZIP é uma das otimizações mais simples e que trazem mais recompensas, porém, muitas pessoas não a fazem. Em geral, todos navegadores modernos suportam compressão GZIP em todas as solicitações HTTP (*HyperText Transfer Protocol*). Para que o recurso funcione nos navegadores, é necessária uma configuração no servidor (GRIGORIK, 2020).

Caching de conteúdo dinâmico

O *caching* de conteúdo dinâmico é uma técnica utilizada para otimizar consultas à base de dados. Existem diversas aplicações para fornecer esse tipo de serviço, sendo as mais utilizadas o *Memcache* e o *EHCACHE*. Basicamente, quando uma consulta é feita a uma base de dados, essa query é analisada por meio de alguma aplicação de *caching* dinâmico, e, caso a informação requisitada esteja em *cache*, é automaticamente entregue ao serviço que solicitou. Caso contrário, a consulta é feita diretamente à base de dados. Essa técnica foi utilizada pela equipe do *Twitter* por meio da ferramenta *Memcache*, tendo reduzido o tempo de latência sentido pelos usuários (PEREIRA, 2011).

Caching de conteúdo estático

Na técnica de *caching* de conteúdo estático o objetivo é atender às requisições dos clientes de uma maneira otimizada. Existem algumas ferramentas que disponibilizam este tipo de serviço, entre elas se destacam as aplicações *Nginx*, *Squid* e *Varnish*. O funcionamento deste tipo de aplicação se baseia em receber requisições dos clientes verificando se a informação solicitada está armazenada em *cache*. Caso a informação estiver em *cache*, a resposta é encaminhada ao cliente; do contrário, o serviço solicitará os dados ao servidor responsável, conforme estabelecido na configuração da aplicação de *caching*, e logo entregue ao cliente. Estes dados passam a ser armazenados em *cache* para consultas futuras (PEREIRA, 2011).

Balanceamento de carga

O balanceamento de carga é uma técnica bastante utilizada quando o objetivo é distribuir o processamento de dados entre dois ou mais servidores, evitando que haja uma sobrecarga em um único servidor. Além disso, é também utilizada para alta disponibilidade, permitindo que mesmo em um ambiente que exista falha em um de seus servidores, outro possa assumir, impedindo que os serviços fiquem indisponíveis. As políticas de balanceamento mais comuns são *roundrobin*, *leastconn* e *source*. Na política *roundrobin* as requisições são atendidas de maneira sequencial, ou seja, cada solicitação é atendida por um servidor de cada vez. Na política *leastconn*, as requisições são processadas pelo servidor que tiver menos conexões ativas em dado momento. Na política *source*, os pedidos são atendidos com base no endereço IP (*Internet Protocol*) do cliente, o que permite lidar com questões de processamentos de acordo com a região geográfica (PEREIRA, 2011).

Web cache

Em um cenário de aplicação *web* é muito comum o alto volume de tráfego de dados entre servidor e cliente. Muitas vezes, este volume de informações pode ocasionar atraso na comunicação de dados, resultando em lentidão no uso de uma aplicação. Ao utilizar o recurso *web cache*, poderá ocorrer uma redução no número e tamanho de solicitações HTTP, fazendo com que as páginas carreguem com muito mais velocidade em virtude da diminuição de requisições feitas pelo *browser* (BERGANTIN, 2014).

Servidores *web* usam uma data de expiração no cabeçalho (*Expires Header*) das respostas HTTP, com o objetivo de comunicar aos *browsers* o intervalo de tempo que certo componente deverá ficar armazenado em *cache* evitando, assim, com que o componente

não seja solicitado novamente, considerando o prazo de validade estipulado pela aplicação (YAHOO, 2020).

Minificar o Javascript e CSS

Minificar o código *Javascript* e *CSS* é uma prática que se destina a remover de um código-fonte todos os caracteres desnecessários, como comentários, espaços em branco, entre outros. A partir disso, o arquivo que contém o código-fonte tem seu tamanho reduzido, melhorando o tempo de resposta de uma aplicação (BERGANTIN, 2014).

Em uma pesquisa com dez principais *sites* dos Estados Unidos da América, a minificação alcançou uma redução de tamanho em média de 25% do seu tamanho original. Existem várias maneiras de minificar o código fonte, por meio de *softwares*, *serviços* e *sites*, entre outros (YAHOO, 2020).

■ TRABALHOS RELACIONADOS

O trabalho apresentado por Pereira (2011) aborda um estudo de caso sobre técnicas de infraestrutura que visam a melhorar a performance de aplicações *web*. O estudo contemplou técnicas tais como o balanceamento de carga, *caching* de conteúdo estático e *caching* de conteúdo dinâmico.

O objetivo de comparar essas técnicas foi o de avaliar o impacto gerado em aplicações *web*, mais especificamente no tipo de aplicação *CMS* (*Content Management Systems*). Estas aplicações são usadas pelas organizações de forma a melhorar a comunicação entre usuários, reduzir a replicação de dados e controlar o acesso aos mesmos. Além disso, os *CMS* também são utilizados como plataforma de suporte ao negócio das organizações. Neste estudo o autor buscou duas aplicações do tipo *CMS* para serem estudadas e aplicadas as técnicas relacionadas à performance. Como critério principal para a escolha das duas aplicações, foi usado como referência o número de *downloads* mensais, além do volume mensal de consultas capturadas pelo *Google* (versão inglesa), entre junho de 2008 e junho de 2009. A partir dos critérios usados para a escolha das soluções *CMS* a serem utilizadas no estudo de caso, foram escolhidas as ferramentas *Joomla* e *Drupal*, sendo as mais usadas pelo público.

Foram usados dois ambientes de teste, sendo um o Cenário Base, utilizado para avaliar a performance das aplicações antes de introduzir qualquer técnica de otimização de performance, e um segundo ambiente denominado Cenário Avançado, onde foram adicionados outros recursos, tais como um servidor para *cache* de conteúdo estático, um serviço de balanceamento de carga e serviços de *cache* de conteúdo dinâmico. Também foi aumentado

o número de servidores *Web*, permitindo assim testar o efeito de balanceamento de carga pelos servidores.

Para comparar e auferir o ganho de performance pela utilização das várias técnicas, foram selecionados módulos comuns entre os CMS *Drupal* e *Joomla*. Os módulos escolhidos foram *RSS*, *Poll*, *Search*, *Index Page* e uma página com conteúdos. Por meio da ferramenta *Apache Bench*, que permite simular vários usuários desempenhando a mesma tarefa e recolher métricas a partir da execução de *workloads* previamente gravados, tornou-se fácil colocar o CMS a ser utilizado por um determinado número de usuários simultaneamente. A comparação teve por base duas métricas comuns neste tipo de sistemas: latência e *throughput*. A latência representa o tempo final que o usuário tem que esperar desde que solicita um pedido até que a resposta fique completa. O *throughput* mede a capacidade do servidor processar N pedidos por determinada unidade de tempo.

De acordo com os testes realizados verificou-se que a técnica de *cache* de conteúdo estático (*Reverse Proxy*) não trouxe benefícios para ambas as aplicações testadas. Já o *cache* para conteúdo dinâmico (*Memcache*) resultou em um ganho de performance no *Joomla*, apenas quando o número de utilizadores a pesquisar por um mesmo conteúdo é elevado. Uma vantagem transversal às aplicações adotadas é a adoção de mecanismos de balanceamento de carga (*Load Balancer*). Estes mecanismos tornam as aplicações mais escaláveis, permitindo que um maior número de utilizadores se beneficie de tempos de resposta mais estáveis.

Em seu trabalho, Lopes (2016) destacou a importância de otimizar páginas *web* para a obtenção de alto desempenho, além da redução no tráfego de internet em redes móveis, visto que grande parte dos acessos às aplicações *web* são oriundas de usuários que possuem algum tipo de dispositivo móvel, como *smartphones* e *tablets*.

O objetivo do trabalho foi o de apresentar como melhorar algumas práticas de programação *web*. Foi exposto um cenário irregular, além de alguns sites de grande porte, visando mostrar algumas práticas aplicadas no ambiente de desenvolvimento de *software*. Além das considerações propostas, o autor teve como objetivo adquirir conhecimentos que permitissem analisar e realizar melhorias em sites. Para otimizar a performance de aplicações *web* o autor utilizou técnicas como *CDN*, compressão *GZIP* e otimização de imagens.

Ao usar a técnica *CDN*, que tem por finalidade acessar conteúdos na Internet a partir de um segmento de rede mais próximo geograficamente, ou com preferências particulares que facilite e torne o acesso mais rápido, obteve-se um ganho considerável. Os testes foram realizados acessando *sites* considerados relevantes pelo autor. Com a utilização do recurso *GZIP*, técnica para comprimir arquivos no servidor antes de enviar ao cliente, houve um ganho de 25,64% no tempo de carregamento da página em que foi realizado o teste.

Quando utilizada a técnica de otimização de imagens por meio da compactação/compressão das mesmas, observou-se uma diferença bastante significativa, que impactou positivamente o tempo de carregamento de uma página. Antes de aplicar qualquer tipo de procedimento para otimizar uma única imagem presente em uma página *web* de teste, o tamanho físico desta era de 1,05MB. Ao requisitar esta página durante os testes realizados, a mesma levou 2,99 segundos para carregar por completo. Após aplicar algumas técnicas de otimização nesta mesma imagem, o seu tamanho físico diminuiu para 155KB, e a página passou a carregar por completo em colocar apenas 1,24 segundos, o que representa uma diferença bastante considerável.

Em seu estudo, Bergantin (2014) apresentou quais as melhores práticas que se combinam nas aplicações móveis a serem desenvolvidas. O trabalho buscou estudar e analisar as boas práticas do ponto de vista dos desenvolvedores e usuários, onde uma *Web App* foi implementada, visando analisar os fatores que incluem interface com o usuário, facilidade de desenvolvimento, desempenho, infraestrutura de comunicação e manutenção. Outro objetivo foi o de avaliar os impactos de boas práticas de otimização de aplicações *web* amplamente utilizadas, uma vez que, em dispositivos móveis algumas práticas podem não ter os benefícios pretendidos, ou, podem até mesmo prejudicar o desempenho.

Entre todos os aspectos citados no estudo de Bergantin (2014), no que diz respeito a boas práticas no desenvolvimento de aplicações *Web*, é interessante destacar as técnicas utilizadas para obtenção de alto desempenho neste cenário. As técnicas foram aplicadas visando melhorar o desempenho e performance no lado do servidor *web*, que possui sua instalação sob uma estrutura com o sistema operacional *Windows 8.1 Pro*, utilizando o *webserver Internet Information Services (IIS)*.

Para otimizar a transferência de arquivos como CSS, HTML e *Javascript* entre servidor e cliente, foi utilizado o recurso GZIP, que por sua vez faz a compressão dos códigos enviados ao navegador do usuário, de forma a aumentar a velocidade do *site*. Para validar os resultados da técnica utilizada, o autor realizou testes em uma página de *login* de sua aplicação *web*. Utilizando a ferramenta do desenvolvedor no navegador *Google Chrome*, foram feitos dois testes, com e sem a compactação do GZIP. Como a tela de *login* é a primeira a ser acessada em uma aplicação, foi possível perceber uma redução considerável em seu tamanho, após utilizar a compactação. Utilizando essa técnica, observou-se uma redução de um quarto do seu tamanho, sendo 26,32% e 500,08KB, resultando em uma economia no tempo de carregamento da página.

Outra técnica utilizada para otimizar a performance da aplicação *web* em questão foi o *web cache*. Por meio desta otimização as requisições do navegador do usuário ao servidor foram reduzidas, impactando diretamente no tempo de carregamento das páginas. Para

validar os resultados o autor testou, utilizando a ferramenta de desenvolvedor disponível no navegador *Google Chrome*, como se comportaria o tempo de carregamento da página, o número de requisições e o tamanho em KB, com e sem o uso do *web cache*. Para tornar o teste ainda mais eficiente, foram utilizados três tipos de conexão de Internet: GPRS a 50 Kbps, 3G a 750Kbps e, por último, a *Wi-Fi*.

Para finalizar, foi usada outra técnica para aumentar a performance na *web*, a minificação de códigos *Javascript* e *CSS*, onde foram removidos caracteres e espaços desnecessários, reduzindo seus tamanhos, aumentando a eficiência no tempo de carregamento. Foram realizados testes utilizando a ferramenta do desenvolvedor disponível no navegador *Google Chrome*.

■ MÉTODO

A solução proposta neste trabalho envolveu a realização de um estudo de caso, que permitiu definir técnicas e ferramentas que pudessem melhorar a performance de aplicações web desenvolvidas pela empresa Tuhl Soluções. Segundo Yin (2001), os estudos de caso são uma metodologia de pesquisa adequada quando se colocam questões do tipo “como” e “por que”. Tais indagações fazem parte do objetivo geral deste trabalho, pois pretendeu-se identificar técnicas e ferramentas que pudessem ser aplicadas na empresa, visando melhorar a performance das aplicações desenvolvidas.

Para o desenvolvimento do estudo de caso foi utilizada uma aplicação *web* desenvolvida pela empresa Tuhl Soluções, mais especificamente um website de um portal de notícias. A ideia foi a de analisar projetos reais e que já estavam em produção, mediante a implementação das técnicas e ferramentas que foram estudadas para melhorar a performance da aplicação escolhida. Após a implementação das técnicas e ferramentas, foram analisados os resultados obtidos, visando validar a implementação das mesmas.

Técnicas empregadas no trabalho

Para este estudo de caso foram estudadas e implementadas, em uma aplicação *web* da empresa Tuhl Soluções, três técnicas. Os critérios utilizados para a escolha dessas técnicas e ferramentas foram com base nos estudos dos trabalhos relacionados e, também, em virtude da estrutura de aplicações *web* utilizada pela empresa Tuhl Soluções.

Uma das técnicas utilizada foi a compressão GZIP, que foi configurada por meio da ferramenta *Nginx*, no *webserver* da aplicação *web*. Essa técnica foi escolhida por ser de fácil configuração, além de trazer resultados consideráveis na otimização de performance *web* (GRIGORIK, 2020).

Outra técnica estudada e que foi aplicada no servidor da aplicação *web* escolhida, foi a otimização de imagens. Neste passo, foi utilizada uma biblioteca do *framework Python*, a *Pillow*. Por meio dessa ferramenta, todas as imagens que são enviadas ao servidor passam automaticamente por um processamento que faz a compressão em 70%, redimensionamento para 800x660px, conversão para o formato JPEG, além de gerar uma segunda versão da imagem no formato *WebP*. No processo de compressão, a biblioteca *Pillow* elimina metadados desnecessários, além de usar um método de otimização baseado em cores próximas, tudo isso sem deixar que a imagem fique com baixa resolução. Com todos esses processos é possível reduzir o tamanho físico da imagem em uma proporção bastante significativa, podendo ser superior a 80%. É importante pontuar que a imagem *WebP* gerada, apesar de otimizar a performance no acesso, visto que seu tamanho físico é menor que o JPEG que também recebe otimizações, ainda não é compatível com alguns navegadores, como *Mozilla Firefox* e o *Safari*, e por este motivo é necessário oferecer as duas versões de imagem para os usuários (PILLOW, 2020).

Por fim, foi usada a técnica de CDN, possibilitando que o acesso à aplicação *web* fosse mais rápido, visto que esse tipo de serviço replica o conteúdo da aplicação em vários servidores separados geograficamente, para que quando um usuário solicite os dados, receba a resposta do servidor mais próximo. O serviço utilizado foi o *Cloudflare*, o mesmo usado por grandes companhias como *Cisco*, *Digital Ocean* e *Udacity*, entre outras. Escolheu-se o *Cloudflare* por ser um dos mais conceituados no mercado digital, além de ser global e possuir a função de CDN em sua versão gratuita (CLOUDFLARE, 2020).

Aplicação utilizada como Estudo de Caso

A aplicação *web* que foi utilizada para aplicar as técnicas e ferramentas é um *site* no formato de portal de notícias, pertencente ao Jornal Folha do Noroeste, disponível em <<https://www.folhadonoroste.com.br>>. O número de acessos diários a este *site* é relativamente alto, em média 7.196 *pageviews* diários, e totaliza cerca de 40 mil usuários mensais. Esses dados foram coletados a partir de informações disponibilizadas pela empresa Tuhl Soluções, e foram calculados com base no intervalo de 01/03/2018 a 01/06/2018. A solução em questão foi escolhida levando em consideração o alto número de acessos, e também por conter um grande número de imagens.

Por se tratar de um projeto real e que está em produção, foi criado um ambiente idêntico ao da solução para receber estas implementações. Dessa forma foi possível fazer as implementações e testes sem que a aplicação passasse por interrupções e indisponibilidades em seu funcionamento. Esse ambiente de testes contém exatamente as mesmas características de *hardware* que o projeto real.

■ RESULTADOS E DISCUSSÃO

Para validação dos resultados obtidos nos testes aplicados, foi utilizada a ferramenta *DevTools* (Ferramenta do Desenvolvedor) disponível no navegador *Google Chrome*. Por meio das análises realizadas utilizando a ferramenta descrita, foram gerados gráficos e tabelas para comparação dos resultados. Neste processo, foram avaliados três itens que trazem informações importantes referentes à performance de uma aplicação *web*, sendo analisados: 1) o tempo de carregamento, que corresponde ao tempo que uma página leva para carregar completamente no *browser* do usuário; 2) número de requisições, que são todas as solicitações HTTP feitas ao acessar uma página e, por fim, 3) o tamanho físico da página, que é a soma de todos os arquivos carregados em uma página *web*.

Os testes foram feitos a partir de uma conexão de banda larga de 40 Mbits/s. Para obter-se uma precisão maior nos testes foi configurado no *DevTools* do navegador *Google Chrome* (versão 70.0.3538.102 64 bits) uma velocidade máxima de conexão de 5MB/s de download e 2,5MB/s de *upload*. Para cada técnica de otimização usada foram efetuados três testes utilizando a ferramenta *DevTools* e, a partir disso, realizou-se uma média destes resultados.

No lado do cliente, para realizar os testes, utilizou-se um computador *Macbook Pro* com o sistema operacional *macOS Mojave* versão 10.14 e as seguintes configurações de *hardware*: processador *Intel Core i5 2,3 GHz*, 8GB de memória RAM e placa de vídeo *Intel Iris Plus Graphics 640* com 1536 MB.

No lado do servidor, utilizou-se uma máquina com o sistema operacional *Ubuntu Server 16.04 LTS* e as seguintes configurações de *hardware*: processador Intel(R) Xeon(R) CPU E5-2630L v2 @ 2.40GHz, com 2GB de memória RAM. O servidor em questão está localizado em uma estrutura computacional estrangeira, mais especificamente na cidade de Nova Iorque, nos Estados Unidos. O motivo de não estar localizado no Brasil, é em virtude do custo benefício do serviço em nuvem utilizado pela empresa Tuhl Soluções, já que os servidores estrangeiros possuem um poder computacional superior aos serviços ofertados no Brasil. Um ponto negativo de usar uma solução internacional é a latência no acesso das aplicações, que é maior se comparada com uma aplicação hospedada em servidores nacionais.

A página usada nos testes tinha 10 imagens de alta resolução (todas acima de 3000x2500 px), sendo a menor com um tamanho de 892KB e a maior com 8,2MB. Além das imagens, outros tipos de arquivos constituem a página *web*, como códigos CSS, *JavaScript* e HTML. É importante evidenciar que as imagens correspondem ao maior conteúdo da página, em se tratando do tamanho desta página, representando 39,49MB de um total de 40,9MB.

Para organizar e validar os testes, os mesmos foram divididos em blocos, conforme serão apresentados nas próximas subseções.

Teste sem utilizar nenhuma técnica de otimização

Neste primeiro teste, todas as técnicas de otimização foram desativadas na aplicação *web*. Os resultados podem ser visualizados no Quadro 1.

Quadro 1. Teste sem utilizar nenhuma técnica de otimização

Tamanho da página (Mega Bytes)	Número de requisições	Tempo de carregamento (segundos)
40.9 MB	61	32.18s

(Fonte: dos autores, 2021)

Teste utilizando apenas GZIP

Neste segundo teste, foi utilizada apenas a técnica de otimização GZIP. Os tipos de arquivos que foram comprimidos por este recurso foram: texto, *JavaScript*, *CSS*, *SVG (Scalable Vector Graphics)*, *WebP*, *PNG* e *JPEG*. Os resultados podem ser visualizados no Quadro 2.

Quadro 2. Teste utilizando apenas GZIP

Tamanho da página (Mega Bytes)	Número de requisições	Tempo de carregamento (segundos)
40.4 MB	61	31.35s

(Fonte: dos autores, 2021)

Teste utilizando apenas CDN

Neste terceiro teste, foi utilizado apenas o recurso CDN. É importante ressaltar que a aplicação *web* está em um servidor estrangeiro, e que os testes foram feitos a partir de uma rede no Brasil. Os resultados podem ser visualizados no Quadro 3.

Quadro 3. Teste utilizando apenas CDN

Tamanho da página (Mega Bytes)	Número de requisições	Tempo de carregamento (segundos)
40.8 MB	62	16.74s

(Fonte: dos autores, 2021)

Teste utilizando apenas o Pillow para imagens JPG

Neste quarto teste, foi utilizado apenas o recurso para otimização de imagens, o *Pillow*. As imagens foram geradas apenas no formato *JPEG*. Os resultados podem ser visualizados no Quadro 4.

Quadro 4. Teste utilizando apenas o *Pillow* para imagens JPEG

Tamanho da página (Mega Bytes)	Número de requisições	Tempo de carregamento (segundos)
5.2 MB	61	3.38s

(Fonte: dos autores, 2021)

*Teste utilizando apenas o *Pillow* para imagens WebP*

Neste quinto teste, foi utilizado apenas o recurso para otimização de imagens, o *Pillow*. As imagens foram geradas apenas no formato *WebP*. Os resultados podem ser visualizados no Quadro 5.

Quadro 5. Teste utilizando apenas *Pillow* para imagens *WebP*

Tamanho da página (Mega Bytes)	Número de requisições	Tempo de carregamento (segundos)
3.9 MB	61	2.7s

(Fonte: dos autores, 2021)

*Teste utilizando todos os recursos de otimização (CDN, GZIP e *Pillow* para imagens JPG)*

Neste sexto teste, foram utilizados todos os recursos de otimização, como CDN, GZIP e *Pillow* (para imagens JPG apenas). Os resultados podem ser visualizados no Quadro 6.

Quadro 6. Teste utilizando todos os recursos de otimização (CDN, GZIP e *Pillow* para imagens JPG)

Tamanho da página (Mega Bytes)	Número de requisições	Tempo de carregamento (segundos)
5.1 MB	61	2.78s

(Fonte: dos autores, 2021)

Com a utilização do *Pillow* para a compressão e redimensionamento das imagens no formato JPEG, foi possível perceber uma redução considerável no tamanho físico das imagens. Como exemplo, destaca-se uma das imagens que passou de 3.3MB para 177KB.

*Teste utilizando todos os recursos de otimização (CDN, GZIP e *Pillow* para imagens WebP)*

Neste último teste, foram utilizados todos os recursos de otimização, como CDN, GZIP e *Pillow*, para imagens *WebP*. Os resultados podem ser visualizados no Quadro 7.

Quadro 7. Teste utilizando todos os recursos de otimização (CDN, GZIP e *Pillow* para imagens *WebP*)

Tamanho da página (Mega Bytes)	Número de requisições	Tempo de carregamento (segundos)
3.9 MB	61	2.36s

(Fonte: dos autores, 2021)

Para melhor comparar as análises, no Quadro 8 são apresentados todos os resultados obtidos nos testes realizados.

Quadro 8. Agrupamento dos resultados obtidos nos testes realizados utilizando diferentes técnicas de otimização

Tipo de técnica de otimização aplicada	Tamanho da página (Mega Bytes)	Número de requisições	Tempo de carregamento (segundos)
Nenhuma	40.9 MB	61	32.18s
GZIP	40.4 MB	61	31.35s
CDN	40.8 MB	62	16.74s
<i>Pillow</i> - JPEG	5.2 MB	61	3.38s
<i>Pillow</i> - WebP	3.9 MB	61	2.7s
Todas as técnicas (<i>Pillow</i> – JPEG)	5.1 MB	61	2.78s
Todas as técnicas (<i>Pillow</i> – WebP)	3.9 MB	61	2.36s

(Fonte: dos autores, 2021)

Entre todas as técnicas de otimização que foram utilizadas, é possível perceber que o recurso *Pillow* para imagens *WebP* foi o que trouxe maiores ganhos nos resultados, baixando o tempo de carregamento da página de 32.18 segundos para 2.7 segundos.

É interessante analisar que a técnica de CDN, quando utilizada de forma individual, representou uma redução de 47,98% no tempo de carregamento da página. Porém, quando utilizada junto com todas as técnicas (*Pillow* – *WebP*), representou um ganho pouco expressivo.

Por fim, observamos que a técnica GZIP foi a que menos representou benefícios neste cenário, pelo fato de a aplicação *web* escolhida conter pouco código *JavaScript*, *CSS* e *HTML*, que são os tipos mais relevantes para o GZIP.

De modo geral, com a utilização de todas as técnicas de otimização de forma simultânea foi possível obter ganhos bastante expressivos. O tamanho total da página baixou de 40,9MB para apenas 3,9MB, sendo que o tempo de carregamento passou de 32.18s para 2.36s. Com esses resultados, é possível analisar a importância que as otimizações trazem para os usuários que acessam aplicações *web*, além dos benefícios para a empresa que custeia financeiramente o armazenamento e processamento destes dados, tendo em vista a economia no momento de adquirir um servidor para hospedar uma aplicação.

■ CONCLUSÃO

Acredita-se que os objetivos propostos para esse estudo de caso foram alcançados de forma satisfatória. Com o abrangente estudo de técnicas e ferramentas capazes de otimizar a performance de aplicações *web*, tornou-se possível, por meio de vários testes, comprovar que houve melhorias bastante significativas no ambiente no qual foram realizados os testes.

Durante o desenvolvimento desse trabalho, foram enfrentadas dificuldades relacionadas com o ambiente de testes. Parte dessas dificuldades podem ser justificadas devido à intermitência na conexão de Internet, sendo necessário reproduzir o mesmo teste três vezes para, então, obter-se uma média dos resultados. Existiram momentos em que a rede em que os testes eram realizados se encontrava extremamente lenta, por isso os mesmos foram realizados após as 23h (UTC-3), onde a ocorrência de instabilidade é menor devido à diminuição no número de usuários online na rede do provedor de Internet.

Por meio dos testes realizados com base na aplicação *web* escolhida, foram aplicadas técnicas de otimização de performance, como ativação do compressor GZIP no *web server*, utilização de CDN e compressão de imagens com o uso do *Pillow* para arquivos JPEG e *WebP*. Os resultados demonstraram que foi possível constatar ganhos bastante significativos, reduzindo o tempo de carregamento da aplicação em uma proporção superior a 90%.

Acredita-se que este trabalho possa contribuir com a comunidade de desenvolvedores de aplicações relacionadas, tendo em vista que com as técnicas utilizadas no presente estudo de caso comprovou-se uma melhoria na eficiência destas aplicações. Muitas vezes, os profissionais da área não conhecem a importância da utilização de determinadas técnicas, o que resulta na má experiência dos usuários desse tipo de aplicação.

Para trabalhos futuros, podem ser aplicadas técnicas no lado do cliente juntamente com as que foram utilizadas no lado do servidor, como *Leverage Browser Caching*, por exemplo, que é uma técnica para armazenar arquivos no navegador do usuário, evitando que ele busque a mesma informação novamente no servidor. Dessa maneira, acredita-se que é possível aumentar ainda mais a eficiência das aplicações.

■ REFERÊNCIAS

1. BERGANTIN, M. E. C. **Análise de Boas Práticas para o Desenvolvimento de Web Wapps**. 77 p. Trabalho de Conclusão de Curso (Bacharelado em Sistemas de Informação) – Fundação de Ensino “Eurípedes Soares da Rocha” – UNIVEM, 2014.. Disponível em: <http://aberto.univem.edu.br/bitstream/handle/11077/994/Carlos%20Eduardo%20Martinelli.pdf?sequence=1&isAllowed=y>. Acesso em: 22 de jun. 2020.
2. CLOUDFLARE. **The Web Performance & Security Company**. 2020. Disponível em: <https://www.cloudflare.com/>. Acesso em: 27 de jun. 2020.
3. CUTLER, B. **Firefox & Page Load Speed: Part II**. 2010. Disponível em: <https://blog.mozilla.org/metrics/2010/04/05/firefox-page-load-speed-%E2%80%93part-ii/>. Acesso em: 07 de abr. 2020.
4. DANTAS, F. V. **WideWorkWeb: uma metodologia para o desenvolvimento de aplicações web num cenário global**. 167 p. Dissertação de Mestrado – Universidade Federal de Campina Grande, 2003. Disponível em: http://docs.computacao.ufcg.edu.br/posgraduacao/dissertacoes/2003/Dissertacao_VanessaFariasDantas.pdf. Acesso em: 07 de abr. 2020.

5. GOOGLE. Um novo formato de imagem para a Web. **Google Developers**. 2016. Disponível em: <https://developers.google.com/speed/webp/>. Acesso em: 09 de abr. 2020.
6. GOMES, J. **Extreme Web Performance – RSJS 2012**. [Slides da Palestra]. 2012. Disponível em: <https://jaydson.com/extreme-web-performance-rsjs-2012/>. Acesso em: 07 de abr. 2020.
7. GRIGORIK, I. Otimizar codificação e tamanho de ativos baseados em texto destinados a transferência. **Google Developers**. 2020. Disponível em: <https://developers.google.com/web/fundamentals/performance/optimizing-content-efficiency/optimize-encoding-and-transfer?hl-pt-br#text-compression-with-gzip>. Acesso em: 10 de abr. 2020.
8. HERTEL, R. **Melhorando o desempenho do site – Compressão GZIP**. Hostinger. 2017. Disponível em: <https://www.hostinger.com.br/tutoriais/melhorando-o-desempenho-do-site-compressao-gzip/>. Acesso em: 10 de abr. 2020.
9. KEPPELEN, G. **Performance Front-end - Front In**. Maceió. [Slides da Palestra]. 2012. Disponível em: <http://www.slideshare.net/keppelen/performance-frontend-front-inmacei>. Acesso em: 08 de abr. 2020.
10. LINDEN, G. **Make Data Useful**. [Slides da Palestra]. 2008. Disponível em: <http://www.scribd.com/doc/4970486/Make-Data-Useful-by-GregLinden-Amazoncom>. Acesso em: 07 de abr. 2020.
11. LOURENÇO, S. R. R. **Como habilitar a compressão GZIP para melhorar a velocidade do seu site**. 2017. Disponível em: <http://www.dicasmkt.com.br/blog/como-habilitar-compressao-gzip/>. Acesso em: 10 de abr. 2020.
12. LOPES, R. G. I. **Boas Práticas para Redução de Tráfego e Obtenção de Alto Desempenho de Páginas Web**. 41 p. Trabalho de Conclusão de Curso (Tecnólogo em Sistemas de Computação) – Universidade Federal Fluminense, 2016. Disponível em: https://app.uff.br/riuff/bitstream/1/5185/1/TCC_IGOR_GUILHERME_RIBEIRO_LOPES.pdf. Acesso em: 08 de abr. 2020.
13. NILSON, F. WebP, hora de brincar? **Tableless**. 2014. Disponível em: <https://tableless.com.br/webp-hora-de-brincar/>. Acesso em: 09 de abr. 2020.
14. PEREIRA, M. C. A. **Técnicas para melhoria de performance em aplicações CMS**. 71 p. Dissertação de Mestrado – Escola Superior de Tecnologia e Gestão do Instituto Politécnico do Porto, 2011. Disponível em: <http://recipp.ipp.pt/handle/10400.22/11100>. Acesso em: 06 de jun. 2020.
15. PILLOW. **Pillow (PIL Fork) 5.3.0.dev0 documentation**. Pillow – Read the Docs. 2020. Disponível em: <http://pillow.readthedocs.io/en/latest/index.html>. Acesso em: 27 de jun. 2020.
16. ROCHA, Z. **Técnicas para Melhoria de Performance em Aplicações Web no Lado do Cliente**. 41 p. Trabalho de Conclusão de Curso (Graduação em Sistemas de Informação) – Universidade Federal do Estado do Rio de Janeiro, 2013. Disponível em: <http://bsi.uniriotec.br/tcc/textos/201303Rocha.pdf>. Acesso em: 07 de abr. 2020.
17. SCHURMAN, E. **Performance Related Changes and their User Impact**. IMPACT. 2009. [Slides da Palestra] Disponível em: <http://slideplayer.com/slide/1402419/>. Acesso em: 07 de abr. 2020.
18. SOUDERS, S. **Alta Performance em Sites Web**. 19 p. Rio de Janeiro: Alta Books, 2007.
19. YAHOO. Best Practices for Speeding Up Your Web Site. **Yahoo Developer Network**. 2020. Disponível em: <https://developer.yahoo.com/performance/rules.html>. Acesso em: 22 de jun. 2020.
20. YIN, R. K. **Estudo de Caso: planejamento e métodos**. 2. ed. Porto Alegre: Bookman, 2001.

Problemas investigativos nas aulas de matemática da escola básica com o pensamento computacional: um relato de experiência mobilizador à aprendizagem

| Aline Silva de **Bona**
IFR

| Natália Bernardo **Nunes**
IFRS

| Rafaela da Silva **Bobsin**
IFRS

| Anelise Lemke **Kologeski**
IFRS

RESUMO

Diante da falta de inclusão digital no Ensino Básico, o presente trabalho relata algumas das atividades desenvolvidas em um projeto de extensão com estudantes do Ensino Fundamental como público-alvo, aplicado em parceria com instituições de ensino do Litoral Norte do Rio Grande do Sul. O objetivo das oficinas foi contextualizar a Matemática com o Pensamento Computacional, por meio da resolução de problemas investigativos, compartilhando aqui a análise qualitativa e quantitativa de 2 problemas específicos que contemplam, além da introdução a computação, também conceitos da disciplina de Matemática da Escola. A metodologia é um estudo teórico e um estudo de caso. Os resultados mostram um nítido aproveitamento dos alunos nas atividades propostas, chegando a aproximadamente 38% de melhoria, com a devida apropriação do Pensamento Computacional atrelada a resolução de problemas investigativos, intimamente relacionada com tarefas específicas da disciplina de Matemática. Desta forma, a proposta de uma metodologia inovadora se apropria da cultura digital dos estudantes e faz uso dos pilares do pensamento computacional como roteiro de prática de sala de aula.

Palavras-chave: Inclusão Digital, Educação Matemática, Problemas Investigativos, Informática na Educação.

■ INTRODUÇÃO

Desde muito tempo, é comum os estudantes apresentarem um desinteresse em atividades tradicionais de sala de aula. Com essa problemática, ocorre a redução do envolvimento para a realização das mesmas e o interesse dos professores em estudar e procurar por novas metodologias que despertem o interesse dos educandos. Em Matemática, particularmente, existem tendências e metodologias para a sala de aula, atreladas a diferentes recursos e estratégias, conforme Bona (2012), Fiorentini e Lorenzato (2009). E tal cenário se acentua com a pandemia, com o ensino remoto, conforme pesquisa realizada pelo Banco Mundial¹ em 2021, apresentando o Impacto da Pandemia na Educação no Brasil quanto a quantidade de estudantes com o nível de aprendizado abaixo do mínimo: antes da Pandemia era de 50%; sete meses sem escola 59%, treze meses sem escola 70%.

A resolução de problemas de Matemática, que é um cenário usual e natural nessa ciência, inclusive uma tendência na Educação Matemática, está se relacionando cada vez mais a diferentes recursos tecnológicos. Ademais, vem percebendo-se que o essencial não é o recurso, segundo Bona (2016), mas a forma e o meio escolhido pelo professor para explorar a resolução dos problemas, como é o caso dos problemas investigativos a serem apresentados neste trabalho e também no evento científico denominado Simpósio Brasileiro de Informática na Educação - SBIE, em 2020, por Bobsin, Nunes, Kologeski, Bona (2020). Eles proporcionam uma atividade lúdica, diversificada e interessante aos estudantes, e paralelamente permitem a resolução coletiva, com seus pares, permitindo aos estudantes que pensem juntos (ora colaborativa e ora cooperativamente), fazendo, compreendendo e trocando ideias, escrevendo ou registrando apenas uma solução, ou cada um da sua maneira, conforme Bona (2012).

Mais do que os problemas investigativos, um elemento atrativo aos estudantes são as sequências lógicas para os problemas que estão presentes na realidade de cada um, como por exemplo: o passo a passo para construir uma referência a uma pesquisa da escola, ou compreender como pagar uma conta digital em um site seguro, gerando o comprovante, inclusive. Tal sequência é muito utilizada na Matemática, assim como na Informática, como defende Bona (2012), e então se percebe o quanto os pilares do Pensamento Computacional organizam uma metodologia de trabalho, e está presente nas etapas da resoluções de problemas investigativos, e que podem ser gradualmente exploradas, a partir de uma atividade simples até a criação de outros itens e critérios criados e verificados pelos estudantes diante da realidade do problema proposto.

¹ Fonte: <https://g1.globo.com/globonews/estudio-i/video/pandemia-provocou-crise-sem-precedentes-na-educacao-brasileira-aponta-banco-mundial-9357742.ghtml>. Acesso: 21 março de 2021

Mais um elemento importante é a linguagem, pois cada estudante organiza uma forma de escrever sua resolução, e que nem sempre faz uso da linguagem Matemática, mas usa uma organização de pensamento que é muito presente e importante para a Informática, como ao construir um algoritmo, por exemplo, adotando posteriormente uma linguagem de programação específica. Os estudantes perceberem que existe mais de uma maneira de resolver um problema, é como encontrar dois meios para achar uma solução, e cada uma terá critérios, e otimizações, que para a informática é essencial a criação de algoritmos.

O trabalho de Araújo et. al. (2018) destaca o quanto é atrativo aos estudantes compreender de forma lúdica, por meio de jogos, o Pensamento Computacional, e que este contempla a necessidade de compreensão de uma sequência lógica, bem como a organização dos pensamentos em blocos, sendo que nessa ação está implícita a abstração (no sentido de otimizar processos, que difere da abstração do processo de aprendizagem, citada por Bona (2012)), a automação e a análise do seu próprio pensamento através da resolução proposta.

Outro trabalho que colabora com a ideia de explorar problemas do tipo que o estudante possa criar suas hipóteses, testar, reconstruir, e novamente tentar validar com ajustes, segundo seus conhecimentos, como se o estudante fosse um cientista, para a aprendizagem de Matemática, é o trabalho de Leitão e Castro (2018), com o uso da plataforma *Scratch*.

Ambos os trabalhos contemplam a importância de que os problemas atrelados a uma lógica, baseada em tecnologias, fazem parte da vida cotidiana, e que se faz necessário aprender a resolver, além disso, pois a escola não deve delimitar um tipo de problema para explorar um conceito ou a criatividade dos estudantes, e sim deve propor um problema investigativa “aberto” que o estudante possa criar hipótese e delinear critérios para daí definir, encontrar uma solução para sua interpretação, segundo Bona (2016).

Com isso, este trabalho tem como pergunta de pesquisa: Como atrelar o Pensamento Computacional na Escola Básica através dos problemas investigativos que contemplem a Matemática? E objetiva explorar conceitos de Matemática e recursos tecnológicos utilizados na vida cotidiana, através de problemas investigativos de Matemática montados de forma a serem atrativos, investigando o potencial do Pensamento Computacional nas aulas de Matemática da Escola Básica.

Cabe destacar que a Base Nacional Comum Curricular - BNCC (2018) faz referência ao Pensamento Computacional como um componente da Matemática, o que enaltece a importância de se promover formação docente neste sentido aos professores, e cita-se das áreas das exatas, num primeiro momento, para depois expandir aos demais professores, para fins de criar sequência diante dos problemas ancorados em contextos, que por natureza vão abordar mais de uma disciplina da Escola Básica. (BONA, BOBSIN, KOLOGESKI, 2020).

Para tornar viável tal aplicação, um projeto de extensão foi criado com o propósito de estimular e promover o Pensamento Computacional aliado aos problemas investigativos da Matemática, que conta com a participação majoritária de alunos de 9º ano do Ensino Fundamental. Para isso, atividades lúdicas foram planejadas e aplicadas junto ao uso de ferramentas da Tecnologia da Informação e Comunicação, em oficinas, levando a inclusão digital para mais perto dos participantes, oportunizando um momento de estímulo ao desenvolvimento, aprendizado, criatividade e autonomia.

■ RELATO DE CASO

As oficinas propostas visam estimular o desenvolvimento do Pensamento Computacional, trabalhando com aspectos da Matemática e da Programação Básica, além da interpretação de texto e do raciocínio lógico, contemplando a metodologia de Aprendizagem Baseada na Resolução de Problemas, mencionada por Bona (2016) e Schuhmacher *et al.* (2016). As atividades de Pensamento Computacional são de 2 tipos:

- Desplugadas: utilizam apenas atividades concretas sem o uso do computador. Normalmente, ela ocorre na própria escola, especialmente se não há um laboratório de informática disponível. São exemplos de atividades realizadas os jogos “Estacionamento Algorítmico” e a “Programação com Papel Quadriculado”, sendo esta última relacionada ao Problema Investigativo 1 apresentado na Figura 1. Em ambas as atividades, os participantes precisam desenvolver ou seguir algoritmos para realizar o que é solicitado.
- Plugadas: as atividades fazem uso de recursos digitais, ou seja, necessitam do uso de computadores. Diferentes jogos como “A Fazendeira” e o “Labirinto Clássico” são apresentados aos alunos, utilizando a programação em blocos para a devida conclusão das tarefas, sendo o Problema Investigativo 2 oriundo deste último, conforme apresentado na Figura 2. Os jogos apresentam diferentes níveis de dificuldade, permitindo uma evolução gradual ao aluno. Normalmente, essas oficinas ocorrem nas dependências da nossa instituição, oportunizando aos participantes conhecerem as atividades e os espaços oferecidos por ela, além das tarefas realizadas durante a execução da oficina.

Essa divisão das oficinas acontece para atender as necessidades específicas de cada escola, já que nem todas possuem laboratórios de informática ou disponibilidade de deslocamento até a nossa instituição, permitindo assim que elas optem pela realização de apenas uma das oficinas, caso necessário. Cada oficina tem a duração de até 4 horas,

conforme disponibilidade das escolas, e é aplicada por pelo menos 2 membros da equipe de execução do projeto (composta por 6 alunos bolsistas, 2 professoras e um técnico em Tecnologia da Informação).

Agrega-se a pesquisa, que tais atividades foram utilizadas em cursos de formação docente ministrados na modalidade extensão, através do *Google Meet*, por uma das professoras-autoras deste texto, em novembro de 2020, e fevereiro de 2021, com grande interesse pelos professores de matemática e áreas correlatas. Foram atendidos 62 professores de escolas públicas do Litoral Norte do Rio Grande do Sul, em novembro de 2020, e 97 professores, em fevereiro de 2021, sendo majoritariamente da matemática, depois 12 ciências, 9 educação artística, e 10 diversas. Todos os professores acharam muito interessantes as atividades, difíceis num primeiro olhar, depois lógicas e atrativas, e apontaram que devido a elevada carga horária de sala de aula não conseguem se apropriar do conceito de Pensamento Computacional para criar atividades, mas usariam em suas salas de aula, em particular no ensino remoto, se tivessem a sua disposição para adaptação curricular apenas.

As oficinas destinadas aos estudantes sempre são precedidas por um Pré-Teste, que avalia o conhecimento prévio dos participantes, e finalizadas com um Pós-Teste, que serve para comparar se houve alguma evolução no aprendizado deles após a realização das tarefas propostas. Ambos os testes utilizados em cada oficina são compostos por 5 questões, planejadas de acordo com as atividades aplicadas, e são idênticos entre o Pré e o Pós-Teste de cada oficina, para que os alunos possam observar seus erros respondendo corretamente as questões ao final da oficina. Os Problemas Investigativos 1 e 2 analisados neste trabalho são, respectivamente, questões trabalhadas nos Testes das oficinas Desplugadas e Plugadas, e são discutidos mais detalhadamente a seguir.

Contexto dos problemas escolhidos

Inicialmente, as oficinas foram idealizadas como uma forma lúdica de levar a inclusão digital aos estudantes de escolas de Ensino Fundamental. Por isso, foram escolhidos diversos problemas investigativos para serem trabalhados em sala de aula, de forma dinâmica e autônoma, em grupos, e que fizessem os alunos refletirem sobre a importância da compreensão de determinados conceitos da Computação atrelados ao Pensamento Computacional, que estão presentes diariamente em nossas mais diversas atividades do cotidiano. Por isso, o Problema Investigativo 1, que consta na Figura 1, foi planejado para apresentar aos estudantes, de forma desplugada, a ideia de seguir instruções bem definidas para a obtenção de um objetivo, como em um algoritmo, enquanto que o Problema Investigativo 2, que consta na Figura 2, traz o exemplo de jogo digital cujo as instruções precisam ser devidamente organizadas para a conclusão da fase do jogo.

No entanto, o leitor pode se perguntar: qual a relação disso com a vida cotidiana? E a resposta é simples: passos são seguidos a todo momento por quem segue uma receita de bolo, as orientações de localização de um GPS ou por quem organiza tarefas em uma empresa. Sendo assim, o objetivo das atividades realizadas é que o estudante se dê conta de que o Pensamento Computacional nada mais é do que organizar devidamente as tarefas do cotidiano, para sua execução, sem o uso de recursos digitais necessariamente. Outros diversos problemas investigativos foram também utilizados durante as oficinas. Mas por questões de limitação de espaço, decidiu-se apresentar apenas dois deles neste trabalho, para uma reflexão mais aprofundada sobre eles conforme segue, onde são atrelados ao uso da matemática em nosso contexto diário.

Figura 1. Problema Investigativo 1 baseado na Plataforma *code.org*.

PROBLEMA INVESTIGATIVO 1. Considere os 4 comandos e a grade 4x4 apresentados abaixo. Qual a sequência de comandos correta para que a segunda coluna da grade fique completamente preenchida, de acordo com a programação no papel quadriculado? Considere o início da atividade no primeiro quadrado superior da esquerda.

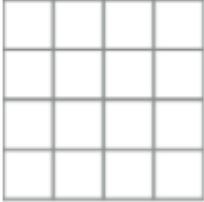
 Mover para a direita.
  Mover para baixo.

 Mover para a esquerda.
  Preencher quadrado.

a)        

b)        

c)        



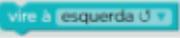
Fonte: Autoria própria (2021).

Figura 2. Problema Investigativo 2 baseado na Plataforma *code.org*

PROBLEMA INVESTIGATIVO 2. Levando em consideração a imagem à esquerda, qual bloco de comando que está faltando e deve ser introduzido na imagem à direita?



a) 

b) 

c) 

Fonte: Autoria própria (2021).

Análise do Problema Investigativo 1 (Oficina Desplugada)

Para solucionar o Problema Investigativo 1 é preciso observar, na Figura 1, a quantidade exata de comandos necessários para que a segunda coluna seja pintada, conforme as instruções fornecidas. Assim, o aluno precisa interpretar de forma matemática o caso, para saber a quantidade de instruções necessárias para executar até concluir o que é solicitado (preencher a segunda coluna da grade). Junto a isso, ele precisa compreender que o ponto de início é o mesmo para todos: o primeiro quadrado superior a esquerda. Sendo assim, a ordem dos comandos também importa para a obtenção do resultado.

Desta forma, fica evidente a relação da resolução do problema com a sua generalização, pois não há nada concreto para apegar-se, exceto por uma figura no papel, e assim entender a localização, os passos, o sentido e a sequência de ações necessárias, evidenciando a necessidade de compreensão e interpretação de um algoritmo, para que o problema permita formalizar a compreensão do Pensamento Computacional, organizado em passos bem definidos.

Matematicamente pode-se resolver por matrizes, dando a resposta em termos que relacionam linha e coluna, e conseqüentemente daí surge a ação, como fazer o passo a passo para a resolução do problema, exigindo abstração (quanto ao processo de aprendizagem, o quanto cada leitura, e testagem, e observação em detalhes proporciona ao estudante um “andar” contínuo na sua lógica de pensar, como se fossem patamares, segundo Bona (2016)) e raciocínio do estudante, pois este precisa exatidão na sequência e saber onde começar.

Diante das opções de resposta apresentadas para o aluno, duas saem completamente fora do escopo esperado, pois uma delas permanece na primeira linha da grade, enquanto a outra permanece na primeira coluna. Para que o aluno se encaminhe para a resposta correta, ele precisa identificar que há um passo necessário para atingir a segunda coluna, através do comando de movimentação dentro da linha, para a esquerda. Somente aí é que ele poderá pensar em como conciliar as instruções para preencher os quadrados e movimentar-se dentro da coluna, utilizando a matemática sem precisar da realização de muito cálculo, fazendo uso de uma operação aritmética simples, em vez de apoiar-se no complexo uso da Álgebra ou da Geometria, por exemplo.

Para contextualizar a validação do Problema Investigativo 1, uma análise gradativa com estudantes participantes das oficinas é realizada, por meio dos depoimentos dos mesmos. Como exemplificação, abaixo constam os depoimentos de dois alunos de Ensino Fundamental, que foram convidados para realização de uma resolução qualitativa:

- Análise do aluno 1 (8º ano): *“Seguindo o enunciado, começo pelo primeiro quadrado superior. Preciso avançar o primeiro quadrado, à direita, e em seguida pintar,*

depois descer e pintar consecutivamente, até chegar ao último quadrado. Por isso é a opção 'b'.”

- Análise do aluno 2 (9º ano): *“Preciso de 8 ações para concluir, e essas ações precisam mudar de linha, não todas junto. Então encontro a resposta por exclusão. Se tiver de justificar, começando de baixo para cima ou de cima para baixo, ou pelos lados, temos que dobrar, pintar, descer, pintar, e assim sucessivamente mais 2 vezes para os dois últimos passos, totalizando 8 passos. Isso pensando no mínimo de passos, pois se puder ter mais passos, podemos ter muitos caminhos, que nem aparecem nas alternativas.”*

Análise do Problema Investigativo 2 (Oficina Plugada)

Espera-se que o estudante observe, na Figura 2, que não precisa saber a quantidade certa de passos necessários para o Zumbi atingir a Flor. Porém ele precisa perceber que existe um teste (“se houver caminho à”) sendo realizado a cada passo efetuado, podendo então mudar o rumo do Zumbi. Tais elementos tornam o processo de resolução mais abstrato, porque o estudante precisa imaginar os movimentos que irão resolver o problema, sem quantificar, explorando a lógica sequencial, e também o sistema posicional (direção, sentido, ponto referencial) como o plano cartesiano, conforme achar necessário.

Por exemplo, se desejar, o aluno pode pensar no Zumbi, que está na posição (0,0), andando no eixo X. Então ele anda (A,0), vira 90° para a esquerda (A,B), sendo que no eixo de A ele ficou parado na horizontal, e apenas subiu em B para chegar na Flor, exigindo um deslocamento no eixo Y (A,C). Assim a ideia do problema é a lógica, e não necessariamente o cálculo da distância que foi percorrida, por exemplo, fazendo uso da lógica matemática e sequencial, construindo habilidades para o Pensamento Computacional através do uso de funções de recursos, exigindo abstração.

Nesse problema, fica evidente os passos da resolução de um problema investigativo, pois há uma exploração inicial, seguida pela formulação de hipóteses, pelos testes das hipóteses e pela validação, quando o aluno encontra de fato a alternativa com lógica. Paralelamente, relacionando ao Pensamento Computacional, se o estudante não compreende a imagem de comandos e sua inserção hierárquica de execução, ele não consegue nem criar a hipótese, então ora a matemática está apoiando o Pensamento Computacional e vice-versa, embora tudo dependa sempre do referencial do estudante quanto ao processo de abstração para a devida resolução do problema.

Na validação do Problema Investigativo 2, abaixo constam, como exemplos, os depoimentos dos mesmos dois estudantes mencionados na seção 4.2:

- Análise do aluno 1 (8º ano): “Como o comando manda avançar até ter um caminho, e se não tiver ele irá virar à esquerda, o próximo comando é avançar. Opção ‘a’.”
- Análise do aluno 2 (9º ano): “Tenho que andar dois passos de um quadrado, daí virar à esquerda se estiver de frente inicialmente, e ando mais 8 passos para chegar na flor. E logicamente preciso virar à esquerda. O caminho é muito curto para pensar em repetição, mas entendo a função pois ela aparece num jogo do celular, e é importante pois executa mais rápido.”

Análise dos Dados

Desde o ano de 2017 o projeto já atendeu mais de 440 alunos da rede pública de ensino. Contudo, decidimos apresentar um estudo de caso com as turmas do ano de 2019, que trabalharam efetivamente com os problemas investigativos apresentados, totalizando 237 alunos, majoritariamente do 9º ano do Ensino Fundamental de escolas públicas municipais e estaduais, dos quais 169 deles participaram das Oficinas Desplugadas, e 169 das Oficinas Plugadas, sendo que alguns destes participantes tiveram a oportunidade de participação nas duas oficinas. Com isso, pretende-se detalhar e apontar reflexões sobre as questões abordadas. Diversos problemas similares foram apresentados durante as oficinas, mas por questões de limitação, apenas 2 deles foram escolhidos para serem apresentados neste trabalho. Na Tabela 1 apresenta-se quantitativamente o percentual dos estudantes que acertaram os problemas antes e depois das oficinas, bem como a amostragem considerada.

Tabela 1. Resultados obtidos no Pré e no Pós-Teste das oficinas.

	Acertos Pré	Acertos Pós	Amostragem
Problema Investigativo 1 - Oficina Desplugada	63,31%	83,43%	169 alunos
Problema Investigativo 2 - Oficina Plugada	17,16%	55,03%	169 alunos

Fonte: Autoria própria (2021).

O Problema Investigativo 1 aborda uma situação mais próxima do real e do concreto, e é possível perceber que os alunos conseguem desenvolvê-lo mais facilmente, melhorando em até 19,88% a compreensão e resolução dele após a realização da oficina, sendo que o resultado obtido no Pré-Teste já é bastante satisfatório (63,31% de acertos). Já o Problema Investigativo 2 relaciona conceitos mais abstratos, que fazem uso de Programação Básica, exigindo uma maior capacidade de compreensão do estudante. Por isso, no Pré-Teste, o resultado obtido foi de apenas 17,16% de acertos, pois foi o primeiro contato de muitos estudantes com a organização de comandos e instruções em formato de um algoritmo específico de programação.

Após o término da oficina, observou-se um progresso nos resultados de até 37,87%, mostrando assim que o nível de dificuldade e abstração necessário foi satisfatoriamente atingido por muitos alunos, aumentando consideravelmente para a devida compreensão e resolução da atividade após a realização da oficina oferecida. Contudo, para elevar o número de acertos além de 55,03%, acreditamos na necessidade de mais encontros e mais horas de atividades, para que todos os educandos possam atingir o mesmo nível de rendimento nas oficinas oferecidas, já que alguns alunos apresentaram maiores dificuldades no desenvolvimento das atividades propostas.

■ DISCUSSÃO

Problemas Investigativos na matemática

Fazer uso de problemas em matemática que sigam uma metodologia, ou seja, contextualizados, aproxima a disciplina das rotinas da vida pessoal e profissional. Essa metodologia, quando atrelada à ação de investigar, leva em consideração os conhecimentos dos estudantes e suas percepções, hipóteses e compreensões sobre o problema a ser resolvido, além da valorização implícita da cultura digital da atual geração dos estudantes. Se prioriza a autonomia dos estudantes, porque para validar suas hipóteses, será necessário pesquisar e encontrar meios e formas para tal e saber argumentar com os colegas para ampla compreensão. Nessa cadeia de pensamentos e do processo de fazer e compreender, o estudante desenvolve habilidades sequenciais e lógicas, além dos conceitos necessários para solucionar o problema (BONA, 2016).

A ação de investigar, para Ponte, Brocardo e Oliveira (2006), é procurar conhecer o que não se sabe, é descobrir relações entre objetos Matemáticos conhecidos ou desconhecidos de maneira a procurar identificar suas propriedades e o processo de criação, e contempla quatro passos: (1) identificar o problema a resolver; (2) formular hipóteses; (3) realizar testes; e (4) argumentar/validar/provar, seja sozinho ou em grupo. Ainda, segundo Bona e Souza (2015), a ação de investigar implica em compreender e buscar soluções justificadas para os problemas, e então contemplam elementos como: mobilizar os estudantes a aprender; a aprendizagem decorre da ação dos estudantes; a prática docente baseada no diálogo e criatividade dos estudantes, pelo fato único de que cada estudante terá uma lógica de pensamento e assim uma solução, e compreender essa multiplicidade de soluções é potencializar a capacidade de aprender.

Aponta-se ainda que os itens citados assim como passos para resolver um problema de matemática está intimamente relacionado com os pilares do pensamento computacional

- decomposição, reconhecimento de padrões, abstração e algoritmo, segundo Bona, Bobsin, Kologeski (2020), e Wing (2010).

Ao resolver um problema, se faz uso de representações, sejam escritas, desenhadas ou gráficas. Assim como a matemática tem uma forma de se comunicar com sua linguagem simbólica, a Informática também tem, segundo os Parâmetros Curriculares Nacionais de matemática (PCN, 1999), e ao se pensar em Pensamento Computacional implicitamente se contempla uma ou mais formas de representar a mesma solução do problema, devido a diferentes pessoas que solucionam, e também aos diferentes recursos tecnológicos que podem ser explorados.

Cabe destacar que a ação investigativa atrelada às tecnologias digitais proporciona aos estudantes uma interatividade (BONA, 2016), e que segundo Bona (2012) maximiza o processo de aprendizagem pelo fato das interações serem além de colaborativas, mas cooperativas, no momento em que se realiza o passo a passo em conjunto, e então, a sequência lógica da resolução de problemas investigativos de matemática se aproxima do Pensamento Computacional, segundo a variedade de contextos que se fazem necessários (FIORENTINI e LORENZADO, 2007), (D'AMBROSIO, 1996).

Pensamento Computacional

Para Silva (2019) a matemática contempla o Pensamento Computacional, que por sua vez, segundo Wing (2010), é compreendido como o método, a maneira, a forma de pensar presente na abstração/lógica de um problema, bem como suas possíveis resoluções de modo que estas sejam representadas de uma forma que pode ser testada, executada, logicamente por um processador, ancorada em artefatos computacionais.

Considerando a disciplina de matemática na Escola Básica, o objetivo torna-se desenvolver habilidades como lógica sequencial, a construção de ideias generalizadas, como os algoritmos das operações com números reais, interpretar e representar, além das construções, segundo o PCN (PCN, 1999) e a Base Nacional Comum Curricular (BNCC, 2018). Com isso, torna-se evidente que a fusão entre o Pensamento Computacional e a matemática não se trata de qualquer atividade de matemática, mas de atividades do tipo problemas investigativos, que podem contemplar (mas não apenas): lista de problemas, projetos interdisciplinares, projetos de pesquisa, demonstrações, aplicações em outras áreas do conhecimento, como jogos ou atividades onde segue-se um padrão, onde o estudante possa se identificar, explorar e propor uma solução por ele testada e válida.

Atividades do tipo “siga o modelo”, que expressam uma mera cópia e pouca construção dos estudantes, com pouca abstração, não contemplam a proposta deste trabalho e não

permitem atrelar a autonomia, a aprendizagem e o desenvolvimento do aluno aos recursos propostos pela relação entre o Pensamento Computacional e a matemática.

Relatos de autores comprovam que a relação entre o Pensamento Computacional e a matemática está interligada, e pode facilitar a atuação docente no desenvolvimento das atividades, pois desenvolve a facilidade e o domínio em sua aplicação, bem como a integração das atividades com o currículo existente e as vivências cotidianas dos estudantes (MARQUES *et. al.* 2017). Riboldi e Reichert (2019) mostram como o uso da plataforma *Scratch* pode auxiliar no ensino de funções matemáticas e no uso do Pensamento Computacional em sala de aula, proporcionando um momento mais dinâmico e interessante aos alunos. Os autores mostram como o uso dessa ferramenta vai de encontro a BNCC (BNCC, 2018) e como o entendimento dos participantes sobre os conceitos abordados melhorou significativamente ao fim das atividades realizadas.

Já no trabalho de Leitão e Castro (2018), é também relatado o uso do *Scratch*, mas em uma oficina para docentes e futuros profissionais que serve para desenvolver novos conceitos a serem utilizados nas aulas de matemática, desenvolvendo um raciocínio a fim de passar o sistema de ensino instrucionista para um de aprendizagem autoral com esta plataforma, já que, segundo os resultados apresentados pelo autor, a cada tentativa, maior é o desenvolvimento de habilidades com o Pensamento Computacional, que por sua vez auxilia na resolução de problemas da disciplina de matemática.

Contudo, o Pensamento Computacional é, por enquanto, um conteúdo escolar não foi formalizado como disciplina dentro dos currículos da Escola Básica, embora esse processo de curricularização esteja se tornando alvo de discussões entre os membros da comunidade brasileira de Informática na Educação. Por isso, como já dito anteriormente, a ideia deste trabalho consiste em atrelar os problemas investigativos de matemática ao Pensamento Computacional para que isso seja viável dentro das escolas. E dada a sequência nas pesquisas, agora com professores, desde 2020, também, apresenta-se bons resultados quanto ao interesse como citados anteriormente, de forma cada vez mais adaptada à modalidade ensino remoto, pois o ensino híbrido, de acordo com a pesquisa cita acima do Banco Mundial, está muito distante da realidade brasileira. E o resultado do interesse dos professores em cursos com a temática proposta pelo trabalho aqui relatado já é um primeiro passo para a mudança da metodologia em sala de aula em prol de cada vez mais mobilizar a aprendizagem e a importância da Escola, do estar na Escola, do Aprender com os Colegas, do pensar sobre problemas investigativos planejados com muitos conceitos curriculares fundamentais à vida pessoal e profissional de todo cidadão.

■ REFERÊNCIAS

1. ARAÚJO, L.; SILVEIRA, H. U. C. ; MATTOS, M., 2018. Ensino do pensamento computacional em escola pública por meio de uma plataforma lúdica. Anais dos Workshops do Congresso Brasileiro de Informática na Educação, p. 589. ISSN 2316-8889. Disponível em: <<https://www.br-ie.org/pub/index.php/wcbie/article/view/8284>>. Acesso em: 25 mar. 21. doi:<http://dx.doi.org/10.5753/cbie.wcbie.2018.589>.
2. BONA, A. S., 2012. Espaço de Aprendizagem Digital da Matemática: o aprender a aprender por cooperação. Tese (Doutorado). Programa de Pós-Graduação em Informática na Educação. Porto Alegre: UFRGS.
3. BONA, A. S., 2016. Aulas Investigativas e a construção de conceitos de matemática: um estudo a partir da teoria de Piaget. Curitiba: CRV.
4. BONA, A. S.; SOUZA, M. T. C. C., 2015. Aulas investigativas e a construção de conceitos de matemática: um estudo a partir da teoria de Piaget. Revista Psicologia USP, v.26, n.2, p. 240-248. Disponível em: <<https://www.scielo.br/pdf/pusp/v26n2/0103-6564-pusp-26-02-00240.pdf>>. Acesso em: 25 mar. 21.
5. BONA, A. S. D.; BOBSIN, R. S; KOLOGESKI, A. (2020) Contextualizando a Matemática em Oficinas com o Pensamento Computacional. In: FREITAS, Patricia Gonçalves de; MELOO, Roger Goulart. Educação em foco: tecnologia digital e inovação em práticas de ensino. Rio de Janeiro, RJ: e-publicar. <https://storage.googleapis.com/production-hostgator-brasil-v1-0-2/102/248102/ZJ2LQxgL/f6cea92d7433494bb6c6a5c33d88e49a?fileName=TECNOLOGIAS%20DIGITAIS%20E%20INOVACAO.pdf>. Março 2020
6. BOBSIN, R. S. NUNES, N. B. KOLOGESKI, A. BONA, A. S. D. O Pensamento Computacional presente na Resolução do Problemas Investigativos de Matemática na Escola Básica . Anais XXXI do Simpósio Brasileiro de Informática na Educação - SBIE - Trilha 5, Online, 2020. Disponível em: <https://sol.sbc.org.br/index.php/sbie/article/view/12903/12757>. Acesso em: 27 mar.21
7. BRASIL, 2018. Ministério da Educação; Secretaria de Educação Básica; Conselho Nacional de Educação. Base Nacional Comum Curricular (BNCC): educação é a base. Brasília: MEC; SEB; CNE; 2018. Disponível em: <<http://basenacionalcomum.mec.gov.br/a-base>>. Acesso em: 25 mar. 21.
8. BRASIL, 1999. Ministério da Educação; Secretaria de Educação Básica; Conselho Nacional de Educação. Parâmetros Nacionais da Educação (PCN): Ensino Médio. Brasília: MEC; SEMTEC; 4v; 1999.
9. D´AMBROSIO, U., 2016. Educação Matemática: da teoria a práxis. Coleção Perspectivas em Educação Matemática. Campinas, SP: Papirus.
10. FIORENTINI, D.; LORENZATO, S., 2017. Investigação em educação matemática: percursos teóricos e metodológicos. Campinas: Autores Associados.
11. LEITÃO, D.; CASTRO, J. B., 2018. Construção de Recursos Digitais de Matemática: uma experiência de autoria com o Scratch. Anais dos Workshops do Congresso Brasileiro de Informática na Educação, p. 510. ISSN 2316-8889. Disponível em: <<https://www.br-ie.org/pub/index.php/wcbie/article/view/8276>>. Acesso em: 25 mar. 21. doi:<http://dx.doi.org/10.5753/cbie.wcbie.2018.510>.

12. MARQUES, M.; CAVALHEIRO, S.; FOSS, L.; AVILA, C.; BORDINI, A., 2017. Uma Proposta para o Desenvolvimento do Pensamento Computacional Integrado ao Ensino de Matemática. Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação - SBIE), p. 314. ISSN 2316-6533. Disponível em: <<https://br-ie.org/pub/index.php/sbie/article/view/7560/5356>>. Acesso em: 25 mar. 21. doi:<http://dx.doi.org/10.5753/cbie.sbie.2017.314>.
13. PONTE, J. P.; BROCARD, J.; OLIVEIRA, H., 2006. Investigações matemáticas na sala de aula. Belo Horizonte, MG: Autêntica.
14. RIBOLDI, S. M. O.; REICHERT, J. T., 2019. A linguagem de programação Scratch e o ensino de funções: uma possibilidade. Anais do XXV Workshop de Informática na Educação (WIE 2019). Disponível em: <<https://br-ie.org/pub/index.php/wie/article/view/8497>>. Acesso em: 25 mar. 21.
15. SCHUHMACHER, E.; ROPELATO, D.; SCHUHMACHER, V., 2016. O Desenvolvimento do Pensamento Computacional no Ensino Médio por meio de Ambientes de Programação. XIV International Conference on Engineering and Technology Education. Salvador/Brasil, p. 239-243.
16. SILVA, L. C. L., 2019. A relação do Pensamento Computacional com o ensino de Matemática na Educação Básica. Dissertação (mestrado) - Universidade Estadual Paulista (Unesp), Faculdade de Ciências e Tecnologia, Presidente Prudente. Disponível em:< https://repositorio.unesp.br/bitstream/handle/11449/191251/silva_lcl_me_sjrp.pdf?sequence=5&isAllowed=y>. Acesso em: 25 mar. 21.
17. WING, J. M., 2010. Computational thinking: what and why?. Disponível em: <http://www.cs.cmu.edu/~CompThink/resources/TheLinkWing.pdf>. Acesso em: 25 mar. 21.

Pyralelo: um sistema para suporte à programação visual de aplicações paralelas de alto desempenho

| André Ricardo Dantas Bezerra **Landim**
UFCG

| Guilherme Álvaro Rodrigues Maia
Esmeraldo
IFCE

RESUMO

A computação paralela é uma técnica que busca dividir as tarefas entre os processadores para aumento de desempenho do sistema. Contudo, desenvolver novas aplicações paralelas não é uma tarefa simples, pois deve-se identificar os pontos de paralelismo da aplicação e expressá-los em linguagens de programação lineares. Este artigo apresenta uma nova ferramenta que permite a modelagem visual e, com isso, é possível abstrair o paralelismo da aplicação, além de gerar automaticamente o código necessário à paralelização. Resultados preliminares mostram que, com a utilização da ferramenta proposta, é possível aumentar a produtividade no desenvolvimento de novas aplicações paralelas de alto desempenho.

Palavras-chave: Alto Desempenho, Aumento de Produtividade, Computação Paralela, Geração de Código, Modelagem Visual.

■ INTRODUÇÃO

Com o desenvolvimento das tecnologias de integração de circuitos, foi possível reduzir o tamanho do transistor e, com isso, os processadores passaram a ficar mais complexos (BOUJILA et al., 2011). Atualmente, é muito comum encontrar processadores com 2 ou mais núcleos para processamento paralelo. Nesse contexto, surgiu então a programação paralela, a qual é uma técnica que se caracteriza pela divisão de tarefas entre vários processadores, obtendo-se, desta forma, um ganho de desempenho na sua execução (GEBALI, 2011).

Contudo, criar aplicações paralelas não é uma tarefa simples e exige certo nível de experiência dos programadores (MATTSON, SANDERS E MASSINGILL, 2005). Assim, diversas técnicas foram criadas para auxiliar o desenvolvimento de novas aplicações paralelas. A maioria baseia-se na técnica de Dividir para Conquistar, onde o princípio básico é dividir o problema em subproblemas e solucioná-los individualmente, agrupar as soluções, até compor a solução do problema (CORMEN et al., 2012). Baseadas nela, estão as técnicas de divisão funcional (paralelismo de tarefas) e de divisão de domínio (paralelismo de dados) (PACHECO, 2011). Na divisão funcional, também conhecida como Decomposição em Pipeline, o programa paralelo tem suas tarefas divididas entre processos para resolver o problema de forma colaborativa. Na divisão por domínio, ou Decomposição de Dados, os dados do problema são divididos entre processos, de forma que eles realizam a mesma quantidade de operações nos seus respectivos conjuntos de dados. No final, os resultados são correlacionados de forma que façam sentido para a solução do problema (PALACH, 2014).

Mattson, Sander e Massingill (2005) citam que o desenvolvimento de uma nova aplicação paralela se resume em são quatro as fases: exposição dos pontos de concorrência no problema, estruturação do algoritmo, projeto do programa paralelo e implementação do programa paralelo, utilizando ambientes de programação paralela, como o MPI (GABRIEL et al., 2004). Contudo, o suporte desses ambientes é inerente ao uso de uma linguagem de programação e, segundo Lee e Webber (2003), há dois problemas com essa abordagem: o primeiro refere-se à linearidade do código fonte, não permitindo uma representação intuitiva dos fluxos paralelos de controle; e complexidade de paralelismo, onde o programador deve identificar e controlar o paralelismo, o compartilhamento de dados, protegendo o acesso aos recursos compartilhados e garantir que as seções críticas do código sejam devidamente delimitadas. Lee e Webber (2003) fazem um apelo ao uso de “linguagens visuais”, pois permitem representar fluxos paralelos e sua sintaxe pode ser adaptada para implicitamente abordar várias questões relacionadas ao paralelismo e controle de concorrência. Trabalhos como CODE 2.0 (NEWTON e BROWNE, 1992) e KAIRA (BÖHM et al., 2014), apresentam ambientes para a criação visual de novas aplicações paralelas. Nesses trabalhos, as aplicações são criadas com a utilização de modelos de grafos, os quais, segundo Newton e Browne

(1992), permitem expressar o paralelismo em alto nível de abstração. Ainda de acordo com os autores, ao se utilizar grafos, os arcos podem representar fluxos de trocas de dados entre os vértices, que podem representar computações básicas sequenciais.

Este artigo apresenta uma nova ferramenta, chamada Pyralelo, para suporte a criação visual de novas aplicações paralelas. A ferramenta proposta inclui suporte a um novo modelo de grafo direcionado, onde o conjunto de vértices podem representar recursos compartilhados (conjuntos de dados de entrada e saída) e computações básicas sequenciais ou paralelas. As arestas, assim como nas abordagens da literatura, representam fluxos de trocas de dados.

O restante do artigo é dividido da seguinte maneira: Na próxima seção, apresenta-se o modelo de grafo proposto bem como a ferramenta proposta para suporte a criação de aplicações paralelas. A Seção 3 apresenta um estudo de caso e alguns resultados preliminares. Por fim, as conclusões e trabalhos futuros são demarcados na Seção 4.

■ MÉTODO

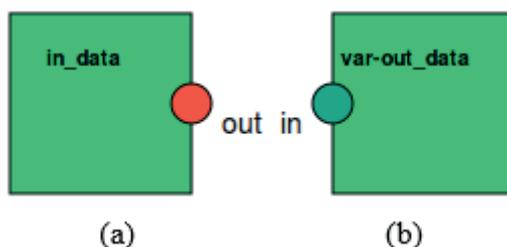
Esta seção é subdividida em três partes: inicialmente, apresenta-se o modelo de grafo proposto para modelagem visual de novas aplicações paralelas; em seguida, é descrito o modelo de paralelização adotado; e, por fim, apresenta-se a ferramenta de suporte à modelagem e execução.

MODELO DE GRAFO

Neste trabalho, a estrutura utilizada para modelar aplicações paralelas é um grafo direcional (CORMEN, 2012), cujos vértices podem representar elementos de dados e elementos de processamento, bem como as arestas representam os fluxos de dados entre elementos de dados e de processamento.

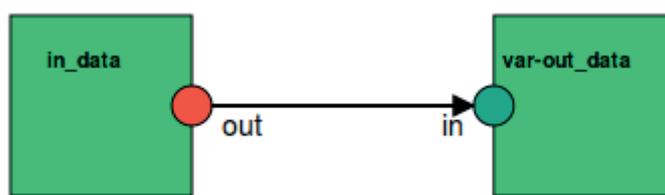
Um vértice de elemento de dados pode consistir de variável, estrutura de dados e arquivo em disco. Um elemento de dados pode ser de entrada (dados a serem processados) ou de saída (dados resultantes de processamento). A Figura 1 apresenta dois exemplos de vértices de elementos de dados, sendo um de entrada (a) e outro de saída (b), respectivamente.

Figura 1. Exemplos de vértices de entrada (a) e saída (b), IFCE, 2015.



O fluxo de dados entre vértices do modelo de grafo adotado é realizado no sentido da esquerda para direita. Com isso, vértices com interface à direita, círculo vermelho no vértice da Figura 1 (a), por exemplo, consistem de vértices de dados de entrada. Já na Figura 1 (b), a interface à esquerda, círculo em cor verde, denota que o vértice é de saída, ou seja, receberá dados de outro vértice. A figura a seguir ilustra o fluxo de troca de dados entre vértice de entrada e de saída.

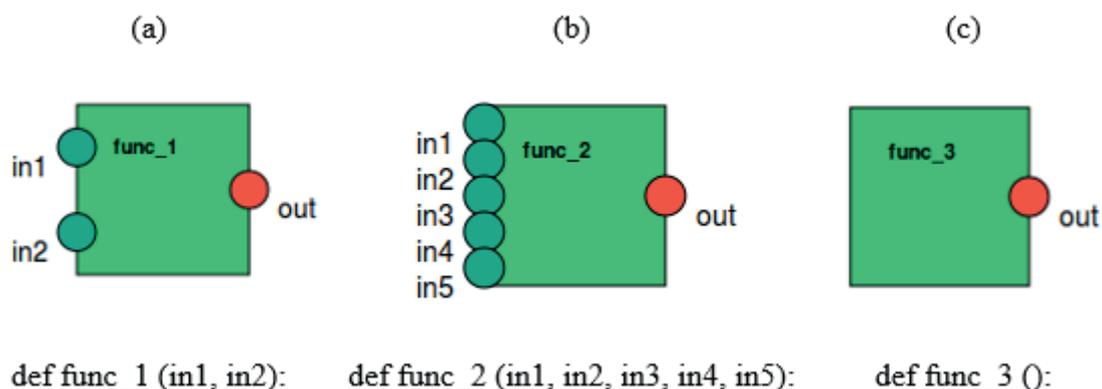
Figura 2. Fluxo de troca de dados entre elementos de entrada e de saída, IFCE, 2015.



Na Figura 2, a aresta, que liga o vértice de entrada (à esquerda) ao vértice de saída (à direita), indica que os dados da entrada serão enviados para a saída.

Os vértices de processamento incluem as instruções necessárias ao processamento de dados, que podem ser computações (funções) básicas sequenciais ou paralelas. A Figura 3 ilustra três vértices representando funções, logo abaixo, com diferentes números de interfaces.

Figura 3. Exemplos de vértices de processamento com diferentes números de interfaces, IFCE, 2015.

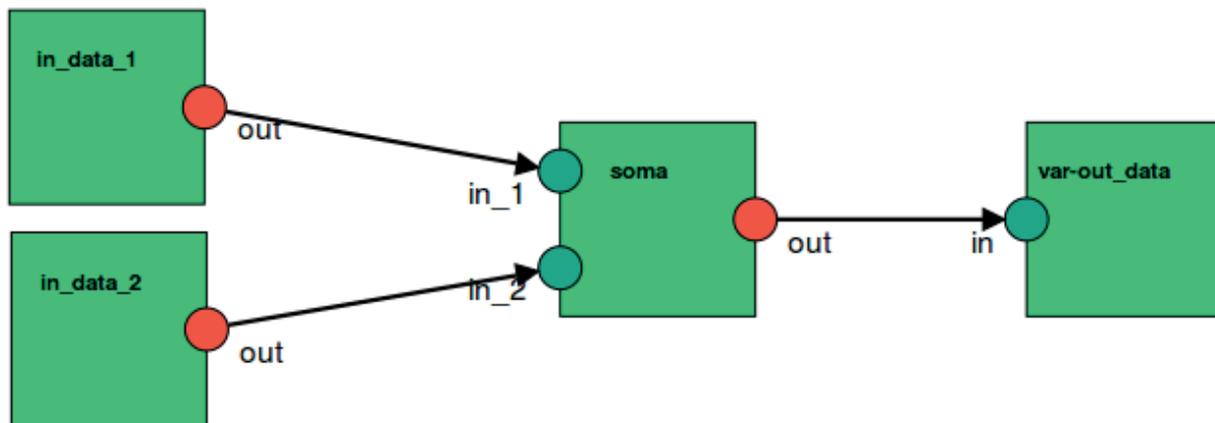


Na Figura 3 (a), vemos um vértice com duas interfaces de entrada de dados. Logo abaixo do vértice, temos a definição da respectiva função. Nela, vemos que os parâmetros de

entrada (in1 e in2) correspondem às interfaces de entrada de dados do vértice. Nas Figuras 3 (b) e (c), temos vértices com cinco e nenhuma interface de entrada, respectivamente.

A figura a seguir mostra um exemplo da modelagem de uma aplicação para adição de dois conjuntos de dados, utilizando o modelo de grafo aqui proposto. Na Figura 4, vemos três vértices de elemento de dados (in_data_1, in_data_2 e var-out_data) e um vértice de processamento (soma). Observando o sentido das arestas direcionais, percebe-se que o fluxo de dados parte dos vértices de dados de entrada para o vértice de processamento e deste para o vértice de dados de saída.

Figura 4. Exemplo de grafo para uma operação de adição, IFCE, 2015.



A subseção a seguir apresenta o modelo proposto para a paralelização das funções representadas por vértices de processamento.

MODELO DE PARALELIZAÇÃO

Neste trabalho, as funções representadas pelos vértices são paralelizadas utilizando o Padrão de Projetos Decorator (GAMMA et al., 2000). Os decorators propostos se baseiam na técnica de divisão de domínio (divisão do conjunto de dados em subconjuntos de mesmos tamanhos para serem processados em paralelo). Com isso, a função definida pelo vértice de processamento é aplicada a cada um desses subconjuntos. Nesse sentido, o projetista da aplicação paralela pode escolher entre processos e threads para realizar as tarefas de processamento paralelo (aplicação das funções aos subconjuntos do problema). Após o processamento dos subconjuntos, as soluções locais são reagrupadas, através de uma função de retorno (conhecida como callback), para compor a solução final.

Para o exemplo da adição, ilustrado na Figura 4, a definição da função soma acrescida de um decorator com processamento paralelo por processos, seria dada da seguinte forma:

1.	@process(4, sum) # decorator
2.	def soma(in_1, in_2): # definicao da funcao
3.	...

Figura 5. Função soma “decorada” com suporte a processamento paralelo, IFCE, 2015.

Na Figura 5, com o uso do decorator proposto (linha 1), as entradas de dados da função soma (in_1 e in_2, declaradas na linha 2) serão divididas em 4 subconjuntos cada. Em seguida, a função soma será aplicada a cada par dos subconjuntos por 4 novos processos. Após a finalização dos processos, os resultados serão agrupados por uma nova soma (função de callback sum, definida no decorator).

A subseção a seguir apresenta a ferramenta proposta para suportar a modelagem visual de aplicações paralelas.

■ RESULTADOS E DISCUSSÃO

Para avaliação da abordagem proposta, elaborou-se um estudo de caso que consiste de uma aplicação paralela para regressão linear simples. Para estimar os parâmetros de uma equação linear foi utilizado o Método dos Mínimos Quadrados (MMQ) (WEISBERG, 2005), dispoendo das seguintes equações:

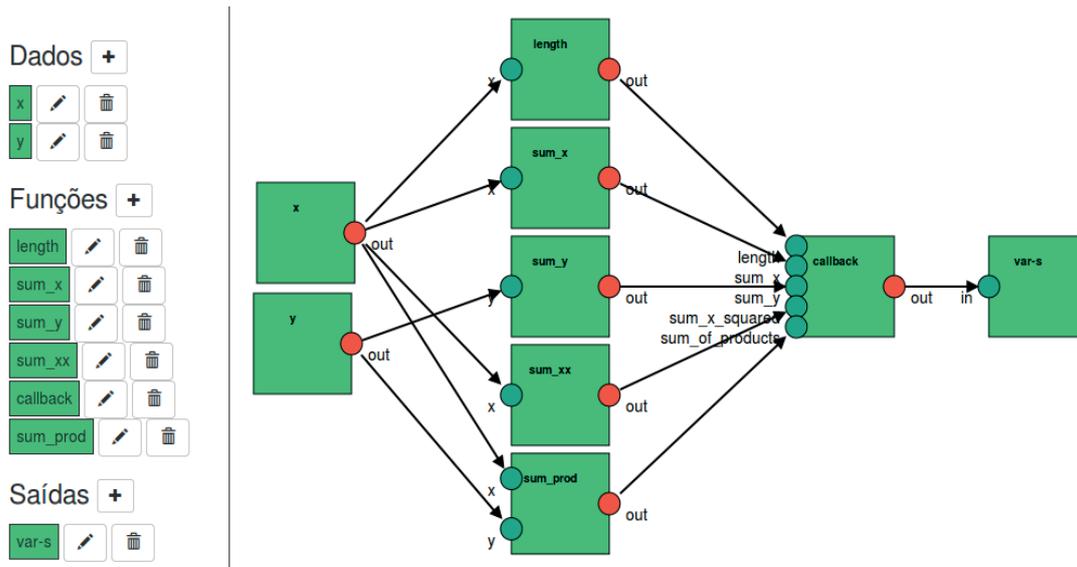
Figura 6. Equações para solução de uma regressão linear simples usando MMQ, IFCE, 2015.

$$a = \frac{n \sum_{i=1}^n x_i \cdot y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2}, b = \frac{\sum_{i=1}^n y_i}{n} - \frac{\sum_{i=1}^n x_i}{n}$$

Na equação da Figura 6, *a* e *b* são, respectivamente, o intercepto e o coeficiente angular do modelo de regressão linear.

Essas fórmulas foram decompostas em função de seus somatórios ($\sum x$, $\sum y$, $\sum xy$ e $\sum x^2$), de forma que cada um deles pudesse ser modelado como um vértice de processamento e, conseqüentemente, ser paralelizado. A Figura 8 ilustra a modelagem gráfica das equações.

Figura 7. Ilustração da modelagem gráfica das equações da regressão linear simples, IFCE, 2015.



Na Figura 7 vemos as bases de dados modeladas como vértices de dados de entrada (x e y), os somatórios como vértices de processamento (sum_x , sum_y , sum_xx e sum_prod) e, por fim as equações do intercepto a , e do coeficiente angular b modeladas como o vértice `callback`. O vértice de saída (à direita) guarda os valores calculados para a e b .

Os resultados exibidos na Tabela 1 mostram que há uma redução no esforço de programação da aplicação paralela, visto que o número de linhas de código informadas pelo projetista é pequeno em relação ao total da aplicação.

Tabela 1. Números de linhas da aplicação de regressão linear, IFCE, 2015.

Código	Linhas de código
Informado pelo projetista (definição de funções)	16
Gerado pelo Pyralelo (elementos de dados, decorators, aplicação)	195
Total	211

Para avaliar o desempenho da aplicação paralela para regressão linear, foi utilizado um conjunto de dados que consiste de amostras de consumo de energia elétrica, realizadas de minuto em minuto, durante 47 meses, em uma das tomadas elétricas de uma casa de família. Ao todo, o conjunto possui 2.075.259 registros com 2 variáveis (hora e voltagem). Para o estudo de caso, as variáveis X e Y correspondem ao instante da medição e a voltagem aferida, respectivamente. A Tabela 2 apresenta os resultados de desempenho para aplicação de regressão linear utilizando a base de dados do estudo de caso, considerando diferentes números de processos para as tarefas paralelas. Percebe-se que, em relação à aplicação sequencial, que inclui apenas um processo, as demais aplicações têm maior de desempenho, destacando-se a de 8 processos.

Tabela 2. Desempenhos sequencial e paralelo para o problema de regressão linear, IFCE, 2015.

Solução	Tempo (s)
Sequencial (1 processo)	1,029
Paralela com 2 processos	0,655
Paralela com 4 processos	0,516
Paralela com 8 processos	0,463

■ CONCLUSÃO

Este artigo apresentou o Pyralelo, uma ferramenta para suportar a criação visual de novas aplicações paralelas. A ferramenta proposta inclui um modelo de grafo, onde é possível modelar elementos de dados e de processamento da aplicação. Na modelagem dos elementos de processamento pode-se escolher computação serial e paralela, por divisão de domínio. Após a modelagem, a ferramenta proposta permite executar e/ou gerar código alvo. Os resultados mostraram que através da modelagem gráfica, é possível reduzir o esforço para a criação de novas aplicações paralelas de alto desempenho.

Os trabalhos futuros incluem a inclusão de novas técnicas para criação de aplicações paralelas, a geração de código alvo em outras linguagens de programação, como C e C++, além da adição de novos modelos de paralelização (e.g. OpenMP e CUDA).

■ AGRADECIMENTOS

Agradecemos ao IFCE *campus* Crato por fomentar esta pesquisa através da concessão de bolsas PIBIC e apoio financeiro, via Edital 008/2014 - Programa de Apoio à Produtividade em Pesquisa (PROAPP).

■ REFERÊNCIAS

1. BÖHM, S., BĚHÁLEK, M., MECA, O., & ŠURKOVSKÝ, M. (2014). Kaira: Development Environment for MPI Applications. In *Application and Theory of Petri Nets and Concurrency* (pp. 385-394). Springer International Publishing, 2014.
2. BOUJILA, A.; ZEPPEFELD, J.; STECHELE, W.; BERNAUER, A.; RINGMANN, O.; ROSENSTIEL, W. and HERKERSDORF, A. Autonomic System on Chip Platform. In *Organic Computing - A Paradigm Shift for Complex Systems*. Springer, 2011.
3. CORMEN, T. H.; LEISERSON, C. E.; RIVEST, R. L.; STEIN, C. *Algoritmos*, 3a. Ed. Elsevier, 2012.
4. GABRIEL, E.; FAGG, G. E.; BOSILCA, G.; ANGSKUN, T.; DONGARRA, J. J., SQUYRES, J. M., SAHAY, V.; KAMBADUR, P.; BARRETT, B.; LUMSDAINE, A.; CASTAIN, R. H.; DANIEL, D. J.; GEBALI, F. *Algorithms and Parallel Computing*. John Wiley & Sons, 2011.

5. GAMMA, E.; HELM, R.; JOHNSON, R.; VLISSIDES, J. Padrões de Projeto: Soluções Reutilizáveis de Software Orientado a Objetos. 5a. Ed. Bookman, 2000.
6. LEE, P. A.; WEBBER, J. Taxonomy for Visual Parallel Programming Language. Technical Report Series-University of Newcastle upon Tyne Computing Science, 2003.
7. MATTSON, G. T.; SANDERS, A. B., MASSINGILL L. B. Patterns for Parallel Programming. Addison-Wesley, Boston, MA, 2005.
8. NEWTON, P.; BROWNE, J. C. The CODE 2.0 graphical parallel programming language. In Proceedings of the 6th international conference on Supercomputing (pp. 167-177). ACM, 1992.
9. PACHECO, P. S. An Introduction to Parallel Programming. [S.I.]: Morgan Kaufmann Publishers, 2011.
10. PALACH, J. Parallel Programming with Python. Packt Publishing Ltd, 2014.
11. WEISBERG, S. Applied Linear Regression, Willey and Sons, Third Edition. Wiley, 2005.

Robótica em tempos de pandemia: uma abordagem usando programação em blocos

| Wilson Rogério Soares e **Silva**
IFPA

| Klévia Letícia **Almeida**
IFPA

| Natalle do Socorro da Costa **Freitas**
UNIASSELVI

RESUMO

Com a pandemia a sala de aula vem se reinventando e a internet se mostrou uma alternativa para continuar os estudos. Com as aulas presenciais suspensas, as escolas tiveram que se adaptar. A metodologia tradicional de ensino presencial teve que dar lugar ao remoto, e encontrar formas de continuar trabalhando os conteúdos da sala de aula. Um desses conteúdos é a robótica. O uso da robótica em sala de aula já é uma realidade. É possível desenvolver o raciocínio de diversas disciplinas usando montagem e programação de robôs. Para codificá-los pode-se usar diversas linguagens de programação, muitas delas usando linhas de códigos. Nesse sentido, esse trabalho traz uma abordagem para analisar como a programação em blocos facilita a inserção da robótica na educação. Para isso, foi desenvolvido um experimento com alunos do ensino técnico subsequente de informática e relatado suas percepções e seus aprendizados sobre o tema.

Palavras-chave: Robótica, Programação, Blocos, Pandemia.

■ INTRODUÇÃO

A Robótica Pedagógica foca na aprendizagem indutiva e na descoberta guiada. Ele utiliza uma combinação de ferramentas de software e equipamentos robóticos educacionais, permitindo que os alunos resolvam problemas de forma divertida e ao mesmo tempo gerem algum aprendizado (RANGEL, 2013). E o professor atua como um guia no processo de aprendizagem. O professor é responsável por oferecer mentoria e apoio às equipes.

A robótica e o pensamento computacional são ferramentas ideais para o desenvolvimento da pedagogia da ciência, tecnologia, engenharia e matemática (STEAM) (PLAZA, 2018). Essa metodologia, permite que tenhamos uma forma de trabalhar a robótica em sala de aula. O ensino da robótica permite aos alunos projetar, construir e programar robôs usando componentes como motores, sensores, rodas, eixos, vigas, conectores e engrenagens (CHAUDHARY, 2016).

Essas abordagens permitem aos alunos experimentem e vivenciem o pensamento científico de maneira interpretativa e reflexiva. Alunos resolvem problemas e se beneficiam de um aprendizado interdisciplinar. Essa prática estimula o raciocínio lógico dos alunos. Por fim, os alunos são agrupados por equipes e participam de desafios, executando diversas tarefas, como: resgatar vítimas em um cenário fictício, realizar lutas de robôs, robôs que resolvem labirintos dentre outras atividades lúdicas.

Quando os robôs estão montados, eles precisam de uma inteligência. Quem faz isso, é a programação. A programação é feita de diversas formas. Existe um tipo de programação, conhecida como blocos, em que o aluno, monta trechos de blocos, para informar o robô o que fazer. Na última década, há um interesse crescente na programação baseada em blocos que visa apoiar o aprendizado precoce e o ensino de programação para crianças (ZAMIN, 2018). Ambientes de programação baseados em blocos são muito populares para introduzir crianças à programação (MILNE, 2019).

Os ambientes de programação de blocos incluem: blocos de sensores, blocos de ação e blocos lógicos. Ao conectar esses blocos, as crianças podem programar estruturas que interagem com o ambiente (WYETH, 2020). Os ambientes de desenvolvimento integrados em que são usados são chamados ambientes de programação baseados em bloco (BBPEs). BBPLs permitir o desenvolvimento de um programa de computador arrastando e soltando (KRALEVA, 2019).

A programação por linhas de códigos já é usada em diversos abordagens de programação de robôs. Nesse trabalho se buscou compreender como a programação em blocos facilita a inserção da robótica na educação. Para isso, foi desenvolvido uma formação, usando tecnologias digitais, com alunos do ensino técnicos subsequente de informática e relatado suas percepções sobre o tema.

■ OBJETIVO

A proposta tem como objetivo geral analisar como a programação em blocos facilita a inserção da robótica na educação. Como objetivos específicos, destaca-se: analisar o aprendizado e a capacidade de resolução de problemas dos alunos, como a programação em blocos é percebida por eles e trazer a discussão sobre trabalhar robótica de forma remota.

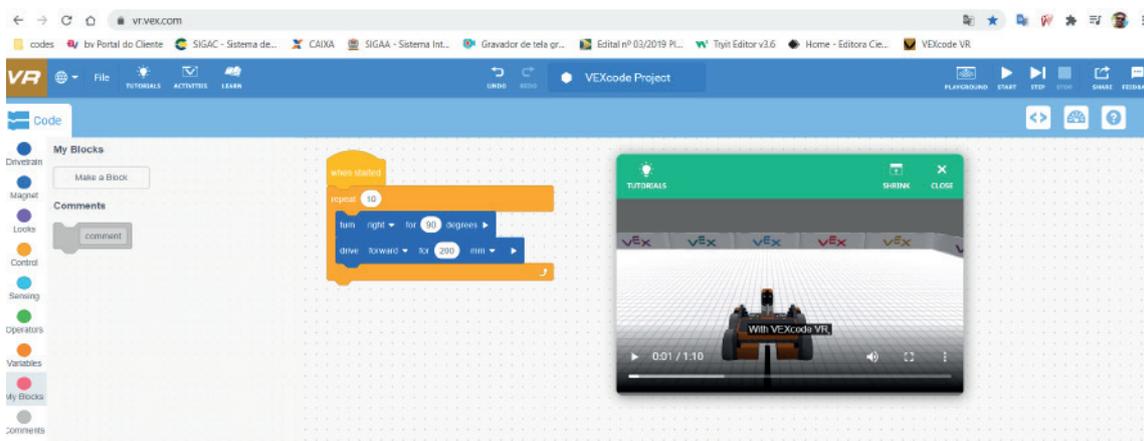
■ MÉTODOS

Esta pesquisa usa uma abordagem qualitativa, que apresenta resultados obtidos por meio de uma pesquisa a campo, que utiliza método de estudo de caso para a coleta e análise dos dados.

Devido a pandemia, as aulas foram suspensas e a abordagem desse trabalho foi adaptada, sendo realizada de forma remota. Dessa forma, foi necessário reinventar a forma tradicional de praticar robótica, assim, foram usadas ferramentas disponíveis na web que possibilitam praticar robótica sem uso físico de um robô.

Para isso, dentre várias plataformas para esse fim, foi escolhida a plataforma virtual de robótica a *vr.vex*. Que é um ambiente de programação de blocos usados para programar robôs diretamente no navegador de internet, disponível em várias plataformas. Essa plataforma conta com todos os comandos necessários para criação de programas para robôs. A plataforma usa o conceito de *drag and drop*, que permite que ao usuário arrastar e soltar os blocos de códigos para a área de programação. Isso permite fácil interação com o ambiente de programação do robô.

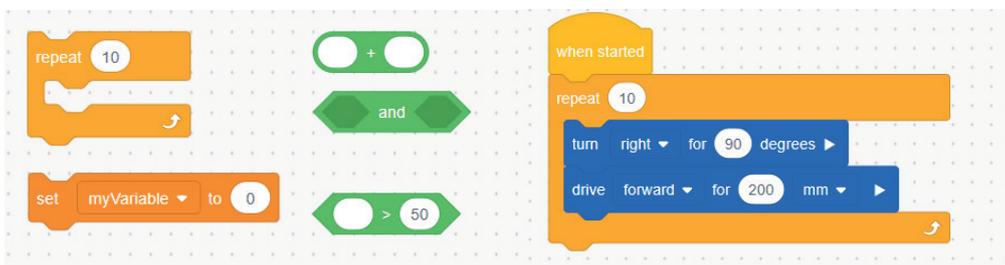
Figura 01. Ambiente da plataforma



Também está disponível todos os comandos usados por uma linguagem de programação tradicionais: blocos com variáveis, blocos de controle, blocos de laços de repetição,

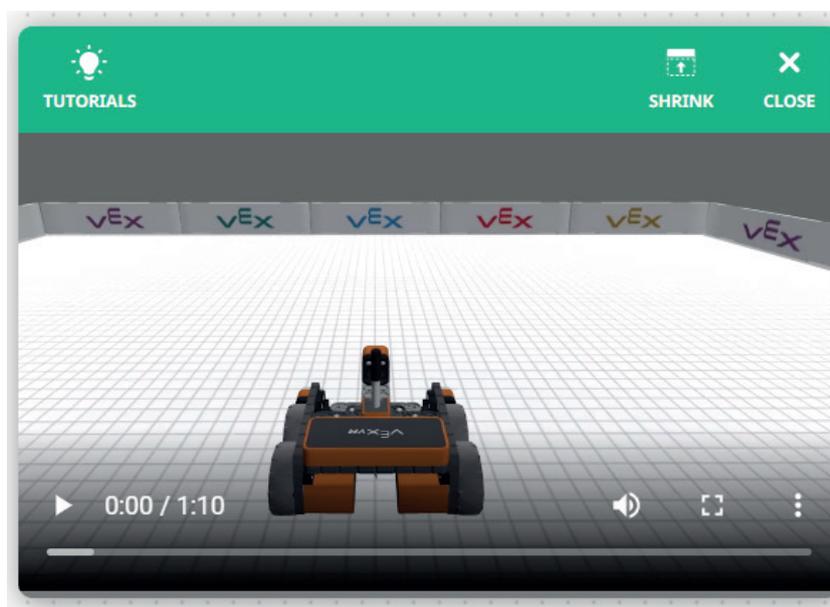
blocos de funções, etc. E também blocos de controle específicos para robótica, como: blocos para mover o robô, blocos para parar o robô, blocos de sensores, etc.

Figura 02. Blocos programáveis de robótica



Para testar a programação do robô é possível visualizar diversos ambientes, com diversas câmeras, que permite ver o robô em 2d e 3d se locomovendo pelo ambiente.

Figura 03. Testar e visualizar os movimentos do robô



Os alunos escolhidos para a pesquisa são da turma do curso técnico subsequente de informática, do Campus Avançado Vigia, que pertence ao IFPA (Instituto Federal de Educação, Ciência e Tecnologia do Pará). Participaram da pesquisa 26 (vinte e seis) alunos, que estão no último semestre do curso, e já cursaram algoritmo, programação e programação web. Todos realizaram as práticas de forma remota usando computadores, tablets e smartphones com acesso à internet.

Para desenvolver esse experimento, foram necessárias algumas etapas, são elas: realizar aula remota, promover uma competição entre os estudantes e responder ao questionário sobre a prática.

Na primeira etapa foi realizado uma aula remota introdutória sobre programação de robôs usando blocos, que irá abordar conceitos básico de um robô, quanto a sensores de percepção, atuadores do robô, mobilidade no ambiente e sua autonomia na tomada de decisão.

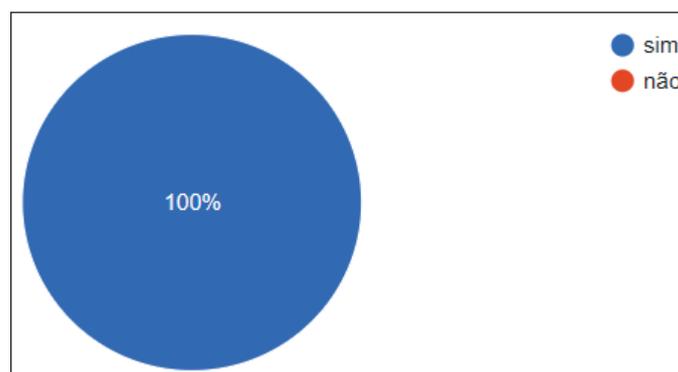
Na segundo etapa foi promovido uma atividade de competição entre os alunos. Os mesmos foram divididos em equipes e tinham como tarefa resolver um desafio. O desafio foi programar um robô autônomo, capaz de se locomover pelo ambiente, desviar de obstáculos e encontrar objetos específicos, possibilitando assim, ganhar pontos.

E na terceira etapa foi coletar, por meio de questionário, a percepção dos alunos sobre a o experimento. Assim os alunos foram convidados a preencher um formulário web *google forms*, com perguntas sobre a prática de programar um robô usando blocos. As perguntas foram: (a) conseguiu entender a aula? (b) o que você achou da plataforma? (c) quanto aos blocos de controle você conseguiu entender as funcionalidades apresentadas? (d) você prefere programar o robô usado blocos ou linhas de códigos? (e) você ficou interessado em aprender mais sobre programação em blocos e robótica? Suas respostas foram então armazenadas e analisadas.

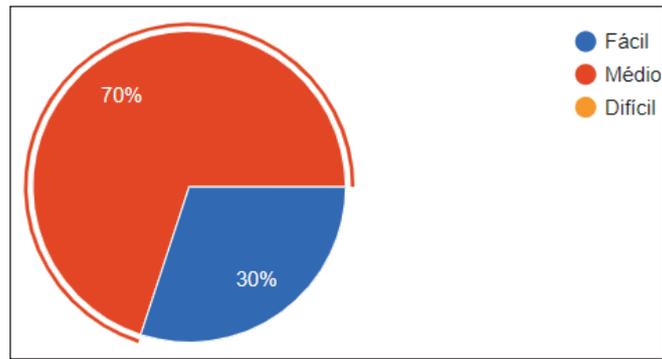
■ RESULTADOS

Após aplicação da prática os dados foram analisados e apresentados em gráficos para facilitar sua compreensão.

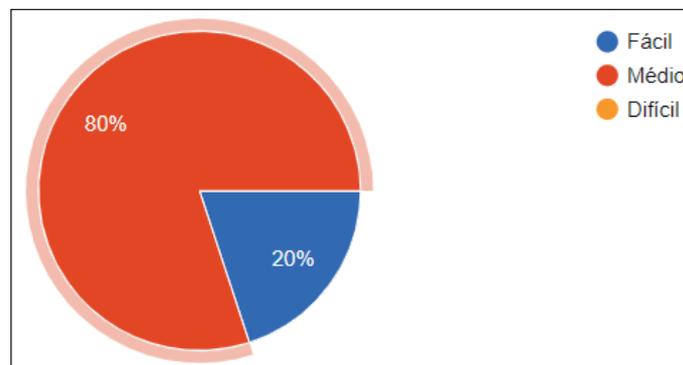
Com relação a primeira pergunta, **(a) conseguiu entender a aula?** Todos os alunos responderam que conseguiram entender o objetivo da prática. A compreensão de todos é importante para desenvolver as atividades propostas.



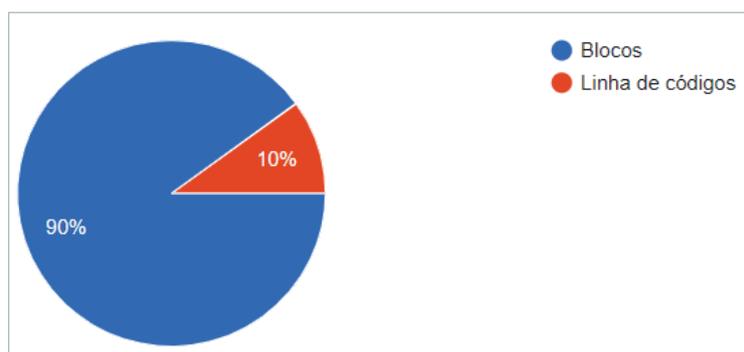
O ato de programar um robô, envolve uma carga mental considerável e tempo para aprendizado das ferramentas. Assim, levando em consideração que foi a primeira vez que os alunos usaram a plataforma web, a pergunta **(b) o que você achou da plataforma?** Teve a maioria, que correspondem a 70 %, achando com média complexidade e demais alunos achando fácil.



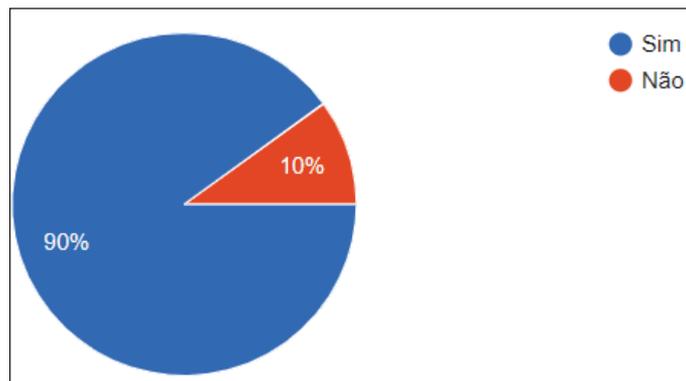
Da mesma forma a pergunta **(c) quanto aos blocos de controle você conseguiu entender as funcionalidades apresentadas?** Uma parte destacou como média complexidade, os demais acharam fácil.



A comparação da linha de comando e blocos na pergunta **(d) você prefere programar o robô usado blocos ou linhas de códigos?** A maioria de 90% respondeu que prefere o uso nesse momento usar blocos. Fortalecendo a ideia de que a forma visual de blocos, moldes e figuras, se mostra mais simples para os alunos.



E com relação a última pergunta, **(e) você ficou interessado em aprender mais sobre programação em blocos e robótica?** A maioria dos alunos que sim e se mostraram motivados a continuar aprendendo.



■ DISCUSSÃO

No decorrer da prática realizada é possível perceber o entusiasmo dos alunos em aprender e realizar os desafios lançados. O que leva a perceber a aderência do tema no processo de aprendizagem programação em blocos.

O ambiente web escolhido, mesmo sendo uma atividade remota, permite ao professor, acompanhar as etapas de aprendizagem do aluno, perceber quais alunos estão com dificuldades nas tarefas, perceber quais alunos estão conseguindo evoluir e assim de forma geral contribuir com o aprendizado coletivo em tempo real.

A aplicação da robótica, com essa roupagem remota, permite com que os alunos e professores possam manter as atividades educacionais e estimula o aprendizado dos alunos, possibilitando continuar estudando robótica fora da sala de aula.

A programação em blocos, apresentada aos alunos, se mostra mais simples para uso na robótica do que usar linhas de códigos. Mesmo sendo programação, os blocos, são interpretados pelos alunos, como uma forma mais simples de programar um robô. E desperta nele o desejo de continuar aprendendo sobre o tema de robótica.

■ CONCLUSÃO / CONSIDERAÇÕES FINAIS

O uso da robótica é mais um atrativo para alunos em todas as idades escolares, pois possibilita interação, práticas e raciocínio lógicos.

Dentre as abordagens de programação dos robôs, tem-se muitas possibilidades e avanços. A citada e trabalhada nessa pesquisa, a programação em blocos, mostra um bom potencial para iniciar a programação de robôs em sala de aulas e assim melhorar seu aprendizado. Uma vez que os resultados demonstraram uma aderência aos jovens alunos que participaram da pesquisa.

A pesquisa atual, que usou uma abordagem remota, com uso de ferramentas web, se mostrou uma alternativa educacional, em tempos de pandemia, para continuar praticando robótica e diminuindo a problemática de não ter aulas presenciais nas escolas públicas.

Por fim, os resultados demonstram que a programação em blocos é um importante instrumento para inserção da robótica na educação, possibilitando aumentar os conhecimentos, habilidades e interesses dos alunos em práticas que envolvem programação e robótica.

Para trabalhos futuros, pretende-se realizar experimentos em outros cursos do IFPA, com o objetivo de avaliar o aprendizado dos alunos com relação ao uso da robótica aplicada em sala de aula.

■ AGRADECIMENTOS

Agradeço a PROPPG/IFPA em possibilitar acesso aos editais e ao Campus Avançado Vigia pelo apoio nas pesquisas.

■ FINANCIAMENTO

Agradeço ao CNPQ/CAPES pelo apoio financeiro na pesquisa.

■ REFERÊNCIAS

1. GOOGLE FORMS. Disponível em < <https://docs.google.com/forms/u/0//>> acessada em 22 de fevereiro de 2021.
2. Kraveva, Radoslava & Krlev, Velin & Kostadinova, Dafina. (2019). A Methodology for the Analysis of Block-Based Programming Languages Appropriate for Children. JCSE, Volume 13. pp.1-10. 10.5626/JCSE.2019.13.1.1.
3. L. R. Milne and R. E. Ladner, "Position: Accessible Block-Based Programming: Why and How," 2019 IEEE Blocks and Beyond Workshop (B&B), Memphis, TN, USA, 2019, pp. 19-22, doi: 10.1109/BB48857.2019.8941230.
4. N. Zamin, H. Ab Rahim, K. S. Savita, E. Bhattacharyya, M. Zaffar and S. N. Katijah Mohd Jamil, "Learning Block Programming using Scratch among School Children in Malaysia and Australia: An Exploratory Study," 2018 4th International Conference on Computer and Information Sciences (ICCOINS), Kuala Lumpur, Malaysia, 2018, pp. 1-6, doi: 10.1109/ICCOINS.2018.8510586.
5. O. J. C. Rangel, A. K. I. Batista, M. R. E. Salazar and T. C. I. Pinzón, "Application of pedagogical robotics in learning for children with special needs APRoLeCSNe," 2013 8th Computing Colombian Conference (8CCC), Armenia, Colombia, 2013, pp. 1-6, doi: 10.1109/ColombianCC.2013.6637533.
6. P. Plaza, E. Sancristobal, G. Carro, M. Castro, M. Blazquez and F. García-Loro, "Multiplatform Educational Robotics Course to Introduce Children in Robotics," 2018 IEEE Frontiers in Education Conference (FIE), San Jose, CA, USA, 2018, pp. 1-9, doi: 10.1109/FIE.2018.8658513.

7. P. Wyeth and H. C. Purchase, "Programming without a computer: a new interface for children under eight," Proceedings First Australasian User Interface Conference. AUIC 2000 (Cat. No.PR00515), Canberra, ACT, Australia, 2000, pp. 141-148, doi: 10.1109/AUIC.2000.822080.
8. V. Chaudhary, V. Agrawal, P. Sureka and A. Sureka, "An Experience Report on Teaching Programming and Computational Thinking to Elementary Level Children Using Lego Robotics Education Kit," 2016 IEEE Eighth International Conference on Technology for Education (T4E), Mumbai, India, 2016, pp. 38-41, doi: 10.1109/T4E.2016.016.
9. VRVEX. Disponível em <<https://vr.vex.com/>> acessada em 20 de fevereiro de 2021.

Sistema especialista para o domínio do licenciamento ambiental: Estudo de caso com Shell Expert Sinta

| José Damião de **Melo**
IFS

| Ubirajara Rodrigues **Xavier**
ADEMA

| Alvaro Luis Enrique Adriazola **Uribe**
UFBA

| Valdenice de Jesus **Melo**
IFBA

| Daniel Antonio de Jesus **Melo**
UFS

| Patrick Henrique da Silva **Brito**
UFAL

| Mikaele Costa **Mendonça**
UFS

| Evandro de Barros **Costa**
UFAL

RESUMO

Objetivo: Modelagem e construção de um protótipo de sistema especialista no domínio do licenciamento ambiental, com a utilização do Shell Expert Sinta utilizando técnicas de inteligência artificial. **Metodologia:** A construção propõe uma máquina de inferência, telas e menus, do tratamento probabilístico das regras de produção e da utilização de explicações sensíveis ao contexto da base de conhecimento modelada. Suas características principais são: Utilização do encadeamento para trás com utilização de fatores de confiança para a representação do conjunto de informações que o especialista possui, construída computacionalmente através de fatos e regras **Resultados:** modelagem e construção de um sistema especialista no domínio do Licenciamento Ambiental, quanto às ações executadas por Órgãos Estaduais de Meio Ambiente, especificamente no enquadramento dos empreendimentos licenciáveis no âmbito da gestão do meio ambiente. **Conclusão:** O sistema se aplica a classificação do potencial poluidor degradador e tipo de licenciamento a ser aplicado as atividades passíveis de licenciamento ambiental. O protótipo apresentou um desempenho satisfatório e verificamos que a ferramenta Shell Expert SINTA se mostra adequada para a construção de sistemas especialistas.

Palavras-chave: Sistemas Especialistas, Expert Sinta, Inteligência Artificial, Licenciamento Ambiental.

■ INTRODUÇÃO

A utilização inicial do termo Inteligência Artificial é atribuída a John MacCarthy, durante um workshop no Dartmouth College em 1956, evento que contou com a participação de Marvin Minsky, Newell e Simon, entre outros, onde a expressão Inteligência Artificial foi o termo adotado para designar o tipo de inteligência construída pelo homem para dotar a máquina de comportamentos inteligentes (Russel, 2008).

Considerando a abrangência do tema podemos elencar uma definição possível para Inteligência Artificial como o estudo e a modelagem de sistemas e dispositivos que visem um comportamento que permita simular as capacidades humanas de decidir, raciocinar, tomar decisões e resolver situações em que existam problemas.

Parte importante do campo da Inteligência Artificial, os sistemas baseados em conhecimento e em particular os sistemas especialistas, categoria de sistemas que resolvem problemas em um determinado domínio onde a presença de um especialista humano é necessária para sua solução, através da utilização de seu conhecimento tácito, explicitado na forma de regras.

São esses sistemas a base deste estudo, onde propomos a modelagem de um sistema especialista de domínio específico e sua validação mediante a construção de um protótipo aplicável ao domínio do Licenciamento Ambiental, especificados para as ações executadas por Órgãos Estaduais de Meio Ambiente - OEMA, emulando as ações do especialista no enquadramento dos empreendimentos licenciáveis.

A estrutura básica de um Sistema Especialista consta de três elementos: a base de conhecimento, responsável por representar o conhecimento do especialista através de regras, fatos e heurísticas; o motor de inferência é o elemento essencial aos sistemas especialistas, nele são aplicadas as regras obtidas na base de conhecimento para a solução dos problemas e a interface com o usuário, responsável por fazer a interface entre o usuário e o sistema. (Mendes, 1999)

Para Mendes (1999), existem alguns pré-requisitos que precisam ser atendidos para que o uso de sistemas especialistas seja adequado. A seguir elencamos alguns desses pontos, que de forma alguma são exaustivos quanto à temática:

- ✓ (...existência peritos que dominem o segmento do conhecimento que encerra o problema, pois é exatamente esse conhecimento que será o responsável direto pela resolução do problema; ...
- ✓ ...existência de tarefas que requeiram conhecimento de detalhes que, se esquecidos, provocam a degradação do desempenho;
- ✓ existência de tarefas que demonstrem grandes diferenças entre o desempenho dos melhores e dos piores peritos;
- ✓ escassez de mão de obra especializada sobre o conhecimento requerido para a solução do problema...)

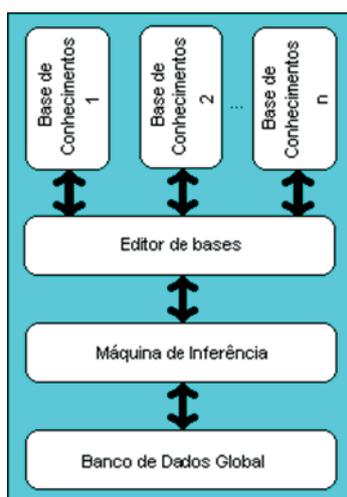
O Shell Expert SINTA é fruto do trabalho do Grupo Sistemas Inteligentes Aplicados do Laboratório de Inteligência Artificial – LIA, um projeto conjunto da Universidade Federal do Ceará e da Universidade Estadual do Ceará. O Objetivo do Shell é a construção automática de sistemas especialistas através da utilização de técnicas de inteligência artificial, regras de produção e probabilidade para representar o conhecimento.

■ MÉTODO

A construção de sistemas especialistas se dá pelo uso da máquina de inferência, construção automática de telas e menus, do tratamento probabilístico das regras de produção e da utilização de explicações sensíveis ao contexto da base de conhecimento modelada. Suas características principais são: Utilização do encadeamento para trás; Utilização de fatores de confiança; Ferramentas de depuração; Possibilidade de incluir ajudas on-line para cada base(Silva, et al, 2020).

Na figura 1 temos a representação esquemática da arquitetura do Shell, onde a base de conhecimentos é a representação do conjunto de informações que o especialista possui, construída computacionalmente através de fatos e regras; o editor de bases permite a construção das bases; a máquina de inferência responde pelas deduções a serem efetuadas sobre a base de conhecimentos e o banco de dados global apresenta quais são as evidências apontadas pelos usuários do sistema especialista durante uma consulta. O Shell Expert SINTA objetiva simplificar ao máximo as etapas de criação de um sistema especialista.

Figura 1. Estrutura Simplificada do Shell Expert Sinta



Fonte: Expert sinta, 2011.

Sobre a base de conhecimento é importante destacar que no Expert SINTA o projetista deverá especificar as variáveis, regras, perguntas, variáveis objetivo e informações adicionais.

Quando cada um destes conjuntos de atributos encontra-se definido, torna-se possível a utilização do sistema especialista.

O Licenciamento Ambiental é um Instrumento da Política Nacional do Meio Ambiente, estabelecida pela Lei nº 6.938/1981, cuja principal função é conciliar o desenvolvimento econômico com a conservação do meio ambiente. O Conselho Nacional do Meio Ambiente – CONAMA, órgão consultivo e deliberativo do Sistema Nacional do Meio Ambiente - SISNAMA, aprovou a Resolução CONAMA nº 237/97, onde está definido que o licenciamento ambiental é o “procedimento administrativo pelo qual o órgão ambiental competente licencia a localização, instalação, ampliação e a operação de empreendimentos e atividades utilizadoras de recursos ambientais, consideradas efetiva ou potencialmente poluidoras ou daquelas que, sob qualquer forma, possam causar degradação ambiental, considerando as disposições legais e regulamentares e as normas técnicas aplicáveis ao caso.” (CONAMA, 2008)

No âmbito do Estado de Sergipe, a estrutura normativa que regulamenta o licenciamento ambiental é a resolução do Conselho Estadual do Meio Ambiente - CEMA 06/2008, que faz a definição de potencial poluidor degradador e a tipologia dos empreendimentos, classificando o primeiro como Baixo(B), Médio(M) ou Alto(A), e definindo os tipos de licenças como Prévia (LP), concedida na fase preliminar do planejamento do empreendimento ou atividade, aprovando sua localização e concepção, atestando a viabilidade ambiental e estabelecendo os requisitos básicos e condicionantes a serem atendidos nas próximas fases de sua implementação.

A Licença de Instalação (LI) autoriza o início da instalação do empreendimento ou atividade de acordo com as especificações constantes dos planos, programas e projetos executivos aprovados e Licença de Operação (LO), que autoriza a operação da atividade, obra ou empreendimento, após a verificação do efetivo cumprimento das exigências das licenças anteriores (LP e LI), bem como do adequado funcionamento das medidas de controle ambiental, equipamentos de controle de poluição e demais condicionantes determinados para a operação. (CEMA, 2006)

A estratégia para a construção do Protótipo com o Shell Expert SINTA para o domínio de licenciamento ambiental teve como meta a obtenção de dois resultados essenciais na fase de enquadramento dos empreendimentos: a tipologia de licença a ser solicitada e qual o potencial poluidor degradador do empreendimento.

Outro ponto importante no escopo deste trabalho está relacionado à definição de porte do empreendimento, que variam entre micro, pequeno, médio, grande e excepcional porte. Deriva-se dessa condição de porte, em associação com a atividade desenvolvida, o potencial poluidor e degradador da atividade que está passível de licenciamento. Neste contexto, as

seguintes variáveis são definidas como objetivo para o Shell: área, atividade, Estágio, Tipo de Licença a ser Emitida e Potencial Poluidor Degradador.

Tabela 1. Conjunto de Regras presentes no protótipo de sistema especialista

Nome	Regra
Regra 1	SE area <= 150 ENTÃO porte = Micro
Regra 2	SE area > 150 E area <= 1000 ENTÃO porte = Pequena
Regra 3	SE area > 1000 E area <= 5000 ENTÃO porte = Media
Regra 4	SE area > 5000 E area <= 10000 ENTÃO porte = Grande
Regra 5	SE area > 10000 ENTÃO porte = Excepcional
Regra 6	SE atividade = COMERCIO E SERVICOS OU atividade = AGROPECUARIA E atividade <> INDUSTRIA E Estagio <> Produção E porte = Pequena OU porte = Micro E porte = Grande E porte = Excepcional ENTÃO Potencial Poluidor Degradador = Baixo
Regra 7	SE atividade = AGROPECUARIA OU atividade = AQUICULTURA OU atividade = IMOBILIARIAS E Estagio <> Produção E porte <> Pequena ENTÃO Potencial Poluidor Degradador = Medio
Regra 8	SE atividade <> DESCONHECIDO E porte = Grande OU porte = Excepcional ENTÃO Potencial Poluidor Degradador = Alto
Regra 9	SE atividade = INDUSTRIA E porte <> DESCONHECIDO ENTÃO Potencial Poluidor Degradador = Alto
Regra 10	SE atividade <> DESCONHECIDO OU atividade = INDUSTRIA E porte = Excepcional OU porte = Grande ENTÃO Potencial Poluidor Degradador = Alto
Regra 11	SE Estagio = Planejamento E porte <> DESCONHECIDO ENTÃO Tipo de Licença a ser Emitida = Previa
Regra 12	SE Estagio = Produção E porte <> DESCONHECIDO E Potencial Poluidor Degradador <> DESCONHECIDO ENTÃO Tipo de Licença a ser Emitida = Operação
Regra 13	SE Estagio = Construção E porte <> DESCONHECIDO E Potencial Poluidor Degradador <> DESCONHECIDO ENTÃO Tipo de Licença a ser Emitida = Instalação

Fonte: Autores

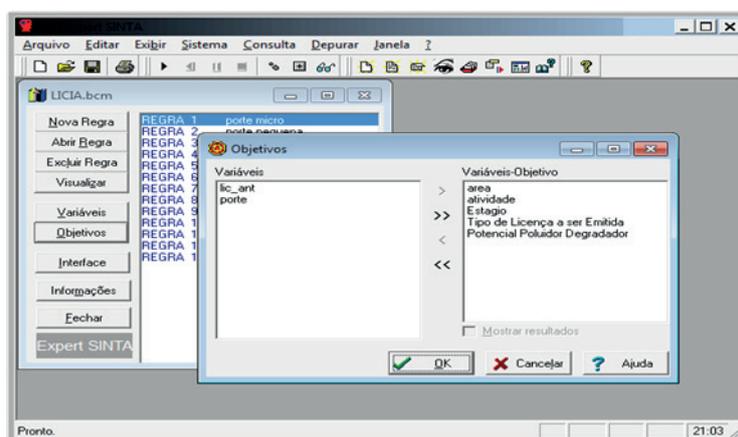
As regras utilizadas no protótipo seguiram o modelo Conectivo-Atributo-Operador-Valor. Foram criadas treze regras no sistema para atender ao escopo do Sistema Especialista, conforme tabela 1 acima, visando emular a ação do especialista humano do domínio, bem como quais seriam as variáveis que necessitariam de participação do usuário na resposta as perguntas, criando a interface com o usuário e as informações gerais do sistema. É solicitado ao usuário que forneça as informações particulares ao caso que o sistema deverá analisar, fornecendo a área ocupada pelo empreendimento, sua atividade operacional e o estágio do empreendimento econômico.

Tabela 2. Variáveis presentes no sistema especialista

Variáveis	Valores	Tipo
estagio	planejamento produção construção	univalorada
area		numérica
atividade	comercio e servicos imobiliarias_agropecuaria indústria aquicultura	univalorada
lic_ant	Lógica (T/F)	univalorada
porte	pequena grande micro media excepcional	multivalorada
potencial poluidor degradador	medio alto baixo	univalorada

Fonte: Autores

Figura 2. Interface para Criação do Sistema Especialista no Shell Expert Sinta



Fonte: Autores

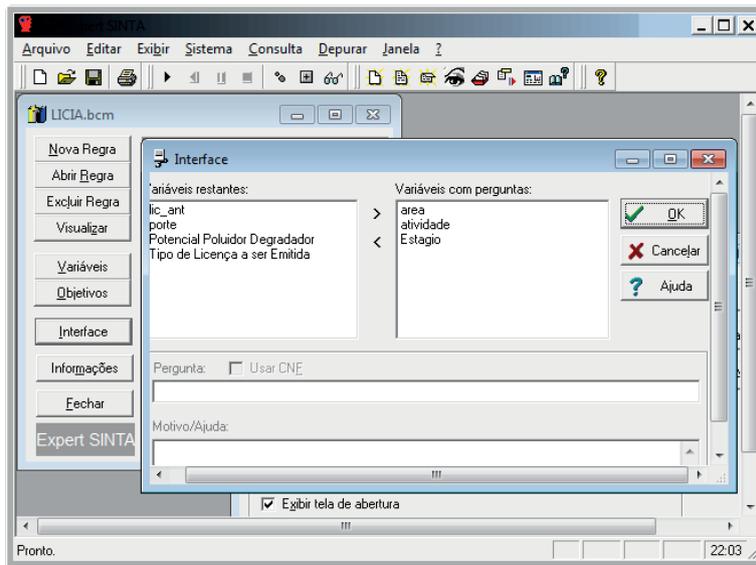
As regras utilizadas no sistema devem seguir o modelo Conectivo - Atributo- operador - valor, as quais estão especificadas a seguir:

Quadro 1. Conjunto de Regras

<p>Regra 1 SE area <= 150 ENTÃO porte = Micro CNF 100%</p>
<p>Regra 2 SE area > 150 E area <= 1000 ENTÃO porte = Pequena CNF 100%</p>
<p>Regra 3 SE area > 1000 E area <= 5000 ENTÃO porte = Media CNF 100%</p>
<p>Regra 4 SE area > 5000 E area <= 10000 ENTÃO porte = Grande CNF 100%</p>
<p>Regra 5 SE area > 10000 ENTÃO porte = Excepcional CNF 100%</p>
<p>Regra 6 SE atividade = COMERCIO E SERVICOS OU atividade = AGROPECUARIA E atividade <> INDUSTRIA E Estagio <> Produção E porte = Pequena OU porte = Micro E porte = Grande E porte = Excepcional ENTÃO Potencial Poluidor Degradador = Baixo CNF 100%</p>
<p>Regra 7 SE atividade = AGROPECUARIA OU atividade = AQUICULTURA OU atividade = IMOBILIARIAS E Estagio <> Produção E porte <> Pequena ENTÃO Potencial Poluidor Degradador = Medio CNF 100%</p>
<p>Regra 8 SE atividade <> DESCONHECIDO E porte = Grande OU porte = Excepcional ENTÃO Potencial Poluidor Degradador = Alto CNF 100%</p>
<p>Regra 9 SE atividade = INDUSTRIA E porte <> DESCONHECIDO ENTÃO Potencial Poluidor Degradador = Alto CNF 100%</p>
<p>Regra 10 SE atividade <> DESCONHECIDO OU atividade = INDUSTRIA E porte = Excepcional OU porte = Grande ENTÃO Potencial Poluidor Degradador = Alto CNF 100%</p>
<p>Regra 11 SE Estagio = Planejamento E porte <> DESCONHECIDO ENTÃO Tipo de Licença a ser Emitida = Previa CNF 100%</p>
<p>Regra 12 SE Estagio = Produção E porte <> DESCONHECIDO E Potencial Poluidor Degradador <> DESCONHECIDO ENTÃO Tipo de Licença a ser Emitida = Operação CNF 100%</p>
<p>Regra 13 SE Estagio = Construção E porte <> DESCONHECIDO E Potencial Poluidor Degradador <> DESCONHECIDO ENTÃO Tipo de Licença a ser Emitida = Instalação CNF 100%</p>

Finalmente foram definidas as variáveis que necessitariam de participação do usuário na resposta as perguntas, definidas com relação a área, atividade e estágio do empreendimento, criando assim interface para a interação com o usuário do sistema especialista e as permitindo a o armazenamento dessas informações mais gerais no sistema.

Figura 3. Definição de Variáveis com perguntas ao usuário

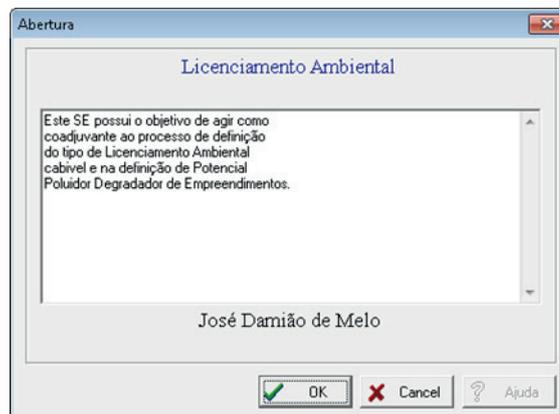


Fonte: Autores

■ RESULTADOS E DISCUSSÃO

Nesta seção vamos apresentar os principais resultados aferidos com a validação do protótipo, bem como derivações e vertentes possíveis que podem advir de sua aplicação no domínio proposto.

Figura 4. Informações gerais do Sistema Especialista



Fonte: Autores

Para executar uma consulta ao sistema, o usuário deverá acionar o botão iniciar, conforme destacado na figura 4. Após iniciar a consulta, a tela de abertura do sistema é apresentada, em seguida é solicitado ao usuário que forneça as informações particulares ao caso que o sistema deverá analisar, fornecendo a área ocupada pelo empreendimento, sua atividade operacional e o estágio do empreendimento econômico.

Figura 5. Formulário de coleta de dados - Área

Licenciamento Ambiental

Informe a area em metros quadrados
(Coloque na faixa abaixo de valores)

Valor: 1000

OK ? Por que?

Fonte: Autores

Para exemplificar a ação do sistema especialista, assumiremos que pretendemos enquadrar um empreendimento a licenciar que possui 1000 metros quadrados, com atividade prevista na área industrial e que atualmente se encontra na fase de planejamento. São exibidos tantos formulários de coleta de dados quantos forem necessários para atender ao conjunto de regras do sistema.

Figura 6. Resultados – Tipologia de Licença

Resultados

Tipo de Licença a ser Emitida

Valor	CNF (%)
Previa	100

Fechar Ajuda

Resultados / Histórico / Todos os valores / 0 sistema /

Fonte: Autores

Munido das informações fornecidas pelo usuário, o sistema fornecerá os dois resultados que são buscados em nosso protótipo: o tipo de licença a ser emitida e o potencial poluidor degradador da atividade a ser licenciada.

Figura 7. Resultados – Potencial Poluidor Degradador

Resultados

Potencial Poluidor Degradador

Valor	CNF (%)
Alto	100

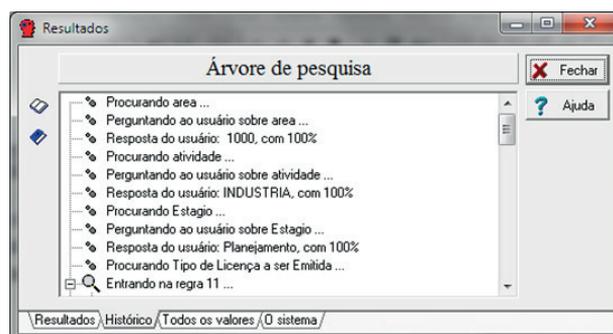
Fechar Ajuda

Resultados / Histórico / Todos os valores / 0 sistema /

Fonte: Autores

Além dos resultados, também está disponível para o usuário o histórico, que apresenta a árvore de pesquisa para os resultados obtidos, disponibilizando todos os valores que estão armazenados nas variáveis e todas as regras que fazem parte do sistema especialista.

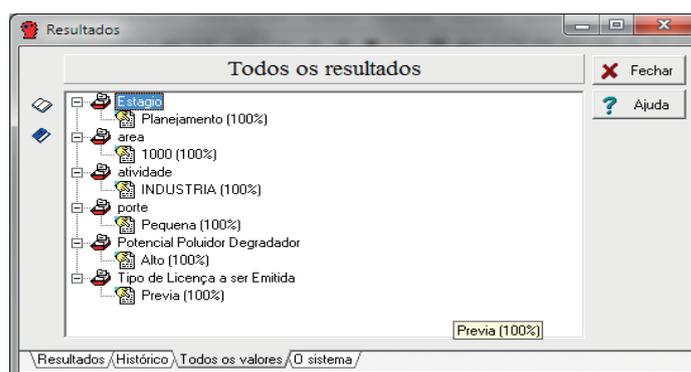
Figura 8. Resultados – Árvore de pesquisa



Fonte: Autores

Estas informações permitem uma análise global de seu funcionamento, bem como visualmente pode-se acessar as principais condicionantes para os resultados obtidos, ampliando a capacidade do usuário do sistema, elencando ainda para o especialista do domínio ou para o analista todo o processo decisório envolvido.

Figura 9. Resultados – Síntese de Resultados



Fonte: Autores

■ CONCLUSÃO

O presente capítulo apresentou um protótipo de sistema especialista para o domínio do licenciamento ambiental, fruto de pesquisa que buscou automatizar as tarefas de definição de tipologia de licença e do potencial poluidor degradador, ações típicas do cotidiano do analista ambiental atuando em organizações estaduais de meio ambiente, utilizando o Shell Expert SINTA.

Estas ações do ponto de vista empírico do processo são executadas por analistas ambientais em um processo denominado de enquadramento do licenciamento do empreendimento econômico ao licenciamento ambiental, nosso protótipo se mostrou eficiente em executar essa ação de forma mediada pela computação, capturando a expertise do especialista do domínio do domínio e representando seu conhecimento mediante o conjunto de

regras e a estrutura do sistema especialista que é processada pelo motor de inferência, atuando assim, em emulação ao comportamento do especialista do domínio.

O sistema prototípico foi aplicado para validação, no sentido de efetuar a classificação do potencial poluidor degradador e tipo de licenciamento a ser aplicado as atividades passíveis de licenciamento ambiental no âmbito do Estado de Sergipe, onde para os objetivos deste trabalho, considerando-se a construção do protótipo, efetuou-se um recorte que permitisse a demonstração e aplicabilidade de técnicas de inteligência artificial e por conseguinte a construção de sistemas especialistas no domínio, de forma mais ampla.

O protótipo apresentou um desempenho satisfatório nos testes e verificamos que a ferramenta Shell Expert SINTA se mostra adequada para a construção de sistemas especialistas e, considerando as limitações impostas ao protótipo, atendeu muito bem a proposta, tendo o protótipo de sistema especialista sido validado através de entrevista com analistas responsáveis pelo enquadramento dos empreendimentos na Administração Estadual do Meio Ambiente – ADEMA, órgão Estadual do Meio Ambiente em Sergipe responsável pelo licenciamento ambiental.

Finalmente, a multidisciplinaridade necessária à análise do processo de licenciamento ambiental por parte de órgãos do poder público se mostrou um domínio bastante interessante e promissor ao uso de elementos e técnicas para o desenvolvimento de sistemas de inteligência artificial e em particular de sistemas especialistas no domínio do licenciamento ambiental.

■ REFERÊNCIAS

1. CEMA, Conselho Estadual do Meio Ambiente. **Resolução 006/2008 de 29 de julho de 2008**, Disponível em < <http://www.adema.se.gov.br/modules/tiny0/index.php?id=31>>. Acesso em 20 de julho de 2011.
2. CONAMA, Conselho Nacional do Meio Ambiente. **Resoluções do Conama: resoluções vigentes publicadas entre julho de 1984 e novembro de 2008**. 2. ed. / Conselho Nacional do Meio Ambiente. – Brasília: Conama, 2008.
3. EXPERT SINTA, “**Manual do Usuário**”; LIA, UFCE. <http://aljmartins.hostmach.com.br/ist/dis/sad/material/16.Sinta_manual.pdf>, acesso em 30 de julho de 2011.
4. MENDES, Raquel Dias. Inteligência artificial: sistemas especialistas no gerenciamento da informação” **Ci. Inf.**, Brasília, v. 26, n. 1, Jan. 1997.
5. RUSSEL, STUART; NORVIG, PETER. **Inteligência Artificial** . 2º ed. São Paulo: Elsevier Campus, 2004.
6. SILVA, Rodes Angelo Batista da, et al. Sistema especialista para diagnóstico do uso do solo em atividades agropecuárias. in: **Braz. J. of Develop.**, Curitiba, v. 6, n. 7, 2020. ISSN 2525-8761. DOI: <https://doi.org/10.34117/bjdv6n7-752>

Software = [(Ideia \cup matemática) \cap
Lógica]

| Shuichi **Murakami**

RESUMO

O computador é uma máquina matemática. Desde a sua origem, a matemática acompanha lado a lado, desde as criações de princípios e métodos que possibilitaram a elaboração conceitual do computador (álgebra booleana), na criação de métodos para agilização (ordenação e indexação), nos métodos para aumentar a segurança na transmissão de dados (criptografia) e uma infinidade de outros métodos e cálculos que têm evoluído cada vez mais trazendo esses benefícios da tecnologia para o nosso cotidiano. Nos dias hodiernos, é impossível imaginar vivermos um mundo sem os computadores, em quase tudo que fazemos, somos cada vez mais dependentes do computador ou das tecnologias derivadas deste. Ele pode estar presente na linha de frente, ou apoiando na retaguarda. Desde um equipamento sobre a mesa, nos veículos mais variados, nos equipamentos médicos e, principalmente no nosso inseparável *smartphone*.

Palavras-chave: Software na Matemática, Matemática no Software, Lógica Matemática no Software”.

■ INTRODUÇÃO

Ao falar de software, particularmente de linguagem de programação, é preciso focar a área de atuação, pois mesmo que essa tecnologia já bastante madura encontra-se em constante evolução, e em via de regra pode proporcionar melhor rendimento para um nicho de aplicações.

Neste trabalho, vamos focar a implicação da Matemática no desenvolvimento de um software que tem por objetivo: primeiro conjunto de softwares, de ser uma ferramenta de apoio aos usuários como facilitador na visualização e compreensão das dinâmicas das funções Matemáticas básicas e, segundo software, utilizando-se dos seus princípios e métodos, poder contribuir para executar as tarefas de pequenos cálculos de forma eficiente empregando os recursos disponíveis no sistema operacional.

Durante as aulas do ensino de funções, nota-se que alguns (ou talvez muitos) alunos têm algum grau de dificuldade no entendimento destas. Em parte talvez por falta de concentração necessária para fazer a conexão entre a equação e o seu resultado ou, talvez devido a maior dificuldade para “enxergar” a parte abstrata da função e os dados resultantes que em geral são apresentados em uma tabela.

- Domínio?
- Imagem?
- Contradomínio?
- Isso é coisa para doidos!

Em geral, essa é a primeira reação que uma grande parte dos jovens estudantes.

- Para que aprender isso? É outra exclamação bastante frequente.
- Onde isso vai ser utilizado? Ora, para usar o meu celular ou meu micro, não preciso nada disso...

Ou ainda, aqueles mais radicais:

- Tinha que matar quem inventou a Matemática! Bem, sem comentários.

Pois é, só que se assim fosse feito, esses radicais (e todos os outros também) teriam que no mínimo voltar às cavernas pois tudo, frisando: T-U-D-O que envolve tecnologia, de alguma forma precisou da Matemática, em maior ou menor grau, desde a bonequinha Barbie e o Ken, os até os modernos robôs e os drones que estão tanto em evidência.

Como já sabemos, a criação do computador teria sido impossível sem o uso da Matemática e, junto com ele, todos os produtos industrializados e até mesmo a produção de alimentos no campo. A lista é tão grande quanto se possa imaginar. Aproveitando o conceito Matemático, essa lista é (quase) infinita, basta conferir...

Se você não adora a Matemática, tudo bem. Nem todos gostam da cor azul, mas daí, a imaginar em viver sem ela, acredite, não dá. A não ser claro, que vá morar no meio do mato ou numa caverna, e sem nenhum produto “da cidade”, principalmente os produtos derivados da tecnologia.

Assim, mesmo que não adore, ao menos deve considerar a Matemática, como uma importante aliada para a sua sobrevivência, tanto no mato e, principalmente na civilização.

■ OBJETIVO

Oportunizar ao estudante, uma breve explanação das aplicações de raciocínio lógico. E para tal, nada melhor que utilizarmos como referência essa máquina lógica, o computador. Afinal ele está presente cada vez mais em nossas vidas. Com isso, não visamos estimular a formação técnica, mas sabemos que o sucesso como cidadão, passará de alguma forma por algum “filtro profissional” em maior ou menor grau. Com vistas a esse horizonte vindouro que já está presente, oferecemos essa contribuição, para quem sabe, possa ajudar a começar a se preparar para a escolha futura para poder exercer a cidadania e se desenvolver. Quem sabe ainda não possa contribuir para a melhora, mesmo que pequena no seu entorno.

Dentre as muitas opções em mãos, tentaremos ajudar a despertar nos jovens estudantes, um maior interesse no estudo da Matemática, demonstrando que ela foi a base fundamental para permitir o desenvolvimento dos softwares tanto o de apresentação como e o da mini-cálculadora. Para isso, contaremos como nasceu a ideia, os desafios que encontramos, e também da grande surpresa de uma descoberta que fizemos durante esse trabalho. Que talvez possa agregar um pontinho na própria Matemática e possa de fato inspirar os jovens.

O primeiro, é um conjunto de softwares gráficos desenvolvidos para ajudar na visualização e na compreensão das dinâmicas das funções matemáticas elementares com um diferencial na apresentação dos gráficos. Em via de regra, devido à alta velocidade de processamento dos computadores modernos, os softwares executam uma função matemática programada num piscar de olhos e, acabam apresentando os resultados “prontos” instantaneamente de forma estática, sem dar a mínima noção das transições dinâmicas que ocorrem do início até o final da função. Dentre tantos, ainda desconhecemos outro que funcione de forma similar, esse aplicativo parece ser um dos poucos, quiçá o único que utilizou essa estratégia.

Acreditamos que muitas pessoas não-iniciadas que não foram apresentados a esses softwares, continuarão sem entender como acontece o passo a passo da evolução da função e, aqueles que conseguirem, acabam memorizando apenas as imagens pontuais ou mesmo as finais, sem a mínima ideia do que ocorre no intervalo.

E isso, poderia fazer falta em algum momento da vida, quando o entendimento de certas dinâmicas forem importantes.

O segundo software é para cálculos simples, como se fosse uma pequena calculadora de mesa, ocupando o mínimo de espaço na tela e da memória de um micro, mas que pode auxiliar sobremaneira nos cálculos, que podem ser realizados de forma simples, sem

a necessidade de redigitação dos dados, tanto nos dados de entrada ou de saída, como se fosse uma extensão do próprio editor.

■ MÉTODOS E FERRAMENTAS

A forma de apresentação será de forma expositiva em slides ou em quadro negro, com algumas ferramentas como: o computador, a Tabela Verdade, para a avaliação de uma questão lógica no programa, os Símbolos de Diagrama de Fluxo utiliza-se os para a traçagem do fluxo da lógica, isto é são empregados quando na definição do fluxo da lógica de programação.

Tabela de Conversão de Hexadecimal - Binário – Decimal que um programador deve conhecer bem, sobretudo aqueles que forem trabalhar com um desses sistemas numéricos diferentes do decimal.

Para facilitar o aprendizado de operações aritméticas em bases diferentes da base decimal, tendo como paradigma o relógio de ponteiro comum, criamos os pseudo-relógios para ajudar na didática de como os sistemas de bases numéricas diferentes trabalham e, têm ajudado bastante para a compreensão de que, apesar dos sistemas numéricos serem de bases diferentes, todos trabalham com o mesmo princípio da soma e do conceito do vai-um.

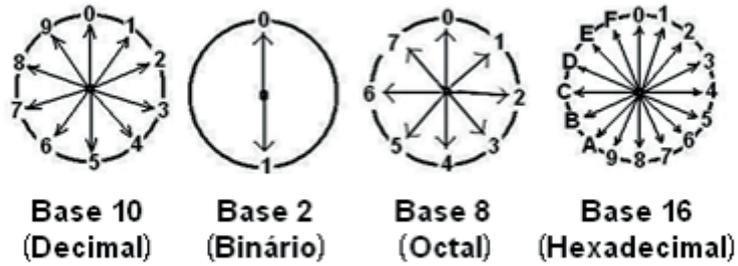
Tabela. Verdade

A	B	$A \wedge B$	$A + B$	$\neg A$	$\neg B$	$A \oplus B$	$\neg(A \wedge B)$	$\neg(A + B)$
V	V	V	V	F	F	F	F	F
V	F	F	V	F	V	V	V	F
F	V	F	V	V	F	V	V	F
F	F	F	F	V	V	F	V	V

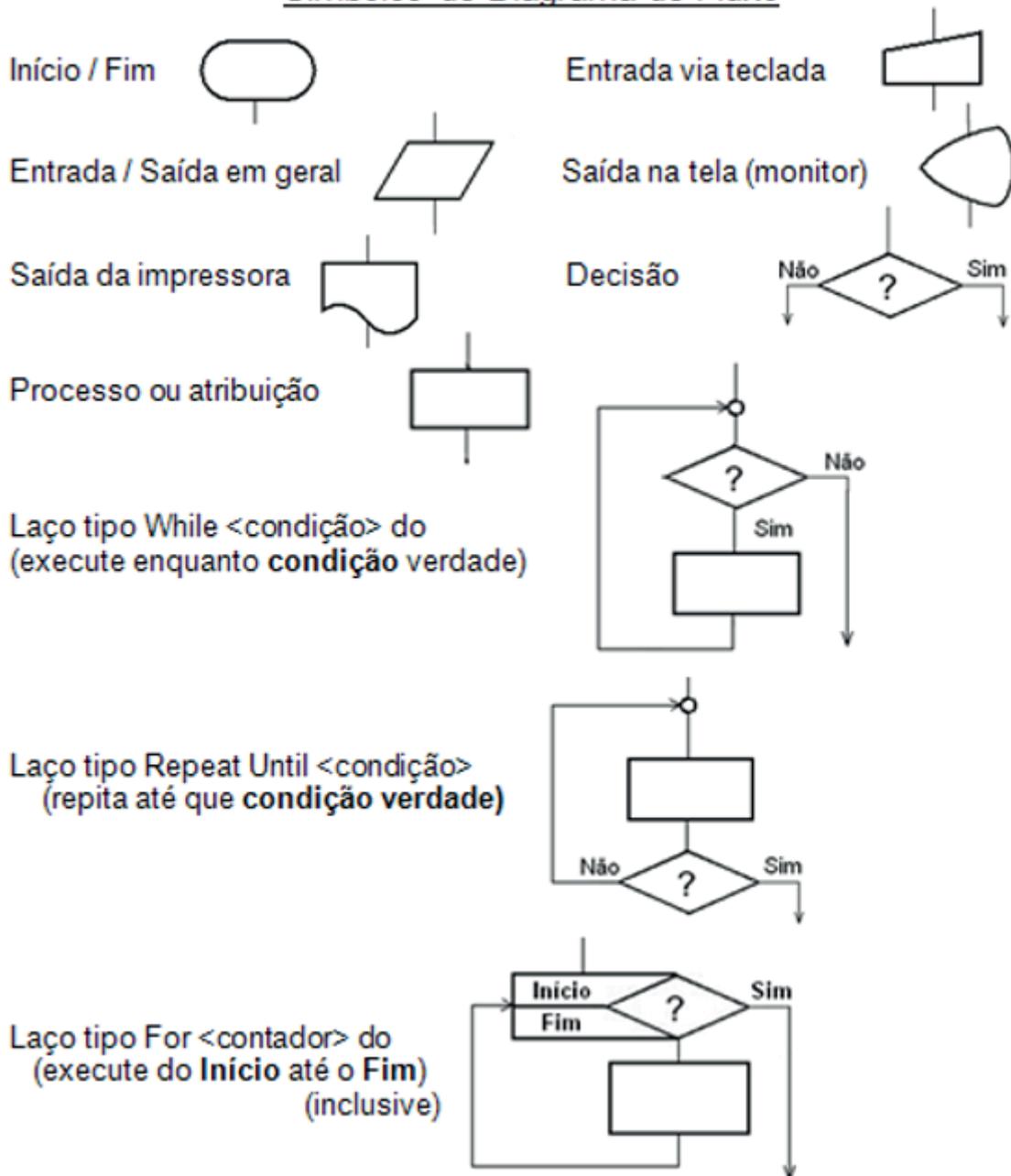
Tabela de Equivalência Numérica de Bases Diferentes

Hexadecimal (base 16)	Binário (base 2)	Decimal (base 10)
0	0000	0
1	0001	1
2	0010	2
3	0011	3
4	0100	4
5	0101	5
6	0110	6
7	0111	7
8	1000	8
9	1001	9
A	1010	10
B	1011	11
C	1100	12
D	1101	13
E	1110	14
F	1111	15

Pseudo-relógios de Sistemas Numerais



Símbolos de Diagrama de Fluxo



A pedra fundamental que possibilitou a invenção do computador foi a Álgebra Booleana, pois sem ela seria impossível a criação dessa maravilhosa máquina que hoje chamamos de micro. Muitos nem imaginam que os primórdios eram grandes, muito grandes, que chegaram a ocupar uma sala inteira...

Como não se trata de curso de programação, é passado aos alunos, de forma abreviada, mas que possam ter uma razoável ideia de como é a rotina de programação de um computador, as etapas e os respectivos usos das ferramentas gráficas apresentadas. A bem da verdade, o que queremos enfatizar é que, não importa se será uma dona de casa, vendedor, médico, engenheiro, administrador ou agricultor. O raciocínio lógico será sempre útil em qualquer situação.

Antes de iniciar a programação de um computador, é feito um roteiro, os mapa(s) para definir os leiautes das telas, com os menus, campos e outros itens a serem exibidos e as suas interações. Após isso, elabora-se os algoritmos ou os diagramas de fluxos, que já podem ser traduzidos diretamente (se já for programador experiente), em linhas de código na linguagem escolhida de programação. Quando se deparar com alguma rotina mais capciosa ou, ficar em dúvida, lança-se mão do uso da Tabela Verdade se for em questão de lógica, ou se for do fluxo de lógica, do Diagrama de Fluxos para avaliar os procedimentos corretos a serem utilizados. Feito os ajustes, a rotina será traduzida (convertida) para a linguagem de programação adotada, compilada e testada (execução). Havendo a necessidade de correção, é interrompido o funcionamento do programa e é aberto a fonte para localizar a falha e corrigido o código-fonte, recompilar e executar um novo teste. Esse processo deve se repetir até que tudo funcione em harmonia conforme planejado.

Apresentamos a seguir, como amostra, uma Rotina do programa de Funções que efetua a plotagem (desenho de pontos) do gráfico da senóide na tela. que é um algoritmo já traduzido para os comandos na linguagem de programação, na sintaxe Pascal.

Rotina do programa de Funções:

```
// By S. Murakami
//-----
Procedure TFormTrigon.PlotarSenóide;
Var
  X,
  Y,
  nXC,
  nYC,
  nXS,
  nYS,
  VEZ : Word;
```

```

INIC_ARCO,
  FIM_ARCO : Extended;
Begin
  ZerarMatrizesLinhas;
  Label10.Left := 660; // Frase = 570
  Label10.Caption := 'Senóide';
  nGRAF := fSENO;
  nKEY := 0;
  While nKEY = 0 Do begin
  With Image1 Do Begin
    Canvas.FillRect(ClientRect);
    TracarMoldura;
  // showmessage('TracarMoldura...');
  VEZ := 0;
  nQUAD := 0;
  INIC_ARCO := 0;
  While (nKEY = 0) and (INIC_ARCO <= 4.71238898)
    Do Begin
      Inc(nQUAD);
      AjustarQuadrante(nQUAD);
      ARCO := INIC_ARCO;
      FIM_ARCO := INIC_ARCO + PASSO;
      While (nKEY = 0) and (ARCO <= FIM_ARCO)
        Do Begin
          SSENO := Sin( ARCO );
          CSENO := Cos( ARCO );
  //   nANGLO := Trunc(ARCO*Pi/180+0.1);
          edtArco.text := FormatFloat('##0.0',
            ARCO*180/Pi);
          If SSENO < 0 then
            edtValor.Font.Color := clRed
          else
            edtValor.Font.Color := clBlue;
          edtValor.text := FormatFloat('0.0000',
            SSENO);
  // Plota o Pixel do Ciclo Trigonométrico.
          nXC := X_CENTRO + Trunc(RAIO*CSENO);
          nYC := Y_INI - Trunc(RAIO*SSENO);

```

```

        Canvas.Pixels[ nXC, nYC ] := clBlack;
// Plota Pixel da senoide.
    Y := Y_INI + Trunc(-RAIO * SSENO );
    X := X_INI + Trunc(RAIO * ARCO);
    If Y > Y_INI then
        Canvas.Pixels[X, Y] := clRed
    Else
        Canvas.Pixels[X, Y] := clBlue;
// Salva os Pixéis das Linhas do Ciclo.
    CapturarBitMapSen(nXC, nYC, X, Y);
    TracarLinhaArco;
    TracarCruzSeno(nXC, nYC, X, Y, clBlue);
//-----
    Application.ProcessMessages;
//-----
    Sleep(nESPERA);
// Restaura os Pixéis das Linhas do Ciclo Trigonométrico.
    ExibirBitMapSen;
    ARCO := ARCO + GRADARC360;
    End; //While (nKEY = 0) and (ARCO <= FIM_ARCO)
    Inc(VEZ);
    INIC_ARCO := INIC_ARCO + PASSO;
    End; //While (nKEY = 0)and(INIC_ARCO<=4.71238898)
    End; // With Image1
End; // While nKEY = 0
    bBtnCossenoide.SetFocus; //Foca o botão: Cossenoide.
End;

```

Como se pode notar, neste tipo de programa, a Matemática permeia intensamente as linhas de código.

Breve histórico do desenvolvimento dos softwares de visualização.

As versões iniciais dos programas de visualização, foram criados na sintaxe BASIC (Beginners All-Purpose Symbolic Instruction Code), no sistema operacional MSDOS®, num ambiente denominado Turbo Basic® da Borland Corporation, mais tarde vendida e para a Power Basic®.

Na versão atual, foram convertidos para no sistema operacional Windows®, na linguagem **Pascal**, no ambiente de desenvolvimento livre, chamado Lazarus®. Todos os

executáveis dessa série, são fornecidos gratuitamente mediante uma solicitação por e-mail (através do qual repassamos o link para baixar do Google Drive).

O software da mini-calculadora, foi desenvolvida em Lazarus, este é um programa utilitário com objetivo de facilitar e melhorar a eficiência daqueles que precisam efetuar pequenos cálculos durante as digitações de relatórios ou mesmo de planilhas dentro de editores de texto simples (bastante comum como em lista de compras de mercado).

Os programas de visualização: para testar as minhas expectativas, pedi à minha esposa (que era professora de matemática) que, quando chegasse o momento adequado, me agendasse uma apresentação, então levaria o micro para rodar os softwares de visualização. E assim fizemos algumas vezes. Consideramos que os resultados foram bastante satisfatórios e promissores, os sorrisos nos seus rostos denunciavam as revisões mentais que faziam ao visualizarem aquilo que a professora já tinha lhes ensinado. Tenho a impressão que aquelas visualizações só que de forma dinâmica, devem ter ajudado a fixar aquilo que tinham visto no quadro-negro.

Penso que o mais, importante e gratificante de ter desenvolvido esse software de aparência simples, foi a surpresa que ele nos reservou durante os testes de consistência (comprovação). E mais, principalmente porque o desenvolvimento desse software, forneceu um argumento irrefutável que repassamos a outros estudantes, docentes e pais, de que não podemos confiar cegamente nas tecnologias (que é senso comum da maioria das pessoas, sobretudo dos estudantes), Precisamos lembrar que atrás de toda a tecnologia, existe um ser humano. E o ser humano sempre está sujeito a falhas.

■ RESULTADOS

As apresentações dos softwares tanto de visualização de funções como a mini-calculadora, transcorreram satisfatoriamente como o esperado e, acreditamos ter alcançado o objetivo.

Surpresa muito interessante foi reservada durante os testes do pequeno programa da mini-calculadora.

Foi durante os testes do software de cálculos que detectamos uma falha de operação em algumas calculadoras ou emuladoras de calculadora científicas que foram utilizados para a conferência dos cálculos.

Esse erro pôde ser detectado porque, sempre durante a codificação de um software ou de uma rotina que envolvem cálculos, temos por hábito, fazemos o que é chamado no jargão da área de programação, o conhecido “teste de mesa”, onde são criadas as amostras seletivas de cálculos e, todos estes são calculados “à mão” e conferidos com o do programa para assegurarmos os resultados, eliminando ou, ao menos minimizando a probabilidade de erros.

Eis que durante esses cálculos de conferência com o uso de calculadoras científicas, surgiu uma situação interessante: detectamos que algumas calculadoras científicas mesmo as de renome (relacionadas na tabela em Resultados), “não conseguiram” calcular diretamente pela hierarquia de operações matemáticas, algumas das potências que já tínhamos “calculados à mão” (rotinas utilizando um algoritmo básico de matemática elementar).

Os resultados dos testes efetuados estão compilados na Tabela de Testes de Cálculos de Potências na seção de RESULTADOS a seguir.

Finalmente pudemos ter absoluta certeza de que os erros de fato, estavam nos programas das calculadoras e emuladores ao invés do algoritmo do nosso despretensioso aplicativo.

Com isso, tentamos entrar em contato com as empresas para alertá-los, mas sem sucesso. Para não dizer de todo, em um destes, o atendente foi um tanto rude.

Tabela de Testes de Cálculos de Potências

Base	Expoente	Resultados Esperados (C0)	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
-0,5	-2,4	5,278	Não	Não	Não	Não	5,278	5,278	Não	Não	5,278	5,278
-0,5	-0,2	-1,148	-1,148	-1,148	Não	-1,148	-1,148	-1,148	Não	Não	-1,148	-1,148
-0,5	0,2	-0,8705	-0,8705	-0,8705	Não	Não	-0,8705	-0,8705	Não	Não	-0,8705	-0,8705
-0,5	1,4	-0,3789	Não	Não	Não	Não	-0,3789	-0,3789	Não	Não	-0,3789	-0,3789
-0,5	2,4	0,1894	Não	Não	Não	Não	0,189	0,189	Não	Não	0,189	0,189
-0,3	0,4	0,6178	Não	Não	Não	Não	0,6178	0,6178	Não	Não	0,6178	0,6178
-0,3	-0,4	1,6186	Não	Não	Não	Não	1,6186	1,6186	Não	Não	1,6186	1,6186
-0,2	-0,4	1,9036	Não	Não	Não	Não	1,9036	1,9036	Não	Não	1,9036	1,9036
-0,4	-1,4	-3,6067	Não	Não	Não	Não	-3,6067	-3,6067	Não	Não	-3,6067	-3,6067
-0,4	2,4	0,1109	Não	Não	Não	Não	0,1109	0,1109	Não	Não	0,1109	0,1109
-2	-0,4	0,7578	Não	Não	Não	Não	0,7578	0,7578	Não	Não	0,7578	0,7578
-2	-2,4	0,1894	Não	Não	Não	Não	0,1894	0,1894	Não	Não	0,1894	0,1894
-2	-3,4	-0,0947	Não	Não	Não	Não	-0,0947	-0,0947	Não	Não	-0,0947	-0,0947
-2	3,4	-10,556	Não	Não	Não	Não	-10,556	-10,556	Não	Não	-10,556	-10,556
-2	-1,4	-0,3789	Não	Não	Não	Não	-0,3789	-0,3789	Não	Não	-0,3789	-0,3789
-2	-2,4	0,1894	Não	Não	Não	Não	0,189	0,189	Não	Não	0,189	0,189
-2	1,4	-2,639	Não	Não	Não	Não	-2,639	-2,639	Não	Não	-2,639	-2,639
-2	2,4	5,278	Não	Não	Não	Não	5,27	5,27	Não	Não	5,27	5,27
-2	-0,2	-0,8705	-0,8705	-0,8705	Não	-0,8705	-0,8705	-0,8705	Não	Não	-0,8705	-0,8705
-2	0,2	-1,148	-1,148	-1,148	Não	Não	-1,148	-1,148	Não	Não	-1,148	-1,148
0	0	Não	Não	Não	Não	1	1	Não	Não	Não	Não	Não
Base	Expoente	Resultados Esperados	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10

Calculadoras e Emuladores empregados nos Testes

- | | | |
|------------------------------------------------------------------|---------------------------------------------------------|------------------------------------------|
| C0: MCalc.exe mini-programa de cálculos do autor | C1: HP-10S calculadora científica | C2: Casio fx-82MS calculadora |
| C3: HP-12C calculadora financeira (modelo dourado) | C4: Windows XP calculadora interna | C5: Windows 7 calculadora interna |
| C6: Casio fx-9860G-SD emuladora da calculadora científica | C7: HP-15C emuladora da calculadora científica | |
| C8: HP-35S emuladora da calculadora científica | C9: HP-300S emuladora da calculadora científica) | |
| C10: RealCalc – Aplicativo de cálculo no Android | | |

■ DISCUSSÃO

Ao depararmos com os erros de cálculos nas calculadoras científicas (não executavam cálculo factível) e, como o mini-programa de cálculos conseguia calcular corretamente, não restou outra alternativa exceto a de conferirmos toda a rotina relativa à potenciação, linha a linha procurando por falhas. Ao final, todos os comandos estavam em ordem e funcionando perfeitamente. Mas antes de jogar a “culpa” pela falha nas calculadoras (afinal algumas eram de marcas renomadas), resolvemos conferir todos os livros da minha esposa (que era professora de matemática à época), para verificar como eram ensinados as operações de potenciação, cerca de cinco a seis coleções. Após essa compilação, conferidas as anotações, constatamos que havia uma lacuna nas explicações pertinentes às frações e sobretudo à potenciação.

Após exaustivos testes, pudemos concluir que naquelas calculadoras não tinham sido aplicados os conceitos fundamentais dessa operação e em seu lugar, muito provavelmente, o programador não muito detalhista, deve ter aplicado algum algoritmo otimizado que acabava provocando os erros detectados sob certas condições. Isso foi comprovado nos testes efetuados em várias calculadoras e emuladores, científicas ou não.

Dessa forma, decidimos ousar um pouco e criar um conjunto de regras que pudesse abarcar todas as operações de potenciação que apresentamos a seguir.

Sem dúvida, estaremos abertos a críticas e sugestões para aprimorá-las.

A pergunta que não quer calar: e se simplesmente tivéssemos desistido de acreditar na nossa lógica e desistido de investigar até encontrar a prova da falha dessas calculadoras?

Será que algum técnico ou engenheiro não teria precisado de um cálculo daquele tipo? Se sim, como teria contornado? Me lembro de um matemático da era das navegações que montou a tabela de senos para navegação e que se suicidou por não ter certeza se as suas tabelas estavam corretas...

Como pudemos comprovar que “nem tudo que reluz é ouro”.

Mais uma vez, o velho ditado estava absolutamente certo...

Mas nem tudo está perdido, esse fato serve como prova cabal e corrobora ainda mais, a importância do conhecimento dos princípios fundamentais da matemática básica.

REGRAS AMPLIADAS PARA CÁLCULOS DE POTÊNCIAS E RAÍZES

Por : Shuichi Murakami

Recordando:

$$a^b = x$$

Onde: **a** = base **b** = expoente **x** = potência

$$a^{b/c} = x \quad \rightarrow \quad x = \sqrt[c]{a^b}$$

Onde: **a^b** = radicando **x** = raiz **c**-ésima de **a^b**

CÁLCULO DA POTÊNCIA POR LOGARITMO

Limitação: somente são aplicáveis para logaritmandos positivos. Isto é, para os casos em que as bases das potências são positivas.

Exemplo:

$$5^3 = x \quad \rightarrow \ln 5^3 = \ln x \quad \rightarrow 3 \cdot \ln 5 = \ln x \quad \rightarrow x = e^{3 \cdot \ln 5}$$

$$(-2)^3 = x \quad \rightarrow \ln (-2)^3 = \ln x \quad \rightarrow 3 \cdot \ln (-2) = \ln x \quad \rightarrow \text{Sem solução.}$$

NB.: $(-2)^3 = -8$ \rightarrow Ainda assim, **não** é possível calcular por logaritmo.

OBS.: como vemos, o método com o uso de logaritmos, restringe enormemente os cálculos de potências, o que pode ocorrer na linguagem de programação Pascal (por exemplo), onde o logaritmo é utilizado para os cálculos de potências. Nessas situações, para contornar essa deficiência, o programador deve estar consciente e lançar mão dos fundamentos básicos da Matemática para levar a cabo, os cálculos de potências. Caso contrário, muitos cálculos que são possíveis deixarão de ser calculados pelos programas de computador, e isso vale também para as calculadoras científicas.

REGRAS PARA O CÁLCULO DE POTÊNCIAS DE NÚMEROS REAIS:

1. Quando a base (**a = 0**).

R: é possível calcular **somente** quando o expoente (**b > 0**) **x = 0**

2. Quando a base (**a > 0**) e o expoente (**b > 0**)

R: é possível calcular por qualquer método: **x > 0**

3. Quando a base (**a > 0**) e o expoente (**b < 0**)

R: é possível calcular por qualquer método: $x > 0$

3.1) Expoente não-irracional:

$$x = 0,5^{-1/3} \quad x = \frac{1}{0,5^{1/3}} \quad x = \frac{1}{\sqrt[3]{0,5}} \quad \mathbf{x = 1,259921}$$

3.2) Expoente irracional:

$$x = 2^{-\sqrt{2}} \quad x = \frac{1}{2^{\sqrt{2}}} \quad \mathbf{x = 0,3752142}$$

4. Quando a base ($a < 0$) e o expoente ($b > 0$):

R: não é possível calcular por logaritmo devido à **base negativa: ($a < 0$)**.

4.1) Quando o expoente **b** for inteiro:

4.1.1) Quando o expoente **b** for par. $x > 0$

4.1.2) Quando o expoente **b** for ímpar $x < 0$

4.2) Quando o expoente **b** for **irracional**, **não** é possível calcular.

Exemplo: $x = (-2)^{\sqrt{2}}$ **R:** não há solução em Reais.

4.3) Quando o expoente **b** for **racional**:

$$x = (-0,5)^{1/3} \quad x = \sqrt[3]{(-0,5)} \quad \mathbf{x = -0,7937005}$$

5. Quando a base ($a < 0$) e o expoente ($b < 0$)

R: não é possível calcular por logaritmo devido à **base negativa: ($a < 0$)**.

5.1) Quando o expoente **b** for **inteiro**:

5.1.1) Quando o expoente **b** for **par** $x > 0$

5.1.2) Quando o expoente **b** for **ímpar** $x < 0$

5.2) Quando o expoente **não** for **irracional**:

$$5.2.1) \quad x = (-2)^{-0,5} \quad x = \frac{1}{(-2)^{0,5}} \quad x = \frac{1}{(-2)^{1/2}} \quad x = \frac{1}{\sqrt{-2}}$$

$$5.2.2) \quad x = (-2)^{-1,5} \quad x = \frac{1}{(-2)^{1,5}} \quad x = \frac{1}{(-2)^{3/2}} \quad x = \frac{1}{\sqrt[2]{(-2)^3}}$$

R: quando o **denominador da fração irredutível do expoente** for **par**, **não** é possível calcular

$$5.2.3) \quad x = (-2)^{-0,4} \quad x = (-2)^{-4/10} \quad x = (-2)^{-2/5}$$

$$x = \frac{1}{(-2)^{2/5}} \quad x = \frac{1}{\sqrt[5]{(-2)^2}} \quad x = \frac{1}{\sqrt[5]{4}} \quad \mathbf{x = 0,7578}$$

$$5.2.4) \quad x = (-2)^{-1,4} \quad x = \frac{1}{(-2)^{1,4}} \quad x = \frac{1}{(-2)^{7/5}} \quad x = \frac{1}{\sqrt[5]{(-2)^7}}$$

$$\mathbf{x = -0,378929}$$

R: quando o **denominador da fração irredutível do expoente** for **ímpar**, é possível calcular.

5.3) Quando o expoente for **irracional**, **não** é possível calcular.

$$\text{Exemplo (1): } x = (-2)^{-\left(\frac{\sqrt{2}}{4}\right)} \quad x = \frac{1}{(-2)^{\left(\frac{\sqrt{2}}{4}\right)}}$$

$$\text{Exemplo (2): } x = (-2)^{-\left(\frac{\sqrt{2}}{3}\right)} \quad x = \frac{1}{(-2)^{\left(\frac{\sqrt{2}}{3}\right)}}$$

$$\text{Exemplo (3): } x = (-2)^{-\sqrt{2}}$$

$$\text{Exemplo (4): } x = (-2)^{\sqrt{2}}$$

$$\text{Exemplo (5): } x = (-2)^{-0,31257393827\dots(\text{irracional})}$$

Conclusão: se a potência tiver base **negativa** e o seu expoente for uma fração, esta deverá ser obrigatoriamente uma **fração irredutível**. **Jamais** poderá ser uma fração **equivalente**, **exceto** quando for mantida a mesma **relação** de **par-ímpar** entre o **numerador** e o **denominador** do expoente. Caso o expoente seja uma **fração imprópria**, deverá ser transformada em uma **fração mista** e, analisar a exequibilidade pela **parte fracionária**.

Exemplo:

$$(-2)^{1/2} = \sqrt{(-2)} \quad \text{não é possível calcular no conjunto dos Reais.}$$

Mas vejamos que, se utilizássemos uma fração **equivalente** no expoente, por exemplo: $\frac{2}{4}$. Poderíamos calcular perfeitamente o $(-2)^{2/4}$ como $\sqrt[4]{(-2)^2}$ (se não simplificarmos a fração do expoente) e, o resultado, seria igual a 1,142136... Mas esta é uma solução incorreta.

Nota-se que, pela proposição inicial o radicando era (-2) e o índice da raiz igual a 2 (par) (**não existe raiz real de índice par de radicandos negativos**).

Neste caso, apesar da razão ser mantida, a **relação par-ímpar** entre o **numerador** e o **denominador** foi alterada e, portanto, não devemos prosseguir os cálculos sob pena de obtermos resultados equivocados.

Por outro lado, se utilizássemos no expoente, uma fração equivalente, por exemplo: $(-2)^{\frac{3}{6}}$ onde, tanto a razão como a **relação par-ímpar** entre o **numerador** e o **denominador** seriam mantidas, o resultado será idêntico ao do cálculo da potência original.

Ou seja, neste caso, podemos concluir igualmente que não será possível calcular a raiz de $\sqrt[6]{(-2)^3}$ no conjunto dos Reais. **c.q.d.**

“Ceux qui ne savent pas qu’ils marchent dans l’obscurité, ne verront jamais la lumière.”

Bruce Lee

■ CONSIDERAÇÕES FINAIS

Muitas vezes, os alunos têm dificuldade para perceber a beleza e a importância dos conceitos Matemáticos talvez por serem às vezes por demais abstratas e, assim, aqueles que não persistem um pouco mais, deixam de usufruir muitos de seus benefícios. Em casos extremos, podendo até ser vital, por negligenciar algum dos princípios fundamentais da Matemática ou, da sua irmã, a Física.

Espero poder contribuir para ajudar a despertar, sobretudo nos jovens, a apreciar e quem sabe, vir a gostar a aprender esta ciência maravilhosa chamada Matemática.

Essa deve ser a meta dos educadores (ainda que estejamos fora de sala de aula) é também nosso, esse desejo de contribuir, de socializar pois assim, acreditamos que um dia, poderemos ajudar a formar mais cidadãos de bem, conscientes do seu papel e sedentos pelo prazer de aprender e sobretudo, de compartilhar.

Fora as aulas de reforço, as feiras de matemática parecem ser as melhores oportunidades para ajudar a disseminar esse despertar.

Nota Importante: creio que se não tivéssemos persistido no seu desenvolvimento, dificilmente teríamos descoberto as falhas insuspeitáveis das **“calculadoras científicas”** apesar de marcas mundialmente consagradas e muito menos das lacunas existentes nas regras de potenciação.

Gostaríamos de deixar registrado o nosso endosso junto com tantos outros, às autoridades da educação, que revisem as grades curriculares e que fizessem um esforço para incluírem a **Lógica de Programação** nas grades curriculares o quanto antes.

Além de não ter contra-indicações, ajudará enormemente a desenvolver o raciocínio lógico, independentemente da carreira ou da opção profissional futura do jovem estudante.

Possibilitará aos egressos, poderem integrarem-se mais rapidamente e desenvolverem-se neste universo de expansão tecnológica cada vez mais acelerada e inexorável.

Eles precisarão muito disso e, todos, de uma forma ou outra, serão beneficiados.

A concretização de ambos somente foram possíveis graças aos conhecimentos dos conceitos fundamentais que aprendemos nas aulas do ensino básico de Matemática....

■ AGRADECIMENTOS

Agradeço a Deus por me dar essa possibilidade e, a todos que de uma forma ou outra, foram mestres em algum momento de minha vida provendo-nos de luz nas brumas de nossos e medos e incertezas, indicando-nos o Norte e, a tantos outros, que nos deram suporte em momentos difíceis. Também àqueles que compartilhamos alguns sonhos e ideias para um dia melhor.

Todos vocês fizeram a diferença.

Tenham certeza: os seus ensinamentos não foram em vão.

A todos, o meu sincero reconhecimento e eterna gratidão.

Sinto-me imensamente gratificado por esta oportunidade.

“Um país se faz com homens e livros”.

Monteiro Lobato

■ REFERÊNCIAS

1. **Apostila de Operações de Números Binários, Octais e Hexadecimais**, Lanterna na popa, Campos Novos /SC.
2. **Apostila de Lógica**, Lanterna na popa, Campos Novos /SC.
3. CIARCIA, Steve – **Construa seu próprio Microcomputador**; McGraw-Hill, 1984, São Paulo
4. COLLINS, William J. – **Programação Estruturada com Estudos de Casos em Pascal**, McGraw-Hill, 1988, São Paulo
5. ELENBERG, Jordan – **O Poder do Pensamento Matemático**, Zahar, 2015, São Paulo
6. FLANDERS, Harley – **Scientific Pascal**; Bikhäuser, Michigan University, 1996, Boston, EUA
7. MENDELSON, Elliott - **Álgebra Booleana e Circuitos de Chaveamento**, McGraw-Hill, 1977, São Paulo
8. SANTOS, Jean Patrick Figueiredo dos - **Desenvolvendo Aplicações com Lazarus**, Agbook, 2011

Sites na Internet consultados:

9. <http://revistaeducacao.uol.com.br/textos/211/aposta-no-futuroo-ensino-de-programacao-tem-se-espalhado-como-330266-1.asp> “**Ensino da programação é a aposta de colégios em todo o mundo**” Postado em Novembro de 2014 pela Luciana Alvarez Acesso em: 12/07/2015
10. <http://www.computacaonaescola.ufsc.br/> “**Computação nas escolas**” Acesso em: 12/07/2015
11. <http://idgnow.com.br/ti-pessoal/2015/04/07/projeto-do-parana-quer-levar-ensino-de-programacao-a-escolas-do-brasil/>
12. “Projeto no Paraná quer levar o ensino da programação a escolas do Brasil” Postado em: 07/04/2015 por Carla Matsu Acesso em: 12/07/2015

SOFTWARE LIVRE NA GESTÃO DE TI: Um Estudo de Caso Aplicado ao Campus Paraíso do Tocantins do IFTO

| Kálita Fernandes da Silva **Fonteles**

| Gislaine Pereira **Sales**

RESUMO

Este trabalho investiga a adoção de um software livre para gestão da TI em uma instituição pública, através de um estudo de caso aplicado ao Instituto Federal de Educação, Ciência e Tecnologia Campus Paraíso do Tocantins. Para tanto, foram realizados o levantamento dos processos de TI existentes na Coordenação de Gestão de Tecnologia da Informação (CGTI) da Instituição, o que tornou possível a identificação das suas necessidades. O estudo e a identificação de boas práticas de gestão de TI a serem utilizadas em conjunto com o software, como uma forma de melhorar e ampliar os resultados obtidos, a análise comparativa dos softwares de gestão de TI identificados na literatura, com o intuito de selecionar o software que melhor atendesse aos requisitos da pesquisa, além de identificar e configurar a melhor infraestrutura para a implantação do software e realizar o acompanhamento e treinamento dos usuários após a sua instalação com a elaboração de um manual para auxiliar o usuário e aula expositiva para sanar eventuais dúvidas sobre o software. Os resultados mostram que o software livre pode ser utilizado para gerir Tecnologia da Informação de Instituições Públicas atendendo de forma satisfatória as suas necessidades e ainda contribuindo para agregar valor aos seus serviços.

Palavras-chave: Gestão de TI, Software Livre, Melhores Práticas de Gestão de TI.

■ INTRODUÇÃO

Por muito tempo as organizações mantiveram seus negócios com pouco ou nenhum apoio da TI. No entanto, com todos os avanços tecnológicos e diante de um mundo globalizado a realidade hoje é muito diferente. Hoje a TI é o diferencial competitivo dentro das organizações (CRISTO; FERREIRA; FURTADO, 2013).

No IFTO campus Paraíso do Tocantins foi detectado um aumento considerável no número de alunos, de laboratórios, de computadores, e conseqüentemente o aumento de serviços e equipamentos de TI para atender as novas demandas de apoio ao ensino. Desta forma, foi observada a necessidade de implantação de um software de gestão de TI que pudesse apoiar a Coordenação de Gestão de Tecnologia da Informação (CGTI), desta instituição, na gestão dos recursos e serviços de TI, fazendo com que a TI possa agregar mais valor à instituição, no que diz respeito ao ensino. Foi levado em consideração o fato de o IFTO ser uma instituição pública, e como tal procura seguir a orientação do governo federal no que refere à adesão de software livre para sistematizar os seus processos de trabalho (BRASIL, 2012).

Neste cenário, a opção por um software livre torna-se uma solução viável, já que o mesmo possui as mesmas funcionalidades de um software proprietário, além de código-fonte aberto e licença sem custo na maioria das vezes, o que possibilita a alteração e adaptação às necessidades da organização, reduzindo assim o tempo e o custo de implantação.

Assim, percebendo a importância da gestão de TI e como afirma Campos (2006) o constante aumento do uso de software livre, este trabalho objetiva investigar se o uso desse tipo de software vai apoiar a implementação de uma gestão da TI mais eficiente na instituição investigada. Para tanto, realizou-se o mapeamento dos processos de TI da CGTI para a identificação dos seus requisitos, e ainda, o estudo comparativo entre softwares livres de gestão de TI identificados durante a pesquisa, com intuito de identificar o que melhor atendesse às necessidades da instituição. Além disso, foi estudada e configurada a melhor infraestrutura de TI para a implantação do software, identificadas as melhores práticas de gestão de TI a serem implementadas com auxílio do software, e por fim analisados os resultados obtidos com a implantação do software no que diz respeito à gestão de TI na instituição, propondo ações corretivas quando necessário.

■ MATERIAIS E MÉTODOS

Esta pesquisa possui caráter exploratório, com uma forma de abordagem quali-quantitativa.

Segundo Gil (2002, p.41) as pesquisas exploratórias “têm como objetivo proporcionar maior familiaridade com o problema, com vistas a torná-lo mais explícito ou a construir hipóteses”. Neste aspecto, a presente pesquisa é classificada como exploratória, pois se utiliza como procedimentos técnicos de coleta de dados, o estudo de caso, a revisão bibliográfica e o *survey*, que propõe um questionário aos usuários dos serviços de TI na instituição investigada para medir a satisfação dos mesmos no que diz respeito à sua gestão de TI. O questionário foi aplicado aos técnicos, professores e bolsistas, no total de 25 pessoas. Para o embasamento teórico e revisão de literatura, foram utilizados como fontes de coleta, dados do tipo bibliográfico, disponíveis em livros e em artigos científicos disponíveis na Internet.

Quanto à forma de abordagem quali-quantitativa, o trabalho utilizou dados qualitativos na avaliação das características da solução aplicada e das práticas utilizadas e seu significado, e quantitativo para valorar o nível de satisfação no ambiente pesquisado.

Foi utilizada a técnica *Business Process Modeling Notation* (BPMN), ou Notação de Modelagem de Processos de Negócio, para a realização do mapeamento de processos em conjunto com a ferramenta *Bizagi Process Modeler*. E, realizadas entrevistas com o coordenador da CGTI para obter informações necessárias para o mapeamento de processos. Para este trabalho foi utilizada a entrevista parcialmente estruturada que é guiada pela relação de pontos de interesse que o entrevistador vai explorando ao longo de seu curso (GIL, 2002, p.117). Para a seleção do software foram utilizadas como pré-requisitos as seguintes características: ser software livre; ser software de gestão de TI; e atender aos requisitos da instituição que foram obtidos através do mapeamento de processos. Esses pré-requisitos foram aplicados a essa pesquisa por se entender que são totalmente necessários para a identificação de um software que realmente venha atender as necessidades da Instituição.

■ RESULTADOS E DISCUSSÃO

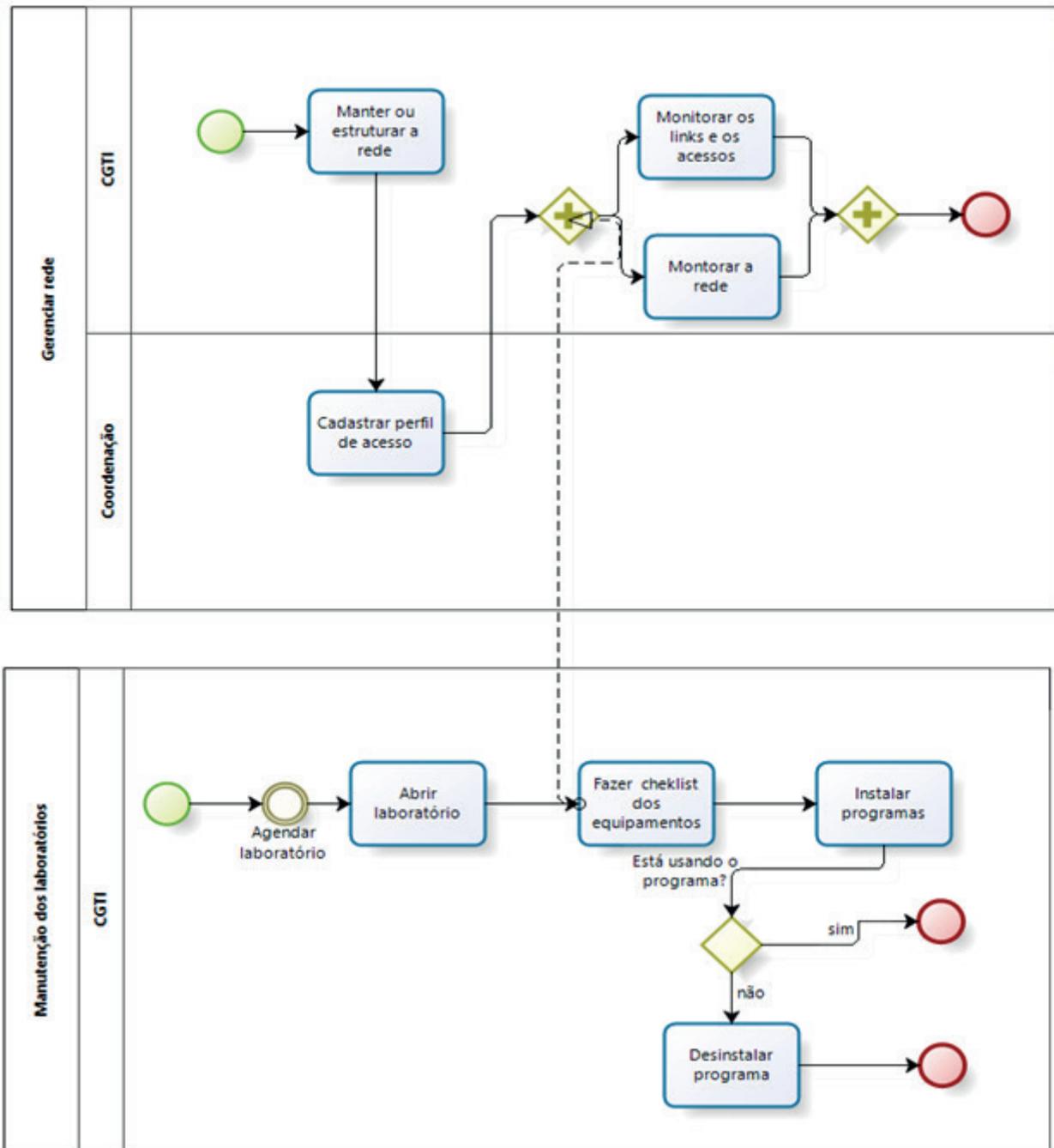
Mapeamento dos Processos de TI da CGTI do IFTO Campus Paraíso – TO

O mapeamento dos processos da CGTI foi realizado com o intuito de apresentar todos os processos existentes e a forma que são realizados atualmente. Este mapeamento buscou identificar as tarefas que podem ser realizadas de forma mais rápida e automática através do uso de um software de gestão de TI.

A partir da identificação dessas tarefas foram listados os requisitos mínimos que o software deveria atender para ser selecionado. Esses requisitos foram utilizados para o estudo comparativo dos softwares. Devido à limitação de páginas desta formatação não serão apresentados aqui todos os processos mapeados, mas os três processos apresentados são os que permitem identificar o maior número de requisitos.

De acordo com a figura 1, nos processos gerenciar rede e manutenção dos laboratórios foi possível identificar os requisitos para gestão de inventário de hardware e software, e gestão de reservas para que o agendamento do uso de laboratórios e equipamentos fosse feito através do sistema.

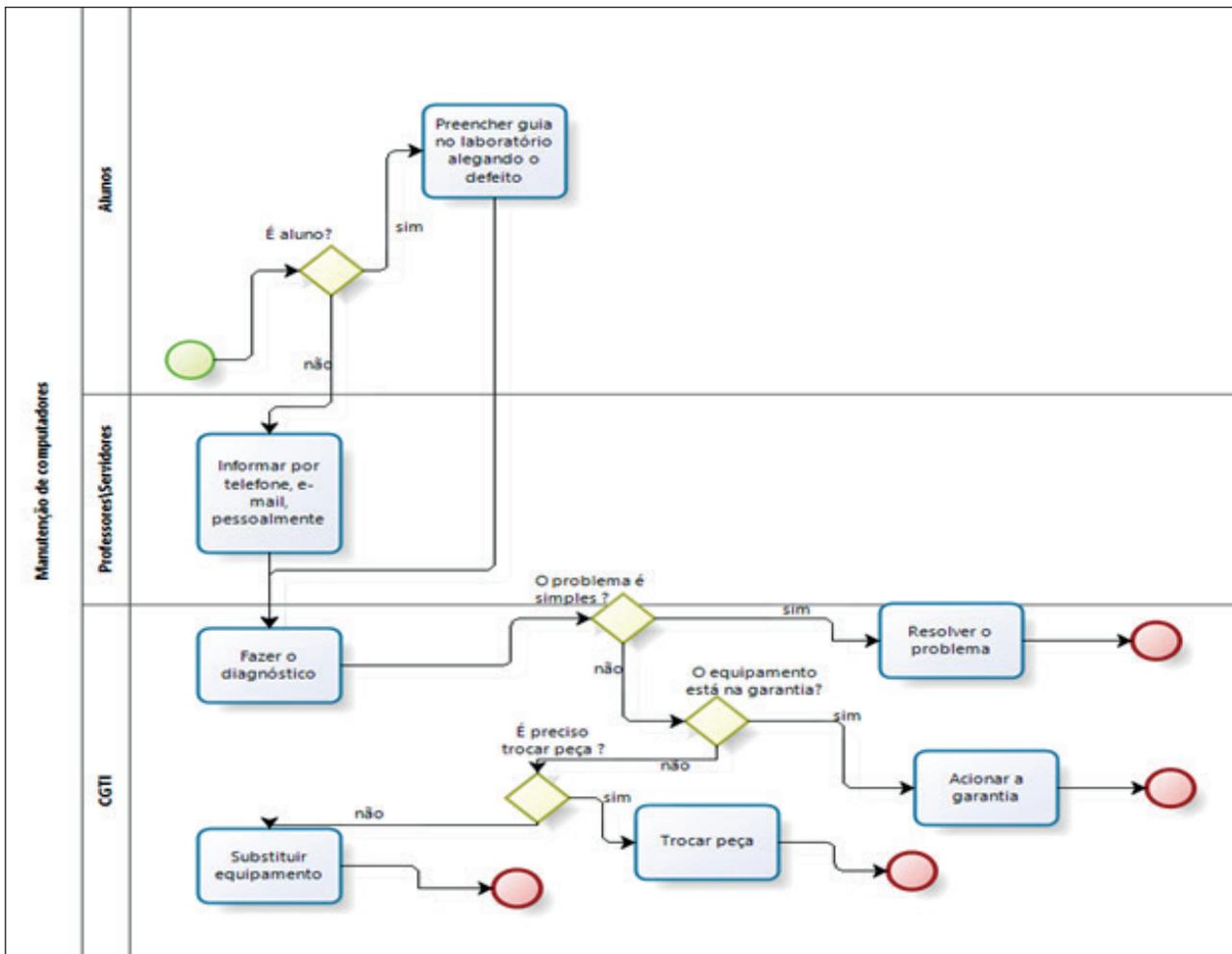
Figura 1. Processos Gerenciar rede e Manutenção dos laboratórios. IFTO, 2014.



A figura 2 mostra o processo de manutenção de computadores, o qual envolve várias atividades. Nesse processo podem-se identificar quatro requisitos: gestão de chamados e *help desk* que permite ao usuário solicitar o atendimento e à coordenação de TI obter as informações pertinentes a cada chamado; gestão de inventário de hardware e software; atualizações automáticas de informações e configurações de hardware e software; e os relatórios

gerenciais que permite a CGTI obter informações sobre os equipamentos, softwares estão instalados nos computadores, garantia, entre outros, auxiliando assim à tomada de decisão e resolução de problemas.

Figura 2. Processo Manutenção de computadores. IFTO, 2014.



■ ANÁLISE COMPARATIVA DOS SOFTWARES

Com a realização do mapeamento de processos foi possível identificar os requisitos, que são as características que os softwares deverão possuir para sanar as necessidades da Instituição. O software no qual se identificou todas essas características ou o maior número delas foi o escolhido.

A tabela 1 apresenta a comparação dos softwares. Como se pode observar, o CACIC atende a 3 requisitos, o OCOMON atende a 4 requisitos e o OCS – NG atende a 2, mas somente o GLPI abrange 5 dos 6 requisitos e se combinado ao OCS – NG atende a todos os requisitos. Por esse motivo concluiu-se que ele é o software que melhor atendia as necessidades da CGTI.

Tabela 1. Análise Comparativa dos Softwares. IFTO, 2014.

REQUISITOS	CACIC	OCOMON	OCS - NG	GLPI
Gestão de chamados	Não	Sim	Com o GLPI	Sim
Gestão de inventário de hardware e software	Sim	Sim	Sim	Sim
Atualização automática de informações e configurações de hardware e software	Sim	Não	Sim	Com o OCS NG
Gestão de Reservas	Não	Não	Com o GLPI	Sim
Geração de relatórios gerenciais	Sim	Sim	Com o GLPI	Sim
Help Desk	Não	Sim	Com o GLPI	Sim

■ INSTALAÇÃO DO GLPI

Para efetuar a instalação e configuração do GLPI foram observados alguns requisitos, os quais são:

- Apache 2;
- PHP 5 ou superior;
- MySQL;
- HTML e CSS.

Com a instalação do GLPI foi possível observar na prática a eficiência de suas funcionalidades, como exemplo o *Help desk*, que permite uma gestão de chamados eficiente, ou seja, permite ao usuário abrir chamados no sistema para solicitar algum serviço/atendimento ou informar um incidente. Ao gestor de TI, permite fazer todo o acompanhamento do chamado desde sua abertura até o seu encerramento, fornecendo gráficos e relatórios dos chamados contendo informações valiosas para uma gestão eficiente como, por exemplo: quantos chamados foram abertos em um determinado período, quais as origens dos chamados e quantos estão em situação “abertos”, “solucionados”, “fechados”, entre outros.

Na funcionalidade de gestão de inventário todas as características dos equipamentos conectados à rede da instituição são armazenadas em uma base de dados, o que permite ao administrador da rede ter uma visão completa dos ativos de TI e tenha à sua disposição informações importantes como: hardwares e softwares instalados, licença dos softwares, espaço de armazenamento utilizado em disco, entre outras (INDEPNET, 2014).

Após a instalação do sistema, com todas as funcionalidades relacionadas à gestão de chamados funcionando corretamente e à sua utilização durante um determinado período de tempo, foi elaborado um questionário para medir o nível de satisfação dos usuários com

relação aos serviços de TI oferecidos e para avaliar também qual a importância em adicionar novos serviços ou otimizar os já existentes através do uso das demais funcionalidades do GLPI que ainda deveriam ser implantadas. Em relação aos serviços já implantados o questionário media o nível de satisfação entre “muito bom”, “bom”, “regular” e “ruim”. Sobre a implantação de novos serviços, o questionário media o nível de importância de serem disponibilizados na instituição em “muito bom”, “bom”, “não faria diferença” e “ruim”.

Com a aplicação do questionário pode-se observar que a implantação do módulo de gestão de chamados trouxe resultados positivos, pois 48% dos que responderam ao questionário avaliaram os serviços em geral como bom e 36% consideraram a abertura de chamados como muito bom. Quanto à implantação de novos serviços, como empréstimos e reserva de equipamentos, inventário, custos, base de conhecimento com informações sobre solução de problemas e incidentes ocorridos, entre outros, ao analisar os resultados do questionário ficou evidente que os usuários consideram importante à implantação de novos serviços, mas do ponto de vista destes usuários, alguns serviços se mostraram mais importantes que outros. Então, foi priorizada a implantação dos que foram considerados mais importantes. A questão sobre registro de custos de hardware e software e, cadastro e gerenciamento de contatos, fornecedores, clientes e contratos, por exemplo, receberam a pontuação “não faria diferença” por 32% e 36% respectivamente dos que responderam ao questionário, portanto foi dado prioridade para os outros serviços e esses ficaram para uma possível implantação futura.

Com isso foi implantada a base de conhecimento do GLPI, onde foi realizado o levantamento dos problemas mais recorrentes na Instituição, e assim foram elaborados pequenos textos de fácil compreensão para auxiliar o usuário na resolução desses problemas corriqueiros.

O módulo de empréstimos é uma parte do GLPI voltada para a reserva de itens. Através desta funcionalidade tem-se a possibilidade de fazer um gerenciamento de reservas adequado, obtendo informações sobre quais itens podem ser reservados, quais estão reservados e os seus detalhes como data da reserva, hora, duração, entre outras. Inicialmente todos os itens que poderiam ser reservados foram cadastrados no sistema com seus dados básicos de identificação, e depois disponibilizados para reserva.

A ativação do módulo de empréstimos permitiu aos professores e servidores fazerem suas reservas sem precisar sair de casa, pois o GLPI é um sistema web e oferece a facilidade de ser acessado de qualquer lugar, desde que se tenha acesso à Internet. Pensando em como facilitar ao máximo a inclusão do sistema no dia a dia dos usuários foi elaborado um manual e oferecido treinamento para sanar eventuais dúvidas sobre o uso do sistema.

O manual do usuário foi elaborado visando abranger todas as atividades que um usuário comum realiza no GLPI entre elas: abrir chamado, fazer reserva, acessar a Base de conhecimento, etc. O manual traz o passo a passo de como realizar essas atividades com várias ilustrações para facilitar a compreensão do usuário. O manual se encontra na base de conhecimento do GLPI.

O GLPI pode ser acessado através do site do Campus Paraíso do Tocantins do IFTO clicando em suporte técnico ou pelo link: <https://paraiso.ifto.edu.br/suporte/> e em seguida e só entrar com o seu usuário e senha e ter acesso a todas as funcionalidades oferecidas pelo sistema.

■ O GLPI E O USO DE BOAS PRÁTICAS: ITIL®

ITIL® – *Information Technology Infrastructure Library* – é um conjunto de livros que contém recomendações e boas práticas para gerenciamento de serviços de TI (FREITAS, 2010, p.57). Esta biblioteca foi desenvolvida pelo governo britânico no final da década de 1980, e em meados da década de 1990 foi reconhecida mundialmente como um padrão de fato para gerenciamento de serviços (CARVALHO; CLEMENTI; SORTICA, 2004). De acordo com Freitas (2010) no ITIL® V3 é apresentado o ciclo de vida do serviço ITIL® que é dividido em cinco ciclos, onde cada ciclo é detalhado em um livro, os quais são: *Service Strategy* (Estratégia de Serviço); *Service Design* (Desenho de Serviço); *Service Transition* (Transição de Serviço); *Service Operations* (Operação de serviço); e *Continual Service Improvement* (Melhoria Continuada de Serviço).

Ao estudar os cinco ciclos da ITIL®, foi possível perceber que o ciclo “Operação de Serviço” é mais voltado para gerenciamento e suporte de serviço e que o GLPI oferece mecanismos para a implantação de praticamente todos os seus processos e funções, os quais são:

- **Gerenciamento de Eventos:** gera alertas ou notificações de um serviço de TI ou Item de Configuração.
- **Gerenciamento de Incidentes:** através do GLPI podem-se executar todas as atividades de gerenciamento de Incidentes.
- **Cumprimento de Requisição:** o GLPI oferece mecanismos para a realização de todo o processo de cumprimento de requisição.
- **Gerenciamento de Problemas:** com o GLPI é possível fazer a Detecção, o Registro, a Categorização, e a Priorização do Problema. É possível também realizar a Investigação e Diagnóstico, o registro do Erro Conhecido através da base de conhecimento (FAQ), o registro da Resolução do Problema, do Fechamento do Problema e a Revisão de Problemas Graves.

- **Central de Serviço:** o GLPI pode ser um forte aliado para a Central de Serviço, podendo ser o único ponto de contato para o usuário e auxiliando a restauração da operação o mais rápido possível oferecendo informações precisas e em tempo real.
- **Gerenciamento Técnico:** a documentação e o Gerenciamento dos conhecimentos em TI podem ser feito através da base de conhecimento (FAQ) do GLPI.
- **Gerenciamento da Operação de TI:** através do GLPI é possível realizar o gerenciamento de manutenção e infraestrutura de TI para que a entrega dos serviços seja feita na data acordada, e contribui para que seja mantida a qualidade dos serviços.

Com isso, conclui-se que o GLPI contribui significativamente para a adoção de boas práticas de gestão de TI, o que auxilia a Instituição a alinhar a TI aos seus objetivos de negócio e contribui para a melhoria dos serviços oferecidos. Pois segundo Carvalho, Clementi e Sortica (2004) os serviços de suporte do ITIL® auxiliam no atendimento das necessidades do cliente, apoiando, desta forma, aos seus objetivos de negócios.

E, ainda, de acordo com Magalhães e Pinheiro (2007) a ITIL® promove uma abordagem qualitativa para o uso econômico, efetivo, eficaz e eficiente da infraestrutura de TI, com o intuito de obter vantagens para a organização tanto em termos de redução de custos pelo aumento da eficiência na entrega e suporte dos serviços de TI quanto de aumento da capacidade da organização de gerar receita, permitindo que a área concentre seu esforço em novos projetos para o atendimento à estratégia de negócio da organização.

■ CONCLUSÕES

A importância da Gestão da Tecnologia da Informação vem sendo cada vez mais percebida pela alta administração das organizações, as quais estão investindo mais na área e em contrapartida cobrando mais resultados positivos. A gestão da TI propicia uma visão mais ampla do todo e concentra seus esforços para atingir os objetivos estratégicos da organização, contribuindo de forma significativa para o sucesso da mesma.

O software livre hoje já é bem aceito e suas liberdades adicionais reconhecidas como um fator crucial na hora de optar entre um software proprietário e um software livre. As sociedades desenvolvedoras de software livre estão se aperfeiçoando constantemente, o que resulta em um produto de mais qualidade, com um número maior de funcionalidades e com uma interface cada vez mais amigável ao usuário.

Com o mapeamento dos processos foi possível conhecer melhor os processos da Instituição investigada e estabelecer os requisitos que os softwares deveriam atender para

sanar as suas necessidades. A análise comparativa dos softwares tornou possível a identificação e escolha do software que atendia ao maior número de requisitos da Instituição.

Como a implantação de um software muda a rotina de quem vai passar a trabalhar com ele, sendo o mesmo desconhecido por parte dos usuários, a elaboração do manual e treinamento com intuito de auxiliar o usuário nesse processo, foi essencial para diminuir os impactos negativos dessa mudança.

Assim, a adoção do software livre de Gestão de TI vem se mostrando eficiente e capaz de suprir as necessidades da Instituição investigada. E, como o GLPI possibilita a implantação da ITIL®, à sua adoção potencializou ainda mais os resultados obtidos. Pois agora a Instituição passa a contar com um software de Inventário para itens tecnológicos com informações precisas e detalhadas sobre os equipamentos conectados a rede e de *Help Desk* para suporte ao usuário proporcionando maior rapidez na detecção e resolução de problemas. Os professores e servidores podem fazer reservas de materiais ou dos laboratórios através do software sem precisar sair de casa. E ainda, sendo utilizado seguindo as recomendações da ITIL. Tudo isso só tem agregado mais valor e contribuído para que seja oferecido ao Campus um serviço de TI com cada vez mais qualidade e rapidez.

Como proposta de trabalhos futuros sugere-se o estudo comparativo da produtividade da CGTI entre os períodos anterior e posterior a implantação do software para se concluir efetivamente se a adoção do software aumentou a produtividade do local.

■ REFERÊNCIAS

1. BRASIL. (2012) – Planejamento Cisl 2013 – 2014. Disponível em: <<http://www.softwarelivre.gov.br/planejamento-2013-2014>>. Acesso em: 22 de Outubro de 2014.
2. CARVALHO, Tereza Cristina M. B.; CLEMENTI, Sergio; SORTICA, Eduardo Almansa (2004). Governança de TI: comparativo entre COBIT e ITIL®. Disponível em: <http://www3.fsa.br/LocalUser/gestaoti/Ativ09_CLEMENTI_2004_Governança_de_TI_-_Comparativo_entre_Cobit_e_ITIL®.pdf>. Acesso em 14 de Abril de 2014.
3. CRISTO, Wallace Borges; FERREIRA, João Carlos P.; FURTADO João Paulo C. (2013). Gestão de TI nas Organizações Contemporâneas. <http://revistapensar.com.br/tecnologia/pasta_upload/artigos/a23.pdf> Acesso em 14 de Abril de 2014.
4. CRUZ, Fábio Rodrigues (2013). Introdução a Modelagem de Processos utilizando BPMN. Disponível em: <http://www.fabiocruz.com.br/wpcontent/uploads/2013/03/paper_BPMN.pdf> Acesso em 19 de Junho de 2014.
5. FREITAS, Marcos André dos Santos (2010). Fundamentos do Gerenciamento de serviços de TI: preparatório para a certificação ITIL® V3 Foundation. Rio de Janeiro: Brasport, 2010.
6. GIL, Antonio Carlos (2002). 1946- Como elaborar projetos de pesquisa/Antônio Carlos Gil. - 4. ed. - São Paulo: Atlas, 2002.

7. INDEPNET (2014). O PROJETO GLPI. Disponível em:<<http://www.glpi-project.org/?page=contact&lang=en>> Acesso em 19 de Maio de 2014.
8. “ITIL® é uma marca registrada do The Cabinet Office no Reino Unido e em outros países”.
9. MAGALHÃES, Ivan Luizio; PINHEIRO, Walfrido Brito (2007). Gerenciamento de Serviços de TI na Prática Uma abordagem com base na ITIL® Inclui ISO/IEC 20.000 e IT Flex®. São Paulo: Novatec, 2007.

Uma abordagem ubíqua para acompanhamento clínico de pacientes

| Alexandre Renato Rodrigues de **Souza**
IFRS/ UFPEL

| João Ladislau Barbará **Lopes**
IFSUL

| Fabrício Neitzke **Ferreira**
IFSUL

| Leonardo Costa **Reichow**
LIFEMED

| Ana Marilza Pernas **Fleischmann**
UFPEL

| Adenauer Correa **Yamin**
UFPEL

RESUMO

Objetivo: Este artigo propõe uma arquitetura de software direcionada à Computação Ubíqua (UbiComp), explorando a infraestrutura provida pela Internet das Coisas, para prover ciência da situação atual de pacientes em ambiente hospitalar. **Resultados:** Como contribuição central, este trabalho provê o acompanhamento da evolução dos sinais vitais dos pacientes, permitindo que o médico possa observar se está sendo atingido o efeito desejado com a administração dos medicamentos. Na perspectiva da arquitetura desenvolvida, esse efeito procurado pelos médicos nos parâmetros vitais é denominado de Meta Terapêutica. A avaliação da arquitetura foi feita através de um estudo de caso, destacando as tecnologias empregadas e os protótipos desenvolvidos. **Conclusão:** A concepção de uma arquitetura para o acompanhamento dos sinais vitais em função da medicação, se mostra uma ferramenta oportuna para avaliar a resposta do paciente ao tratamento, bem como para enviar alertas aos profissionais de saúde quando os valores dos sinais vitais indicarem uma resposta inesperada ao medicamento. Este pronto aviso à equipe médica de que a Meta Terapêutica não foi atingida, constitui uma das contribuições centrais do EXEHDA-TG para agilização das práticas relacionadas ao tratamento de pacientes, com potencial contribuição para que o tempo de internação seja reduzido.

Palavras-chave: Saúde Ubíqua, Ciência de Situação, Meta Terapêutica, Internet das Coisas.

■ INTRODUÇÃO

Na Ubicomp, os sistemas computacionais devem ser capazes de reagir às mudanças do estado das diferentes variáveis contextuais de seu interesse, as quais devem ser coletadas em ambientes largamente distribuídos (SÁNCHEZ GUINEA; NAIN; LE TRAON, 2016). Nesse sentido, os recentes avanços científicos e tecnológicos na área da Internet das Coisas (*Internet of Things – IoT*) têm proporcionado o uso de sensores em larga escala, os quais constituem fontes geradoras de informações contextuais para as aplicações Ubíquas conscientes do contexto (SEZER; DOGDU; OZBAYOGLU, 2018].

O suporte para obtenção da situação dos pacientes é baseado na percepção de sinais vitais dos mesmos, a qual é provida por meio de dados sensorizados. Os sensores coletam dados de contexto dos pacientes, provendo suporte para que o sistema de Computação Ubíqua possa fornecer serviços que reajam de acordo com o contexto. Nesse sentido, a ciência de situação é uma forma promissora de ampliar os modelos tradicionais de monitoramento e alerta, e que pode auxiliar a desenvolver aplicações mais robustas e capazes de melhor atender às necessidades da comunidade em um ambiente hospitalar (TEMDEE; PRASAD, 2018).

Desta forma, este trabalho propõe a concepção de uma arquitetura de software para um ambiente ubíquo direcionada a prover ciência da situação atual de pacientes. Essa arquitetura, denominada EXEHDA-TG, acompanha a evolução dos sinais vitais, permitindo que o médico possa confirmar se está sendo atingido o efeito desejado com a administração dos medicamentos prescritos. Esse comportamento desejado pelos médicos nos parâmetros vitais é, na arquitetura proposta, chamado de Meta Terapêutica. Como contribuição central, a arquitetura desenvolvida automatiza a aquisição e viabiliza a avaliação dos sinais vitais com base no conhecimento dos médicos, os quais especificam a Meta Terapêutica a ser perseguida. O EXEHDA-TG possibilita o envio de alertas aos profissionais de saúde em uma perspectiva ubíqua, provendo mobilidade no acesso às informações médicas.

A arquitetura proposta tem como premissa sua integração ao *middleware* EXEHDA (*Execution Environment for Highly Distributed Applications*). O EXEHDA provê uma arquitetura de software que visa criar e gerenciar um ambiente ubíquo, bem como promover a execução, sob este ambiente, das aplicações da computação ubíqua (LOPES et al. 2014).

■ MÉTODO

O EXEHDA-TG mantém a proposta organizacional do *middleware* EXEHDA, sendo composto de dois tipos principais de servidores: (i) Servidor de Integração: responsável pela relação com o meio através de sensores e atuadores; e (ii) Servidor de Avaliação: responsável pelo armazenamento e processamento das informações contextuais referentes aos

pacientes, coletadas do ambiente ubíquo provido pelo EXEHDA, que explora a infraestrutura provida pela IoT.

O ambiente ubíquo gerenciado pelo EXEHDA tem uma organização celular, onde cada célula pode empregar vários Servidores de Integração, sendo que cada um provê dados a um Servidor de Avaliação, e pode coletar dados de vários leitos. Uma célula pode atender um ou mais hospitais e, por sua vez, um hospital de grande porte pode ter mais de uma célula.

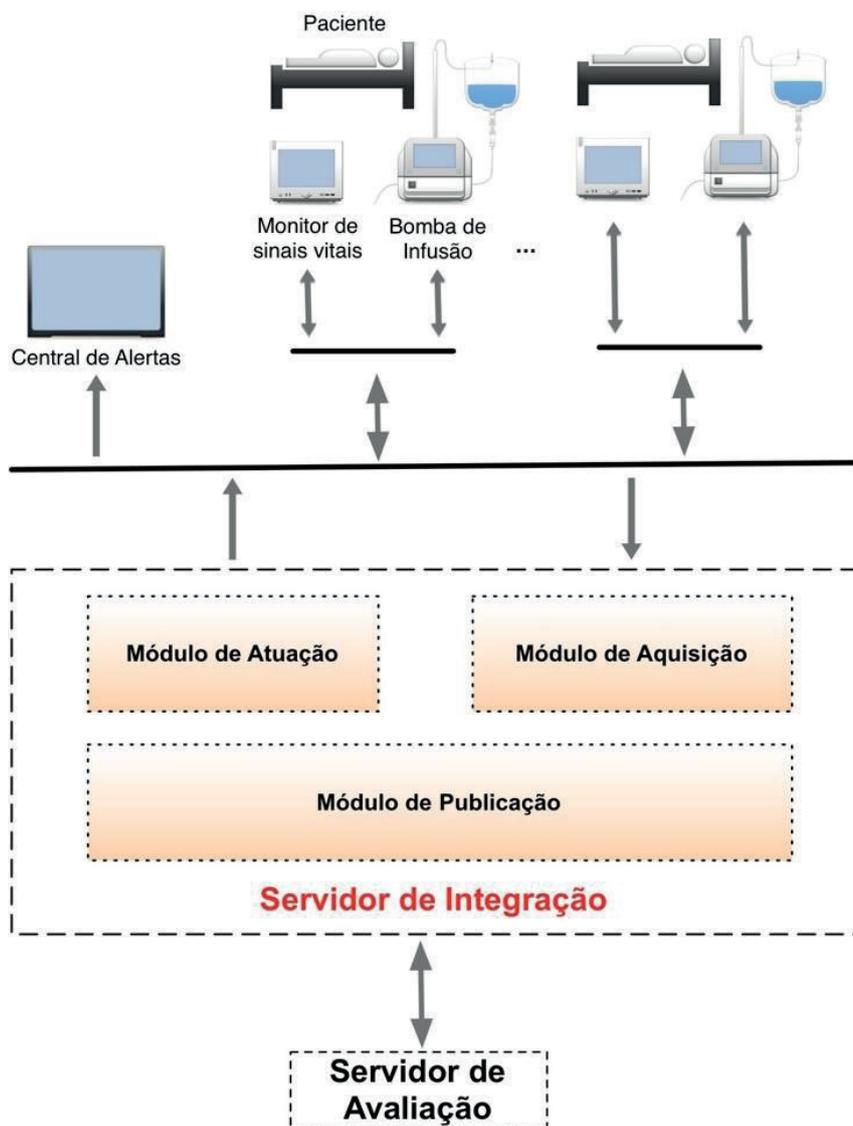
O EXEHDA-TG tem como objetivo fornecer uma solução capaz de: (i) coletar de forma automatizada informações contextuais dentro do ambiente hospitalar; (ii) processar estas informações contextuais coletadas; e (iii) prover alternativas de atuação sobre o meio e permitir o envio de alertas a dispositivos móveis.

A descrição do EXEHDA-TG, feita a seguir, está organizada a partir de seus servidores e respectivas funcionalidades, sendo realizadas as necessárias associações entre os mesmos, e com os outros serviços do *middleware* EXEHDA.

Servidor de Integração

O **Servidor de Integração** foi concebido para prover as funcionalidades necessárias para interfaceamento com os equipamentos eletromédicos, viabilizando no meio ubíquo a aquisição de dados através de sensores, bem como os pertinentes acionamentos de atuadores. Na Figura 1 é apresentada a arquitetura desse servidor.

Figura 1. Servidor de Integração do EXEHDA-TG



O **Módulo de Aquisição** é responsável pela coleta e tratamento dos dados de sensoriamento das bombas de infusão e monitores de sinais vitais. Este módulo realiza a interface para captura das informações dos sensores dos equipamentos eletromédicos, considerando os protocolos de comunicação necessários para cada tipo de dispositivo. Este módulo emprega as facilidades do EXEHDA de utilizar a IoT como estratégia de intercomunicação. A leitura dos sensores é realizada de acordo com o agendamento definido pelas Metas Terapêuticas. Os dados coletados são tratados em etapas, através das quais é feita sua normalização e a definição de quais dados devem ser efetivamente publicados no Servidor de Avaliação.

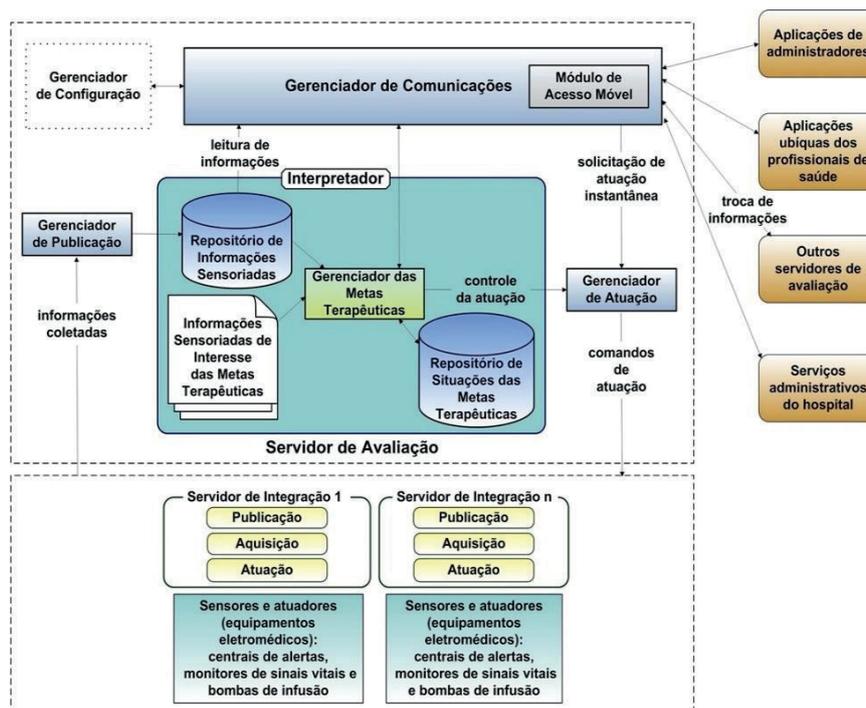
O **Módulo de Publicação** é responsável por coordenar o principal fluxo de dados entre os Servidores de Integração e o Servidor de Avaliação, promovendo a publicação dos dados coletados e garantindo uma persistência local dos mesmos quando a comunicação entre estes servidores estiver temporariamente indisponível.

O **Módulo de Atuação** tem como função acionar os dispositivos dos equipamentos eletromédicos responsáveis por apresentar alertas visuais e sonoros aos profissionais de saúde, tais como: *displays* de cristal líquido (LCD - *Liquid Crystal Display*), *buzzers*, altofalantes, LEDs (*Light-Emitting Diode*), etc. Portanto, as mensagens apresentadas nas telas dos equipamentos eletromédicos são consideradas um tipo específico de atuação.

Servidor de Avaliação

O **Servidor de Avaliação** foi desenvolvido para ser o responsável por receber valores de contexto coletados dos pacientes e dos equipamentos eletromédicos, armazenar os mesmos e disparar avaliações situacionais solicitando atuações e alertas sempre que necessário. A Figura 2 mostra uma visão geral desse servidor, onde é apresentada a sua relação com os Servidores de Integração, aplicações de administradores, aplicações ubíquas dos profissionais de saúde, outros servidores de avaliação e serviços administrativos do hospital.

Figura 2. Servidor de Avaliação do EXEHDA-TG



Os módulos que compõem o Servidor de Avaliação são responsáveis pela ciência de situação, desde a aquisição de dados contextuais até o acionamento de atuadores e o envio de alertas. A seguir está a descrição destes módulos.

O **Gerenciador de Configuração**, através das aplicações de administradores, provê as funcionalidades necessárias para realizar a parametrização e gerenciamento do Servidor de Avaliação, considerando as características dos sensores e atuadores dos equipamentos eletromédicos. Este módulo armazena, em um repositório próprio, a programação das

bombas de infusão, definições das Meta Terapêuticas e os parâmetros de aquisição para os dispositivos eletromédicos.

O **Gerenciador de Comunicações** é composto por um barramento ESB (*Enterprise Service Bus*) e um Módulo de Acesso Móvel, os quais possuem as funcionalidades necessárias para atendimento das aplicações distribuídas dos usuários. As aplicações ubíquas dos profissionais de saúde e outros Servidores de Avaliação utilizam o Gerenciador de Comunicações para solicitar dados de sinais vitais, *status* e parâmetros dos equipamentos eletromédicos e para acionar atuadores. Essas aplicações ubíquas dos profissionais de saúde estão em execução nos diferentes nodos que integram a estrutura computacional ubíqua provida pelo *middleware* EXEHDA.

Com a intenção de prover acesso em deslocamento ao EXEHDA-TG foi concebido o **Módulo de Acesso Móvel**, o qual está organizado em dois blocos.

Bloco de Exibição de Informações Contextuais: executa no dispositivo móvel e disponibiliza ao usuário relatórios, tanto gráficos como textuais, acerca das informações contextuais de seu interesse. A comunicação do bloco com o Servidor de Avaliação acontece em duas etapas: (i) é feita uma inspeção das informações de contexto que estão sendo tratadas; e (ii) são solicitados dados específicos para determinada informação contextual em um intervalo de tempo. As funcionalidades estão organizadas em três módulos: (i) Módulo de Relatório: trata da exibição das informações contextuais no dispositivo móvel; (ii) Módulo de Acesso: consiste de um servidor com total acesso ao Repositório de Informações Sensoriadas; e (iii) Módulo de Requisição: tem por objetivo solicitar informações contextuais ao Módulo de Acesso, a pedido do Módulo de Relatório. **Bloco de Tratamento de Alertas Proativos:** tem por objetivo sinalizar o usuário quando da ocorrência de eventos de seu interesse através de notificações proativas. As funcionalidades estão organizadas em dois módulos: (i) Módulo de Alertas, executa no dispositivo móvel, disponibilizando alertas ao usuário, empregando para isso o mecanismo nativo de notificação da plataforma móvel. Esta opção tem como característica propiciar ao usuário um gerenciamento integrado da natureza dos alertas praticados pelo seu dispositivo móvel; e (ii) Módulo de Distribuição, executa no mesmo equipamento do Servidor de Avaliação, em regime de operação ininterrupta. Este módulo opera mantendo os alertas produzidos pelas diferentes regras contextuais, tratadas pelo Interpretador. Seu acesso pelos dispositivos móveis ocorre através do Módulo de Comunicação, o qual, através de um barramento ESB, propicia acesso as diferentes funcionalidades do Servidor de Avaliação.

O **Gerenciador de Publicação** é responsável pelo suporte à coleta de sinais vitais, *status* e parâmetros das bombas de infusão e monitores de sinais vitais, capturados através do Servidor de Integração. Essas funcionalidades são implementadas através do barramento ESB do Servidor de Avaliação.

O **Gerenciador de Atuação** é responsável pelo acionamento dos atuadores dos equipamentos eletromédicos. Esse componente tem como função disparar no ambiente ubíquo ações que mudem o estado dos equipamentos eletromédicos, apresentando alertas visuais e sonoros aos profissionais de saúde, sempre que solicitado pelos demais módulos do Servidor de Avaliação.

O **Interpretador** tem como principal função realizar as tarefas de gerenciamento e identificação das informações contextuais, utilizando para isto dados especificados no componente **Informações Sensoriadas de Interesse das Metas Terapêuticas**. A tarefa de identificação é processada sobre os dados coletados através dos sensores e tem o intuito de produzir informações de contexto de mais alto nível. A seguir são apresentados os componentes deste módulo.

Repositório de Informações Sensoriadas: armazena os dados de contexto coletados pelos sensores dos equipamentos eletromédicos. A estrutura prototipada traduz a organização da arquitetura do middleware EXEHDA, contemplando as relações entre as aplicações, componentes, sensores, ambientes e os contextos de interesse. Neste repositório estão armazenados os sinais vitais dos pacientes e os dados de sensoriamento das bombas de infusão e monitores, obtidos através do Módulo de Aquisição do Servidor de Integração, mantendo dessa forma um registro dos contextos que podem ser avaliados pelo Gerenciador da Meta Terapêutica. As informações de contexto capturadas são registradas neste repositório juntamente com seu *timestamp* da coleta, *timestamp* de registro e Servidor de Integração de origem. As informações deste repositório podem ser acessadas pelas aplicações ubíquas dos profissionais de saúde através do Gerenciador de Comunicações.

Repositório de Situações: armazena os resultados das avaliações realizadas pelo Gerenciador da Meta Terapêutica sobre as informações contextuais oriundas do Repositório de Informações Sensoriadas. Este repositório mantém um histórico de quando e quais Metas Terapêuticas foram atingidas. As informações deste repositório podem ser acessadas pelas aplicações ubíquas dos profissionais de saúde através do Gerenciador de Comunicações.

Gerenciador das Metas Terapêuticas: é um mecanismo de inferência baseado em regras, o qual produz contextos deduzidos de acordo com as Metas Terapêuticas definidas pelos profissionais de saúde. O Gerenciador das Metas Terapêuticas processa regras do tipo Evento-Condição-Ação para identificação de situações, utilizando técnicas baseadas em Lógica Temporal e Espacial (ALLEN; FERGUSON, 1994; SALAS, 2017).

Os parâmetros destas regras são definidos pelos profissionais de saúde, levando em consideração as faixas de valores e o tempo esperado como meta para os sinais vitais dos pacientes em função da administração de um medicamento. Quando um dado coletado é publicado no Repositório de Informações Sensoriadas, o Gerenciador das Metas Terapêuticas

consulta o componente Informações Sensoriadas de Interesse das Metas Terapêuticas para identificar quais regras devem ser avaliadas. Os resultados destas avaliações são então registrados no Repositório de Situações das Metas Terapêuticas.

■ RESULTADOS E DISCUSSÃO

Estudo de Caso

Esta seção apresenta as tecnologias utilizadas na concepção do EXEHDA-TG, bem como o estudo dos cenários de uso empregados na avaliação de suas funcionalidades. O estudo de caso desenvolvido contempla tarefas referentes ao sensoriamento e a coleta de informações contextuais, avaliação e notificação dos dados de contexto aos demais serviços do *middleware*.

A bomba de infusão e o monitor de sinais vitais utilizados neste estudo de caso possuem comunicação de dados via *Wi-Fi*. Com isso, é possível remotamente parametrizar os dados de uma nova infusão, fazer a leitura dos sinais vitais coletados, verificar o estado dos sensores/alarmes e o andamento da infusão.

Foram utilizadas plataformas computacionais baseadas em Linux para viabilizar, no ambiente ubíquo, a aquisição de dados através de sensores e o acionamento de atuadores. Como Servidor de Integração foi utilizada a *Raspberry Pi* modelo B, cujas características principais são apresentar tamanho reduzido, baixo consumo de energia e custo reduzido, quando comparada com desktop convencionais.

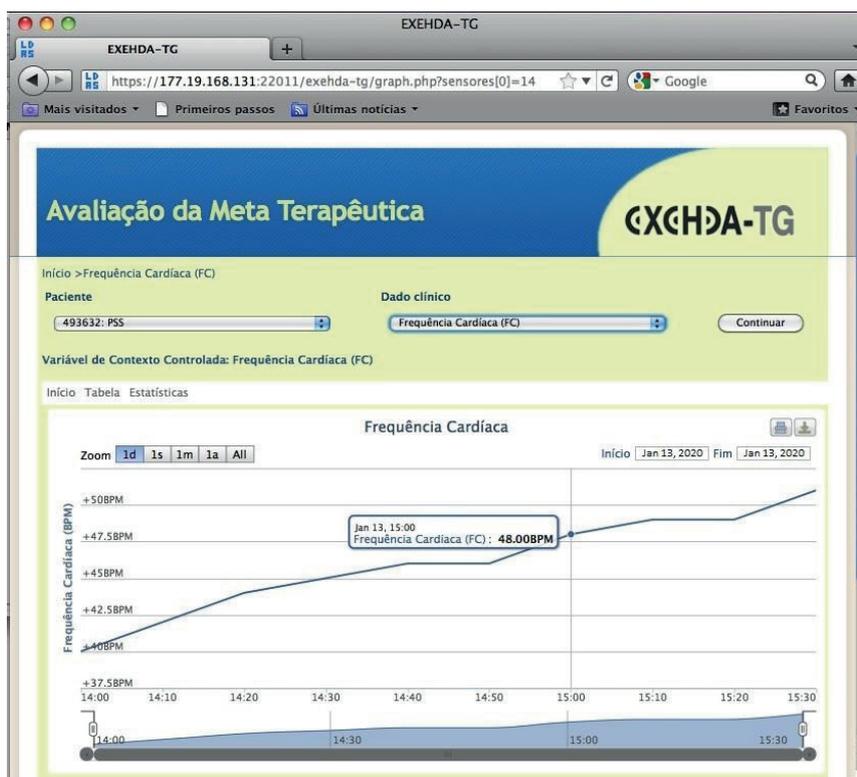
O Servidor de Avaliação é responsável por receber valores de contexto coletados dos pacientes e dos equipamentos eletromédicos, armazenar os mesmos e disparar avaliações situacionais, solicitando atuações e gerando alertas sempre que necessário. Para isto, neste estudo de caso foi utilizado um hardware com processador Intel E3400- 2.6GHz de dois núcleos, com memória RAM de 4Gb, com o Sistema Operacional Ubuntu Server.

Neste estudo de caso foram desenvolvidas duas aplicações com o objetivo de acompanhamento e avaliação das Metas Terapêuticas: a primeira é utilizada com navegadores Web e a segunda é direcionada para a plataforma Android.

Através da aplicação para navegadores é possível: (i) consultar os dados de contexto coletados pelos sensores dos equipamentos eletromédicos e armazenados no Repositório de Informações Sensoriadas do Servidor de Avaliação, visualizando-os através de gráficos e tabelas; (ii) visualizar relatórios textuais com os dados coletados da última semana; (iii) acompanhar a evolução de mais de um sinal vital em um mesmo gráfico, permitindo compará-los; e (iv) realizar análises estatísticas dos dados coletados. Na tela inicial da aplicação para navegadores é possível fazer a seleção de pacientes e os dados clínicos de interesse

do profissional de saúde. Na Figura 3 é apresentado um gráfico típico da evolução dos sinais vitais. Através do mesmo é possível acompanhar a variação dos sinais vitais em função da medicação administrada ao paciente. Posicionando o *mouse* no gráfico é detalhado o valor e o momento (data e horário) de coleta do mesmo. Também é possível visualizar mais de um sinal vital simultaneamente em um mesmo gráfico, sendo viabilizada assim a comparação entre os mesmos.

Figura 3. EXEHDA-TG: acompanhamento gráfico de sinais vitais



Através da opção da interface Web do EXEHDA-TG pode-se visualizar os sinais vitais de um paciente coletados na última semana. Estas informações são apresentadas em colunas, uma para cada dia da semana, onde também é possível ver os valores médio, máximo e mínimo de cada dia. Na opção "Estatística" (vide Figura 4) é apresentada uma funcionalidade da aplicação Web que viabiliza o cruzamento de dados contextuais envolvendo múltiplos sinais vitais a partir de diferentes regras. Esse recurso permite a adição, remoção e edição de regras e parâmetros.

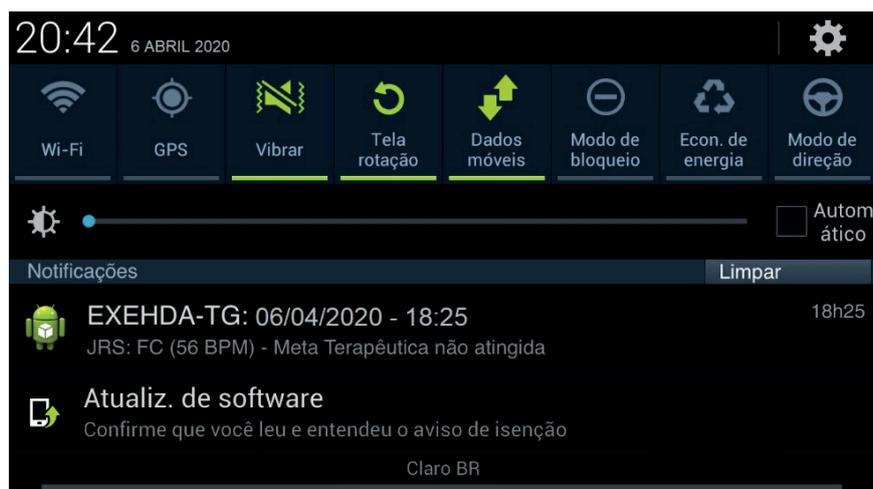
Figura 4. EXEHDA-TG: informações estatísticas do quadro clínico de pacientes



A prototipação do Módulo de Acesso Móvel é constituída de duas partes que interoparam: uma aplicação móvel e outra que processa junto ao Servidor de Avaliação. A parte móvel é voltada para *smartphones* e/ou *tablets* que empregam o sistema operacional Android. Na tela de abertura da aplicação é possível selecionar um paciente, o dado clínico de interesse e a forma (gráfica ou textual) desejada para visualizar o mesmo. Na tela de informações gráficas da aplicação móvel é apresentado o gráfico com a evolução dos sinais vitais, onde é possível selecionar o período de tempo desejado de exibição (uma hora, um dia, uma semana ou um ano). Ao clicar na curva do sinal vital é apresentado o valor e o momento (data e horário) de coleta do mesmo. O eixo vertical do gráfico é ajustado de forma automática, levando em consideração os valores mínimos e máximos a serem plotados.

Para disponibilização dos alertas ao usuário foi utilizado o mecanismo nativo de notificação da plataforma Android (vide Figura 5). Esta opção tem como característica propiciar ao usuário um gerenciamento integrado da natureza dos alertas praticados pelo seu dispositivo móvel.

Figura 5. EXEHDA-TG: envio de alertas sobre o estado clínico de pacientes



Trabalhos Relacionados

A discussão dos trabalhos relacionados abrangeu os seguintes projetos: MTM- CC (KOUTKIAS et al. 2010), HeartCycle (CHOUVARDA et al. 2011), SIMAp (LEITE, 2011) e PA-DSS (GUNAWARDANE et al. 2009). Esta discussão teve por base aspectos decorrentes das premissas consideradas para concepção do EXEHDA-TG: arquitetura distribuída ou centralizada, armazenamento e consulta dos dados de contexto, suporte ao tratamento de regras, regras baseadas no conhecimento de médicos especialistas, verificação da resposta ao tratamento medicamentoso, interface com equipamentos eletromédicos e abordagem siga-me para envio de alertas. Estes aspectos são considerados na comparação entre os trabalhos relacionados apresentada na Tabela 1.

Tabela 1. Comparação com os trabalhos relacionados

Características	Projetos				
	MTM-CC	HeartCycle	SIMAp	PA-DSS	EXEHDA-TG
Armazenamento e consulta sobre os dados de contexto	não	sim	não	não	sim
Suporte ao tratamento de regras	sim	não	sim	sim	sim
Regras baseadas no conhecimento de médicos especialistas	sim	não	sim	sim	sim
Verificação da resposta ao tratamento medicamentoso	sim	sim	não	não	sim
Interface com equipamentos eletromédicos	não	sim	sim	sim	sim
Abordagem siga-me para envio de alertas	não	não	sim	sim	sim

Todos os projetos estudados possuem arquitetura distribuída, e de forma análoga ao EXEHDA-TG, contemplam esforços de processamento dos dados contextuais em diferentes posições da arquitetura. O armazenamento dos dados de contexto é uma funcionalidade presente em todos os projetos. Porém, apenas o EXEHDA-TG disponibiliza facilidades para consulta posterior dos dados coletados.

Com exceção do projeto HeartCycle, os demais provêm suporte ao tratamento de regras. O EXEHDA-TG diferencia-se dos trabalhos relacionados por permitir a definição das Metas Terapêuticas de forma personalizada através do uso de regras, viabilizando seu emprego para diversos propósitos. Assim, no EXEHDA-TG a parametrização da Meta Terapêutica é feita pelos médicos, os quais podem criar regras para serem informados, através de alertas, caso os valores dos sinais vitais satisfaçam alguma condição que coloque o paciente em risco ou caso algum sintoma não desejado seja identificado.

Os projetos SIMAp e PA-DSS não consideram a evolução dos sinais vitais em função do tratamento medicamentoso de forma personalizada por paciente. O EXEHDA-TG diferencia-se por gerar alertas quando os sinais vitais não estão evoluindo em função da medicação, conforme o esperado pelo médico em cada caso clínico.

No projeto MTM-CC os sinais vitais são coletados através de uma rede de sensores biomédicos sem fio. Por sua vez, o EXEHDA-TG, de forma análoga aos projetos HeartCycle, SIMAp e PA-DSS, coleta os sinais vitais dos pacientes através de monitores. Entretanto, o EXEHDA-TG diferencia-se destes trabalhos por também possuir suporte para interface com bombas de infusão, permitindo avaliar e armazenar os alarmes gerados pelas mesmas.

No projeto MTM-CC os alertas são enviados a um servidor central, enquanto no EXEHDA-TG estes alertas são enviados tanto para uma central de alertas no ambiente hospitalar, como para os dispositivos móveis dos profissionais de saúde. Com isso, o EXEHDA-TG diferencia-se por suportar a abordagem siga-me, a qual permite ao usuário iniciar aplicações e acessar dados a partir de qualquer lugar.

■ CONCLUSÃO

A concepção de uma arquitetura para o acompanhamento dos sinais vitais em função da medicação, se mostra uma ferramenta oportuna para avaliar a resposta do paciente ao tratamento, bem como para enviar alertas aos profissionais de saúde quando os valores dos sinais vitais indicarem uma resposta inesperada ao medicamento.

Desta forma, através do recebimento destes alertas, o médico pode reavaliar a prescrição dos medicamentos ou prescrever um novo fármaco para que o tratamento tenha a resposta desejada. Este pronto aviso à equipe médica de que a Meta Terapêutica não foi atingida, constitui uma das contribuições centrais do EXEHDA-TG para agilização das

práticas relacionadas ao tratamento de pacientes, com potencial contribuição para que o tempo de internação seja reduzido.

Dentre outros, na continuidade da pesquisa, os seguintes aspectos deverão ser considerados em trabalhos futuros: (i) realizar a validação clínica no ambiente hospitalar, analisando o desempenho do EXEHDA-TG em uma situação cotidiana; (ii) explorar o emprego de regras de processamento contextual que utilizem outros mecanismos de inferência de mais alto nível, ampliando as opções de inferência sobre os dados coletados; e (iii) dar continuidade aos procedimentos de integração do EXEHDA-TG com os diferentes serviços e funcionalidades do *middleware* EXEHDA, na perspectiva de prover soluções ubíquas explorando a infraestrutura provida pela IoT.

■ AGRADECIMENTOS

O presente trabalho foi realizado com apoio do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001 e Instituto Federal do Rio Grande do Sul (IF-RS).

■ REFERÊNCIAS

1. ALLEN, J. F.; FERGUSON, G. Actions and Eventis in Interval Temporal Logic. *Journal of Logic and Computation*, 1994.
2. CHOUVARDA, I.; GKONTRA, P.; KOKONOZI, A.; SEMERTZIDIS, P.; CAFFAREL, J.; and MAGLAVERAS, N. Novel Approaches for Medication Compliance and Effectiveness Analysis and Support in Cardiovascular Disease Patients. In *33rd Annual International Conference of the IEEE EMBS*, 2011.
3. GUNAWARDANE, T. S. F. W.; KOGGALAGE, R.; Rodrigo, R.; and RAJAPAKSE, S. Patient Alert and Decision Support System. In *ACTEA 2009: 2009 International Conference on Advances in Computational Tools for Engineering Applications*, 2009.
4. KOUTKIAS, V. G.; CHOUVARDA, I.; TRIANTAFYLLIDIS, A.; MALOUSHI, A.; GIAGLIS, G. D.; and MAGLAVERAS, N. A Personalized Framework for Medication Treatment Management in Chronic Care. *IEEE Transactions On Information Technology In Biomedicine*, 2010.
5. LEITE, C. R. M. Arquitetura Inteligente Fuzzy para Monitoramento de Sinais Vitais de Pacientes: Um Estudo de Caso em UTI. Tese de doutorado, Universidade Federal do Rio Grande do Norte - Programa De Pós-Graduação em Engenharia Elétrica e Computação, 2011.
6. LOPES, J.; SOUZA, R.; GEYER, C.; COSTA, C.; BARBOSA, J.; PERNAS, A.; and YAMIN, A. A middleware architecture for dynamic adaptation in ubiquitous computing. *Journal of Universal Computer Science*, 20(9):1327–1351, 2014.

7. SALAS, Eduardo, ed. Situational awareness. Routledge, 2017.
8. SÁNCHEZ GUINEA, A.; NAIN, G.; LE TRAON, Y. A Systematic Review on the Engineering of Software for Ubiquitous Systems. J. Syst. Softw., New York, NY, USA, v.118, n.C, p.251–276, Aug. 2016.
9. SEZER, O. B.; DOGDU, E.; OZBAYOGLU, A. M. Context-Aware Computing, Learning, and Big Data in Internet of Things: A Survey. IEEE Internet of Things Journal, IEEE, v.5, n.1, p.1–27, Feb 2018.
10. TEMDEE, P.; PRASAD, R. Context-Aware Communication and Computing: Applications for Smart Environment. Switzerland: Springer International Publishing, 2018.

Uma experiência em ensino de programação de computadores com alunos da rede pública usando a linguagem *Scratch*

I Wendell Bento **Geraldes**
IFG

I Ernane Rosa **Martins**
IFG

I Ulisses Rodrigues **Afonseca**
IFG

RESUMO

Este artigo apresenta um relato de experiência sobre um curso de extensão em programação de computadores utilizando a linguagem Scratch. O curso foi realizado no Instituto Federal de Goiás campus Luziânia, com base nos planos de ensino do Portal Programaê! e também na iniciativa Code Club Brasil. Um questionário foi aplicado no início e no final do curso para avaliar o grau de desenvolvimento da capacidade dos estudantes em resolver problemas. Os resultados não mostraram uma diferença significativa no número de erros e acertos do primeiro para o segundo questionário. No entanto uma análise mais detalhada mostra um aumento na capacidade de resolução de problemas lógicos e gramaticais. Os estudantes mostram também um forte interesse em continuar os estudos em programação de computadores.

Palavras-chave: Scratch, Programação.

■ INTRODUÇÃO

Um elemento fundamental para a sociedade é a educação, em seu papel de criar competências e desenvolver habilidades que possam ser úteis aos seus cidadãos e promover sua inserção no mundo do trabalho e desenvolvimento pessoal. Com o rápido desenvolvimento tecnológico é preciso que a educação ajude o aprendiz a se adaptar a estas mudanças de forma a garantir que ele não só saiba utilizar mas também criar novos aparatos e melhores tecnologias.

Nos últimos anos várias iniciativas foram criadas para auxiliar o processo de ensino-aprendizagem de programação de computadores. Entre elas, merecem destaque a plataforma *Code.org*¹ criada pelos irmãos Ali e Hadi Partovi e o *Code Club World*² criado pela designer Clare Sutcliffe e a desenvolvedora Linda Sandivik. Estas iniciativas fortalecem a ideia de que a grande maioria das profissões do século 21 exigem uma compreensão da ciência da computação e principalmente o conhecimento da programação como relata o *Model Curriculum for K-12 Computer Science*. [Silva et al. 2015]

Analisando a literatura sobre o assunto, é possível verificar diversos artigos científicos que tratam desta questão, onde 60% destes abordam a programação no contexto da educação superior, 22% no ensino médio, 11% no ensino fundamental e apenas 3% das pesquisas destina-se ao ensino técnico. Outro dado relevante é que a maioria dos estudos nesta área é aplicado nas universidades (aproximadamente 63%) e 34% nas escolas. [Silva et al. 2015]

Ao verificar os estudos realizados no ensino de programação no ensino fundamental e técnico, optou-se por direcionar a pesquisa a este grupo em especial. Desta forma este trabalho apresenta um relato de experiência de um curso de extensão em programação de computadores com foco no ensino fundamental e médio, realizado no Instituto Federal de Goiás Campus Luziânia, com base nos planos de ensino do portal Programaê! e também na iniciativa Code Club Brasil.

Este artigo encontra-se organizado da seguinte forma. A seção 2 apresenta os procedimentos metodológicos. A seção 3 apresenta os resultados obtidos. A seção 4 apresenta uma discussão acerca dos resultados. A seção 5 apresenta as considerações finais e sugestões de trabalhos futuros.

1 <https://code.org/>

2 <https://www.codeclubworld.org/about/>

■ PROCEDIMENTOS METODOLÓGICOS

Participantes e formato do curso

O curso foi realizado no primeiro semestre de 2018 no Instituto Federal de Goiás campus Luziânia. As atividades tiveram duração de 3 horas diárias, de segunda a sexta-feira, durante 2 semanas, perfazendo um total de 30 horas. Participaram do curso 30 estudantes, de 15 a 19 anos, do 9.º ano do ensino fundamental, 1.º ano do Ensino Médio e do 1.º semestre do Ensino Superior da rede pública de Luziânia/GO. A seleção se deu através do boletim escolar e de acordo com a média geral das notas dos estudantes, respectivamente em ordem decrescente da maior para a menor nota. As vagas restantes e não preenchidas foram distribuídas para estudantes do ensino médio regular e ensino superior.

As aulas foram realizadas em um laboratório de informática contendo 30 microcomputadores com acesso a internet. Um professor conduziu as atividades com os estudantes. A atuação deste foi direcionada na apresentação dos conceitos básicos da computação através de atividades lúdicas e sem o uso do computador. Também foram utilizadas aulas on-line em plataforma Web para os conceitos fundamentais da programação de computadores. O curso foi totalmente baseado nos planos de aula do portal **Programaê!** que possui três diferentes categorias:

1. **Para além do computador:** onde os alunos aprendem de forma abstrata as estruturas básicas do Pensamento Computacional, como por exemplo, loops, condicionais, funções, variáveis e lógica de programação. As atividades são desplugadas, isto é, não precisam do computador para serem realizadas. Em grande parte as fichas de aula pertencem a plataforma *Code.org* e estão traduzidas para o Português do Brasil. Cada ficha traz um resumo da aula, os objetivos, materiais necessários para realizar a atividade proposta, a preparação da aula, um vocabulário de palavras utilizadas dentro do contexto da aula, uma introdução proposta para que o professor/instrutor possa tratar o assunto, a atividade em si e ajustes que dependem da série em que os alunos estão inseridos na escola.
2. **Praticando e aprendendo:** onde os alunos aprendem estruturas básicas do Pensamento Computacional, tais como loops, condicionais, funções e variáveis, e também elementos da lógica matemática aplicada a programação. As atividades desta categoria também podem ser encontradas na plataforma *online Code.org* e são compostas por vídeos explicativos e exercícios práticos usando blocos de programação para resolver problemas.
3. **Narrativas e projetos:** onde os alunos aprendem a desenvolver projetos utilizando

a linguagem *Scratch*. [Telefônica 2018]

Para avaliação dos resultados foi utilizado um questionário contendo 10 (dez) questões, 9 (nove) de múltipla escolha e 1 (uma) para avaliar a percepção dos estudantes quanto a programação de computadores os ajuda em outra matéria. O questionário foi obtido no trabalho realizado por Zanatta (2015) sendo aplicado integralmente. [Zanatta 2015]

Atividades Desenvolvidas

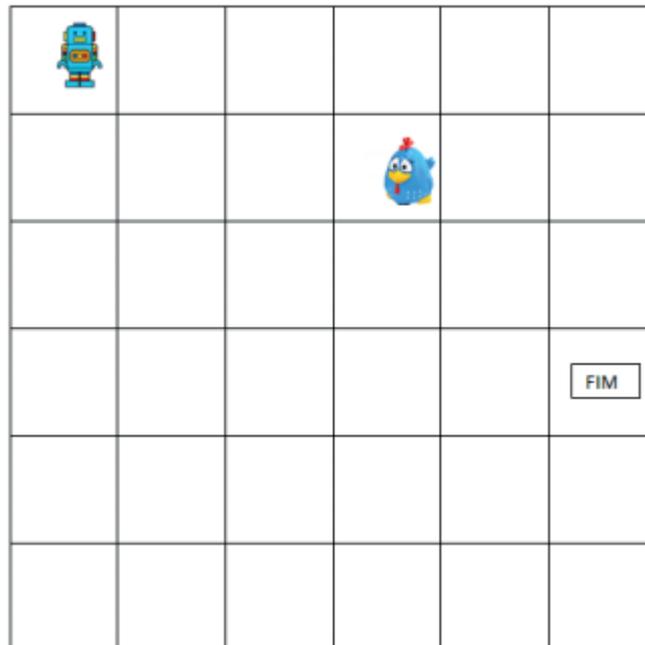
As atividades do curso foram realizadas seguindo a seguinte ordem de planos de aula: i) *Para além do computador*, ii) *Praticando e aprendendo*, iii) *Narrativas e projetos*. As atividades desplugadas encontradas no plano de aula *Para além do computador* eram seguidas de atividades plataforma *online Code.org* encontradas no plano de aula *Praticando e aprendendo*.

As aulas onde foram utilizadas atividades de Computação Desplugada foram realizadas de forma alternada às atividades na plataforma *online*. No primeiro dia de curso, os estudantes participaram de uma atividade chamada simulando um robô, onde um deles faria o papel do robô e o outro o papel do programador. Nesta atividade algumas instruções foram combinadas previamente, como por exemplo, Frente (n) passos, Trás (n) passos, Vire a direita 90 graus, Vire a esquerda 90 graus, Pegue o objeto e Solte o objeto.

O piso do pátio da escola foi marcado com uma fita crepe de maneira que fosse possível visualizar uma tabela de 6 linhas e 6 colunas (Figura 1). Os estudantes foram desafiados a dizer quais instruções seriam necessárias para executar a ação desejada, qual o caminho mais curto para o robô atingir o objetivo e quantas instruções seriam necessárias para realização da ação em menor tempo possível.

Na aula 2 sobre decodificação binária foi utilizado um instrumento pedagógico chamado decodificador binário que consiste em uma tabela onde as letras do alfabeto são representadas por um código formado por 8 pequenos quadrados que podem ser brancos ou pretos como mostra a Figura 2. Na aula 3, chamada de Labirinto a atividade na plataforma *online* consiste em 20 exercícios que promovem o aprendizado de algoritmos – sequência de instruções. Os exercícios são pequenos desafios em um labirinto que tem como objetivo reconhecer algumas estruturas de programação como sequências, ciclos (loops) e condições.

Figura 1. Aula 1 - simulando um robô



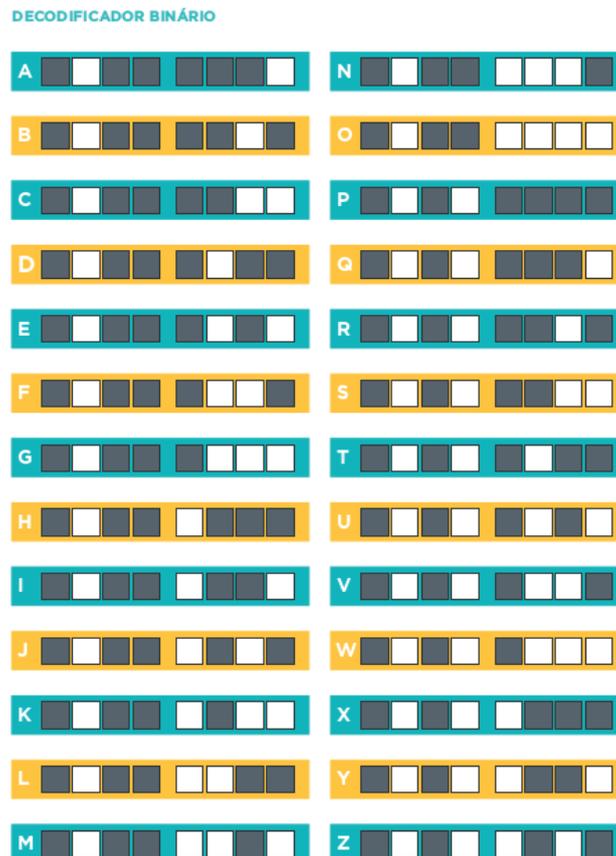
Em cada aula o estudante aprende algum conceito importante da Computação. Na aula 4 por exemplo, eles aprendem sobre quatro etapas do Pensamento Computacional: Decomposição, Padrões, Abstração e Algoritmos. A atividade denominada Criar um monstro, ajuda os estudantes a compreender estes conceitos de maneira fácil e divertida. A proposta é classificar os monstros existentes em um catálogo, depois analisar os padrões existentes entre eles e em seguida abstrair as diferenças para criar uma lista de instruções para recriar um dos monstros do catálogo. Na aula 7, os estudantes são convidados a criar algoritmos que descrevam orientações do mundo real, aprender sobre como pensar em diversas soluções para um mesmo problema e criar soluções mais eficientes para os problemas. A atividade se baseia em um jogo clássico antigo chamado Tangram, que é um quebra-cabeças geométrico chinês. O exercício é realizado em duplas, um estudante fica de costas para o outro, um deles é o computador e o outro o programador. O estudante que assume o papel do programador segura uma folha com imagens e deve descrevê-las sem mostrá-las ao aluno que assume o papel de computador que deve reconstruir a imagem de acordo com o relato do programador.

Na aula 14, os estudantes aprenderam sobre estruturas de programação, como por exemplo, sequências, estruturas de repetição, estrutura condicional e também sobre funções. Nesta atividade foi utilizada a plataforma *online Code.org* na fase 13 – A fazendeira 2 do curso introdutório de 20 horas. São 10 exercícios que envolvem as diferentes estruturas de programação.

Na aula 22, as atividades já são realizadas na plataforma *Scratch* do MIT. Os estudantes aprendem sobre dois conceitos para o desenvolvimento dos projetos no *Scratch*: eventos e paralelismo. Na atividade 1 estes conceitos são apresentados utilizando exemplos do dia

a dia e em seguida no ambiente. Na atividade 2 os alunos são convidados a abrir o projeto “Esconde Esconde” e o professor pode substituir o bloco “quando clicar em” por “quando a tecla for pressionada” exemplificando o conceito de eventos. Em seguida os alunos são desafiados a criar um programa com eventos em paralelo, para ajudá-los o professor pode utilizar o projeto “Aula 2 – Pega Pega”.

Figura 2. Decodificador Binário



Instrumento de coleta de dados

A fim de avaliar o desempenho dos estudantes antes e depois do curso de programação de computadores, utilizou-se o questionário de Zannata (2015) sem nenhuma adaptação. O questionário foi estruturado da seguinte forma: (a) *questões de lógica envolvendo números*; (b) *questões de matemática*; (c) *questões de lógica envolvendo preposições lógicas*; (d) *questões de interpretação de texto*. A última questão foi discursiva com objetivo de conhecer a percepção do estudante quanto a importância da programação em seu cotidiano escolar.[Zanatta 2015]

O item (a) foi composto de questões como a seguinte:

1 - TACO está para ATCO como 7683 está para:

- a) 3678 b) 6783 c) 8376 d) 7837**

Esta questão avalia a capacidade do estudante em relacionar as palavras com os números.

O item (b) foi composto de questões como a seguinte:

6 - Adriana vai fazer esta subtração: $650 - 38$. O resultado dessa operação sera?

- a) 299 b) 399 c) 631 d) 641

Esta é uma questão que avalia o conhecimento matemático do estudante.

O item (c) foi composto de questões como a seguinte:

3 - Uma taça está para os cereais como um envelope está para:

- a) O carteiro c) A carta b) A caixa de correio d) O selo

Esta é uma questão que avalia o pensamento lógico.

O item (d) foi composto de questões como a seguinte:

Figura 3. Imagem utilizada na questão 07

PASSAGEM DE ÔNIBUS		
TERMINAL RODOVIÁRIO N° 6 5 7 8 9 Belo Horizonte — MG		6 5 7 8 9
de: BELO HORIZONTE para: SÃO PAULO		BH/SP
DATA 22/05/99	AGENTE José Cintra	VIAÇÃO LUXOR
POLTRONA 22	HORÁRIO 23h30 min	Prefixo 008954 KM 590,8
ÔNIBUS LEITO	PREÇO R\$ 98,70	via do passageiro
ATENÇÃO, USUÁRIO Mantenha sempre em seu poder esta passagem.		

7 - Observe a imagem acima e responda: o passageiro vai iniciar a viagem: (Figura 3).

- a) À noite b) À tarde c) De madrugada d) Pela manhã

Esta questão avalia a capacidade de interpretação de textos do estudante.

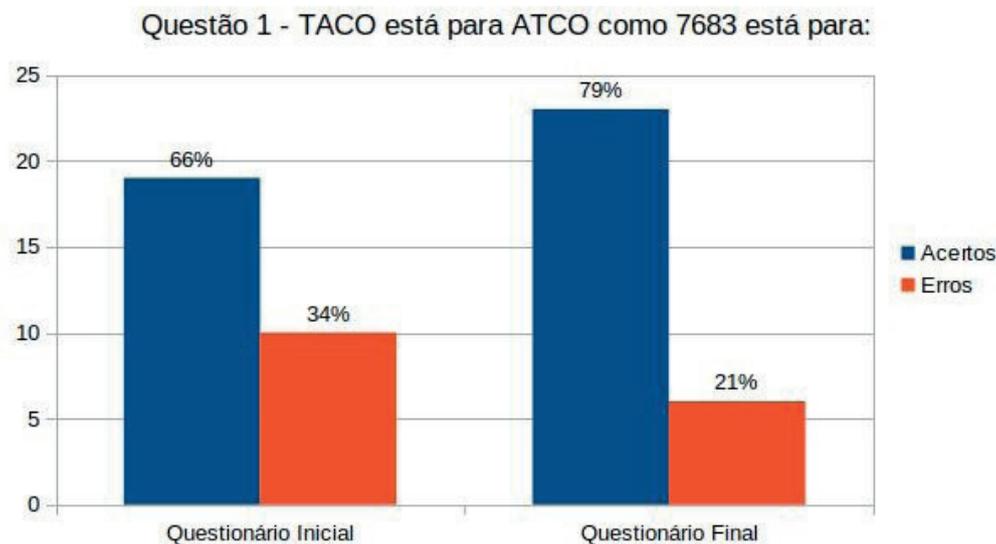
Os questionários foram aplicados no primeiro dia do curso e também no último dia do curso. O primeiro questionário foi respondido por 29 estudantes, 21 (72%) do sexo masculino e 8 (28%) do sexo feminino. Com relação a escolaridade, 13 (46%) dos respondentes são do Ensino Fundamental, 10 (36%) do Ensino Médio e 5 (18%) do Ensino Superior. O mesmo número de estudantes responderam o segundo questionário ao final do curso.

■ RESULTADOS

O resultado da avaliação dos erros e acertos dos questionários aplicados na primeiro dia e no último dia de aula foi confrontado. No primeiro questionário houve um total de 217 (83%) de acertos e 44 (17%) de erros. No segundo questionário houve um total de 219 (84%) de acertos e 42 (16%) de erros. Esta pequena diferença encontrada nos resultados obtidos não foi suficiente para afirmar que houve alguma melhora efetiva na capacidade dos estudantes em resolver problemas.

Por outro lado, uma análise mais detalhada em cada questão mostrou que, por exemplo, na primeira questão do questionário, que tem como foco a lógica, 66% dos estudantes responderam corretamente no questionário aplicado no início do curso, este percentual subiu para 79% no final do curso (Figura 4).

Figura 4. Questão 1 do questionário de avaliação



Na nona questão, que tem como foco a gramática houve uma melhora no percentual de acertos, no questionário aplicado no início do curso, 76% dos estudantes acertaram esta questão, e no questionário aplicado ao final do curso este percentual subiu para 83%.

A décima questão do primeiro questionário fez a seguinte pergunta ao estudante: “Você acredita que aprender a programar lhe ajudará em outras matérias ? “ e no segundo questionário “Você acredita que aprender a programar lhe ajudou em outras matérias ? “.

A análise das respostas mostram que ao iniciar o curso de programação com *Scratch*, 83% dos estudantes consideravam que o curso lhes ajudaria em outras matérias, enquanto 17% tinham dúvidas a respeito. No questionário aplicado no último dia do curso o percentual de estudantes que responderam afirmativamente a esta questão subiu para 90%, enquanto 7% ainda tinham dúvidas a respeito. Apenas 1 (3%) estudante afirmou que o curso não lhe ajudaria em outras matérias.

Este resultado mostra que houve um aumento no interesse dos estudantes em aprender mais sobre programação de computadores.

■ DISCUSSÃO

Durante as atividades, notou-se uma grande motivação dos estudantes em relação as aulas de Computação Desplugada que foram realizadas intercaladas com as aulas *on-line*. O motivo deste entusiasmo parece estar relacionado com o fato de que a aula de

Computação Desplugada é uma técnica que visa ensinar os fundamentos da Computação de forma lúdica, sem o uso de computadores, sem distrações e detalhes técnicos em demasia. [Vieira et al. 2013]

Um fato importante e relevante sobre as aulas de Computação Desplugada, é que as mesmas em sua maioria são realizadas em duplas ou pequenos grupos. Uma avaliação deste curso, mostrou que esta característica fez com que houvesse uma forte interação entre os estudantes e seus pares.

As atividades *online* foram realizadas individualmente, através de pequenos jogos de blocos, os estudantes eram desafiados a resolver problemas em cada aula. Estas atividades focam na autonomia do estudante e mostram uma característica instrucionalista de aprendizagem, pois não dependem de um tutor ao lado do aprendiz para lhe ensinar.[Valente 2019]

Outro fato interessante observado durante o curso, está relacionado ao uso da linguagem *Scratch*. Optou-se por utilizar os roteiros para construção de jogos da iniciativa *Code Club Brasil*. Cada roteiro estabelece instruções passo a passo para a criação de jogos simples, mas com vários conceitos da programação de computadores envolvidos, como por exemplo, o uso de estruturas de repetição e variáveis. Nesta etapa do curso os estudantes também mostraram bastante interesse pelas aulas, alguns foram além das instruções fornecidas, modificaram os projetos, melhorando e aperfeiçoando os jogos.

■ CONCLUSÃO

A identificação do problema, o entendimento e reflexão acerca de suas causas e a busca de soluções são essenciais no cotidiano atual. Estas habilidades podem e devem ser incentivadas no ambiente escolar. A elaboração de propostas de atividades que busquem motivar os estudantes a desenvolver este aprendizado são sempre bem vindas.

Assim, ressalta-se a importância de se pensar e ampliar a oferta de atividades que favoreçam o desenvolvimento de competências como a cooperação mútua, autonomia e o Pensamento Computacional. Para além da formação de profissionais de programação, é fato de que a sociedade atual precisa de indivíduos criativos e capazes de encontrar soluções para problemas do presente e de um futuro próximo.

Segundo Resnick et. al (2009) programar amplia o alcance do que você pode aprender, ajuda a criar estratégias de resolução de problemas e design para domínios além da Computação. A programação oferece oportunidades para refletir sobre o seu próprio pensamento e até pensar em si. [Resnick et al. 2009]

A escolha dos planos de aula e objetos de aprendizagem mostrou-se adequada para atender as necessidades do curso. A alternância entre as aulas de Computação Desplugada e *online* foram bem recebidas pelos estudantes, porém é preciso avaliar a ordem e talvez a

própria relevância das atividades *online*, onde os conceitos podem ser introduzidos através da linguagem *Scratch*.

As atividades desplugadas, proporcionaram aos estudantes vivenciar e compreender as quatro etapas do Pensamento Computacional: decomposição, padrões, abstração e algoritmo e também desenvolver habilidades para o trabalho em equipe durante a resolução de problemas.

Também foram utilizados neste curso os guias para projetos de jogos com *Scratch* da iniciativa *Code Club Brasil*, a linguagem simples e o formato passo a passo dos projetos, proporcionou um aprendizado eficiente e motivou o trabalho em grupo.

Os resultados obtidos através da avaliação quantitativa dos questionários não foram suficientes para demonstrar se o curso proporcionou aos estudantes uma melhora na capacidade de resolver problemas. No entanto, uma análise mais detalhada das respostas as questões mostrou uma evolução considerável no pensamento lógico e na gramática.

Sugere-se para trabalhos futuros, melhorar o instrumento de avaliação e aperfeiçoar os planos de ensino para que possam desenvolver o aprendizado dos conceitos fundamentais da Computação e das habilidades inerentes ao Pensamento Computacional.

■ REFERÊNCIAS

1. Resnick, M., Maloney, J., and Monroy-Hernandez, A. (2009). *Scratch: Programming for all. ACM*, 52(11):60–67.
2. Silva, R. d., Medeiros, T. J., Medeiros, H., Lopes, R., and Aranha, E. (2015). Ensino-aprendizagem de programação: uma revisão sistemática da literatura. *Revista Brasileira de Informática na Educação*, 23(1).
3. Telefônica, F. (2018). Programaê! brasil. Disponível em: [http:// fundacaotelefonica.org.br/projetos/programae/](http://fundacaotelefonica.org.br/projetos/programae/) Acesso em 21 maio 2018.
4. Valente, J. A. (2019). Informática na educação: instrucionismo x construcionismo. Disponível em: <http://www.educacaopublica.rj.gov.br/biblioteca/tecnologia/0003.html> Acesso em 6 maio 2018.
5. Vieira, A., Passos, O., and Barreto, R. (2013). Um relato de experiência do uso da técnica Computação desplugada. *Anais do XXI WEI*, pages 670–679.
6. Zanatta, A. C. (2015). Programação de computadores para crianças - metodologia do code club brasil. Monografia, Universidade Federal de Santa Catarina.

Uso de inteligência artificial para auxiliar o combate à obesidade: estudo de caso em Python

| Edjair Aguiar **Gomes Filho**
UNIVASF

| Mateus Amorim **Silva**
UNIVASF

| Leonardo Corsino **Campello**
UNIVASF

| Ricardo Argenton **Ramos**
UNIVASF

| Bráulio Gonçalves **Leal**
UNIVASF

RESUMO

A obesidade, doença crônica não transmissível, é considerada uma epidemia pela OMS, sendo uma doença relacionada a maus hábitos alimentares. Com o avanço da tecnologia, técnicas têm sido desenvolvidas para auxiliar o combate e a prevenção da obesidade. A inteligência artificial é um dos focos no desenvolvimento de novos métodos que objetivam auxiliar o combate à obesidade. Modelos de redes neurais, campo de estudo dentro da inteligência artificial, têm sido propostos devido à sua capacidade de reconhecer padrões e classificar dados complexos. Este trabalho teve como objetivo principal o desenvolvimento de uma rede neural artificial em Python, utilizando as técnicas de *Feedforward* e *Backpropagation*. Através de experimentos, foram coletados dados quantitativos de 10 indivíduos sobre seus hábitos alimentares em cada uma das refeições diárias, além das medidas de peso corporal e circunferência abdominal, durante um período de 30 dias. A pesquisa foi realizada em Petrolina (PE) e Juazeiro (BA), com pessoas de idade entre 18 e 50 anos.

Palavras-chave: Obesidade, Inteligência Artificial, Redes Neurais, Algoritmo.

■ INTRODUÇÃO

A obesidade é uma doença definida como o acúmulo excessivo de gordura no organismo (OLIVEIRA, 2003). Ela é classificada como uma doença crônica não-transmissível (DCNTs), caracterizando-se por sua evolução lenta, ser assintomática e apresentar causas multifatoriais, sendo algumas delas: genéticas, metabólicas, sociais, comportamentais e/ou culturais (TAVARES, 2010). Segundo Dias et al. (2017), uma das formas de diagnosticar a obesidade é utilizando o índice de massa corporal (IMC), calculado como a razão da massa corporal pela altura ao quadrado.

A Organização Mundial de Saúde considera a obesidade como uma epidemia mundial e a define em três graus de gravidade, baseado no IMC: grau I (moderado excesso de peso), grau II (obesidade leve ou moderada) e grau III (obesidade mórbida) (WHO, 2018). Embora não exista uma fórmula para combater o aumento gradativo de pacientes com obesidade, algumas medidas vêm sendo tomadas no Brasil buscando minimizar a taxa de pessoas obesas no país. Segundo Dias et al. (2017), as políticas de debate sobre promoção da saúde estão sendo abordadas como forma de instruir a população sobre alimentação adequada e saudável, apesar de não tratar diretamente sobre obesidade.

No ramo da tecnologia, diversos métodos vêm sendo desenvolvidos a fim de auxiliar no combate à epidemia. Dentre essas técnicas, a inteligência artificial (IA) vem sendo aplicada e vista no ramo da saúde como solução inovadora, justamente por incorporar os benefícios da tecnologia e da psicologia, atendendo diferentes necessidades de saúde por meio de várias mídias (STEPHENS et al., 2019). É esperado por especialistas da área da computação que os algoritmos de inteligência artificial possam auxiliar os profissionais de saúde, fornecendo dados clinicamente relevantes em tempo real e garantindo a qualidade da informação através de sistemas inteligentes. Segundo Reenita Das (2016), até 2025, espera-se que a IA seja utilizada em 90% dos sistemas nos EUA e em 60% dos hospitais e companhias de seguros globais. A utilização da IA na saúde pode ser justificada pela obtenção de resultados mais precisos para os pacientes, custos reduzidos de tratamento e planos de tratamento centrados no paciente (BOURGUET et al., 2013).

A inteligência artificial está sendo utilizada no desenvolvimento de novos métodos que objetivam analisar dados estatísticos, aprender e reconhecer padrões e classificar qualquer tipo de dados (BENEDIKTSSON, 1990). Dentro os campos de estudo da inteligência artificial, as redes neurais têm sido alvidradas devido à sua capacidade de reconhecer padrões e classificação de dados complexos. Os modelos de rede neural têm vantagem sobre os métodos estatísticos por serem livres de distribuição e não serem necessários conhecimentos prévios sobre as distribuições estatísticas das classes nas fontes de dados, a fim de aplicar esses métodos para classificação (FLECK et al., 2016).

■ OBJETIVO

Discutir o uso da inteligência artificial na prevenção da obesidade a partir da análise de resultados obtidos através do processamento de variáveis quantitativas de alimentação. Um modelo de rede neural é implementado para auxiliar indivíduos adultos na previsão do peso corporal e circunferência abdominal, baseando-se nas informações de alimentação diária, contribuindo para a tomada de decisão em relação à qualidade dos alimentos ingeridos nas refeições. Os dados coletados no experimento foram, então, utilizados no algoritmo de inteligência artificial desenvolvido, apoiado nas técnicas de redes neurais, com o intuito de reconhecer padrões e tentar estimar o peso corporal futuro de indivíduos adultos baseado nos pesos das refeições realizadas diariamente. Isso torna possível separar o objetivo do trabalho em duas questões: A primeira questão é a interpretação e a comparação entre os resultados da previsão realizada pela rede neural e os valores reais obtidos na coleta de dados; A segunda questão é a interpretação e discussão desse resultado, a fim de avaliar a eficiência da rede neural montada, a relevância e o impacto das variáveis utilizadas no controle da alimentação do grupo de indivíduos analisado no estudo de caso.

■ MÉTODOS

COLETA DOS DADOS

Para a realização da coleta dos dados da pesquisa foram selecionados 10 indivíduos para participar da pesquisa durante 30 dias. Durante este período, coletou-se os seguintes dados: peso das três principais refeições diárias (café da manhã, almoço e jantar), peso corporal diário e circunferência abdominal (medidos no início do dia, ao acordar, antes de ir ao banheiro). Como instrumento para a realização das medidas foram fornecidas individualmente balanças digitais com precisão de duas casas decimais e fitas métricas. A pesquisa foi realizada em Petrolina, estado de Pernambuco, e Juazeiro, estado da Bahia. Cada indivíduo participante desta pesquisa foi responsável pela própria coleta dos dados, sendo estes coletados nas próprias residências ou locais de trabalho dos envolvidos.

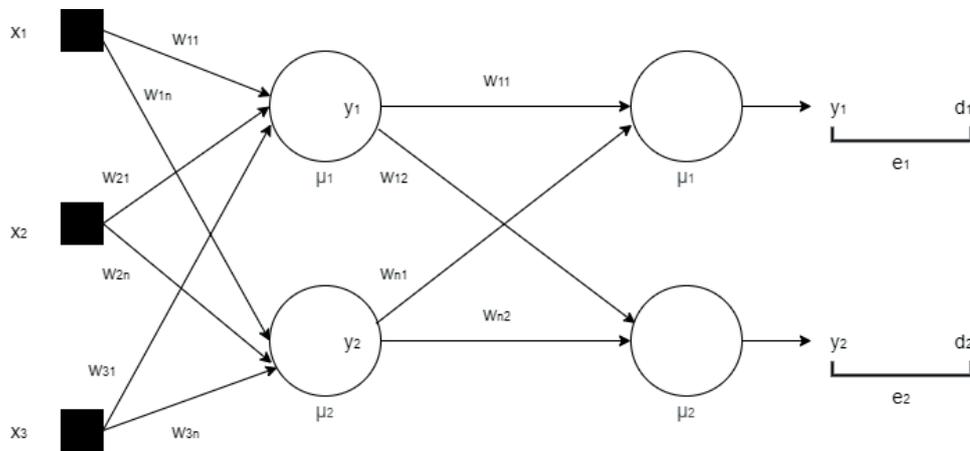
DESENVOLVIMENTO DO ALGORITMO

A implementação do código-fonte foi feita na linguagem de programação Python, versão 3.7, que contempla as funcionalidades necessárias para desenvolver o algoritmo. O ambiente de desenvolvimento foi o Spyder (Anaconda 3).

A rede neural proposta neste trabalho foi desenvolvida implementando-se as técnicas de redes neurais *Feedforward* e *Backpropagation*, com função sigmoide comum para uso

na aprendizagem da rede. Além disso, a arquitetura da RNA utilizada foi de duas camadas (oculta e saída), conforme ilustra a Figura 1.

Figura 1. Arquitetura da rede neural artificial implementada.



Fonte: Autor

Durante a inicialização do programa, são executadas as primeiras declarações, referentes à configuração da rede neural. Essa relação de declarações inclui: número de épocas de treinamento, quantidade de dados que serão usados para treinamento da rede, taxa de aprendizado e a quantidade de neurônios utilizados em cada uma das camadas de neurônios (entrada, camada oculta e saída). A Tabela 1 mostra a correspondência de variáveis usadas.

Tabela 1. Variáveis de inicialização da rede.

Nome da variável	Descrição
<i>numEpocas</i>	Número de épocas de treinamento da rede
<i>q</i>	Quantidade de dados de entrada usados para treinamento
<i>neta</i>	Taxa de aprendizado da rede
<i>m</i>	Quantidade de neurônios na camada de entrada
<i>N</i>	Quantidade de neurônios na camada intermediária
<i>L</i>	Quantidade de neurônios na camada de saída
<i>peso1, peso2, peso3</i>	Dados das refeições, extraídos no arquivo texto
<i>bias</i>	Parâmetro polarizador de valor 1
<i>Xb</i>	Vetor de entrada da rede
<i>W1 e W2</i>	Matrizes de pesos da rede neural artificial

Os valores iniciais dos pesos para cada conexão são gerados de maneira aleatória, utilizando a função *random* da biblioteca *numpy*. Esses valores são alocados matricialmente, de forma que as matrizes são inicializadas com tamanhos correspondentes à quantidade de neurônios nas camadas. Após isso, os dados de entrada são carregados a partir de um arquivo texto, que contém as informações coletadas durante o experimento, e armazenados

nas variáveis correspondentes, sendo as variáveis *peso1*, *peso2* e *peso3* responsáveis por receber os valores de entrada – os pesos das refeições. A variável *d* é um vetor com os dados de saída desejáveis. O último passo das configurações iniciais é concatenar os dados de entrada com o *bias* da rede – elemento utilizado para aumentar o grau de liberdade dos ajustes dos pesos. Essa concatenação é salva na variável *Xb*, que servirá como vetor de entrada da RNA.

Em seguida, chama-se a função responsável por executar o *Backpropagation*: o sinal de erro calculado na saída da rede é propagado no sentido reverso, camada a camada, e ao final deste processo os pesos são ajustados de acordo com a regra de correção de erro. Uma vez que a rede passou pelo processo de aprendizagem, a mesma deve ser testada e validada para confirmar se o resultado obtido é suficiente. O teste foi realizado utilizando a primeira parte dos dados de entrada, correspondente aos 30 primeiros valores do arquivo texto, e retornando o valor do erro médio obtido na camada de saída. Os dados foram submetidos ao algoritmo de rede neural pouco a pouco, para que fosse possível avaliar o comportamento e aprendizagem da rede.

■ RESULTADOS

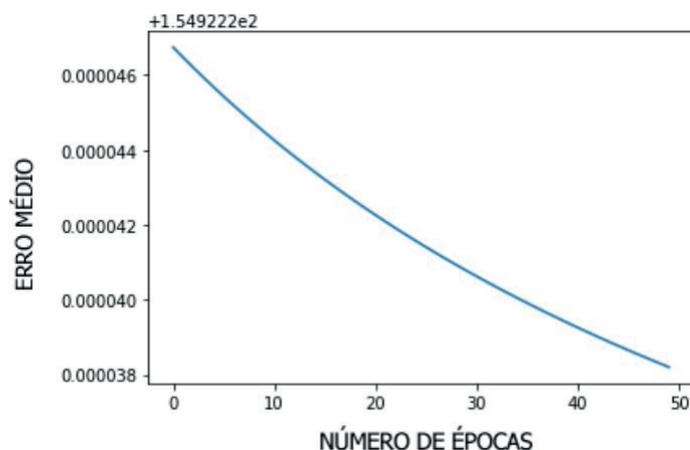
Implementada a rede neural, realizou-se seu treinamento utilizando os valores descritos na Tabela 2.

Tabela 2. Variáveis utilizados para o treinamento.

Nome da variável	Valor utilizado no treinamento
numEpocas	50
q	30
neta	0.05
N	10

Na Figura 2, é mostrado o gráfico do comportamento do erro médio durante a etapa de treinamento, desenvolvido dentro do próprio algoritmo utilizando a biblioteca *matplotlib*.

Figura 2. Erro médio em relação ao número de épocas do treinamento.



Fonte: Autor

A partir dos resultados obtidos, foi possível calcular a diferença entre o valor real e o estimado pela RNA para as variáveis de peso corporal e circunferência abdominal. A Figura 3 exibe o teste de erro, calculado durante a fase de teste através de uma subtração entre os valores dos dados de saída obtidos e valores de saída desejáveis. A menor diferença calculada teve valor de 0.00 e a maior de 3.0 (em kg).

Figura 3. Valores do teste de erro da fase de treinamento.

```
[ [-0.9 -2. ]  
 [-1.6 -2.8]  
 [-1.6 -2. ]  
 [-1.7 -2. ]  
 [-1.1 -1.5]  
 [-1.5 -0.5]  
 [-1.6 -3. ]  
 [-1.4 -1.5]  
 [-1.7 -2. ]  
 [-1.3 -2. ]  
 [-1.4 -3. ]  
 [-2. -1.5]  
 [-1.2 -1.5]  
 [-1.3 -3. ]  
 [-1.6 -1. ]  
 [ 0.3 0. ]  
 [-0.2 0. ]  
 [-0.5 -1. ] ]
```

Fonte: Autor

■ DISCUSSÃO

Após realizados os testes com diferentes configurações do algoritmo, foi possível perceber que, quanto mais dados inseridos na rede para aprendizagem, maior foi o número de épocas necessárias para convergir e atingir o resultado satisfatório. O algoritmo foi capaz de convergir apenas com quantidade relativamente baixas de dados.

À medida que se ampliou a quantidade de dados utilizada para treinamento, foi possível notar a dificuldade da rede em convergir, apresentando grande quantidade de valores discrepantes (*outliers*), principalmente para a variável de peso corporal, o que é natural devido ao metabolismo do ser humano mudar constantemente. Portanto, apesar da rede neural funcionar em relação ao processamento de dados, faz-se necessário fornecer mais informações na camada de entrada para tornar possível a convergência da rede e reconhecimento de padrões, uma vez considerada a grande heterogeneidade da natureza dos dados.

■ CONCLUSÃO / CONSIDERAÇÕES FINAIS

A partir do trabalho desenvolvido, evidencia-se a importância do uso de novas tecnologias visando a prevenção e o combate à obesidade, afirmada como doença crônica em aumento constante entre indivíduos. Além disso, foi possível atingir o desenvolvimento de um novo algoritmo de rede neural artificial em linguagem Python, com arquitetura de duas camadas ocultas e as técnicas de *Feedforward* e *Backpropagation*, para utilização no processo de aprendizagem do algoritmo.

Entretanto, conclui-se também que ainda é possível e necessário coletar dados de novas variáveis que afetam diretamente o metabolismo do corpo humano e influenciam, principalmente, nos valores de peso corporal, a fim de produzir com maior exatidão a estimativa do peso corporal futuro dos indivíduos. Através das pesquisas realizadas, entende-se que o presente trabalho teve grande potencial de importância para a sociedade em geral, pois disponibiliza mais uma ferramenta que pode contribuir no combate à obesidade, facilitando no controle da alimentação adotada por cada indivíduo.

■ REFERÊNCIAS

1. BENEDIKTSSON, Jon A.; SWAIN, Philip H.; ERSOY, Okan K. **Neural network approaches versus statistical methods in classification of multisource remote sensing data**. IEEE Transactions on Geoscience and Remote Sensing, Vancouver, Canada, v. 28, n. 4, p.540-552, jul. 1990.
2. BOURGUET, Jean-rémi et al. **An artificial intelligence-based approach to deal with argumentation applied to food quality in a public health policy**. Expert Systems With Applications, [s.l.], v. 40, n. 11, p.4539-4546, set. 2013.
3. DIAS, Patricia Camacho et al. **Obesidade e políticas públicas: concepções e estratégias adotadas pelo governo brasileiro**. Cad. Saúde Pública, Rio de Janeiro, v. 33, n. 7, 2017. Disponível em: . Acesso em: 30 mai. 2019

4. FLECK, Leandro et al. **REDES NEURAIAS ARTIFICIAIS: PRINCÍPIOS BÁSICOS**. Revista Eletrônica Científica Inovação e Tecnologia, Medianeira, Paraná, v. 1, n. 13, p.47-57, jun. 2016. Disponível em: . Acesso em: 10 jun. 2019.
5. OLIVEIRA, Cecília L.; FISBERG, Mauro. **Obesidade na infância e adolescência: uma verdadeira epidemia**. Arq Bras Endocrinol Metab [online]. 2003, vol. 47, n.2, pp.107-108. ISSN 1677-9487.
6. REENITA DAS. Forbes. **Five Technologies That Will Disrupt Healthcare By 2020**. 2016. Disponível em: . Acesso em: 06 jun. 2019.
7. STEPHENS, Taylor N. et al. **Feasibility of pediatric obesity and prediabetes treatment support through Tess, the AI behavioral coaching chatbot**. Translational Behavioral Medicine, [s.l.], v. 9, n. 3, p.440-447, 16 maio 2019. Oxford University Press (OUP). <http://dx.doi.org/10.1093/tbm/ibz043>.
8. TAVARES, Telma Braga; NUNES, Simone Machado; SANTOS, Mariana de Oliveira. **Obesidade e qualidade de vida: revisão da literatura**. Revista Médica de Minas Gerais, v. 20, n. 3, p. 359-366, 2010.
9. WHO, World Health Organization. **BMI classification**. World Health Organization, Geneva, 2018. Disponível em: http://www.who.int/bmi/index.jsp?introPage=intro_3.html. Acesso em: 29 mai. 2019.

Utilizando serviços em nuvem para reduzir os custos de desenvolvimento e fornecimento de software

| Mário Santos **Sousa**

| Karlos Kelvin **Santos**
UFCA

| Guilherme Álvaro Rodrigues Maia
Esmeraldo
IFCE

RESUMO

Neste artigo é proposto um meio de desenvolver aplicações que possam ser fornecidas a custos mais acessíveis e dessa forma sua aquisição possa ser realizada por empresas dos mais diversos portes. Para alcançar o objetivo, é **proposta** uma arquitetura de software para a aplicações oferecidas como serviço e acessadas através da internet, onde os clientes utilizam uma única instância da aplicação compartilhando recursos de hardware e software. Propõe-se também, como meio de atingir o objetivo, implantá-la na nuvem a fim de reduzir os custos operacionais da mesma e minimizar as preocupações em torno da escalabilidade da aplicação. Para avaliação do trabalho proposto, foi desenvolvida uma aplicação web voltada para a gestão comercial, capaz de servir vários clientes com pouca infraestrutura. Como resultados, foi possível constatar a redução da complexidade de projeto, redução do tempo de desenvolvimento do isolamento de dados, por meio da utilização de tecnologias adequadas.

Palavras-chave: Desenvolvimento de Software, Fornecimento de Software, Redução de Custos, Serviços de Nuvem, Tecnologias Java.

■ INTRODUÇÃO

Em busca de otimizar a gestão das informações e dos processos, as organizações buscam soluções e ferramentas na Tecnologia da Informação (PLACHTA, 2012). Contudo, pequenas e médias empresas podem não ter acesso a uma solução tecnológica adequada por não possuírem recursos para custear todas as etapas de compra e implantação de um sistema convencional. Nesse ponto surge a seguinte pergunta: Como desenvolver aplicações de qualidade com preços acessíveis à empresas dos mais variados portes?

Como uma possível solução, pode-se apontar os serviços em nuvem: IaaS (Infrastructure as a Service), PaaS (Platform as a Service) e SaaS (Software as a Service). De acordo com Veras (2012), IaaS é a capacidade do provedor de oferecer uma infraestrutura de processamento e armazenamento de forma transparente para o cliente, ao passo que PaaS é o modelo de serviço de nuvem onde os clientes utilizam plataformas de desenvolvimento de terceiros. SaaS por sua vez, pode ser definido como “o modelo em que o aplicativo é oferecido como um serviço aos clientes que o acessam através da Internet” (VELTE; VELTE; ELSENPETER, 2012).

Segundo Velte, Velte e Elsenpeter (2012), uma empresa pode se beneficiar de diferentes formas ao adotar serviços de nuvem. Dentre elas é possível citar a redução de custos, uma vez que a tecnologia é contratada de forma incremental o que garante economia a longo prazo.

A partir da utilização de dois serviços de nuvem, PaaS e o SaaS, este trabalho tem como objetivo apresentar uma abordagem para desenvolver e fornecer software a custos menores, facilitando assim a entrada de aplicações de qualidade em qualquer empresa independentemente do porte da mesma.

Para alcançar esse objetivo, pretende-se pela utilização do PaaS reduzir os custos operacionais da aplicação, minimizando os custos de implantação. Para Velte, Velte e Elsenpeter (2012), PaaS permite ao cliente a implantação de aplicações sem os custos da compra de servidores e do armazenamento dos mesmos reduzindo os custos iniciais do projeto. Porém, apenas reduzir os custos operacionais da aplicação não é suficiente para mitigar o valor de venda, portanto é necessário ainda reduzir os custos que os clientes devem arcar durante a contratação de um software.

Para reduzir os custos inerentes à contratação de um software por parte dos clientes optou-se em desenvolver e oferecer a aplicação seguindo o modelo de aplicações SaaS que pode ser definido como: “Software implementado como um serviço hospedado e acessado pela Internet” (CHONG; CARRAO, 2006). Assim, por ser um serviço, o cliente paga apenas pelo que utilizar, garantindo baixo custo e a diminuição dos riscos na aquisição de software, ou seja, o preço de uma aplicação SaaS varia de acordo com a quantidade de funcionalidades

oferecidas, como por exemplo cópias de segurança e restaurações de informações. Com isso, é possível que empresas de menor porte, que necessitam de menos funcionalidades, usem a mesma aplicação que empresas de grande porte.

Veras (2012) aponta que, diferentemente de softwares tradicionais que são instalados no próprio cliente, soluções SaaS possuem uma única instância, o servidor do provedor. Com isso, ao contratar soluções SaaS os clientes não precisam adquirir infraestrutura para executar a aplicação, como servidores, computadores e equipamentos para interconectá-los, o que reduz os gastos durante a contratação do serviço.

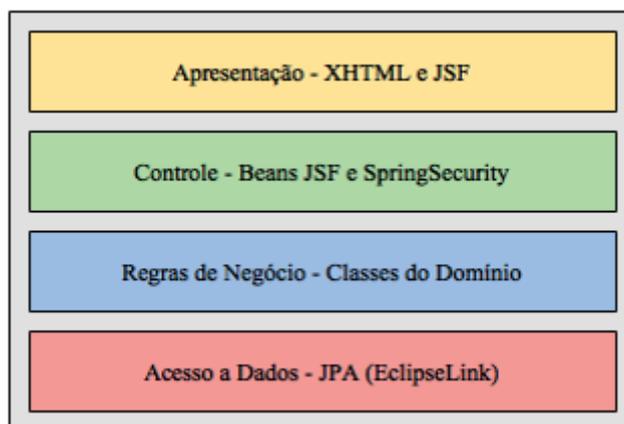
Para avaliar a abordagem proposta foi desenvolvido um software seguindo o modelo multi-tenant para sistemas SaaS que segundo Neto et al (2012 apud HARRIS e AHMED, 2011) é o tipo de solução SaaS onde uma única aplicação é compartilhada por todos os inquilinos (tenants), ou seja, todos os usuários do serviço acessarão uma única instância da aplicação, essa única instância se comportará diferente para cada cliente, baseando-se nas configurações fornecidas por cada inquilino. Como todos os clientes acessarão a mesma instância da aplicação, ela deve fornecer um isolamento dos dados de cada um. O isolamento de dados em aplicações multi-tenant é talvez a característica mais importante desse tipo de solução e é necessário garantir que nenhum dado seja exposto para clientes que não sejam os respectivos proprietários.

Em soluções multi-tenant é possível realizar esse trabalho de diversas formas, uma vez que existe mais de uma abordagem para separação e organização de dados dos clientes no banco de dados. Essas abordagens são: banco de dados separados; banco de dados compartilhado e esquemas separados; e banco de dados e esquema compartilhados (VERAS, 2012). Na abordagem de banco de dados e esquema compartilhados, os dados de todos os clientes ocupam as mesmas tabelas. Dessa forma, a separação de dados deve ser realizada de forma lógica, no código da aplicação, o que é apontado, por Veras (2012) e outros autores, como fator para aumento na complexidade da aplicação, tornando o esforço inicial de desenvolvimento muitas vezes maior que o esperado e, conseqüentemente, encarecendo o projeto. Essa abordagem foi adotada para o desenvolvimento da aplicação citada anteriormente, pois a mesma requer menores investimentos em hardware. Segundo Chong, Carraro e Wolter (2007), isso ocorre pelo fato de que das três abordagens, a última permite servir um maior número de inquilinos por servidor, o que contribui para a diminuição do valor do serviço.

■ MÉTODO

A abordagem proposta se baseia em uma arquitetura de software dividida em camadas (LARMAN, 2007). A Figura 1 apresenta a organização das camadas da abordagem proposta e as tecnologias utilizadas em cada uma delas.

Figura 1. Arquitetura Proposta.



Fonte: Próprios autores (2015).

A primeira camada, denominada camada de apresentação, é a responsável por apresentar aos clientes os dados da aplicação. Abaixo dela, há a camada de controle, que se relaciona com a segurança da aplicação e tratamento de requisições oriundas da camada de apresentação, incluindo atividades de controle de acesso e de permissões. A camada de regras de negócio representa subsistemas que possuem e/ou representam conceitos e regras do domínio da aplicação e por fim, a camada de acesso a dados (DAO) a qual tem responsabilidades diretamente vinculadas a atividades em banco de dados.

Para implementação das camadas apresentadas na Figura 1, utilizou-se, na camada de apresentação, páginas web desenvolvidas com XHTML e componentes JSF; o controle de acesso é realizado pelo Spring Security - um framework baseado na especificação JAAS do Java EE -; a camada de controle é realizada por Beans do JSF (essas classes são responsáveis por tratar eventos nas páginas web e acessam a camada que representa as regras de negócio da aplicação, que por sua vez acessa a camada de acesso a dados responsável por qualquer ação que envolva o banco de dados).

Para camada de acesso a dados foi utilizado o framework EclipseLink que segue a especificação Java Persistence API (JPA), essa especificação provê um modelo de persistência para o mapeamento objeto-relacional, ou seja, o mapeamento de objetos de software para o modelo de banco de dados relacional. Assim, a JPA determina como escrever objetos Java para que eles possam representar o mapeamento e fornece as interfaces que realizam o trabalho entre a aplicação e o banco de dados. Para Lascano (2008), o uso da JPA alivia

o trabalho do programador, uma vez que não é necessário lidar manualmente com dados persistentes e sim delegar essa função a um componente específico.

O fato de a abordagem multi-tenant não ser especificada na JPA, deixa cada framework livre para implementar a funcionalidade segundo seus critérios, isso implicou em um estudo de dois frameworks para identificar o mais apropriado para o uso, tendo em vista as especificidades do projeto e principalmente a arquitetura de dados escolhida para o mesmo.

Foram analisados os dois principais frameworks presentes no mercado, EclipseLink e o Hibernate (ambos são de uso gratuito). A versão disponível do Hibernate no momento do estudo não suporta a abordagem com banco de dados e esquemas compartilhados, ao passo que o EclipseLink suporta a abordagem escolhida. Esse fator determinou a escolha do framework EclipseLink para compor a camada de acesso aos dados e realizar o mapeamento objeto-relacional do projeto.

A implementação do JPA EclipseLink se encarrega de realizar a separação lógica dos dados nas operações com o banco de dados. Durante a execução da aplicação, é necessário criar estruturas que serão responsáveis pela separação lógica dos dados. Essas estruturas encapsulam toda a complexidade da tarefa, restando ao desenvolvedor criar as classes e configurá-las com anotações específicas do EclipseLink. A utilização de tal framework reduziu o tempo e a complexidade para o desenvolvimento do isolamento de dados, o que era apontado por Veras (2012) como fator para aumento no tempo de desenvolvimento em arquiteturas com banco de dados e esquemas compartilhados.

Após o desenvolvimento da aplicação a mesma foi implantada em uma plataforma oferecida como serviço, o OpenShift, mantido pela RedHat. Nessa plataforma é possível criar um ambiente em poucos minutos com servidores de aplicação java e/ou web containers, banco de dados e até mesmo criar um ambiente de integração contínua.

No OpenShift não é possível realizar configurações a nível de sistema operacional, dessa forma o cliente não precisa se preocupar com otimizações desse nível, de contrapartida é possível alterar as configurações das tecnologias escolhidas para o ambiente reforçando a proposta de PaaS onde o foco do cliente se mantém na aplicação. Essa plataforma foi escolhida como provedor PaaS por ser uma plataforma simples de configurar e trabalhar, além do seu acesso gratuito.

■ RESULTADOS E DISCUSSÃO

Para avaliar a abordagem proposta, implementou-se uma aplicação de gestão comercial com suporte a multi-tenant.

Ao utilizar a aplicação pela primeira vez é necessário cadastrar um novo tenant. Essa opção representa a aquisição do serviço pelas empresas, porém sem a necessidade de

pagamento e termos de contrato, visto que não se faz necessário a existência desse tipo de rotina. Porém em sistemas em produção essas rotinas além de logicamente muito importantes representam um meio pelo qual novos clientes podem testar o sistema, pois é comum que fornecedores SaaS ofereçam períodos de teste para a novos usuários a fim de que eles possam verificar se o sistema atende às suas necessidades.

Quando um novo inquilino é cadastrado um novo registro é criado na tabela empresa, essa tabela armazena todos os inquilinos do sistema. A partir do momento em que usuário está “logado” no sistema ele pode operá-lo sem preocupação com exposição dos seus dados, pois não é possível que outro inquilino acesse dados que não pertencem a ele próprio. Para demonstrar o isolamento de dados do sistema foi desenvolvido um cadastro de produtos. Como foi exposto anteriormente, a aplicação faz uso de um único banco de dados e um único esquema, sendo assim, os dados dos clientes estão fisicamente juntos o que torna necessário a separação lógica deles, por meio de filtros na programação. A imagem abaixo demonstra a tabela produto a qual é usada para armazenar os produtos de todos os inquilinos.

Figura 2. Tabela Produto.

	id integer	empresa_id character varying(31)	descricao character varying(255)	preco double precision
1	20	8	soda	3.8
2	19	8	coca-cola	3
3	23	9	mouse	12
4	24	10	notebook	1200
5	25	10	monitor	700
6	27	11	Produto Teste	2000
7	28	11	Produto Teste 001	10.54

Fonte: Próprios autores (2015).

É importante notar que cada produto está associado a um único cliente por meio da coluna empresa_id. O valor dessa coluna é utilizado pelo EclipseLink como parâmetro para qualquer ação desenvolvida no sistema, servindo como parâmetro para qualquer ação no banco de dados. Quando um usuário acessa a interface de gerenciamento de produtos, a aplicação apresenta apenas aqueles que estão associados à mesma empresa do usuário logado.

Analisando a aplicação desenvolvida, percebe-se um alto grau de complexidade. Isso pode ser comprovado a partir do uso do utilitário LocMetrics. Esse utilitário contabiliza dados da aplicação como o total de linhas de código (LOC), linhas comentadas do código, arquivos de código e de diretórios, entre outras informações. Com a análise realizada com o LocMetrics foi identificado que a Média de Complexidade Ciclomática de McCabe (MCCABE, 1976) da aplicação é igual a 206. Essa medida expõe o número de possíveis caminhos que o fluxo de execução da aplicação pode percorrer através da sequência de instruções do código fonte. De acordo com Charney (2005), qualquer aplicação cujo código obtém a média superior a 51, pode ser considerada de alta complexidade. Porém mesmo a aplicação possuindo

alto grau de complexidade, o trabalho despendido para sua construção foi minimizado a partir do uso de frameworks adequados. A Tabela 1 apresenta outros resultados do uso do LocMetrics na aplicação.

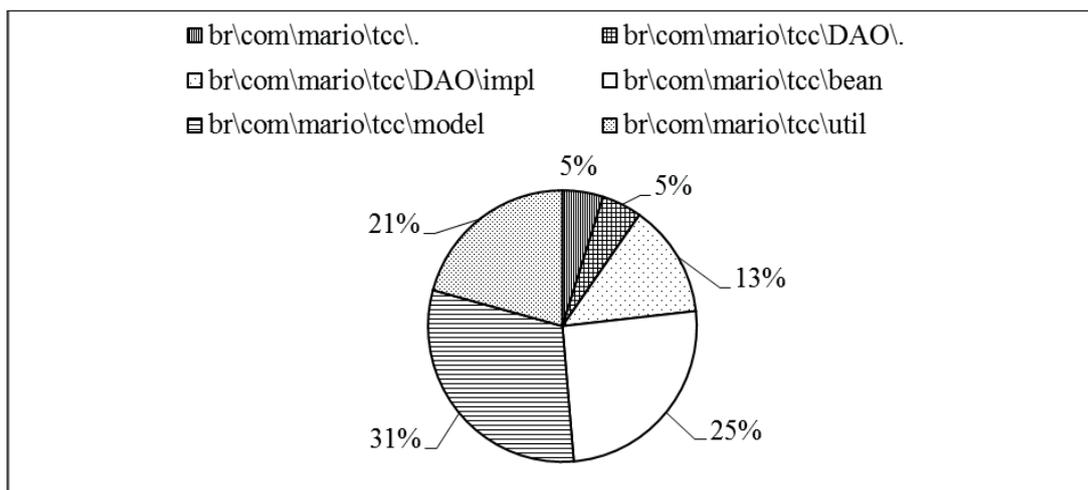
Tabela 1. Resultados do LocMetrics.

Métrica	Total
Arquivos de Código (source files)	34
Diretórios	39
Linhas de Código (LOC)	2129
Linhas de Código Físico e Executável (SLOC-P)	1498
Linhas de Código Lógico e Executável (SLOC-L)	1106
Média de Complexidade McCabe (MVG)	206

Fonte: Próprios autores (2015).

O gráfico seguinte, gerado a partir de informações da análise executada pelo LocMetrics, demonstra em quais diretórios há uma maior concentração de código.

Figura 3. Concentração de código por diretório.



Fonte: Próprios autores (2015).

Esses diretórios representam para a aplicação os pacotes (packages) utilizados para organização do código. O pacote que possui maior número de linhas de código é o br.com.mario.tcc.model, o qual inclui as classes de domínio da aplicação como entidades persistentes e classes que possuem as regras de negócio da aplicação. Os cálculos para essas porcentagens são baseados no SLOC-P, que é a quantidade restante da subtração das linhas em branco e comentários do total de linhas de código.

■ CONCLUSÃO

O presente trabalho abordou a proposta de utilizar os serviços de nuvem, PaaS e SaaS como alternativa para diminuição dos custos de fornecimento, bem como de desenvolvimento de soluções tecnológicas.

Os resultados mostraram que o tempo e a complexidade de desenvolvimento de aplicações SaaS com uma abordagem compartilhada podem ser reduzidos quando escolhidas as tecnologias e práticas corretas para o desenvolvimento. A diminuição no esforço e, conseqüentemente, no tempo e no custo do projeto deve-se também à utilização de ferramentas de alta produtividade. É válido citar como fundamental no desenvolvimento do sistema, e que contribuiu no desempenho do projeto, o framework EclipseLink. Ele permitiu agilizar o desenvolvimento do isolamento de dados, encapsulando a complexidade envolvida neste processo, que é exatamente o ponto onde estava previsto o maior esforço de programação. A redução do tempo no desenvolvimento do isolamento de dados resulta também na redução dos custos de produção de aplicações desse tipo, conseqüentemente o barateamento do projeto, o que pode apoiar a redução dos preços finais da solução.

É fundamental para o desenvolvimento de qualquer software, a preocupação com sua arquitetura e o requisito de escalabilidade. Desenvolver uma aplicação do tipo SaaS, não seria diferente, visto que software como serviço são fornecidos pela internet e possuem naturalmente a expectativa de muitos acessos, uma vez que estão acessíveis em escala global.

Com a abordagem proposta, pode-se reduzir os custos no desenvolvimento e na comercialização de software oferecendo-o como serviço, além da possibilidade de se obter soluções de alta qualidade e comercializadas a custos menores.

O fornecimento de software como serviço abre a possibilidade de alterar o contexto dos mercados para aquisição de software, havendo, em um primeiro cenário, a redução nos valores de custo de soluções preexistentes, e que não são vendidas como serviço; ou a contratação de novas soluções SaaS com valores mais acessíveis.

É possível concluir também que o consumo de plataformas como serviço beneficiam o projeto dando aos clientes a capacidade de focar no produto final, evitando preocupações secundárias como aquisição de servidores e configurações dos mesmos, além de que o uso de PaaS beneficia a economia a longo prazo e tem a capacidade de propiciar um escalonamento mais fácil da aplicação. De forma geral o uso de serviços de nuvem traz vários benefícios para empresas optantes por esse modelo, inclusive para fornecedores de outros serviços de nuvem.

Vale lembrar que a aplicação do estudo de caso deste artigo não está totalmente concluída, foram desenvolvidas apenas funcionalidades essenciais para demonstrar os conceitos inerentes ao trabalho. O mesmo equivale a escolha do OpenShift como plataforma,

pois durante o desenvolvimento da aplicação não foi realizado estudos para determinar que provedor SaaS seria ideal para o projeto, uma vez que não seria necessário realizar essa tarefa imediatamente, pois o aplicativo não está efetivamente em produção. Como trabalhos futuros pretende-se avaliar provedores PaaS e integrar, ao desenvolvimento da aplicação, soluções que visam o aumento da qualidade, como integração contínua e testes.

■ REFERÊNCIAS

1. ANTHONY T. V.; TOBY J. V.; ELSENPETER R. **Cloud Computing: Computação em Nuvem Uma abordagem Prática**. 2. ed. Rio de Janeiro: Alta Books. 2012. p. 173 – 191. cap. 9.
2. CHARNEY, R. Programming Tools: **Code Complexity Metrics**. 2005. Disponível em <<http://www.linuxjournal.com/article/8035>>. Acesso em: 15 mai. 2015.
3. CHONG F.; CARRARO G.; WOLTER R. **Arquitetura de dados para múltiplos inquilinos**. Microsoft Corporation. 2006. Disponível em <<http://msdn.microsoft.com/pt-br/library/aa479086.aspx>> Acesso em: 15 mai. 2014.
4. CHONG F.; CARRARO G. **Estratégias de arquitetura para cauda longa (Long Tail)**. Microsoft Corporation. 2006. Disponível em <<http://msdn.microsoft.com/pt-br/library/aa479069.aspx>> Acesso em: 15 mai. 2014.
5. LARMAN, Craig. **Utilizando UML e padrões: uma introdução à análise e ao projeto orientados a objetos e ao desenvolvimento iterativo** / Craig Larman. 3. ed. Porto Alegre : Bookman. 2007.
6. LASCANO, Jorge Edison. **JPA implementations versus pure JDBC**. 2008. Disponível em: <http://www.espe.edu.ec/portal/files/sitiocongreso/congreso/c_computacion/PaperJPAversusJDBC_edisonlascano.pdf>. Acesso em: 29 jun. 2015.
7. NETO, Josino Rodrigues et al. **Software as a Service: Desenvolvendo Aplicações Multi-tenancy com Alto Grau de Reuso**. 2012.
8. PLACHTA, C. A tecnologia no Suporte à Gestão da Informação e aos Processos de Negócios Inteligentes. In: **Gestão da informação, inovação e inteligência competitiva**. 1. ed. Rio de Janeiro: Editora Saraiva. 2012. p. 91 – 117. cap. 6.

SOBRE O ORGANIZADOR

Prof. Dr. Ernane Rosa Martins

Doutor em Ciência da Informação com ênfase em Sistemas, Tecnologias e Gestão da Informação, na Universidade Fernando Pessoa, em Porto/Portugal. Mestre em Engenharia de Produção e Sistemas, possui Pós-Graduação em Tecnologia em Gestão da Informação, Graduação em Ciência da Computação e Graduação em Sistemas de Informação. Professor de Informática no Instituto Federal de Educação, Ciência e Tecnologia de Goiás - IFG (Câmpus Luziânia) ministrando disciplinas nas áreas de Engenharia de Software, Desenvolvimento de Sistemas, Linguagens de Programação, Banco de Dados e Gestão em Tecnologia da Informação. Pesquisador do Núcleo de Inovação, Tecnologia e Educação (NITE), certificado pelo IFG no CNPq. Membro do Conselho Editorial da Editora Científica Digital. Membro do Conselho Técnico Científico da Atena Editora. Membro do Corpo Editorial da Pantanal Editora. Membro do Conselho Editorial da Editora Bagai. Membro do Conselho Editorial da Editora e-Publicar. ORCID: <https://orcid.org/0000-0002-1543-1108>. Personal homepage: <https://ernanemartins.wordpress.com>.

ÍNDICE REMISSIVO

A

Algoritmo: 275

Análise: 14, 15, 18, 20, 27, 39, 40, 43, 59, 80, 169, 178, 179, 180, 242

Aprendizagem: 15, 26, 33, 34, 39, 40, 50, 72, 73, 74, 76, 79, 85, 175, 184

Atendimento: 109, 114

Atributos: 63

Avaliação: 16, 17, 20, 47, 51, 54, 56, 57, 76, 85, 115, 119, 250, 251, 252, 253, 254, 255, 256, 258

C

Capacitação: 69, 70

Ciência: 44, 107, 138, 151, 200, 237

Clustering: 15, 19, 25, 45

Código: 193, 261, 290

Computação: 21, 25, 26, 94, 137, 151, 170, 176, 235, 249, 250, 261, 267, 268, 271, 272, 273, 292

Conhecimento: 141, 143, 144, 145, 150, 151

D

Decisão: 36, 37, 151

Desempenho: 80, 81, 82, 111, 153, 170

E

Educação: 5, 6, 7, 9, 25, 26, 44, 45, 54, 84, 85, 172, 173, 183, 184, 185, 200, 237, 273

Ensino: 7

Estudo de Caso: 107, 151, 206

Evolução: 28

Expert Sinta: 209

G

Gerenciamento: 39, 86, 88, 94, 244, 245, 246, 247

Gestão: 50, 94, 109, 110, 143, 144, 145, 150, 151, 170, 237, 238, 242, 245, 246, 292

I

Inclusão Digital: 172

Informática: 26, 45, 85, 94, 172, 173, 174, 182, 183, 184, 185, 273

Inteligência Artificial: 207, 217

L

Licenciamento: 90, 151, 207, 208, 210

M

Mapeamento: 15, 18, 20, 239

Matemática: 105, 172, 173, 174, 175, 184, 185, 219, 220, 221, 226, 230, 233, 234

Mecanismo: 96

Mineração de Dados: 29

Modelagem: 139, 140, 144, 145, 151, 207, 239, 246

N

Nuvem: 292

P

Pandemia: 173, 197

Perfis: 15, 18

Práticas: 169, 170, 237

Problemas: 171, 172, 175, 176, 181, 184, 244

Produtividade: 194

Programação: 15, 26, 39, 175, 180, 185, 197, 234, 235, 273

R

Recursos: 184

Redes: 36, 37, 94, 123, 124, 135, 138, 275

Robótica: 196, 197, 198

S

Saúde: 77, 108, 113, 117, 118, 119, 249, 276, 281

Scratch: 47, 48, 49, 52, 54, 55, 174, 183, 184, 185, 204, 263, 264, 267, 268, 271, 272, 273

Serviços: 247

Sistemas: 39, 58, 94, 109, 114, 121, 169, 170, 207, 209, 223

Software: 37, 94, 195, 218, 219, 237, 262, 285, 292

T

Tecnologias: 54, 71, 84, 85, 109

Tendências: 43

U

Usabilidade: 47, 48, 49, 51

W

Web: 32, 37, 40, 49, 90, 99, 137, 138, 155, 159, 161, 162, 169, 170, 256, 257, 266



follow us



www.editoracientifica.org

contato@editoracientifica.org

ISBN 978-658982602-6



9 786589 826026

VENDA PROIBIDA - ACESSO LIVRE - OPEN ACCESS



editora científica