

# SABEDORIA DAS MULTIDÕES: DA URDIDURA DAS REDES DIGITAIS A TRAMA DA INTELIGÊNCIA SOCIAL

Viviani Rios Kwecko<sup>1</sup>

## 1. Introdução

Este exercício de escrita tangencia a sociedade em suas políticas de mobilização, objetivando demonstrar como a criação de um modelo tecnológico de escuta e interação é capaz de qualificar a relação da população com sua cidade via produção de conhecimento em redes sociais digitais. Vivemos em um mundo interconectado, incorporado a uma rede global sustentada por computadores que funcionam como meios de acesso a milhares de informações. Diariamente usuários do *Twitter*<sup>2</sup>, *Instagram*<sup>3</sup> e *Facebook*<sup>4</sup> geram um enorme volume de dados - sobre o que estão fazendo e sentindo. Os conteúdos não são apenas acessados, mas também produzidos, principalmente em meio às mídias sociais na *Web*. Usuários descrevem seus sentimentos e percepções sobre uma variedade de produtos, instituições, serviços e manifestam-se sobre tópicos específicos de suas vidas. Diante da vasta quantidade de informações que são produzidas e trafegam diariamente na web, a pesquisa nas mídias sociais tem se intensificado nos últimos anos (WOODALL et al., 2017). Pessoas passam a ser produtoras de conteúdo já que os dados disponíveis na internet foram primeiro produzidos por seres humanos digitando, pressionando um botão de gravação, registrando uma imagem digital ou teclando um código de barras. De acordo com Laine e Frühwirth (2010) as empresas responsáveis pelos sites de redes sociais foram as primeiras a desenvolver métodos e técnicas para monitorar as redes visando agrupar perfis de usuários por interesse e afinidades com objetivo de comercialização dessa informação para anunciantes. Ramalho (2013) e Silva (2012) destacam que o ato de monitorar consiste em aplicar diversas metodologias para acompanhar e compreender o que falam de um determinado tema, transformando os dados obtidos em conhecimento - já que o conteúdo resgatado pode ser processado e classificado segundo interesses interpretativos. Laine e Frühwirth (2010) caracterizam um monitoramento de redes

---

<sup>1</sup> Doutora em Educação em Ciências na linha pesquisa em Tecnologias Educativas no Processo de Aprendizagem pela Universidade Federal do Rio Grande (FURG). Pesquisadora-docente de plataformas interdisciplinares entre a tecnologia, processos de criação e Artes Visuais no Instituto de Letras e Artes - ILA/ FURG. Membro do Centro de Robótica e Ciências de Dados iTEC/FURG, coordenadora dos projetos de pesquisa PIER 21 - Laboratório de empreendedorismo criativo e inovação tecnológica; e Percursos culturais, aprendizagens ubíquas e inovação tecnológica/ FURG.

sociais por seu fluxo de processamento: coleta, armazenamento, classificação, categorização, adição de informações e análise de menções públicas online a determinados termos previamente definidos por seus emissores. Identificamos na potencialidade das tecnologias digitais um sistema condutor para esse acontecer sugerindo que, combinadas, a possibilidade em perceber as representações emergentes no discurso coletivo e as manifestações computacionais e ubíquas poderiam estruturar representações sociais reflexivas a partir de uma mediação sistemática das múltiplas inteligências do indivíduo e do sistema. Deste modo nossa pesquisa desenvolveu um modelo para compreender as representações sociais reveladas a partir de um processo de escuta da interação social no *Facebook*, do qual os dados são extraídos, armazenados, transformados e analisados com o objetivo de gerar informações capazes de mobilizar a rede a novos processos reflexivos. Propomos uma "Batea" (sistema) de extração de inteligência social a partir de três etapas: (i) na primeira são utilizadas diferentes técnicas de mineração de dados e Deep Learning para escuta do cidadão através de um modelo de classificação, análise do sentimento e extração dos termos representativos de um conjunto de postagens extraídas da rede social Facebook; e (ii) na segunda etapa, prodcedece a organização dos contextos narrados com a exploração e associação dos conteúdos através de uma Análise Estatística Implicativa (KWECKO et al., 2019)].

Neste artigo detalharemos a primeira etapa na qual concebemos um processo que inicia-se a partir do todo escutado (postagens) com base em classificações dos discursos por eixos de inteligência que também identificam a polaridade de sentimentos expressos na opinião sobre a vida nas cidades. Nas cidades contemporâneas os interesses e as demandas dos cidadãos emergem de diferentes canais disponibilizados para emissão de sua opinião como *e-mails*, ouvidorias, portais de reclamação e em grupos específicos de mídias sociais. Entretanto, faz-se importante diferenciarmos uma coleção de ideias e valores de uma representação social que congrega um conjunto organizado e estruturado de informações, crenças, opiniões e atitudes. Lévy (2007) define esse formato de inteligência como as habilidades distribuídas entre os indivíduos que ao serem mobilizadas efetivam competências a serem usadas em prol de uma coletividade. Para a União Europeia uma *Smart City* (Cidade Inteligente) define-se pelos fluxos de interação estratégica entre o cidadão, a infraestrutura e os serviços de informação e comunicação de forma a dar resposta às necessidades sociais, educacionais e econômicas da sociedade. Para representar uma cidade inteligente um dos índices aplicados no Brasil é o Ranking Connected Smart Cities da Urban Systems que verifica o nível de

inteligência de uma cidade a partir de 11 eixos, (mobilidade, urbanismo, meio ambiente, energia, tecnologia e inovação, economia, educação, saúde, segurança, empreendedorismo e governança) e 73 indicadores.

## 2. Uma Bateia para Inteligência Social - Caminhos metodológicos

O modelo é dividido em duas fases principais: (i) Aprendizado, processo no qual realizamos a adoção de técnicas de mineração de dados e aprendizado profundo de modo a treinar o sistema a identificar o contexto discutido na cidade por meio da obtenção dos termos representativos e classificação das postagens em opiniões associadas aos 11 eixos de cidades inteligentes; a (ii) Inferência, processo no qual apresenta-se um conjunto de postagens ao sistema para realizar a escuta e interação através do reconhecimento de padrões e associações entre os dados via meta-regras de associação e de produção de sentido. O período de extração abrange desde a primeira postagem no grupo até a última mensagem datada em julho de 2017 e o armazenamento das informações ocorreu em um banco de dados sem a identificação individual dos usuários, mantendo apenas a identificação dos sites de origem das postagens.

A **Fase de Aprendizado** está organizada em 4 estágios: criação do *dataset*; obtenção dos termos representativos; sinonimização e o tratamento de classificação dos Eixos de Inteligência e Polaridade de Sentimento. Para tal propõe-se uma ferramenta de anotação desenvolvida em *PHP3*, de modo a permitir que diferentes usuários classifiquem um conjunto de postagens extraídas de grupos do *Facebook* em um dos 11 eixos de cidades inteligentes e nas polaridades de sentimento negativo, positivo e neutro (KWECKO, 2019). Neste sistema as postagens são apresentadas de maneira aleatória, limitadas em 5 procedimentos de anotação. Para obtenção dos termos representativos analisa-se cada postagem do *dataset* iniciando pelo processo de *tokenização*, no qual são descartadas as *stop words* e os termos simples para posterior cálculo da frequência de expressões. Utilizamos o processo de *Part-Of-Speech Tagger*, da ferramenta *Stanford CoreNLP*, para gerar as palavras candidatas que representam os substantivos mais frequentes encontrados nos documentos. Além desse conjunto de palavras, o sistema também compara o conjunto de termos a uma lista de palavras sementes anotadas para cada aspecto com base na literatura e no conhecimento de domínio por especialistas. A partir de todo esse processamento obtemos uma relação de Termos Representativos por aspecto da inteligência da cidade. Após a organização dos conjuntos de termos representativos passamos ao processo de sinonimização. Os termos são reagrupados em

aspectos semânticos imperfeitos de representatividade, ou seja, quando o significado das palavras é apenas semelhante e não idêntico. Para essa fase contamos com consultas à base de sinônimos e o domínio por especialistas na área. O novo subconjunto de termos representativos sinonimizados representam as variáveis de entrada para a fase de Inferência para a etapa de Compreensão do Contexto. Para identificação dos eixos de inteligência e polaridade de sentimento de cada uma das postagens propõe-se o uso de um classificador. Esse classificador é baseado na utilização de técnicas de Aprendizado Profundo (*DL*) e, para tal, é treinado utilizando-se o *dataset* de treinamento contendo as anotações referentes ao eixo de inteligência da cidade e o sentimento relacionado a determinada postagem. A partir desse processamento o classificador aprende a catalogar cada postagem em cada um dos 11 eixos de inteligência da cidade e em uma das 3 polaridades de sentimento.

### **3. Aprendendo a Escutar: Resultados e discussão**

O modelo foi aplicado e testado em um estudo de caso associado ao monitoramento da rede social *Facebook*. Extraímos 267.902 mil comentários associados e 890.524 mil *likes* de 40 páginas e grupos públicos relacionados às informações da cidade do Rio Grande/RS, Brasil, obtendo um total de 42.062 mil mil publicações. O conjunto de dados para treinamento do sistema foi anotado com o auxílio de um sistema *Web* um total de 42.062 mil publicações extraídas, o *dataset* atingiu 1.787 publicações anotadas. De posse do conjunto de postagens anotadas para cada eixo do *Ranking Connected Smart City* inicia-se a etapa de Mineração de Dados, que utiliza como entrada um conjunto de palavras sementes e os documentos da base para o treinamento, bem como obtém-se um conjunto de palavras candidatas. A título de exemplo, obtivemos 514 termos representativos para o eixo economia, bem como os 729 para o eixo educação. No processo de Sinonimização apresenta-se neste recorte do estudo de caso a avaliação do modelo apenas para o eixo Educação. Para o eixo Educação os 729 termos representativos foram sinonimizados em 32 conjuntos de sinônimos imperfeitos. Neste estudo de caso, o processo de sinonimização foi realizado manualmente, entretanto o modelo admite implementações automáticas.

### **4. Referências bibliográficas**

LAINE, Mikko OJ; FRÜHWIRTH, Christian. Monitoramento de mídias sociais: ferramentas, características e implicações. In: **Conferência Internacional de Negócios de Software** . Springer, Berlim, Heidelberg, 2010. p. 193-198.

LÉVY, Pierre. A inteligência possível do século XXI. **Revista FAMECOS**, v. 14, n. 33, p. 13-20, 2007.

MANVILLE, Catriona et al. Mapeamento de cidades inteligentes na UE. 2014.

RAMALHO, José Antonio. **Mídias sociais na prática**. Elsevier Brasil, 2013.

SILVA, Tarcízio et al. Para entender o monitoramento de mídias sociais. **E-book: Creative Commons**, 2012.

KWECKO, Viviani Rios et al. Inteligência Social de uma cidade: escutar e aprender os contextos narrados pelo cidadão em meios as redes sociais. **RENOTE**, v. 17, n. 1, p. 183-192, 2019.

WOODALL, William H. et al. An overview and perspective on social network monitoring. **IISE Transactions**, v. 49, n. 3, p. 354-365, 2017.