



**MARCELO PEREIRA DA SILVA
(ORGANIZADOR)**

COMUNICAÇÃO, MÍDIAS E EDUCAÇÃO 2

Atena
Editora
Ano 2019

Marcelo Pereira da Silva
(Organizador)

Comunicação, Mídias e Educação

2

Atena Editora
2019

2019 by Atena Editora
Copyright © Atena Editora
Copyright do Texto © 2019 Os Autores
Copyright da Edição © 2019 Atena Editora
Editora Chefe: Profª Drª Antonella Carvalho de Oliveira
Diagramação: Geraldo Alves
Edição de Arte: Lorena Prestes
Revisão: Os Autores



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

O conteúdo dos artigos e seus dados em sua forma, correção e confiabilidade são de responsabilidade exclusiva dos autores. Permitido o download da obra e o compartilhamento desde que sejam atribuídos créditos aos autores, mas sem a possibilidade de alterá-la de nenhuma forma ou utilizá-la para fins comerciais.

Conselho Editorial

Ciências Humanas e Sociais Aplicadas

Profª Drª Adriana Demite Stephani – Universidade Federal do Tocantins
Prof. Dr. Álvaro Augusto de Borba Barreto – Universidade Federal de Pelotas
Prof. Dr. Alexandre Jose Schumacher – Instituto Federal de Educação, Ciência e Tecnologia de Mato Grosso
Prof. Dr. Antonio Carlos Frasson – Universidade Tecnológica Federal do Paraná
Prof. Dr. Antonio Isidro-Filho – Universidade de Brasília
Prof. Dr. Constantino Ribeiro de Oliveira Junior – Universidade Estadual de Ponta Grossa
Profª Drª Cristina Gaio – Universidade de Lisboa
Prof. Dr. Deyvison de Lima Oliveira – Universidade Federal de Rondônia
Prof. Dr. Edvaldo Antunes de Faria – Universidade Estácio de Sá
Prof. Dr. Eloi Martins Senhora – Universidade Federal de Roraima
Prof. Dr. Fabiano Tadeu Grazioli – Universidade Regional Integrada do Alto Uruguai e das Missões
Prof. Dr. Gilmei Fleck – Universidade Estadual do Oeste do Paraná
Profª Drª Ivone Goulart Lopes – Istituto Internazionele delle Figlie de Maria Ausiliatrice
Prof. Dr. Julio Candido de Meirelles Junior – Universidade Federal Fluminense
Profª Drª Keyla Christina Almeida Portela – Instituto Federal de Educação, Ciência e Tecnologia de Mato Grosso
Profª Drª Lina Maria Gonçalves – Universidade Federal do Tocantins
Profª Drª Natiéli Piovesan – Instituto Federal do Rio Grande do Norte
Prof. Dr. Marcelo Pereira da Silva – Universidade Federal do Maranhão
Profª Drª Miranilde Oliveira Neves – Instituto de Educação, Ciência e Tecnologia do Pará
Profª Drª Paola Andressa Scortegagna – Universidade Estadual de Ponta Grossa
Profª Drª Rita de Cássia da Silva Oliveira – Universidade Estadual de Ponta Grossa
Profª Drª Sandra Regina Gardacho Pietrobon – Universidade Estadual do Centro-Oeste
Profª Drª Sheila Marta Carregosa Rocha – Universidade do Estado da Bahia
Prof. Dr. Rui Maia Diamantino – Universidade Salvador
Prof. Dr. Urandi João Rodrigues Junior – Universidade Federal do Oeste do Pará
Profª Drª Vanessa Bordin Viera – Universidade Federal de Campina Grande
Prof. Dr. Willian Douglas Guilherme – Universidade Federal do Tocantins

Ciências Agrárias e Multidisciplinar

Prof. Dr. Alan Mario Zuffo – Universidade Federal de Mato Grosso do Sul
Prof. Dr. Alexandre Igor Azevedo Pereira – Instituto Federal Goiano
Profª Drª Daiane Garabeli Trojan – Universidade Norte do Paraná
Prof. Dr. Darllan Collins da Cunha e Silva – Universidade Estadual Paulista
Profª Drª Diocléa Almeida Seabra Silva – Universidade Federal Rural da Amazônia
Prof. Dr. Fábio Steiner – Universidade Estadual de Mato Grosso do Sul
Profª Drª Girlene Santos de Souza – Universidade Federal do Recôncavo da Bahia
Prof. Dr. Jorge González Aguilera – Universidade Federal de Mato Grosso do Sul
Prof. Dr. Júlio César Ribeiro – Universidade Federal Rural do Rio de Janeiro
Profª Drª Raissa Rachel Salustriano da Silva Matos – Universidade Federal do Maranhão
Prof. Dr. Ronilson Freitas de Souza – Universidade do Estado do Pará
Prof. Dr. Valdemar Antonio Paffaro Junior – Universidade Federal de Alfenas

Ciências Biológicas e da Saúde

Prof. Dr. Benedito Rodrigues da Silva Neto – Universidade Federal de Goiás
Prof. Dr. Edson da Silva – Universidade Federal dos Vales do Jequitinhonha e Mucuri
Profª Drª Elane Schwinden Prudêncio – Universidade Federal de Santa Catarina
Prof. Dr. Gianfábio Pimentel Franco – Universidade Federal de Santa Maria
Prof. Dr. José Max Barbosa de Oliveira Junior – Universidade Federal do Oeste do Pará
Profª Drª Magnólia de Araújo Campos – Universidade Federal de Campina Grande
Profª Drª Natiéli Piovesan – Instituto Federal do Rio Grande do Norte
Profª Drª Vanessa Lima Gonçalves – Universidade Estadual de Ponta Grossa
Profª Drª Vanessa Bordin Viera – Universidade Federal de Campina Grande

Ciências Exatas e da Terra e Engenharias

Prof. Dr. Adélio Alcino Sampaio Castro Machado – Universidade do Porto
Prof. Dr. Alexandre Leite dos Santos Silva – Universidade Federal do Piauí
Profª Drª Carmen Lúcia Voigt – Universidade Norte do Paraná
Prof. Dr. Eloi Rufato Junior – Universidade Tecnológica Federal do Paraná
Prof. Dr. Fabrício Menezes Ramos – Instituto Federal do Pará
Prof. Dr. Juliano Carlo Rufino de Freitas – Universidade Federal de Campina Grande
Profª Drª Neiva Maria de Almeida – Universidade Federal da Paraíba
Profª Drª Natiéli Piovesan – Instituto Federal do Rio Grande do Norte
Prof. Dr. Takeshy Tachizawa – Faculdade de Campo Limpo Paulista

Dados Internacionais de Catalogação na Publicação (CIP) (eDOC BRASIL, Belo Horizonte/MG)	
C741	Comunicação, mídias e educação 2 [recurso eletrônico] / Organizador Marcelo Pereira da Silva. – Ponta Grossa, PR: Atena Editora, 2019. – (Comunicação, Mídias e Educação; v. 2) Formato: PDF Requisitos de sistema: Adobe Acrobat Reader. Modo de acesso: World Wide Web. Inclui bibliografia ISBN 978-85-7247-695-9 DOI 10.22533/at.ed.959190910 1. Aprendizagem. 2. Comunicação – Pesquisa – Brasil. 3. Comunicação na educação. I. Silva, Marcelo Pereira da. II. Série. CDD 371.1022
Elaborado por Maurício Amormino Júnior – CRB6/2422	

Atena Editora
Ponta Grossa – Paraná - Brasil
www.atenaeditora.com.br
contato@atenaeditora.com.br

APRESENTAÇÃO

Comunicação, Mídias e Educação constituem campos do saber que se entrecruzam e emolduram, por meio de suas especificidades de objetos e arcabouços teóricos, metodológicos e epistemológicos, fragmentos relevantes da arquitetura na qual a sociedade, as organizações e os atores sociais se assentam.

Nesse sentido, este livro contém um tripé, a saber, Comunicação, Mídias e Educação, que se (im)põe como condição na construção da sociedade e na consolidação da democracia, da participação, do diálogo e da análise crítica que alimenta as possibilidades de compreensão da complexa sociedade na qual estamos imersos.

A Comunicação, as Mídias e a Educação, assim, se apresentam como três campos que materializam múltiplas expectativas, desafios e oportunidades em um tempo no qual emergentes formas de ver, estar e sentir o mundo ressignificam o tecido social, redefinem profissões e produzem outras formas de interação, trocas e socialidades.

Destarte, dividimos esta obra em 2 partes: A primeira problematiza, por meio de diferentes métodos, análises, discussões e epistemes, o universo das Redes e Mídias Sociais da Internet, contendo artigos que tratam dos atores que emergem com o surgimento e a cimentação das redes sociotécnicas, os discursos que circulam no ecossistema virtual e os obstáculos decorrentes dessa ecologia.

A segunda parte engloba artigos que versam acerca das Mídias e do Jornalismo, lançando luz sobre a constituição das mídias sociais da Internet e das mídias de massa, assim como no lugar que o Jornalismo ocupa no contexto pós-moderno. Por meio de diversas discussões, metodologias e problematizações que aprofundam o olhar sobre as Mídias e o Jornalismo, tais artigos pavimentam a estrada por onde caminham, ainda que em sentidos que ora divergem e ora convergem nas interfaces entre Comunicação, Mídias e Educação.

Temos que Comunicação, Mídias e Educação devem ser entendidas e colocadas no centro da existência humana, dado que se tornaram medulares para a construção de uma sociedade mais aberta, justa, empática e sensível às demandas das labirínticas alteridades.

Marcelo Pereira da Silva

SUMÁRIO

PARTE 1: REDES E MÍDIAS SOCIAIS DA INTERNET

CAPÍTULO 1	1
A CELEBRIDADE PELO ESCÁRNIO: GRETCHEN, RISO E REDES SOCIAIS	
Jaciane Freire Santana João Gabriel Lourenço da Silva Santos Fabiana Moraes da Silva	
DOI 10.22533/at.ed.9591909101	
CAPÍTULO 2	11
A FORMAÇÃO DE EDITORIAS DE MÍDIAS SOCIAIS EM REDAÇÕES JORNALÍSTICAS E OS DILEMAS SOBRE O PROFISSIONAL “HÍBRIDO”	
Robson Roque Ivan Satuf	
DOI 10.22533/at.ed.9591909102	
CAPÍTULO 3	24
AMINER.: METADADOS DE PESQUISAS ACADÊMICAS ATRAVÉS DA INTELIGÊNCIA ARTIFICIAL	
Giuliano Carlo Rainatto Genesio Renovato da Silva Neto Jucilene Faria Norberto de Almeida Andrade	
DOI 10.22533/at.ed.9591909103	
CAPÍTULO 4	39
ANÁLISE DISCURSIVA DE PERFIS DE DIGITAIS INFLUENCERS MIRINS	
Elane da Silva Sousa Regysane Botelho Cutrim Alves	
DOI 10.22533/at.ed.9591909104	
CAPÍTULO 5	53
ECOSSISTEMA DA DES-INFORMAÇÃO: TIPOS DE CONTEÚDOS FRAUDULENTOS NAS ELEIÇÕES PRESIDENCIAIS DE 2018	
Kennedy Anderson Cupertino de Souza Marilene Mattos Salles	
DOI 10.22533/at.ed.9591909105	
CAPÍTULO 6	66
FAKENEWS NA ATUALIDADE: ESTUDO DA DISSEMINAÇÃO DE NOTÍCIAS FALSAS COMO RECURSO DE PROPAGABILIDADE	
Luiz Guilherme de Brito Arduino Vânia de Moraes	
DOI 10.22533/at.ed.9591909106	

CAPÍTULO 7	77
JORNALISMO ESPORTIVO E E-SPORTS: UMA ANÁLISE DA COBERTURA JORNALÍSTICA SOBRE A POSSÍVEL INSERÇÃO NOS JOGOS OLÍMPICOS	
Guilherme Fernandes Mota Silva	
Luísa Guimarães Lima	
DOI 10.22533/at.ed.9591909107	
CAPÍTULO 8	88
MUDANÇA DE COMPORTAMENTO DO CONSUMIDOR DE NOTÍCIAS COM O FENÔMENO SEGUNDA TELA	
Suély Zonta	
DOI 10.22533/at.ed.9591909108	
CAPÍTULO 9	98
MÍDIAS DIGITAIS E CAPITAL SOCIAL: UM ESTUDO SOBRE AS AÇÕES DA CI COMUNICAÇÃO E INFORMAÇÃO NO FACEBOOK	
Gabriel Gustavo Carneiro Braga	
Letícia Silva Mendonça	
Carolina Guerra Libério	
DOI 10.22533/at.ed.9591909109	
CAPÍTULO 10	110
O ELEITOR CONECTADO: PERFIL E CONSUMO DE CONTEÚDO NAS ELEIÇÕES 2018	
Maíra Martins Moraes Vitorino	
DOI 10.22533/at.ed.95919091010	
CAPÍTULO 11	125
PARA QUE SERVE UM BOATO NUMA CRISE DEMOCRÁTICA? REFLEXÕES SOBRE OS SINTOMAS, A PARTICIPAÇÃO E A UTILIDADE DOS BOATOS NA CRISE DA DEMOCRACIA BRASILEIRA	
Iasminny Thábata Sousa Cruz	
DOI 10.22533/at.ed.95919091011	
CAPÍTULO 12	138
PÁGINA BOLSOMINIONS ARREPENDIDOS: UMA ANÁLISE QUALI-QUANTITATIVA	
Izailma Jaciara Araujo Costa	
Márcia Inabelly Araújo dos Santos	
DOI 10.22533/at.ed.95919091012	
PARTE 2: MÍDIAS, JORNALISMO E ANÁLISES	
CAPÍTULO 13	147
COMPOSIÇÃO, CORES E SUBJETIVIDADE: ESTUDO DO DISCURSO PREGNANTE EM INFOGRÁFICOS DO “LA NACIÓN DATA” E “ESTADÃO DADOS”	
Kelly De Conti Rodrigues	
Carlos Alberto Garcia Biernath	
Marcelo Pereira da Silva	
DOI 10.22533/at.ed.95919091013	

CAPÍTULO 14	161
A INVISIBILIDADE DO COVEIRO E O JORNALISMO LITERÁRIO: HISTÓRIAS DE PARATINGA	
Tiago Florencio de Abreu Angelita Pereira de Lima	
DOI 10.22533/at.ed.95919091014	
CAPÍTULO 15	170
A QUEDA HISTÓRICA DE ANUNCIANTES NO JORNAL O POPULAR: UMA REFLEXÃO SOBRE O FUTURO DO JORNAL IMPRESSO E SUA CRISE DE FINANCIAMENTO	
Edson Francisco Leite Junior	
DOI 10.22533/at.ed.95919091015	
CAPÍTULO 16	182
A REPRESENTATIVIDADE FEMININA NA SÉRIE <i>ORANGE IS THE NEW BLACK</i>	
Camilla Pessoa Barros Bibiano	
DOI 10.22533/at.ed.95919091016	
CAPÍTULO 17	191
BLOCKCHAIN E JORNALISMO DIGITAL: UMA REFLEXÃO SOBRE O MODELO DE NEGÓCIOS DA EMPRESA THE CIVIL MEDIA COMPANY	
Lucas Rezende Costa Luísa Guimarães Lima	
DOI 10.22533/at.ed.95919091017	
CAPÍTULO 18	202
COTAS UNIVERSITÁRIAS NAS COLUNAS DE CARTA CAPITAL: UM ESTUDO DE REPRESENTAÇÃO SOCIAL	
Elisa Fabris de Oliveira Edinete Maria Rosa	
DOI 10.22533/at.ed.95919091018	
CAPÍTULO 19	214
DO FEMININO AO FEMINISMO: UMA ANÁLISE DE REPORTAGENS NA REVISTA AZMINA	
Thais Martins Rossi Maria Emília Pelisson Manente	
DOI 10.22533/at.ed.95919091019	
CAPÍTULO 20	226
FEMINICÍDIO E A IMPRENSA BRASILEIRA: ANÁLISE DA COBERTURA DIGITAL SOBRE O CASO TATIANE SPITZNER	
Bruna Silvestre Innocenti Giorgi	
DOI 10.22533/at.ed.95919091020	

CAPÍTULO 21	238
IMAGINÁRIO, MULTICULTURALISMO E APOCALIPSE NO FILME CÍRCULO DE FOGO	
Rafael Iwamoto Tosi	
DOI 10.22533/at.ed.95919091021	
CAPÍTULO 22	250
LIMITAÇÕES À DIVERSIDADE SIGNIFICATIVA DE VERSÕES NAS NOTÍCIAS: A COBERTURA D'O GLOBO E DO DIARIO DE PERNAMBUCO NA ELEIÇÃO PRESIDENCIAL DE 2018	
Nathália Carvalho Advíncula	
Maria Clara de Oliveira Martins	
Heitor Costa Lima da Rocha	
DOI 10.22533/at.ed.95919091022	
CAPÍTULO 23	262
O CINEMA E O DUPLO: ANÁLISE MIMÉTICA DO FILME MARINA	
Bárbara dos Santos Oliveira	
Crislene Susane Fernandes Moreira	
Alexandre Bruno Gouveia Costa	
DOI 10.22533/at.ed.95919091023	
CAPÍTULO 24	273
O SENSACIONALISMO E A FUNÇÃO SOCIAL DO JORNALISMO POLICIAL MARANHENSE: UM ESTUDO DO PROGRAMA BANDEIRA 2	
Samantha Kelly Tinôco Araújo	
Alexandre Bruno Gouveia Costa	
DOI 10.22533/at.ed.95919091024	
CAPÍTULO 25	284
<i>PANTHER IS THE NEW BLACK</i> : REPRESENTATIVIDADE E CULTURA NA COMUNICAÇÃO DO FILME PANTERA NEGRA	
Rodrigo Sérgio Ferreira de Paiva	
DOI 10.22533/at.ed.95919091025	
CAPÍTULO 26	297
POR TRÁS DAS GRADES: O SILÊNCIO SOBRE OS PRESÍDIOS FEMININOS NO BRASIL	
Gabriel Barros da Silva Eduardo	
Julia Borsoi de Oliveira	
Natalia Vicente Teixeira	
Maria Emilia Pelisson Manente	
William Silva de Oliveira	
DOI 10.22533/at.ed.95919091026	
CAPÍTULO 27	306
PRODUÇÃO IMAGÉTICA NO CINEMA: CONVERGÊNCIAS REPRESENTATIVAS ENTRE AS PRODUÇÕES DE JEAN-LUC GODARD E LARS VON TRIER	
Marcelo dos Santos Matos	
DOI 10.22533/at.ed.95919091027	

CAPÍTULO 28	316
VALORES-NOTÍCIA NO JORNALISMO AUTOMOTIVO	
Sergio Quintanilha	
DOI 10.22533/at.ed.95919091028	
CAPÍTULO 29	330
UMA REVISÃO NARRATIVA DA MÍDIA E DA SAÚDE MENTAL NA ATUALIDADE	
Paloma da Silva	
Andressa Rosa de Araújo	
Bianca Gonçalves de Carrasco Bassi	
DOI 10.22533/at.ed.95919091029	
CAPÍTULO 30	344
TEORIAS DO JORNALISMO: UMA REVISÃO BIBLIOGRÁFICA PARA DESCONSTRUIR AS <i>FAKE NEWS</i>	
Gabriela Souza Silva	
Mariana Oliveira Santos	
Carmen Regina de Oliveira Carvalho	
DOI 10.22533/at.ed.95919091030	
CAPÍTULO 31	356
RETRATOS E IDENTIDADES DO LICEU MARANHENSE: UMA VIVÊNCIA DA ARTE DA FOTOGRAFIA NO COTIDIANO ESCOLAR	
Elma Vilma Silva Ferreira	
Ellen Lucy Viana	
DOI 10.22533/at.ed.95919091031	
SOBRE O ORGANIZADOR	363
ÍNDICE REMISSIVO	364

AMINER.: METADADOS DE PESQUISAS ACADÊMICAS ATRAVÉS DA INTELIGÊNCIA ARTIFICIAL

Giuliano Carlo Rainatto
Genesio Renovato da Silva Neto
Jucilene Faria
Norberto de Almeida Andrade

RESUMO: Neste artigo, apresentamos um novo mecanismo de busca acadêmica online de pesquisa e baseado em Inteligência Artificial (IA), o AMiner, que visa superar o Google Scholar está expandindo seu corpus de artigos para cobrir cerca de 10 milhões de artigos de pesquisa em administração, ciência da computação e neurociência, entre outras áreas de igual importância. Ele faz parte da segunda geração do sistema ArnetMiner. Desde o seu lançamento em 2008, juntaram-se vários outros mecanismos de buscas acadêmicas baseadas em IA prometendo classificar e classificar trabalhos acadêmicos usando uma compreensão mais sofisticada de seu conteúdo e contexto. Os algoritmos e dados de pesquisa acadêmica do AMiner estão disponíveis para pesquisadores por meio de uma interface de programação de aplicativos (API). A AMiner reuniu um grande conjunto de dados acadêmicos com mais de 180.000.000 de perfis de pesquisadores e 294.000.000 de publicações da Internet até novembro de 2018, juntamente com vários subconjuntos que foram construídos para diferentes fins de pesquisa.

PALAVRAS-CHAVE: Bibliometria, AMiner, Metadados, Inteligência Artificial

ABSTRACT: In this article, we present a new online academic search engine for Artificial Intelligence (AI) search and use, the AMiner, which targets Google Scholar is expanding its article corpus to around 10 million business research articles. , Computer science and neuroscience, among other areas of equal importance. It is part of the second generation ArnetMiner system. Since its launch in 2008, other AI academic search engines have teamed up to classify and classify scholarly works using a more sophisticated understanding of their content and context. Algorithms and academic research data are available to researchers through an application programming interface (API). AMiner has gathered a large academic dataset with more than 180,000,000 researcher profiles and 294,000,000 Internet publications for Brazil from 2018, along with several subsets that were built for different research purposes.

KEYWORDS: Bibliometrics, AMiner, Metadata, Artificial Intelligence

1 | INTRODUÇÃO

A bibliometria é o uso de métodos estatísticos para analisar os dados dos índices

de citação podem ser analisados para determinar a popularidade e o impacto de artigos, autores e publicações específicas. A análise de citações é um método bibliométrico comumente usado que é baseado na construção do grafo de citações., uma representação de rede ou gráfico das citações entre documentos (Wu, 2011).

Muitos campos de pesquisa usam métodos bibliométricos para explorar o impacto de seu campo, o impacto de um conjunto de pesquisadores, o impacto de um determinado artigo ou para identificar documentos particularmente impactantes dentro de um campo específico de pesquisa (Wolfram, 2016).

O ArnetMiner (também conhecido como AMiner) é um serviço online gratuito usado para indexar, pesquisar e extrair grandes dados científicos. O sistema foi projetado para pesquisar e executar operações de mineração de dados em publicações acadêmicas na Internet, usando a análise de redes sociais para identificar conexões entre pesquisadores, conferências e publicações (Nasution, Noah e Saad, 2016).

Isso permite fornecer serviços como: descoberta de especialistas, pesquisa geográfica, análise de tendências, recomendação de revisores, pesquisa de associação, pesquisa de curso, avaliação de desempenho acadêmico e modelagem de tópicos. O ArnetMiner foi criado como um projeto de pesquisa em análise de influência social, ranking de redes sociais e extração de redes sociais (Shah e Pudi, 2019).

O ArnetMiner é comumente usado na academia para identificar relações entre e desenhar correlações estatísticas sobre pesquisa e pesquisadores (Bobadilla et. al., 2019).

Ele atraiu mais de 10 milhões de acessos IP independentes de 220 países e regiões. O produto foi usado na plataforma SciVerse da Elsevier, e em conferências acadêmicas como Academy of Management, ICBMSI (Conferência Internacional sobre Gestão Empresarial e Inovação Social, ICMTI (Conferência Internacional sobre Gestão e Tecnologia da Informação, entre outros (Safder, Hassan e Aljohani, 2018).

O ArnetMiner extrai automaticamente o perfil do pesquisador da web. Ele coleta e identifica as páginas relevantes e usa uma abordagem unificada para extrair dados dos documentos identificados. Também extrai publicações de bibliotecas digitais online usando regras heurísticas (Cabanac, Frommholz e Mayr, 2019).

Integra os perfis dos pesquisadores extraídos e as publicações extraídas. Emprega o nome do pesquisador como o identificador. Uma estrutura probabilística foi proposta para lidar com o problema da ambigüidade do nome na integração. Os dados integrados são armazenados em uma base de conhecimento da rede de pesquisadores (RNKB). Os outros produtos principais da área são o Google Scholar, o Scirus da Elsevier e o projeto de código aberto CiteSeer (Alyami e Assiri, 2018).

A variedade de sites acadêmicos de redes sociais, incluindo o ResearchGate, Microsoft Academic, Google Scholar, Semantic Scholar, ResearchGate e Academia.

edu ganharam grande popularidade ao longo da última década (Baepler e Murdoch, 2010).

O objetivo comum desses sistemas de rede social acadêmica é fornecer aos pesquisadores uma plataforma integrada para consultar informações e recursos acadêmicos, compartilhar suas próprias conquistas e se conectar com outros pesquisadores (Bobadilla et. al., 2019).

Diversas questões dentro das redes sociais acadêmicas foram investigadas nesses sistemas. No entanto, a maioria dos problemas é investigada separadamente por meio de processos independentes (Borkar e Rajeswari, 2013). Como tal, não existe um processo congruente ou uma série de métodos para a mineração de redes sociais acadêmicas diferentes (Cabanac, Frommholz e Mayr, 2019).

Segue Nasution e Noah (2011) a falta de tais métodos pode ser atribuída a dois motivos:

(1) Falta de informação baseada em semântica. As informações de perfil de usuário obtidas apenas do usuário que inseriu suas informações ou extraídas por heurística são, por vezes, incompletas ou inconsistentes. Os usuários não preenchem informações pessoais apenas porque não estão dispostos a fazê-lo;

(2) Falta de uma abordagem de modelagem unificada para mineração efetiva da rede social. Tradicionalmente, diferentes tipos de fontes de informação na rede social acadêmica foram modelados individualmente e, portanto, as dependências entre eles não podem ser capturadas. No entanto, podem existir dependências entre dados sociais. Serviços de busca de alta qualidade precisam considerar as dependências intrínsecas entre as diferentes fontes de informação heterogêneas.

Neste cenário, através da plataforma AMiner, nosso objetivo é responder a quatro questões:

- (1) Como extrair automaticamente o perfil do pesquisador da Web existente?
- (2) Como integrar as informações extraídas (ou seja, perfis e publicações dos pesquisadores) de diferentes fontes?
- (3) Como modelar os diferentes tipos de fontes de informação em um modelo unificado?
- (4) Como fornecer serviços de pesquisa avançados em uma rede construída?

Para responder às perguntas acima, uma série de novas abordagens são implementadas dentro do sistema AMiner. A arquitetura geral do sistema é mostrada na Figura 1.

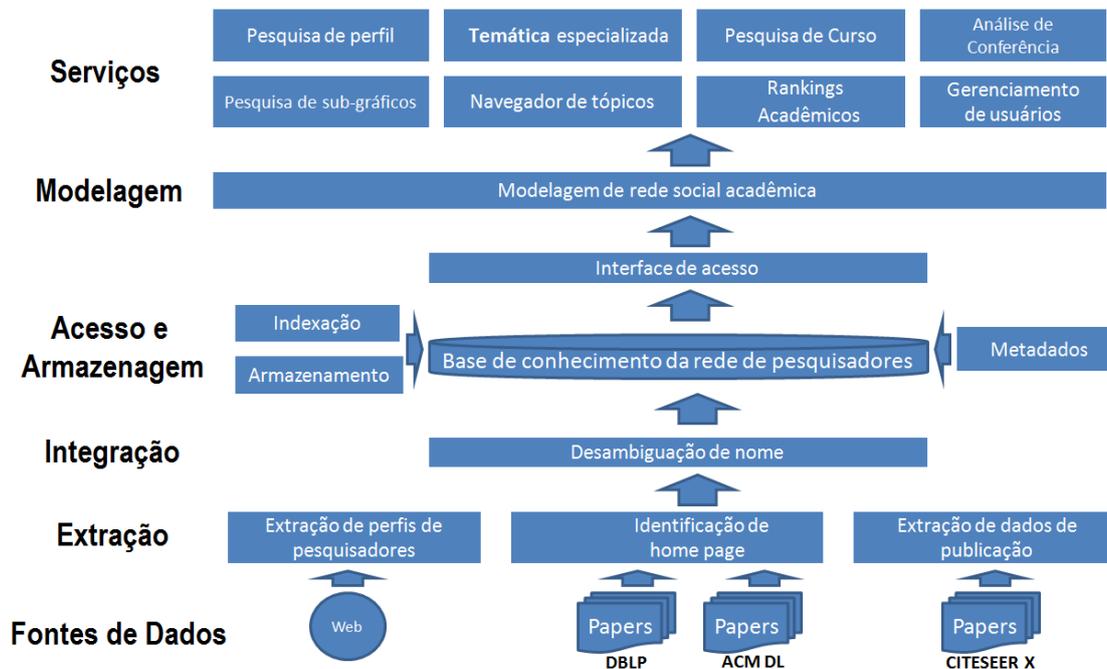


Figura 1: arquitetura geral do sistema AMiner

Conforme (Bobadilla et. al., 2019), o sistema consiste principalmente de cinco componentes:

1) Serviços: Fornece vários serviços com base nos resultados de modelagem: pesquisa de perfil, descoberta de especialistas, análise de conferência, pesquisa de curso, pesquisa de subgráfico, navegador de tópicos, classificações acadêmicas e gerenciamento de usuários.

2) Modelagem: Utiliza um modelo probabilístico generativo para modelar simultaneamente os diferentes tipos de fontes de informação. O sistema estima uma mistura de distribuição de tópicos associada às diferentes fontes de informação.

3) Acesso e Armazenagem: Fornece armazenamento e indexação para os dados extraídos e integrados na base de conhecimento da rede de pesquisadores. Especificamente, para armazenamento, emprega Jena, uma ferramenta para armazenar e recuperar dados ontológicos; para indexação, emprega o método de indexação de arquivos invertidos, um método estabelecido para facilitar a recuperação da informação.

4) Integração: Junta e integra os perfis dos pesquisadores extraídos e as publicações extraídas. O aplicativo emprega o nome do pesquisador como o identificador. Um modelo probabilístico e uma estrutura abrangente foram desenvolvidos para lidar com o problema da ambigüidade do nome na integração. Os dados integrados são então armazenados, classificados e indexados em uma base de conhecimento de rede de pesquisa.

5) Extração: O foco está na extração automática de perfis de pesquisadores da Web. Primeiro, o serviço coleta e identifica as páginas relevantes (por exemplo,

páginas iniciais ou páginas introdutórias) da Web e usa uma abordagem unificada para extrair dados dos documentos identificados. Também extrai publicações de bibliotecas digitais on-line usando regras heurísticas. Além disso, é adotada uma abordagem simples, mas muito eficaz, para o perfil de usuários da Web, aproveitando o poder do big data.

Para vários recursos do sistema, por exemplo, extração de perfis, desambiguação de nomes, modelagem de tópicos acadêmicos, pesquisa de perícia e mineração de redes sociais acadêmicas, propomos algumas novas abordagens para superar as desvantagens que existem nos métodos convencionais (Nasution, Noah e Saad, 2016).

O restante deste trabalho está organizado da seguinte forma. A seção 2 discute os trabalhos relacionados e a seção 3 apresenta nossas abordagens propostas no sistema. A seção 4 mostra algumas aplicações do AMiner. A seção 5 lista os conjuntos de dados que construímos. Finalmente, a Seção 6 faz uma conclusão.

2 | TRABALHOS RELACIONADOS

Anteriormente várias questões nas redes sociais acadêmicas foram investigadas e alguns sistemas foram desenvolvidos conforme veremos a seguir:

Google Scholar: fornece um mecanismo de pesquisa para identificar os hiperlinks de publicações que estão disponíveis publicamente ou podem ser obtidas por meio de bibliotecas institucionais. O Google Acadêmico não é um site de rede social no sentido geral, mas ainda assim se tornou uma plataforma importante para pesquisar recursos acadêmicos, acompanhar as pesquisas mais recentes, promover as próprias conquistas e acompanhar o impacto acadêmico (Ovadia, 2014).

Microsoft Academic: emprega tecnologias de aprendizado de máquina, análise semântica e mineração de dados para ajudar os usuários a explorar informações acadêmicas mais poderosamente (Wolfram, 2016).

Semantic Scholar: é projetado para ser um mecanismo de busca “inteligente” para ajudar os pesquisadores a encontrar melhores publicações acadêmicas mais rapidamente. Em comparação com o Google Acadêmico e a Microsoft Academic, o Semantic Scholar pode destacar rapidamente os artigos mais importantes e identificar as conexões entre eles (Wu, 2011).

ResearchGate: tem o objetivo de conectar pesquisadores geograficamente distantes e permitir que eles se comuniquem continuamente. Os usuários registrados do site têm um perfil de usuário e podem compartilhar sua produção de pesquisa, incluindo artigos, dados, capítulos de livros, patentes, propostas de pesquisa, algoritmos, apresentações e código-fonte de software. Os usuários também podem acompanhar as atividades de outras pessoas e participar de discussões com eles (Parack, Zahid e Merchant, 2012).

Academia.edu: é um site de rede social acadêmica com fins lucrativos. Ele permite que seus usuários criem um perfil, compartilhem seus trabalhos, monitorem seu impacto acadêmico, selecionem áreas de interesse e sigam a pesquisa que evolui em campos específicos (Safder, Hassan e Aljohani, 2018).

Embora a maioria dos sistemas acima tenha integrado uma quantidade gigantesca de recursos acadêmicos e fornecido meios abundantes de pesquisa e consulta de funções de rede social, eles não realizaram análise sistemática de nível semântico ou mineração (Shah e Pudi, 2019).

Conseqüentemente, nosso objetivo principal é fornecer uma abordagem de modelagem unificada para obter uma compreensão maior e mais profunda da conexão semântica em redes acadêmicas grandes e heterogêneas, compostas por autores, artigos, conferências, periódicos e organizações. Como resultado, o sistema pode fornecer pesquisa especializada e pesquisa centrada no pesquisador.

3 | METODOLOGIA

Nesta seção, apresentamos em detalhes os desafios da mineração de dados de redes sociais acadêmicas (Bobadilla et. al., 2019) através do sistema AMiner e apresentamos os métodos e soluções.

3.1 Extração de Perfil

Definimos o esquema do perfil do pesquisador, estendendo a ontologia da FOAF, como mostra a Figura 2. No esquema, 24 propriedades e duas relações são definidas.

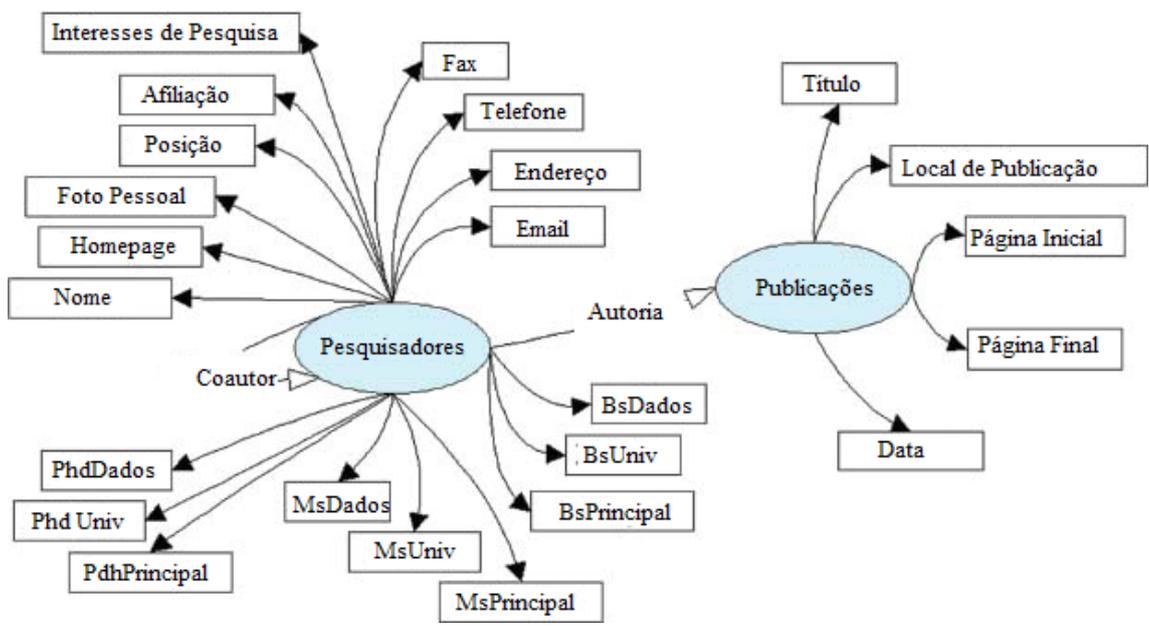


Figura 2: 24 propriedades e 2 relações, adaptado de Obadia (2014)

Certamente não é uma tarefa trivial extrair a rede de pesquisa da Web. Os pesquisadores de diferentes universidades, institutos ou empresas têm diferentes

modelos de página e perfil e feeds de dados. Portanto, um método de extração ideal deve considerar o processamento de todos os tipos de modelos e formatos (Ovadia, 2014). Cabanac, Frommholz e Mayr (2019) propõem a abordagem em três etapas:

1) Identificação de página relevante: Dado o nome de um pesquisador, primeiro obtemos uma lista de páginas da Web por um mecanismo de pesquisa (a API do Google é usada) e depois identificamos a página inicial ou a página de introdução usando um classificador. Definimos um conjunto de recursos, como se o título da página contém o nome da pessoa e se o endereço da URL (em parte) contém o nome da pessoa e se emprega SVM para a classificação.

2) Pré-processamento: Separa-se o texto em tokens e atribuímos tags possíveis a cada token. Os tokens formam as unidades básicas e as páginas formam as seqüências de unidades na etapa de marcação a seguir.

3) Marcação: Dada uma seqüência de unidades, determinamos a seqüência de tags correspondente mais provável usando um modelo de marcação treinado. O tipo de tag corresponde à propriedade definida na Figura 2. Definimos cinco tipos de tokens (palavra padrão, palavra especial, token de imagem, termo e sinal de pontuação) e usamos heurística para identificar tokens na Web. Depois disso, atribuímos várias tags possíveis a cada token com base no tipo de token e, em seguida, um modelo de CRF treinado é usado para encontrar a melhor atribuição de tag com a maior probabilidade.

Recentemente, revisitamos o problema do perfil do usuário da Web no Big Data e propomos uma abordagem simples, mas muito eficaz, chamada MagicFG, para criar perfis de usuários da Web, aproveitando o poder do Big Data (Safder, Hassan e Aljohani, 2018). Para evitar a propagação de erros, a abordagem integra a identificação de página e a extração de perfil em uma estrutura unificada (Alyami e Assiri, 2018). A Figura 3 dá um exemplo de perfil de pesquisador:

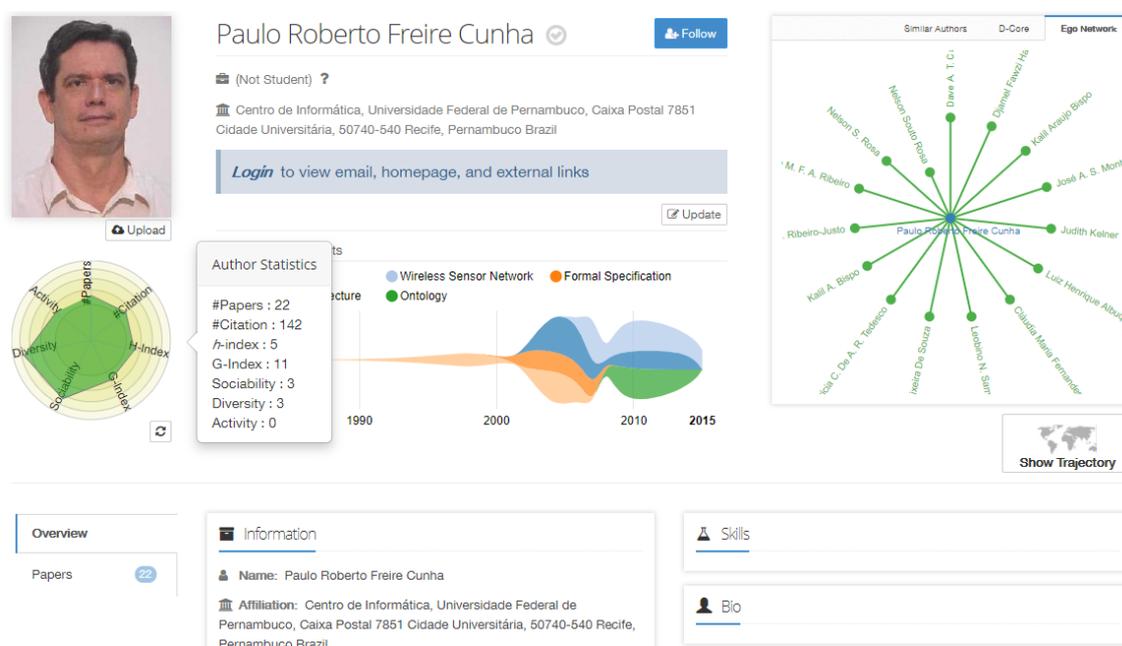


Figura 3: exemplo de perfil de pesquisador

Para melhorar o desempenho do perfil, apresentamos o conceito de credibilidade contextual e também apoia a incorporação do conhecimento humano (Wu, 2011). Nasution, Noah e Saad (2016), define o conhecimento humano como declarações lógicas de Markov e as formaliza em um modelo de gráficos de fatores. O método MagicFG foi implantado no sistema AMiner para criar perfis de milhões de pesquisadores (Bobadilla et. al., 2019).

3.2 Desambiguação do Nome

Coletamos mais de 200 milhões de publicações de bibliotecas de dados on-line existentes, incluindo EBSCO, JSTOR, DOAJ e outras. Em cada fonte de dados, os autores são identificados por seus nomes. Para integrar os perfis do pesquisador e os dados de publicação, usamos o nome do pesquisador e o nome do autor como o identificador. Conforme Safder, Hassan e Aljohani (2018) este processo tem inevitavelmente o problema ambíguo.

O problema de desambiguação é lançado ao atribuir uma tag a cada papel, com cada tag representando um pesquisador real (Ovadia, 2014). Mais recentemente, propôs-se uma estrutura adicional abrangente para abordar o problema da desambiguação de nomes. (Alyami e Assiri, 2018).

O método foi agora implantado na AMiner para lidar com o problema da desambiguação de nomes na escala de bilhões, o que demonstra sua eficácia e eficiência (Shah e Pudi, 2019).

3.3 Modelagem de Tópicos

Na pesquisa acadêmica, a representação do conteúdo de documentos de texto, interesses de autores e temas de conferências é uma questão crítica de qualquer abordagem (Borkar e Rajeswari, 2013). No entanto, essa representação não pode utilizar as dependências “semânticas” entre palavras.

Infelizmente, modelos de tópicos existentes, como a Indexação semântica latente (pLSI) probabilística, a Alocação de Dirichlet Latente (LDA) e o modelo Autor-Assunto não pode ser aplicado diretamente ao contexto da pesquisa acadêmica (Parack, Zahid e Merchant, 2012). Isso ocorre porque eles simplesmente não podem capturar todas as dependências intrínsecas entre documentos e conferências (Alyami e Assiri, 2018).

Uma abordagem unificada de modelagem de tópicos é proposta para modelar simultaneamente características de documentos, autores, conferências e dependências entre eles. Para simplificar, usamos conferência para designar conferência, diário e livro no modelo.

O modelo proposto é chamado modelo Autor-Conferência-Tópico (ACT). Mais especificamente, diferentes estratégias podem ser empregadas para modelar as distribuições de tópicos e, conseqüentemente, os modelos implementados podem ter

diferentes capacidades de representação do conhecimento (Cabanac, Frommholz e Mayr, 2019).

3.4 Pesquisa especializada

Ao procurar recursos acadêmicos e formular uma consulta, o usuário procura encontrar autores com conhecimentos específicos e trabalhos e conferências relacionados às áreas de interesse da pesquisa (Ovadia, 2014).

No sistema AMiner, apresentamos uma estrutura de pesquisa de especialização em nível de tópico. Diferentemente dos tradicionais mecanismos de busca da Web que realizam a recuperação e a classificação no nível do documento, estudamos o problema de pesquisa de especialização em nível de tópico em relação a redes heterogêneas distintas (Wolfram, 2016).

Um modelo de tópico unificado, chamado Rastreamento de Citação de Tópico (RCT), é proposto para modelar simultaneamente aspectos tópicos de diferentes objetos na rede acadêmica (Safder, Hassan e Aljohani, 2018).

Com base nos modelos de tópicos aprendidos, investigamos o problema de pesquisa de especialização em três dimensões: classificação, análise de rastreamento de citações e pesquisa de gráfico de tópicos. Especificamente, Shah e Pudi (2019) propõem um método de caminhada aleatória em nível de tópico para classificar diferentes objetos. Na análise de traços de citações, procuramos descobrir como um estudo influencia seu estudo de acompanhamento (Wolfram, 2016).

A Figura 4 dá um exemplo resultado de especialistas encontrados para a consulta "Social Network Analysis":

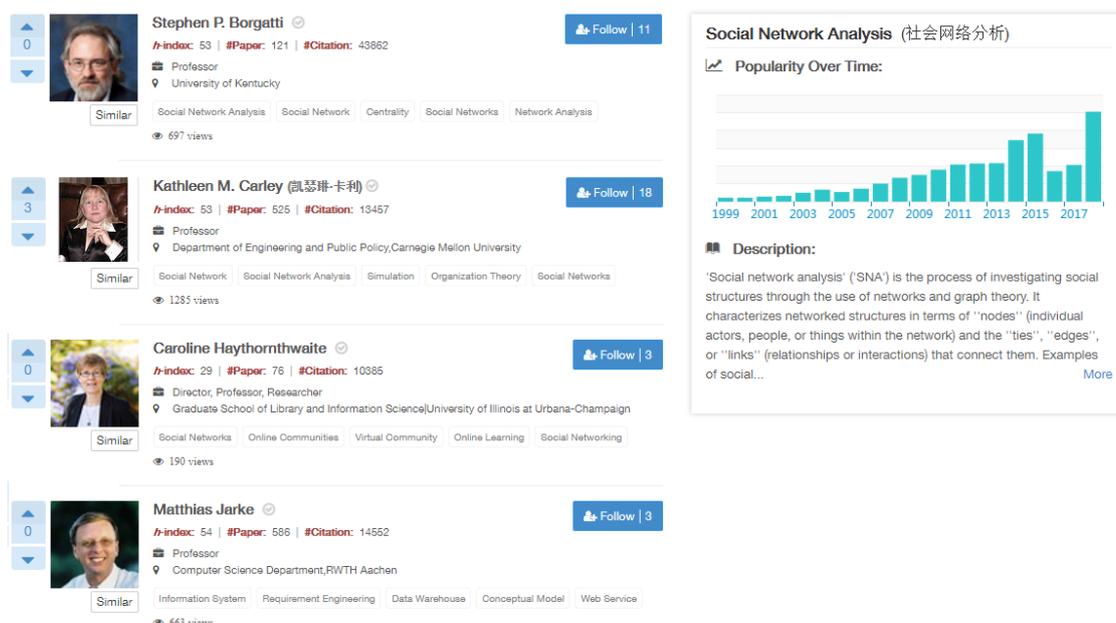


Figura 4: exemplo resultado de especialistas encontrados para a consulta "Social Network Analysis"

Na análise de traços de citações, procuramos descobrir como um estudo

influencia seu estudo de acompanhamento (Wolfram, 2016). Finalmente, há uma função de busca de gráficos tópicos, com base na modelagem de tópicos e análise de rastreamento de citações (Bobadilla et. al., 2019).

3.5 Mineração da Rede Social Acadêmica

Com base no sistema AMiner, esse conjunto de funções de mineração de redes sociais acadêmicas centradas em pesquisadores inclui análise de influência social, mineração de relacionamento, recomendação de colaboração, análise de similaridade e evolução da comunidade (Cabanac, Frommholz e Mayr, 2019).

Análise de Influência Social: Em grandes redes sociais, as pessoas são influenciadas por outros por vários motivos. Propomos um modelo de Propagação por Afinidade de Tópicos (TAP) para diferenciar e quantificar a influência social. O TAP pode obter resultados de qualquer modelagem de tópicos e da estrutura de rede existente para executar a propagação de influência no nível do tópico (Safder, Hassan e Aljohani, 2018).

Recentemente foi projetado uma estrutura de ponta a ponta que chamada de DeepInf para o aprendizado de representação de características e para prever a influência social. Cada usuário é representado por uma sub-rede local na qual ele está embutido. Uma rede neural gráfica é usada para aprender a representação da sub-rede que, por sua vez, efetivamente integra os recursos específicos do usuário e as estruturas de rede (Alyami e Assiri, 2018).

Mineração de Relacionamento Social: Inferir o tipo de relacionamento social entre dois usuários é uma tarefa muito importante na mineração de relacionamento social. Propomos um framework de dois estágios denominado Modelo Gráfico de Fatores Probabilísticos com restrição de Tempo (TPFG) para inferir relações de conselheiro-assessor na rede de coautores (Wolfram, 2016).

A ideia principal é alavancar um modelo de gráfico de fatores probabilísticos com restrição de tempo para decompor a probabilidade conjunta dos conselheiros desconhecidos sobre todos os autores (Parack, Zahid e Merchant, 2012).

Além disso, desenvolveu-se uma estrutura denominada TranFG para classificar o tipo de relações sociais entre diferentes recursos heterogêneos. A estrutura incorpora teorias sociais em um modelo de gráfico de fator, que efetivamente melhora a precisão de prever os tipos de relações sociais em uma rede de destino ao emprestar conhecimento de outra rede de origem (Alyami e Assiri, 2018).

Análise de similaridade: Estimar a similaridade entre os vértices é uma questão fundamental na análise de redes sociais (Ovadia, 2014). Propôs-se um método à base de amostragem para estimar os topo- k vértices similares (Bobadilla et. al., 2019). O método baseia-se na nova ideia do método de amostragem por caminhos aleatórios conhecido como Panther (Nasution, Noah e Saad, 2016).

Dada uma rede particular como ponto de partida, o Panther gera aleatoriamente vários caminhos de um comprimento pré-definido e, em seguida, a similaridade entre dois vértices pode ser modelada como estimativa da possibilidade de que os dois vértices apareçam nos mesmos caminhos (Cabanac, Frommholz e Mayr, 2019).

Recomendação de colaboração: Colaborações interdisciplinares geraram um enorme impacto na sociedade. No entanto, geralmente é difícil para os pesquisadores estabelecerem essas colaborações entre domínios (Wu, 2011). Analisamos os dados de colaboração entre domínios de publicações de pesquisa e propôs-se um modelo de Aprendizado de Tópicos entre Domínios (ATD) para recomendação de colaboração (Wolfram, 2016).

Para lidar com conexões esparsas, o ATD consolida as colaborações entre domínios existentes por meio de camadas de tópicos, em vez de utilizar camadas de autor. Isso alivia o problema de escassez (Alyami e Assiri, 2018).

Para lidar com conhecimentos complementares, o ATD modela distribuições de tópico de domínios de origem e de destino separadamente, bem como a correlação entre domínios. Para lidar com a assimetria de tópicos, o ATD apenas modela tópicos relevantes para a colaboração entre domínios (Bobadilla et. al., 2019).

Evolução Comunitária: Como as redes sociais são bastante dinâmicas, é interessante estudar como as pessoas nas redes formam clusters diferentes e como os vários clusters evoluem com o tempo. Estudamos a co-evolução de mineração de objetos multi-tipados em um tipo especial de rede heterogênea, chamada de rede de estrelas. Em seguida, examinamos como os objetos multi-tipificados influenciam uns aos outros na evolução da rede (Cabanac, Frommholz e Mayr, 2019).

Propõe-se uma evolução baseada no Modelo de Mistura de Processos Hierárquicos Dirichlet que detecta a co-evolução de objetos multi-tipados sob a forma de uma evolução de cluster multi-tipificada em redes estelares dinâmicas. Um algoritmo de inferência eficiente é fornecido para aprender o modelo proposto (Safder, Hassan e Aljohani, 2018).

4 | APLICAÇÃO

O AMiner foi desenvolvido para fornecer serviços abrangentes de pesquisa e mineração para redes sociais de pesquisadores. Neste sistema, concentrou-se em: (1) criar um perfil baseado em semântica para cada pesquisador, extraindo informações da Web distribuída; (2) integrar dados acadêmicos (por exemplo, os dados bibliográficos e os perfis dos pesquisadores) de múltiplas fontes; (3) pesquisando com precisão a rede heterogênea; (4) analisar e descobrir padrões interessantes da rede social pesquisadora construída. As principais funções de pesquisa e análise no AMiner são resumidas na seção seguinte.

Pesquisa de perfil: Digite um nome de pesquisador (por exemplo, Paulo

Roberto Freire). O sistema retornará o perfil baseado em semântica criado para o pesquisador usando técnicas de extração de informações. Na página de perfil, as informações extraídas e integradas incluem: informações de contato, foto, estatísticas de citação, avaliação de desempenho acadêmico, interesse de pesquisa (temporal), histórico educacional, gráfico social pessoal, financiamento de pesquisa e registros de publicações incluindo informações de citação e os documentos que são atribuídos automaticamente a vários domínios diferentes (Shah e Pudi, 2019).

Temática especializada: Insira uma consulta (por exemplo, análise de redes sociais). O sistema retornará especialistas neste tópico. Além disso, o sistema irá sugerir a melhor conferência e os principais trabalhos sobre este tópico. Existem dois algoritmos de classificação: VSM e ACT. O primeiro é semelhante ao modelo de linguagem convencional e o segundo baseia-se no nosso modelo de Autor-Conferência-Tópico (ACT). Os usuários também podem fornecer feedbacks para os resultados da pesquisa (Safder, Hassan e Aljohani, 2018).

Análise de conferência: Digite um nome de conferência (por exemplo, KDD). O sistema retorna aqueles que são os pesquisadores mais ativos nesta conferência, bem como os principais trabalhos (Nasution, Noah e Saad, 2016).

Pesquisa de curso: Insira uma consulta (por exemplo, Social network analysis). O sistema retornará aqueles que estão ministrando cursos relevantes para a consulta (Ovadia, 2014).

Pesquisa de subgráficos: Insira uma consulta (por exemplo, mineração de dados). O sistema primeiro informa quais tópicos são relevantes para a consulta (por exemplo, cinco tópicos “Data mining”, “Dados XML”, “Mineração de Dados/Processamento de Consultas”, “Design de Dados/Banco de Dados da Web” e “Web Mining” são relevantes) e, em seguida, exibir o subgráfico mais importante descoberto em cada tópico relevante, ampliado com um resumo para o subgráfico (Shah e Pudi, 2019).

Navegador de Tópicos: Com base em no modelo Autor-Conferência-Tópico (ACT), descobrimos automaticamente 200 tópicos importantes das publicações. Para cada tópico, atribuímos automaticamente um rótulo para representar seus significados. Além disso, o navegador apresenta os pesquisadores mais ativos, as conferências/trabalhos mais relevantes e a tendência de evolução dos tópicos descobertos (Safder, Hassan e Aljohani, 2018).

Graus Acadêmicos: O sistema possui oito medidas para avaliar a realização do pesquisador. As medidas incluem: h-index, citação, tendências, atividades, longevidade, diversidade e sociabilidade. Para cada medida, produzimos uma lista de classificação em diferentes domínios. Por exemplo, pode-se pesquisar aqueles que têm os números de citação mais altos no domínio de “Social network analysis” (Bobadilla et. al., 2019).

Gerenciamento de usuários: Pode-se registrar como um usuário para: (1) modificar as informações do perfil extraído; (2) fornecer feedback sobre os resultados

da pesquisa; (3) seguem pesquisadores da AMiner; e (4) criar uma página da AMiner que pode ser usada para anunciar conferências e workshops, ou recrutar estudantes (Nasution, Noah e Saad, 2016).

5 | CONJUNTO DE DADOS

A AMiner reuniu um grande conjunto de dados acadêmicos com mais de 180.000.000 de perfis de pesquisadores e 294.000.000 de publicações da Internet até junho de 2018, juntamente com vários subconjuntos que foram construídos para diferentes fins de pesquisa. Os detalhes desses subconjuntos são os seguintes e podem ser encontrados em <https://aminer.org/>

Rede de Citação: Os dados de citação são extraídos do EBSCO, JSTOR e outras fontes. O conjunto de dados contém 1.572.277 artigos e 2.084.019 relações de citação. Cada artigo é associado com resumo, autores, ano, local e título. O conjunto de dados pode ser usado para agrupamento com informações de rede e laterais, estudando a influência na rede de citações, encontrando os documentos mais influentes, a análise de modelagem de tópicos, etc.

Rede Social Acadêmica: Esses dados incluem artigos, citação em papel, informações sobre o autor e colaboração do autor. O conjunto de dados contém 1.712.433 autores, 2.092.356 artigos, 8.024.869 relações de citação e 4.258.615 relações de colaboração observadas entre os autores.

Conselheiro-assessor: O conjunto de dados é composto por 815.946 autores e 2.792.833 relações de co-autoria. Para avaliar o desempenho de inferir relações de conselheiro-assessor entre coautores, criamos dados menores sobre a verdade do terreno usando o seguinte método: (1) coletar as informações do conselheiro-assessor (2) rastrear manualmente as informações do conselheiro-orientador a partir das páginas iniciais dos pesquisadores. Finalmente, rotulamos 1.534 relações de coautores das quais 514 são relações de conselheiro-assessor.

Coautor do tópico: É uma rede de coautores baseada em tópicos, que contém 640.134 autores de 8 tópicos e 1.554.643 relações de coautoria. Os oito tópicos são: Mineração de Dados / Regras de Associação, Serviços da Web, Redes Bayesianas / Função de Crença, Mineração Web / Fusão de Informações, Web Semântica / Lógicas de Descrições, Machine Learning, Sistemas de Banco de Dados / Dados XML e Recuperação de Informações.

Autor de artigo de tópico: O conjunto de dados é coletado para fins de recomendação de domínio cruzado, que contém 33.739 autores associados a 5 tópicos, além de 139.278 relações de coautoria. Os cinco tópicos são Mineração de Dados (com 6.282 autores e relacionamentos de 22862 co-autores), Gestão Empresarial (com 9.150 autores e 31851 relações de coautores), Business Intelligence (com 5.449 autores e 27.712 coautores), Visualização (com 5.268 autores) e

19.261 relacionamentos de coautores) e Big Data (com 7.590 autores e 37.592 relacionamentos de coautores).

Citação-tema: É uma rede de citações baseada em tópicos que contém 2.329.760 artigos de 10 tópicos e 12.710.347 relações de citações. Os 10 tópicos são: Mineração de Dados / Regras de Associação, Web Services, Redes Bayesianas / Função de Crença, Mineração Web / Fusão de Informações, Web Semântica / Lógicas de Descrição, Aprendizado de Máquina, Sistemas de Banco de Dados / Dados XML, Reconhecimento de Padrões / Análise de Imagens, Recuperação de Informações e Sistema de Linguagem Natural / Tradução Estatística de Máquinas.

Comunidade do Kernel: É uma rede de coautoria com 822.415 nós e 2.928.360 extremidades não direcionadas. Cada vértice representa um autor e cada borda representa um relacionamento de coautor.

Coautor dinâmico: O conjunto de dados contém 1.768.776 artigos publicados durante o período de 1986 a 2012 com 1.629.217 autores envolvidos. Cada ano é considerado como um registro de data e hora e há 27 registros de data e hora no total. Em cada registro de data e hora, criamos uma borda entre dois autores se eles tiverem coautoria de pelo menos um artigo nos três anos mais recentes (incluindo o ano atual). Nós convertemos a rede de co-autor não direcionada em uma rede direcionada considerando cada borda não direcionada como duas bordas direcionadas simétricas.

Temática especializada: Esse conjunto de dados é uma referência para a descoberta de especialistas, que contém 1.781 especialistas de 13 tópicos.

Pesquisa de associação: Esse conjunto de dados é usado para avaliar a eficácia de abordagens de pesquisa de associação que contém 8.369 pares de autores específicos para nove tópicos. Cada par de autores contém um autor de origem e um autor de destino.

Resultados do Modelo de Tópico para o Conjunto de Dados da AMiner: Há os resultados do modelo ACT no conjunto de dados da AMiner que contém os principais 1.500.000 de artigos e autores de 350 tópicos.

Coautor: Esta é uma rede de coautores no sistema AMiner que contém 1.560.640 autores e 4.258.946 relações de co-autoria.

Desambiguação: Esse conjunto de dados é usado para estudar a desambiguação de nomes em uma biblioteca digital. Ele contém 110 autores e suas afiliações, bem como seus resultados de desambiguação (verdade fundamental).

6 | CONCLUSÃO

Reconhecemos que a AMiner ainda está em fase de desenvolvimento, tanto na escala de recursos quanto na qualidade dos serviços. No entanto, no futuro, vamos explorar métodos inteligentes adicionais para extrair conhecimento profundo de redes

científicas e implantaremos uma estrutura mais conveniente e personalizada para fornecer pesquisa acadêmica e encontrar serviços.

REFERÊNCIAS

Alyami, H. Y., & Assiri, E. A. (2018). Invisible Web and Academic Research: A Partnership for Quality. *International Education Studies*, 11(4), 84-91.

Baepler, P., & Murdoch, C. J. (2010). Academic analytics and data mining in higher education. *International Journal for the Scholarship of Teaching and Learning*, 4(2), 17.

Bobadilla, J., Gutiérrez, A., Patricio, M. Á., & Bojorque, R. X. (2019). Analysis of scientific production based on trending research topics. An Artificial Intelligence case study. *Revista Española de Documentación Científica*, 42(1), 1-16.

Borkar, S., & Rajeswari, K. (2013). Predicting students academic performance using education data mining. *International Journal of Computer Science and Mobile Computing*, 2(7), 273-279.

Cabanac, G., Frommholz, I., & Mayr, P. (2019). Bibliometric-Enhanced Information Retrieval: 8th International BIR Workshop. In *European Conference on Information Retrieval* (pp. 394-399). Springer, Cham.

Nasution, M. K., & Noah, S. A. (2011). Extraction of academic social network from online database. In *2011 International Conference on Semantic Technology and Information Retrieval* (pp. 64-69). IEEE.

Nasution, M. K., Noah, S. A. M., & Saad, S. (2016). Social network extraction: Superficial method and information retrieval. *arXiv preprint arXiv:1601.02904*.

Ovadia, S. (2014). ResearchGate and Academia. edu: Academic social networks. *Behavioral & social sciences librarian*, 33(3), 165-169.

Parack, S., Zahid, Z., & Merchant, F. (2012). Application of data mining in educational databases for predicting academic trends and patterns. In *2012 IEEE International Conference on Technology Enhanced Education (ICTEE)* (pp. 1-4). IEEE.

Safder, I., Hassan, S. U., & Aljohani, N. R. (2018). AI cognition in searching for relevant knowledge from scholarly big data, using a multi-layer perceptron and recurrent convolutional neural network model. In *Companion Proceedings of the The Web Conference 2018* (pp. 251-258). International World Wide Web Conferences Steering Committee.

Shah, T., & Pudi, V. (2019). Mining Intellectual Influence Associations. In *BIR@ ECIR* (pp. 100-111).

Wolfram, D. (2016). Bibliometrics, information retrieval and natural language processing: natural synergies to support digital library research. In *Proceedings of the joint workshop on bibliometric-enhanced information retrieval and natural language processing for digital libraries (BIRNDL)* (pp. 6-13).

Wu, C. J., Chung, J. M., Lu, C. Y., Lee, H. M., & Ho, J. M. (2011). Using web-mining for academic measurement and scholar recommendation in expert finding system. In *2011 IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology* (Vol. 1, pp. 288-291). IEEE.

SOBRE O ORGANIZADOR

Marcelo Pereira da Silva - Pós-doutor em Comunicação pela Universidade Estadual Paulista “Júlio de Mesquita Filho”, na Faculdade de Arquitetura, Artes e Comunicação, desenvolvendo o projeto intitulado: “Ecologia da Comunicação Organizacional – consumidores, instituições e públicos de afinidade nas redes sociais virtuais: interatividade, decepção, convivência e conflitualidade” (2018). Doutor em Comunicação pela Universidade Metodista de São Paulo na linha de pesquisa Comunicação Institucional e Mercadológica, defendendo a tese: “A comunicação corporativa e o discurso do consumidor contemporâneo nos sites sociais de reclamação: decepção e coabitação na rede – desafios e oportunidades” (2016). Mestre em Comunicação Midiática pela Universidade Estadual Paulista “Júlio de Mesquita Filho”, na Faculdade de Arquitetura, Artes e Comunicação, apresentando a dissertação: “Sentidos de Brasil na imprensa argentina – A teia noticiosa do periódico *Clarín* (2009). Bacharel em Relações Públicas pela Universidade Estadual Paulista “Júlio de Mesquita Filho”, na Faculdade de Arquitetura, Artes e Comunicação (2003). Atualmente, é docente permanente do Mestrado Interdisciplinar "Cultura e Sociedade", do Mestrado Profissional de Comunicação e do curso de Relações Públicas da Universidade Federal do Maranhão, Campus São Luís. É diretor da Assessoria de Comunicação da Universidade Federal do Maranhão, coordenando os Núcleos de Relações Públicas e Cerimonial, Rádio e TV, Web Jornalismo e Produção Visual e Publicidade desde agosto de 2018. Coordena o Grupo de Pesquisa ECCOM – Ecologia da Comunicação Organizacional na Universidade Federal do Maranhão. E-mail: marcelosilva_rp@hotmail.com

ÍNDICE REMISSIVO

A

Aminer 36

Análise de discurso 39, 46, 148, 159

Análise quantitativa 259

Anúncio 133, 134, 142, 170, 171, 176, 177, 178, 292

B

Blockchain 191, 192, 198, 199, 200, 201

Boato 125, 127, 128, 130, 131, 132, 133, 134, 135, 136

C

Capital social 9, 98, 99, 100, 101, 104, 107, 109, 192, 200

Celebridade 1, 2, 5, 6, 8, 9, 10

Cinema 23, 189, 238, 239, 240, 262, 264, 265, 271, 272, 283, 284, 285, 286, 288, 291, 293, 294, 295, 306, 307, 308, 310, 311, 312, 313, 315

Cobertura jornalística 77, 82, 83, 86, 250, 324, 327

Comportamento do consumidor 88, 295

Consumo 6, 11, 12, 14, 54, 88, 91, 96, 110, 113, 114, 117, 119, 144, 193, 216, 287, 288, 328, 334, 336, 340, 343

Conteúdo 3, 8, 12, 14, 15, 19, 24, 31, 49, 53, 54, 55, 56, 59, 60, 61, 62, 63, 64, 69, 79, 80, 82, 83, 88, 90, 92, 93, 94, 95, 96, 102, 103, 104, 106, 110, 113, 114, 117, 138, 140, 141, 143, 144, 148, 155, 156, 164, 170, 171, 172, 174, 175, 180, 181, 188, 192, 197, 199, 206, 207, 209, 211, 216, 226, 227, 230, 231, 232, 236, 252, 255, 256, 257, 258, 263, 276, 280, 286, 287, 297, 298, 331, 332, 348, 350, 361

Cotas 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213

Cotidiano 18, 44, 47, 55, 81, 126, 135, 167, 174, 194, 195, 196, 211, 216, 222, 233, 282, 295, 297, 298, 304, 312, 313, 314, 335, 356, 358, 361

E

Eleições 53, 54, 57, 59, 61, 63, 64, 66, 67, 68, 69, 70, 74, 75, 76, 110, 111, 112, 114, 115, 116, 118, 119, 120, 121, 122, 123, 127, 132, 136, 138, 157, 255, 260, 355

E-sports 77, 79, 81, 83, 84, 85, 86

F

Fake News 53, 55, 56, 59, 60, 63, 64, 65, 66, 123, 134, 136, 293, 344, 345, 349, 350, 353, 354, 355

Feminismo 185, 214, 217, 218, 219, 224, 225

Fotografia 70, 73, 262, 289, 313, 325, 356, 357, 358, 359, 361, 362

I

Identidade 39, 45, 46, 49, 50, 51, 52, 129, 163, 207, 208, 211, 212, 229, 249, 263, 264, 287, 294, 356, 357, 358, 361, 362

Imaginário 219, 238, 239, 240, 241, 242, 245, 246, 248, 292

Imprensa 63, 102, 104, 108, 109, 110, 111, 113, 121, 123, 124, 127, 129, 132, 135, 136, 162, 171, 173, 175, 176, 181, 194, 195, 198, 204, 213, 216, 217, 219, 224, 226, 227, 229, 230, 234, 235, 236, 251, 254, 260, 275, 277, 278, 283, 302, 309, 318, 319, 327, 332, 335, 342, 359, 363

Influenciadores digitais 44, 46, 64

Infográfico 147, 149, 150, 153, 154, 155, 156, 157, 158, 159, 291

J

Jornal impresso 14, 18, 102, 103, 104, 170, 171, 172, 174, 175, 178, 179, 194, 320, 350

Jornalismo 11, 12, 13, 14, 15, 17, 18, 19, 20, 21, 22, 23, 55, 56, 64, 68, 77, 78, 79, 81, 82, 86, 87, 88, 91, 92, 94, 95, 96, 109, 113, 118, 121, 138, 146, 159, 160, 161, 162, 163, 168, 169, 170, 171, 174, 179, 180, 181, 182, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 214, 216, 217, 219, 220, 221, 223, 224, 225, 226, 232, 250, 251, 253, 254, 255, 257, 259, 260, 261, 262, 266, 273, 274, 275, 276, 277, 278, 279, 280, 282, 283, 297, 298, 301, 303, 304, 316, 317, 318, 319, 321, 322, 323, 324, 325, 326, 327, 328, 329, 344, 345, 346, 347, 348, 349, 350, 353, 354, 355, 363

Jornalismo automotivo 316, 317, 318, 319, 321, 322, 323, 324, 325, 326, 327, 328

Jornalismo esportivo 77, 81, 82, 86, 87

Jornalismo literário 161, 162, 163, 168, 169

M

Mídia 5, 6, 8, 10, 12, 14, 22, 42, 48, 49, 51, 64, 68, 76, 78, 79, 82, 83, 90, 92, 93, 96, 100, 102, 103, 108, 109, 111, 126, 127, 128, 132, 137, 139, 159, 171, 175, 180, 181, 182, 193, 196, 197, 204, 216, 217, 224, 226, 227, 229, 230, 233, 234, 235, 236, 249, 251, 254, 255, 261, 273, 274, 277, 279, 282, 283, 295, 309, 316, 318, 319, 320, 323, 324, 327, 328, 330, 331, 332, 333, 334, 335, 336, 337, 338, 339, 340, 341, 342, 343, 352, 354, 355, 361

Mídias digitais 14, 79, 88, 98, 100, 102, 188, 297, 298, 302

Multiculturalismo 238, 239, 240, 241, 248, 249

N

Notícias 2, 4, 12, 13, 14, 17, 40, 53, 54, 55, 56, 57, 59, 60, 62, 63, 64, 65, 66, 67, 68, 71, 72, 76, 79, 83, 85, 88, 94, 95, 99, 102, 104, 105, 106, 108, 121, 123, 127, 130, 131, 141, 144, 164, 174, 175, 180, 191, 192, 194, 197, 198, 201, 218, 230, 231, 233, 250, 251, 252, 254, 255, 256, 257, 259, 260, 261, 274, 276, 277, 279, 283, 294, 303, 304, 316, 318, 320, 323, 325, 331, 335, 337, 338, 345, 346, 349, 350, 351, 353, 354

P

Presídio 184, 187, 188, 299

R

Redes sociais 1, 3, 5, 6, 8, 9, 11, 12, 13, 14, 17, 18, 19, 20, 21, 22, 25, 26, 28, 29, 33, 34, 35, 42, 43, 44, 46, 52, 53, 54, 55, 56, 57, 58, 63, 64, 66, 67, 69, 70, 71, 73, 75, 82, 83, 93, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 111, 112, 113, 117, 119, 121, 122, 123, 128, 131, 138, 139, 141, 145, 146, 165, 172, 200, 235, 259, 260, 285, 286, 289, 291, 331, 344, 345, 350, 357, 358, 361, 363

Reportagem 69, 102, 150, 161, 162, 164, 168, 169, 214, 215, 218, 219, 220, 221, 222, 223, 225, 231, 235, 236, 280, 281, 297, 301, 302, 304, 319, 324, 326, 338

Representatividade 116, 182, 184, 188, 189, 203, 211, 259, 260, 284, 285, 286, 287, 288, 289, 290, 291, 292, 293

S

Saúde mental 330, 332, 333, 334, 335, 336, 337

Segunda tela 88, 89, 93, 94, 95, 96

Sensacionalismo 227, 232, 235, 252, 273, 274, 275, 276, 277, 278, 279, 280, 282, 283, 342

Subjetividade 52, 135, 147, 155, 162, 221, 332, 333, 342, 343

V

Valor-notícia 197, 317, 318, 319, 320, 321, 322, 323, 324, 325, 326

Agência Brasileira do ISBN

ISBN 978-85-7247-695-9



9 788572 476959