Renato Fabbri renato.fabbri@gmail.com Osvaldo Novais de Oliveira Junior

chu@ifsc.usp.br

Introdução

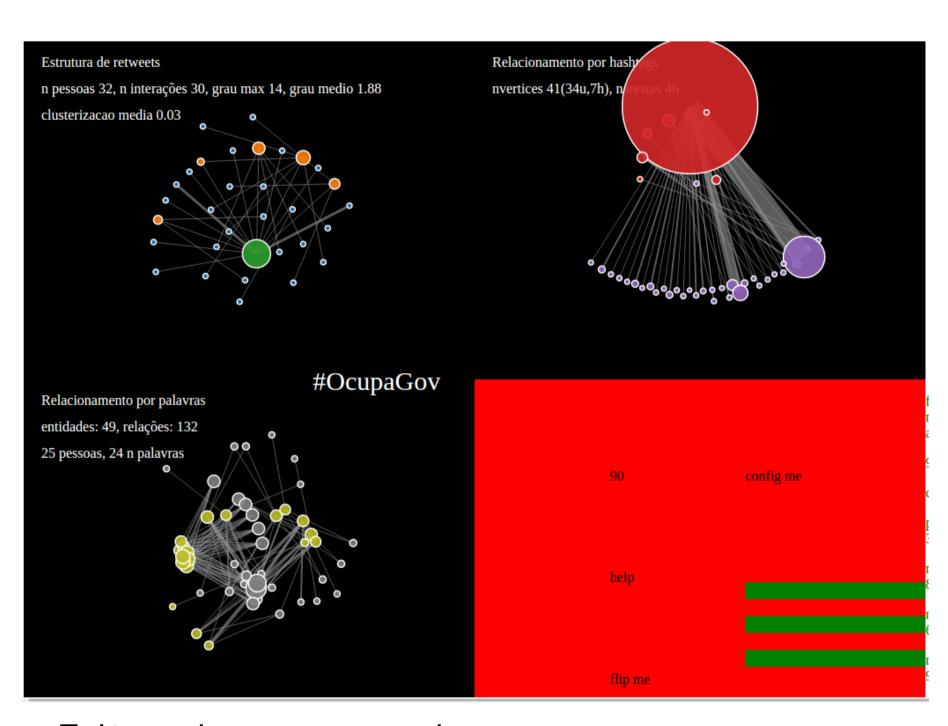
Os rastros das atividades em redes sociais permitem a detecção de comunidades e papéis desempenhados pelos indivíduos[1,2]. No entanto, não há uma abordagem focada no participante, em que as medidas e visualizações são voltadas para conscientizá-lo e capacitar seu aproveitamento dos mecanismos sociais[3]. Para isso, destacam-se as áreas de Redes Complexas, Processamento de Linguagem Natural, Dados Linkados e Participação Social. A primeira é dedicada às relações em rede, a segunda à extração de informação dos textos, a terceira aos metadados e ontologias e a última à ação da sociedade civil nos assuntos públicos. Este arcabouço sustenta a apreensão e modelagem dos processos sociais. Como resultado, há contribuições para formalismos nas dinâmicas sociais, implementações computacionais orientadas ao participante e a escrita de textos acadêmicos[1,2,3]. Este trabalho conflui com uma série de trabalhos realizados por acadêmicos (e.g. IFSC/USP, CCNH/UFABC, Lappis/UnB), por órgãos do governos (SG/PR) e pela sociedade civil (PNUD/ONU), em uma iniciativa conjunta estado-sociedade para um portal federal de participação social (Participa.br).

Materiais

Arquivos públicos de mailing lists (GMANE). Dados de grupos e perfis do Facebook. Tweets sobre hashtags e palavras em específico. Nos casos de desenvolvimentos tecnológicos, como ontologias, streamers e visualizações, os materiais são as APIs, protocolos e conceitualizações.

Métodos

Para os resultados em redes complexas, foi usada a habitual comparação com a Erdös-Renyi com o mesmo número de vértices e arestas[1]. A **estabilidade** foi observada com uma janela deslizante de tamanho fixo de mensagens[2]. A diferenciação da produção textual dos setores conectivos (de hubs, intermediários e periféricos) foi observada através da comparação dos histogramas dos usos de recursos linguísticos nas mensagens, com uma adaptação do método de Kolmogorov-Smirnoff[3]. Para as frentes computacionais, são usadas principalmente as linguagens Python e Javascript para: tarefas numéricas, visualizações, enriquecimento semântico, conexão com APIs. Foi desenvolvido um método computacional de construção de ontologia OWL orientado aos dados. Um método característico deste trabalho é o diálogo com as comunidades online interessadas ou relacionadas ao pesquisador, o que foi por vezes chamado de "física etnográfica" ou "física antropológica" por parceiros. Todo código e documentação produzida é disponibilizada abertamente, geralmente em repositórios git.



Telões de streaming de estruturas sociais para acompanhamento de processos midiáticos.

Resultados

Física:

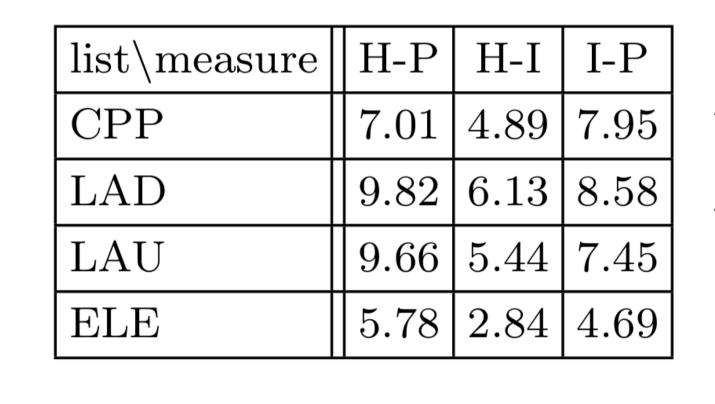
- *) Observada estabilidade nas redes humanas de interação.
- *) Observada diferenciação no uso dos recursos linguísticos quando a conectividade do participante muda.
- *) Experimentos de física etnogrática/antropológica, com hipóteses verificadas.
- *) Resultados não sistematizados, como traços homogêneos em redes livres de escala.

strength 5000 , 10000 15000 5000 10000 15000 Estabilidade na fração de agentes em cada setor conectivo

Computacional:

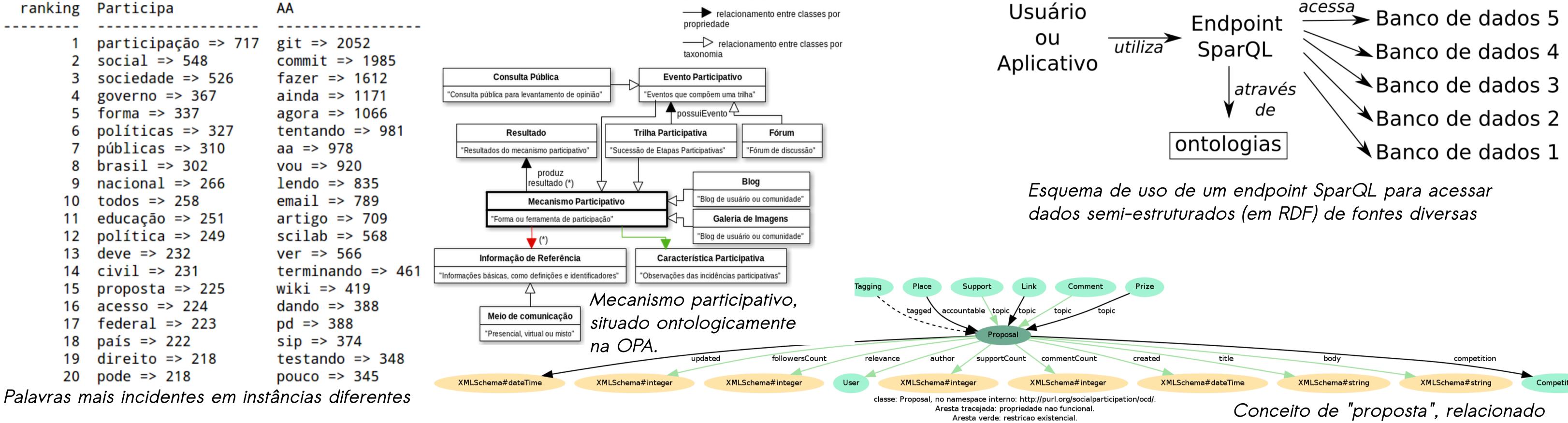
- *) Implementações de streaming de estruturas sociais e recursos em nuvem.
- *) Avanços em dados linkados abertos, com novas ontologias, roteiros de triplificação
- de bancos relacionais, endpoint SparQL e conexão de legados públicos à LOD (Linked Open Data).
- *) Novo método de construção de ontologias orientado aos dados.

- **Usos:** *) Rotinas de coleta e difusão de informação.
- *) Portal federal de participação social, Participa.br
- *) Chancela PNUD/ONU para entregar estes conhecimentos e tecnologias à população.
- *) Telões de monitoramento público em tempo real de redes sociais.
- *) Pesquisa: vários casos e conjunto de dados interessantes. Transbordam implementações, propriedades, hipóteses. Alguns artigos estão prontos, outros tantos rascunhados. As publicações e participação em eventos estão nas conclusões.



Diferenciação, em quatro listas de email, da produção textual dos setores **H**ub, Intermediário e Periférico. Calculado:

$$c(\alpha) = \frac{D_{n,n'}}{\sqrt{(\frac{n+n'}{n.n'})}}$$



Conclusões e Trabalhos Futuros

Embora esta pesquisa tenha sido acolhida por parceiros e iniciativas diversas, os aspectos práticos e multidisciplinares tornam grande o tempo de dedicação necessário. O ideal seria realizar a pesquisa com uma equipe de programadores, linguistas, comunidadores e físicos. Foram publicados artigos sobre esta pesquisa e sobre pesquisas menos centrais para o doutorado. Houve bastante participação em eventos, alguns acadêmicos, como o fechamento do Congresso Internacional de Teoria Crítica (ROMA), outros artísticos e culturais e ainda outros puxados pela Presidência da República. Os feitos acumulam o suficiente para caminhar na direção de uma qualificação. A publicação dos artigos já escritos, e dos rascunhados, e o posterior aprofundamento com parceiros e literatura, deve acompanhar as investidas finais deste doutorado. Para finalização da parte tecnológica, está sendo desenvolvido um sistema de navegação pelos dados interligados, enriquecido através de análises de instâncias e classes, recomendação de recursos a partir de um recurso de referência, explicitação dos possíveis usos dos recursos e métodos de recomendação, e entrega dos métodos como código computacional aberto.

	T					
	2013		2014		2015	
Atividade	1°	2°	1°	2°	1°	2°
créditos	•	•				
rev. de literatura	•	•	•			
dev. modelos atuais		•	•	•	•	
implementação comp.		•	•	•	•	•
escrita de tese					•	•
escrita de artigos	•	•	•	•	•	•
parcerias	•	•	•	•	•	•

Cronograma de atividades ao longo dos semestres





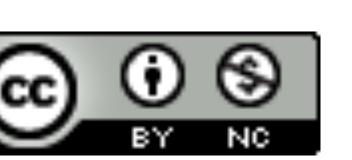
Referências

(1) Easley, David, and Jon Kleinberg. Networks, crowds, and markets: Reasoning about a highly connected world. Cambridge University Press, 2010.

(2) Allemang, Dean, and James Hendler. Semantic web for the working ontologist: effective modeling in RDFS and OWL. Elsevier, 2011.







ontologicamente na OCD