

Do Extrator de Conhecimento Coletivo à Ágora Virtual: desenvolvendo uma ferramenta para democracia participativa

**Tiago Novaes Angelo¹, Cesar José Bonjuani Pagan¹, Romis Ribeiro Faissol Attux¹,
Ricardo Ribeiro Gudwin¹**

¹ Faculdade de Engenharia Elétrica e Computação – Universidade Estadual de Campinas
(UNICAMP) – Campinas – SP – Brazil

{attux,gudwin}@dca.fee.unicamp.br, pagan@dmsi.fee.unicamp.br,
tiagonovaesangelo@gmail.com

***Abstract.** The emergence of ICTs resulted in deep changes in access to information and knowledge, bringing a new perspective to the strengthening of democracy in contemporary societies. In this context, technology appears as promise to rescue a more direct citizen participation in public business. The aim of this paper is to contextualize this new moment in the history of democracy and to present a tool based in natural language processing, whose purpose is to provide foundations for the development of a virtual platform for participatory democracy.*

***Resumo.** O surgimento das TICs acarretou profundas mudanças no acesso à informação e conhecimento, trazendo uma nova perspectiva para o fortalecimento da democracia nas sociedades contemporâneas. Neste contexto, a tecnologia surge como promessa de resgatar uma participação cidadã mais direta nos assuntos públicos. O objetivo deste artigo é contextualizar este novo momento na história da democracia e apresentar uma ferramenta fundamentada no processamento de linguagem natural, cujo propósito é dar bases para o desenvolvimento de uma plataforma virtual de democracia participativa.*

1. Introdução

O desenvolvimento científico-tecnológico tem acarretado rápidas e profundas alterações na sociedade e em seus modos de organização, estabelecendo novas formas de agir, pensar e comunicar [Hall 2006]. Dentre essas novas tecnologias, encontram-se as Tecnologias de Informação e Comunicação (TICs).

A intensa virtualização e o aumento da capacidade de processamento da informação trouxeram à tona novas possibilidades para o exercício ativo e participativo da cidadania, que antes estava restrito devido a limitações tecnológicas. É possível observar, por exemplo, o surgimento de redes sociais e comunidades virtuais onde são promovidos debates com intensa participação e grande fluxo de informação, estabelecendo novos espaços de conscientização e geração de opinião.

No entanto, ainda são poucas as iniciativas de uso destas tecnologias com o objetivo de estabelecer espaços públicos de participação política. Uma das dificuldades é desenvolver métodos de organização e tratamento da informação que garantam a ampla

participação de todos os atores de uma comunidade, seja em consultas públicas ou na coleta de dados para deliberações coletivas.

Neste contexto, este artigo apresenta o “Extrator de Conhecimento Coletivo” (ECC), uma ferramenta desenvolvida a partir de tecnologias de processamento de linguagem natural e mineração de dados capaz de trazer à tona o conhecimento coletivo [Angelo 2014]. Além disso, também serão apresentados os fundamentos de um projeto que pretende desenvolver uma plataforma virtual de participação popular, denominada Ágora Virtual, tendo o ECC como núcleo de processamento de informação.

O presente artigo está estruturado em 3 eixos: o primeiro apresenta um histórico da evolução do conceito de democracia e como a tecnologia emerge como promessa de resgatar seus valores básicos de participação popular. Em seguida, é apresentado o ECC como uma ferramenta de coleta de dados sociais e tratamento da informação coletiva. Por fim, é apresentada a ideia de se utilizar futuramente o ECC como núcleo de uma plataforma virtual de participação popular nos moldes de uma Ágora Virtual.

2. Democracia Participativa: nova demanda do mundo contemporâneo

Foi na Grécia no Século V, mais especificamente na cidade-estado de Atenas, que a democracia passou a fazer parte do pensamento filosófico e político, tendo sido estabelecida ali a primeira sociedade democrática conhecida [Canfora 2008].

A democracia grega pautava-se na intensa atividade do cidadão nos assuntos coletivos, muitas vezes subordinando a vida privada às questões públicas e ao bem comum. Com a queda de Atenas e a ascensão de impérios, estados fortes e regimes militares, os ideais democráticos deixaram de ser uma prática comum, mas foram amplamente difundidos pela Europa principalmente pela República Romana e pelo Império que a ela seguiu, e voltaram à tona no fim da Idade Média e início do período Iluminista a partir de novas leituras da democracia clássica [Held 2006].

Neste processo, umas das principais modificações foi a transferência da participação direta dos cidadãos para um sistema centralizado de representação política. O filósofo e economista inglês John Stuart Mill defendia que a ideia grega de polis não era sustentável numa sociedade numerosa e complexa tal como era a sociedade europeia no século XVI [Held 2006]. Surge então a democracia representativa, na qual o poder da participação política não é mais diretamente exercido pelo cidadão, mas sim por uma figura que o representa e é escolhida pelo voto popular.

A representação política durante séculos foi (e ainda é) a principal forma de atuação de governos democráticos. Porém, seu papel nem sempre se estabeleceu de forma que os reais interesses da população fossem de fato atendidos. Alguns autores [Bennett and Entman 2001][Bucy and Gregson 2001] atribuem a este modelo a causa do desinteresse e descrédito do cidadão nos negócios públicos, uma vez que a cisão entre a esfera civil e política enfraqueceu o controle e a participação cidadã abrindo espaço aos interesses privados e de pequenos grupos e, principalmente, à corrupção.

Porém, o atual desenvolvimento tecnológico traz novas perspectivas para a promoção democrática e para a superação da crise provocada pelo papel da representação, uma vez que a virtualização e a elevada capacidade de processamento da informação tornam possível a reaproximação entre as esferas civil e política com a participativa mais

ativa e direta do cidadão nos assuntos públicos. A este novo movimento social dá-se o nome de Democracia Participativa, que se caracteriza pelo uso das TICs no resgate dos valores clássicos de democracia. Um projeto que tem como objetivo desenvolver tecnologia para democracia participativa é o “Extrator de Conhecimento Coletivo” [Angelo 2014].

3. ECC: uma metodologia para coleta de dados sociais

Focado na proposta de coleta de dados sociais, o ECC busca o conhecimento coletivo decorrente da participação dos membros de uma comunidade através de um conjunto de algoritmos cujo objetivo é apresentar temas e ideias mais relevantes de um banco de dados formado por relatos em linguagem natural coletados durante uma consulta pública. Os fundamentos científicos do ECC foram criteriosamente escolhidos e focaram-se em duas áreas do conhecimento: Redes Complexas e mineração de dados [Angelo 2014].

As etapas que executam esta metodologia foram definidas segundo a abordagem *Knowledge Discovery in Databases* (KDD) [Fayyad et al. 1996]: seleção dos dados, pré-processamento, transformação e mapeamento, mineração de dados e, por fim, interpretação e avaliação do conhecimento extraído.

3.1. Arquitetura do ECC

O ECC é uma heurística que busca coletar as informações mais relevantes de um banco de dados formado por uma grande quantidade de pequenos textos. Seu objetivo é classificar as informações e selecionar as mais representativas extraíndo temas e parágrafos. Para sua implementação, foi proposta uma arquitetura formada por quatro módulos que interagem entre si e com fontes externas, conforme detalhados a seguir.

- **Módulo CRC - Construtor de Rede Complexa:** O primeiro módulo é responsável por receber o banco de dados composto por um conjunto de textos e processá-lo gerando uma rede complexa com pesos baseados na co-ocorrência de palavras. Em um primeiro momento, os relatos escritos em linguagem natural são colocados em um documento único. As palavras contidas neste documento são *tokenizadas*, rotuladas e lematizadas por um lematizador [Stemmer 2013]. O objetivo é simplificar e reduzir o texto, melhorando seu processamento sem perder informação semântica. Eliminadas as *stop-words*, o documento resultante é transformado em um grafo onde os vértices representam as palavras e as arestas indicam a co-ocorrência das mesmas no texto. O processo de construção da rede se dá a partir da leitura de cada parágrafo, sobrepondo um ao outro através do incremento do valor das arestas (peso) ou da criação de arestas conforme a ocorrência de novas combinações de palavras. Por fim, de posse da rede, métricas de centralidade são calculadas.
- **Módulo RNQ - Ranqueador:** Este módulo tem como objetivo determinar a importância de cada nó do grafo a partir do cálculo de métricas de centralidade realizado no módulo anterior, gerando uma lista (ranking) com os vértices em ordem decrescente de valor. A métrica de centralidade escolhida para representar a importância dos nós foi o *betweenness* já que leva em consideração o fato das arestas serem direcionadas e com peso [Newman 2010].
- **Módulo EC - Extrator de Caminhos:** Este módulo é composto por um algoritmo cujo objetivo é encontrar conjuntos de palavras relevantes no contexto da rede,

denominados proto-frases¹, aplicando um algoritmo que parte dos vértices mais significativos (determinado pelo valor de sua métrica de centralidade) e “caminha” probabilisticamente pelo grafo segundo o peso de suas arestas. A entrada deste módulo são as palavras (vértices) escolhidas pelo usuário conforme a pontuação na tabela Ranking. A saída é um conjunto de palavras-chave que formam uma proto-frase, a qual será passada ao próximo módulo.

- **Módulo MF - Mapeador Final:** O último módulo do ECC tem como função extrair o parágrafo no documento de entrada que contém o trecho que melhor coincida com as palavras das proto-frases, tendo assim os parágrafos que melhor representam o coletivo de ideias. Para cada proto-frase uma frase é extraída segundo a quantidade de palavras coincidentes com o documento inicial e a ordem que se apresentam. O usuário recebe como saída os parágrafos que melhor se adaptam à proto-frase correspondente, sendo este, portanto, um representante do conhecimento coletivo.

O Extrator de Conhecimento Coletivo foi implementado em linguagem Java e testado utilizando uma amostra de 150 textos compostos por relatos em linguagem natural. Os autores do estudo concluíram que os resultados obtidos mostraram-se satisfatórios e que a metodologia elaborada atingiu o objetivo de se conhecer a percepção de um coletivo a respeito do que é vivenciado e relatado por seus participantes [Angelo 2014].

A arquitetura proposta e implementada faz parte exclusivamente de um instrumento de processamento da informação. Porém, sua aplicabilidade está condicionada a um contexto mais amplo, que abarca o desenvolvimento de uma plataforma virtual onde usuários poderão interagir tanto inserindo informação quanto recebendo aquilo que está sendo processado. Uma possibilidade para o desenvolvimento desta plataforma é fundamentá-la no conceito de Ágora Virtual [Lévy 2002].

4. Ágora Virtual: uma ferramenta para democracia participativa

A ideia de uma plataforma online para democracia participativa foi inspirada no conceito de Ágora Virtual, que utiliza-se da ideia de ciberdemocracia para expressar o uso das TICs na promoção da democracia [Lévy 2002]. A Ágora Virtual é uma hipótese utópica de plataforma online de democracia direta, a qual explora as potencialidade do ciberespaço na busca de problemas, debates pluraristas, tomada de decisão coletiva e avaliação dos resultados o mais próximo possível das comunidades envolvidas. Para que isto torne-se realidade, é preciso desenvolver ferramentas de filtragem inteligente de dados, navegação em meio a informação, simulação de sistemas complexos, comunicação transversal de forma a favorecer a tomada de decisão em coletivos heterogêneos e dispersos.

Nesta perspectiva, o ECC, como uma metodologia de coleta de dados sociais, pode ser um primeiro passo para o desenvolvimento de uma plataforma de web-democracia nos moldes da Ágora Virtual. Destarte, o próximo passo é desenvolver uma plataforma online onde os participantes de uma comunidade possam se expressar livremente através de relatos escritos em linguagem natural sobre temas específicos. O modelo poderá basear-se em uma rede social onde cada membro da comunidade possuirá seu próprio perfil por onde

¹Proto-frase é uma sequência de palavras oriundas dos vértices da rede que, por sua vez, fazem parte dos textos originais. O algoritmo do módulo EC seleciona essas palavras e as coloca em sequência, formando uma *string*.

poderá acessar aos temas ou perguntas a serem respondidas. Estes relatos seriam processados em tempo real pelo ECC, o qual extrairia o conhecimento coletivo e o apresentaria a todos os usuários em forma de um relatório virtual.

O projeto de criação desta Ágora Virtual envolve a melhoria da tecnologia de processamento da informação do ECC e o desenvolvimento da plataforma e seu estabelecimento na Web. Como aplicação prática, após o amadurecimento da ideia e implementação do sistema, já está em estudo utilizá-la como ferramenta de consulta para coleta dados para o desenvolvimento do Plano Diretor de uma Universidade Pública, da qual toda a comunidade acadêmica poderá participar.

5. Conclusões

O desenvolvimento do ECC e da Ágora Virtual têm como propósito abrir caminhos para superar algumas limitações do nosso modelo de democracia representativa, possibilitando o fortalecimento da democracia participativa.

Ainda são inúmeros os desafios para implementar uma ferramenta de democracia participativa, ultrapassando os limites do desenvolvimento científico-acadêmico. Questões essenciais para o pleno exercício da cidadania como a inclusão digital, igualdade no acesso a informação, liberdade de expressão, educação política, conscientização do cidadão entre outras são condições básicas para o sucesso destes projetos, e devem ser igualmente debatidos e conquistados pela sociedade.

Referências

- Angelo, T. N. (2014). Extrator de conhecimento coletivo: uma ferramenta para democracia participativa. Mestrado, DCA, Faculdade de Engenharia Elétrica, UNICAMP.
- Bennett, W. L. and Entman, R. M. (2001). *Mediated politics: Communication in the future of democracy*. Cambridge University Press, Cambridge.
- Bucy, E. P. and Gregson, K. S. (2001). Media participation a legitimizing mechanism of mass democracy. *New media & society*, 3(3):357–380.
- Canfora, L. (2008). *Democracy in Europe: A History of an Ideology*, volume 5. John Wiley & Sons, New York.
- Fayyad, U. M., Piatetsky-Shapiro, G., Smyth, P., and Uthurusamy, R. (1996). *Advances in knowledge discovery and data mining*. the MIT Press, Cambridge.
- Hall, S. (2006). *A identidade cultural na pós-modernidade*. DPA, Rio de Janeiro.
- Held, D. (2006). *Models of democracy*. Polity, Cambridge.
- Lévy, P. (2002). *Cyberdémocratie*. Odile Jacob, Paris.
- Newman, M. (2010). *Networks: an introduction*. Oxford University Press, New York.
- Silva, S. P. d. (2005). Graus de participação democrática no uso da internet pelos governos das capitais brasileiras. *Opinião Pública*, 11:450–468.
- Stemmer (2013). Disponível em: <http://www.nilc.icmc.usp.br/nilc/tools/stemmer.html>. Acessado em Junho de 2013.