

**FACULDADE DE TECNOLOGIA ALCIDES MAYA
CURSO TÉCNICO EM INFORMÁTICA**

LORENZO POOCH LEUCK

**AS POSSIBILIDADES DO JORNALISMO DE DADOS
INTRODUÇÃO AO TEMA A PARTIR DA API DA CÂMARA DOS DEPUTADOS**

Porto Alegre

2020

LORENZO POOCH LEUCK

AS POSSIBILIDADES DO JORNALISMO DE DADOS
INTRODUÇÃO AO TEMA A PARTIR DA API DA CÂMARA DOS DEPUTADOS

Projeto de Pesquisa apresentado como
requisito parcial para obtenção do título de
Técnico em Informática da Faculdade de
Tecnologia Alcides Maya

Orientador: Prof. Vinicius Possamai

Porto Alegre

2020

LISTA DE SIGLAS

API	Aplication Programming Interface
-----	----------------------------------

SUMÁRIO

1 INTRODUÇÃO	5
1.1 Definição do Tema ou Problema	5
1.2 Delimitações do Trabalho	5
1.3.1 Objetivo Geral	6
1.3.2 Objetivos Específicos	6
1.4 Justificativa	6
2 REVISÃO BIBLIOGRÁFICA	6
2.1 O Que é Jornalismo de Dados?	6
2.2 O Jornalismo de Dados transformando a sociedade	7
2.3 A API da Câmara dos deputados como fonte de dados	9
3 DESCRIÇÃO DA SOLUÇÃO	8
4 METODOLOGIA	9
5 VALIDAÇÃO	10
6 CONCLUSÃO	11
7 CRONOGRAMA	12
8 REFERÊNCIAS BIB	13

1 INTRODUÇÃO

A práxis jornalística do século XXI teve que rever seus modelos editoriais diante da popularização da internet. Embora muitos profissionais da área insistissem na incompatibilidade entre nova e velha mídia, o fator determinante da mudança centrou-se no aspecto convergente das redações (LIMA, 2016). Esse fenômeno sociotécnico trouxe uma série de implicações à prática profissional do jornalista, que vão desde a racionalidade econômica a habilidades práticas e à uma ética filosófica (LEWIS e WUSTLAND, 2015). À medida que a sociedade foi se informatizando, jornalistas e programadores puderam somar o conhecimento prático da profissão à capacidade de contar histórias com a escala e a variedade de documentos digitais disponíveis (FELLE, 2016; DIAKOPOULOS, 2011; GRAY et al, 2012). Nesse contexto emergiu uma geração ansiosa para encontrar reportagens e futuras investigações em bases de dados semi-estruturadas e não estruturadas (BONAQUE, 2016; WU et al, 2012).

1.1 Definição do Tema ou Problema

Para Trasel (2014b), a difusão de ferramentas analíticas intuitivas e a adesão de políticas de acesso à informação contribuíram nesta etapa da cultura jornalística. Lewis (2013) destaca que nela é possível repensar o framework das notícias tradicionais a partir de valores estruturalmente diferentes, mas de certa forma complementares ao jornalismo institucional, como a ética hacker e uma cultura tecnológica fundamentalmente baseada em softwares de código aberto.

1.2 Delimitações do Trabalho

Ainda que o campo de pesquisa selecionado seja multidisciplinar, a contribuição do presente trabalho se limita à comunicação dos resultados a partir da API da câmara dos deputados. A partir de ferramentas de análise baseadas na biblioteca d3.js, será proposto um método que contemple necessidades de cobertura

nas redações. Contudo, o desenvolvimento a implementação destas não será documentado presente trabalho.

1.3.1 Objetivo Geral

Repensar o jornalismo a partir de ferramentas digitais de análise de dados.

1.3.2 Objetivos Específicos

- a) Descrever e contextualizar o Jornalismo de Dados.
- b) Apontar como o Jornalismo de Dados pode transformar as redações e a sociedade.
- c) Analisar as possibilidades do Jornalismo de Dados a partir da API da Câmara dos Deputados.

1.4 JUSTIFICATIVA

A profissionalização do jornalismo de dados requer o contínuo desenvolvimento e uso de ferramentas digitais. Träsel (2014) consta que lidar com dados nas redações requer que repórteres trabalhem lado a lado com cientistas da informação e engenheiros de software, criando databases, modelos estatísticos e visualizações gráficas.

2 REVISÃO BIBLIOGRÁFICA

2.1 O que é jornalismo de dados?

Coddington (2015) divide o surgimento do Jornalismo de Dados em três etapas. A Reportagem Assistida por Computador (RAC), introduziu o pensamento computacional a partir de métodos empíricos e análise estatística (DE SOUZA et al, 2018). Esse quadro de possibilidades, aliado à mentalidade investigativa, aumentou dramaticamente o poder dos repórteres, fazendo com que muitos não dependessem mais de outros profissionais para usar dados (MAYO e LESHNER, 2000). Garrison (1996) consta que entre as ferramentas “altamente sofisticadas” da época estavam *Word*, *Excel* e DOS. Até então, essas eram vistas apenas como facilitadores que não mudavam a essência da profissão (LIMA, 2016).

O jornalismo guiado por dados, por sua vez, foi além da ideia de reportagem investigativa, concentrando-se em técnicas de análise e visualização que ajudam na compreensão de grandes conjuntos de dados (DE SOUZA et al, 2018; RODRIGUEZ et al, 2015). Outra característica é a possibilidade de incorporar narrativas e elementos de interação nas matérias, para que o público possa descobrir fatos por si mesmo (LANGER e ZELLER, 2017; RODRIGUEZ et al, 2015). Na opinião de Lesage e Hackett (2014), as formas participativas do jornalismo de dados configuram “um meio de campo pós-moderno” à noção de busca da verdade. Borges-Rey (2016) aponta que:

Os jornalistas de dados tendem a envolver o público em projetos colaborativos de crowdsourcing compartilhando conjuntos de dados como parte de seus resultados de notícias. Além disso, como os ideais de jornalismo de dados impregnam a cultura noticiosa das redações profissionais, os correspondentes especializados procuram colaborar com as unidades de jornalismo de dados para fornecer solidez e robustez às suas histórias através do uso de evidências numéricas e métodos científicos rigorosos. (BORGES-REY, 2016)

O ingresso de desenvolvedores de *software* e programadores na criação de aplicativos de notícias com multimídia interativa e visualizações de dados (LEWIS e USHER, 2014) pode ter fortalecido esse movimento. Contudo, há de se considerar outro ponto de influência: as forças de mercado que buscam aplicar as novidades da área em modelos de negócios (SPLENDORE, 2016).

Posteriormente o Jornalismo Computacional se constituiu em torno da abstração e automação, produzindo modelos computáveis e algoritmos que classificam informações (DE SOUZA et al, 2018). A necessidade de verificar afirmações em tempo real direcionou parte do campo à checagem de fatos. Embora a automação dessa prática ainda não seja realidade, várias frentes de pesquisa estão tendo avanços consideráveis. (HASSAN, 2017). Wu et al (2014) demonstra essa possibilidade a partir da ferramenta computacional *iCheck*, que afere alegações em bancos de dados estruturados. Já Broussard (2015), destaca o uso de inteligência artificial para acelerar processos de investigação e analisar vários conjuntos de dados de uma só vez.

2.2 O jornalismo de dados transformando as redações e a sociedade

Autores como Francisco-Revilla (2012) e Goasdoué (2013) declararam que as máquinas não tomariam atividades relacionadas à produção e distribuição de conteúdo informativo. Contudo, a complexidade do meio tecnológico acabou provando o contrário. De acordo com Diakopoulos e Koliska (2017), entre as funcionalidades factíveis da área, está o emprego de softwares e algoritmos para cobrir eventos simples e adaptar notas à diferentes dispositivos e pontos de vista. Dessa maneira, a *Associated Press* produz diariamente milhares de notícias sobre esportes e finanças, além disso, simulações e previsões baseadas em dados também se tornaram comuns: o *New York Times*, assim como o *BuzzFeed* usam algoritmos para decidir quando e como o conteúdo noticioso será distribuído (DIAKOPOULOS e KOLISKA, 2017).

Formas de jornalismo tecnologicamente automatizadas prospectam um horizonte de incerteza e precarização de trabalho aos profissionais. Enquanto que, às organizações noticiosas, não só reduzem custos como geram novas fontes de receita (DE SOUZA et al 2018; THURMAN, 2017; LEWIS e WESTLUND, 2015). Muitas empresas de mídia na posição de oligopólios tiveram retornos publicitários pequenos nas últimas décadas devido à fragmentação de audiências e concorrência com as mídias sociais. Esse problema foi resolvido por meio da mineração de dados, ao possibilitar que anunciantes atendessem públicos específicos (LEWIS e WESTLAND, 2015).

A personalização algorítmica do meio digital, junto à atuação de “robôs de notícias” traz sérios problemas para a manutenção do ideal jornalístico em uma sociedade democrática informada. (BUCHER, 2017; LOKOT, 2016). Birnbaum (2015) ressalta que a falta de contexto no consumo de notícias se agravou quando a maioria das informações passou a ser publicada e filtrada a partir de hábitos de navegação. Lesage e Hackett (2014) explanam esse fenômeno ao apontar que a produção de dados no jornalismo não tem o caráter irreduzível e transparente da ciência. Apesar do “regime da objetividade”, dados são tecnologicamente, organizacionalmente e simbolicamente mediados por práticas discursivas. Diakopoulos (2015) corrobora com essa visão, considerando que vieses ideológicos podem se manifestar em modelos computacionais e sistemas de mídia. O próprio público internaliza em sua forma de entender e descrever o mundo uma mistura única de arquétipos míticos e modernos, influenciados pela família, cultura e tradição (BOON, 2017). Na prática, esse arranjo não difere muito dos algoritmos em questão. Diakopoulos (2015) clama transparência e escrutínio ao poder desses sistemas que, nas perspectiva de Lewis e Westland (2015) trazem resultados sociais e políticos adversos à “ecologia do conhecimento comum”.

De acordo com Lima (2016), as empresas jornalísticas brasileiras buscam se manter atualizadas nas tendências internacionais, mas não têm a infraestrutura adequada para o desenvolvimento de novos veículos e tecnologias - assim, jornalistas que não formam a metodologia necessária para trabalhar com Jornalismo de Dados ficam à mercê de agências, empresas de monitoramento e institutos que realizam pesquisas em troca de exposição midiática. Esse cenário traz riscos aos princípios fundamentais, como a imparcialidade, a objetividade e a credibilidade (LIMA, 2016). A autora conta que os projetos envolvendo o Jornalismo de Dados no país ignoram o potencial uso das mídias sociais, utilizando-as apenas como ferramentas de divulgação e repercussão de matérias presentes em seus portais. Não obstante, ela frisa a necessidade de capacitar profissionais para metodologias da Jornalismo de Dados da seguinte maneira:

Por meio delas, as empresas jornalísticas poderiam obter uma fonte bastante rica de pautas de interesse público, bem como uma ampla base de dados a serem trabalhados não somente em ações pontuais [...] como também em outras pautas apresentadas no cotidiano de qualquer cidade. No entanto, o que se percebe são iniciativas espaçadas, oriundas de parcerias entre plataformas de monitoramento, agências e instituições de ensino e pesquisa, que encontram, verificam e transformam informações relevantes em matérias ou artigos, negociadas previamente com as empresas jornalísticas. (LIMA, 2016, p 295)

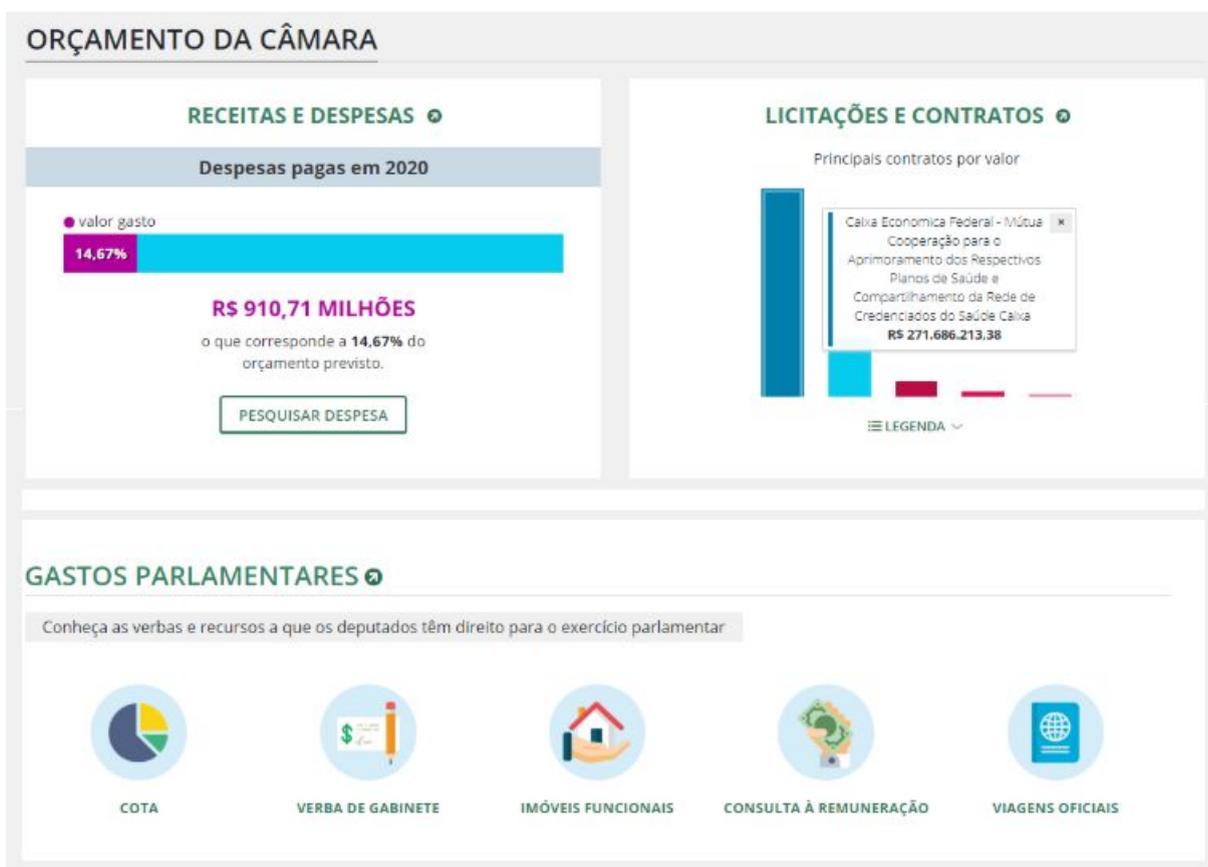
Gynnild (2014) acredita que a inovação do jornalismo não está estritamente ligada aos modelos de negócios ou às visões das empresas de mídia, mas sim ao desenvolvimento de uma mentalidade apta a entender e apresentar questões sociais a partir de *insights* e experiências interdisciplinares.

2.3 API da Câmara dos Deputados como fonte de dados

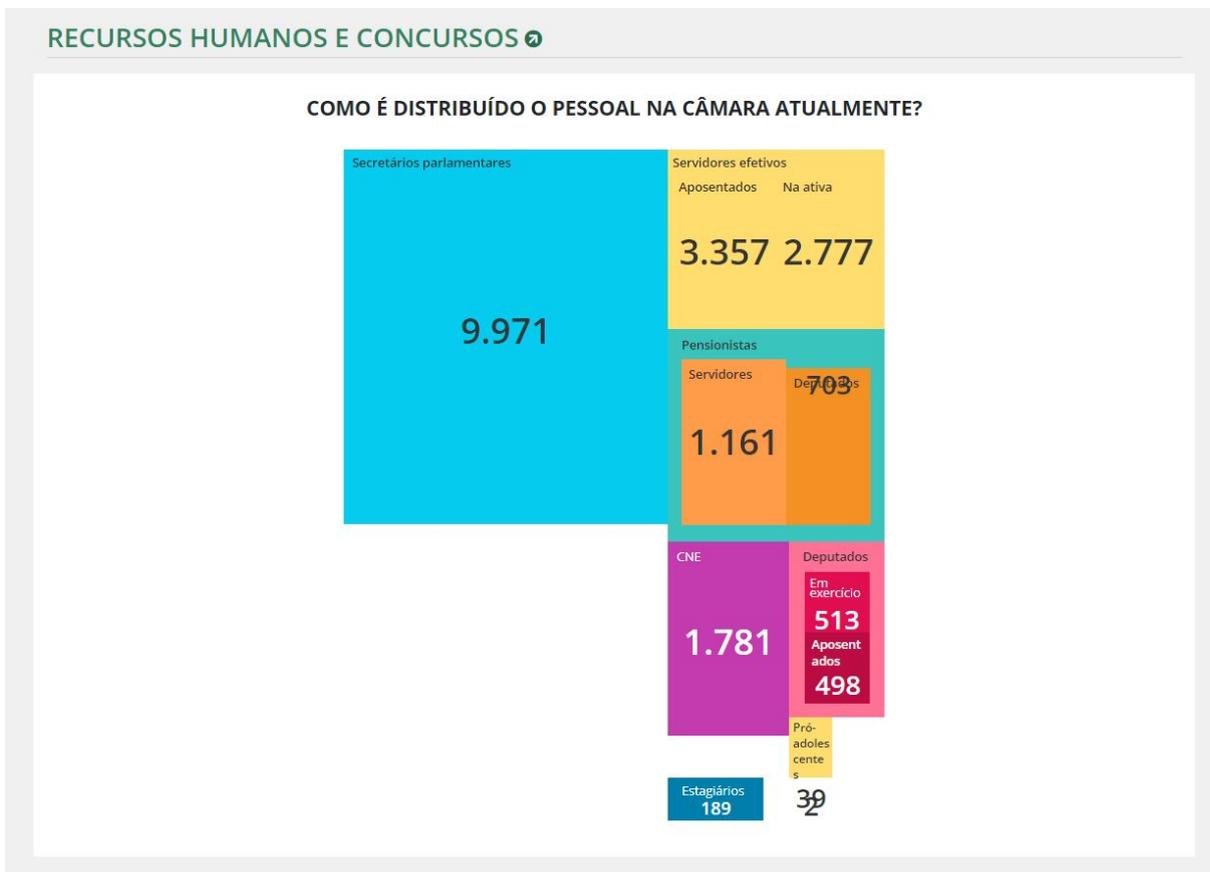
Atualmente a Câmara dos Deputados adere uma política de Dados Abertos. Isso significa, em teoria, que qualquer informação referente ao trabalho dos parlamentares está disponível ao público. No site oficial da Câmara há uma interface de fácil acesso para consulta de informações a partir de um filtro temporal.



Em destaque estão as seções “Áreas temáticas mais votadas”, “Plenário” e “Comissões”. Nos três casos há a opção de “Ver todas”, que leva uma lista dos projetos de leis respectivos.

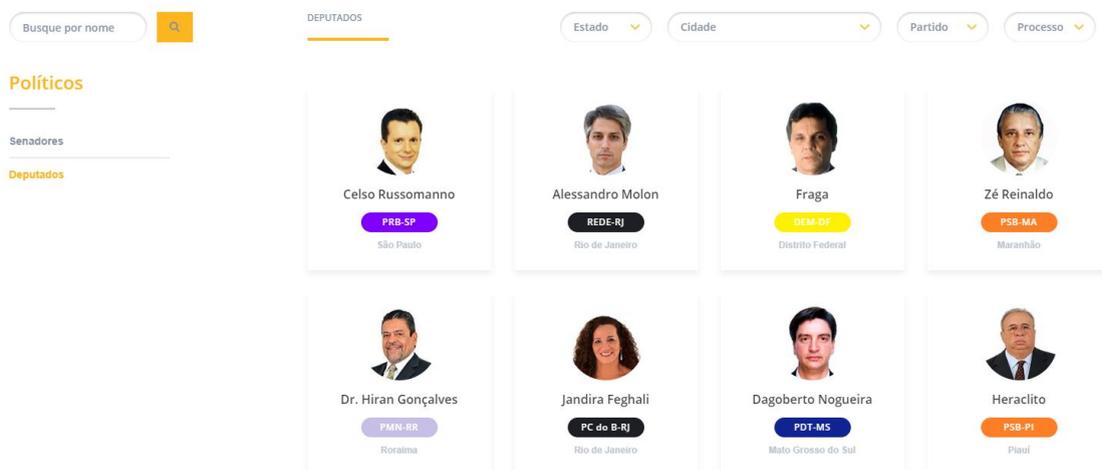


Na seção referente ao gasto de dinheiro público, há mais opções para buscas personalizadas, permitindo que o usuário filtre por deputados, estados e partidos.

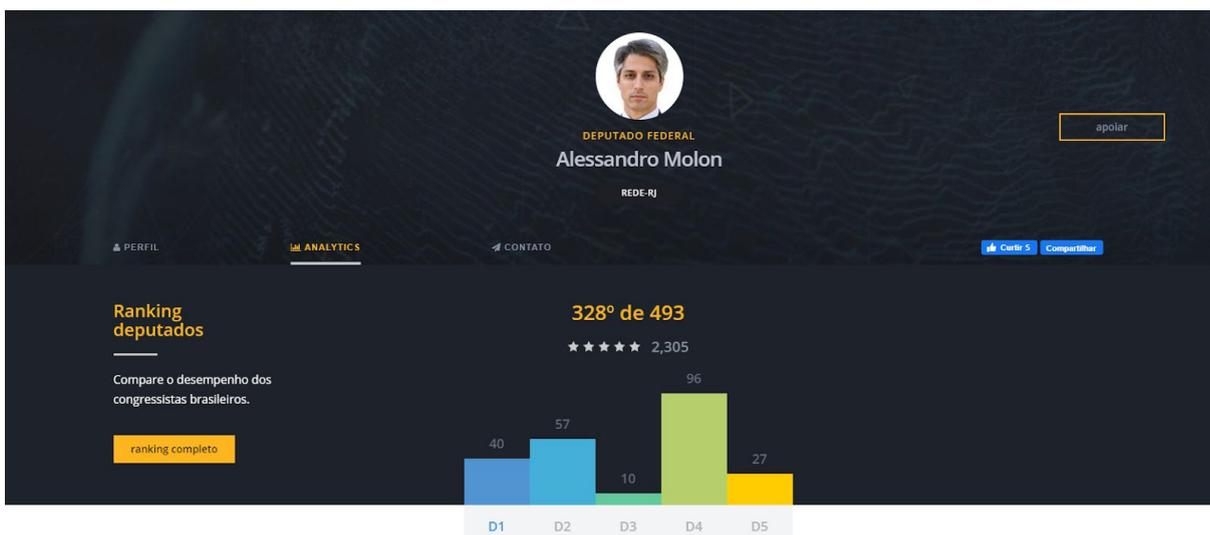


A última seção traz um gráfico no estilo *treemap* com dados sobre a distribuição de “Recursos Humanos e Concursos”. Apesar deste tipo de gráfico ser utilizado para buscas segmentadas, o usuário não pode esmiuçar mais as distribuições apresentadas.

Além desta interface, a Câmara apresenta projetos de terceiros com seus dados. Atualmente 12 projetos são listados. Destes, 10 tem o objetivo de informar o público por meio de *newsletters* e outras formas de conteúdo personalizado. Dois são semelhantes ao propósito deste artigo. O primeiro é o aplicativo Kan, para acompanhar gastos públicos, no site da câmara há um link para seu repositório github. Contudo, o acesso ao aplicativo em si ainda não está disponível. O segundo projeto pertinente aqui é o Atlas Político.



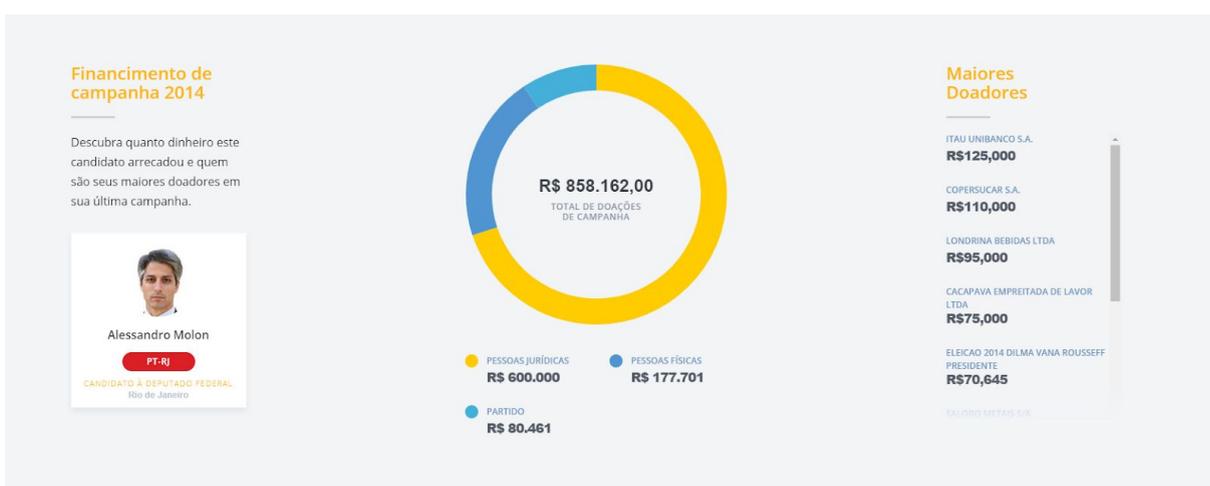
Na tela inicial há uma consulta por parlamentares, onde o usuário pode buscar por nome, e filtrar se é deputado ou senador, qual é seu estado, sua cidade, seu partido e processos em que está inserido.



O site apresenta um sistema de ranqueamento que considera 5 fatores. São estes: Representatividade, Campanha Responsável, Ativismo Legislativo, Debate Parlamentar e Fidelidade Partidária.



Em seguida como o parlamentar se posiciona no mapa ideológico da Câmara, levando em conta suas atividades e projetos de lei.



Também há uma seção referente à natureza das doações de campanha que levam o parlamentar à Câmara ou ao Senado.

Processos

Não encontramos nenhum processo contra Alessandro Molon na base de dados do Atlas Político. Caso essa informação esteja desatualizada, entre em contato via info@atlaspolitico.com

Biografia

Conheça mais sobre o político e sua trajetória pessoal.

NOME COMPLETO

Alessandro Lucciola Molon

PROFISSÃO

DEPUTADO

DATA NASCIMENTO

28/10/1971

GRAU DE INSTRUÇÃO

SUPERIOR COMPLETO

NATURALIDADE

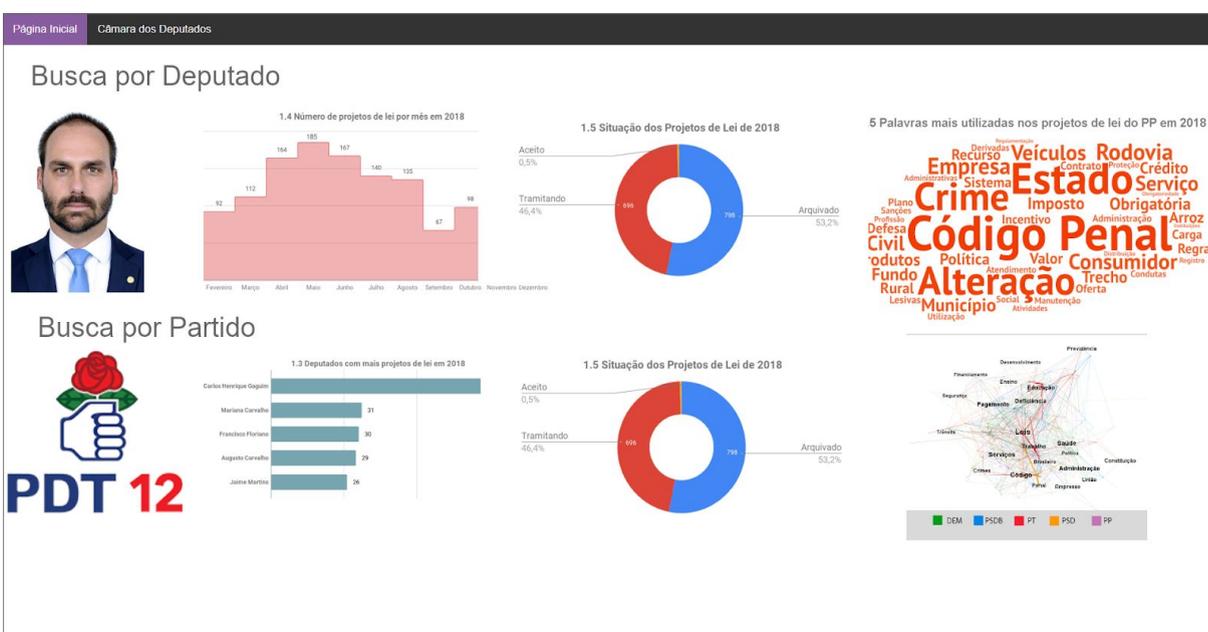
BELO HORIZONTE - MG

O site ainda traz informações sobre processos em que o parlamentar pode estar inserido e sua biografia.



Por fim, há uma seção dedicada ao mapa ideológico, que possui quatro eixos: Esquerda e Direita; Oposição e Governo.

A interface oficial da Câmara dos Deputados e do Atlas Políticos auxiliaram a conceitualizar um modelo de análise próprio.



A interface desse artigo se diferencia das outras propondo mais opções de interação, para que os usuários possam customizar e complexificar suas consultas. E, assim que obterem a informação que precisam, baixar os gráficos ou compartilhá-los em suas redes. No imagem acima, dois exemplos de consulta. Por deputado, com gráfico temporal do número de projetos de lei por mês, gráfico de pizza à situação em que se encontra seus projetos, e por fim as temáticas destes com uma Tagcloud. Por partido, gráfico de barras com os parlamentares com maior número de projetos de lei, gráfico de pizza à situação em que se encontra seus projetos e um grafo representando a conexão entre os temas dos projetos de lei.

3 METODOLOGIA

Técnicas de pesquisa bibliográfica e documental, assim como de análise de conteúdo e de redes sociais integram a metodologia aplicada neste estudo.

A pesquisa bibliográfica será adotada em seu sentido restrito, definido por Stumpf (2005) como “Um conjunto de procedimentos para identificar, selecionar, localizar e obter documentos de interesse para a realização de trabalhos acadêmicos”. A partir deste modelo será empregada uma revisão da literatura nacional sobre análise de redes para mídias sociais a partir dos grupos de estudo UFPEL, LABIC-UFES, UFMG, IBPAD, FGV e Davint-PUCRS, e do congresso Brasnam.

Segundo Moreira (2005), a pesquisa documental funciona como expediente eficaz para contextualizar fatos, situações e momentos, conseguindo desta maneira introduzir novas perspectivas sem deixar de respeitar a substância original dos documentos. Esta técnica será usada para documentar a frequência e a forma no uso da ARMS por parte da grande mídia, levando em conta “quando” e “porque” a visualização foi feita por jornalista do veículo ou profissional de fora.

O presente trabalho também se valerá da análise de conteúdo. Esta metodologia, de acordo com Bardin (1977), possui uma “flexibilidade de aplicação” que não se limita ao campo linguístico. Este aspecto é um aspecto decisivo a este trabalho de pesquisa, já que o conteúdo de mídia social a ser analisado, além de texto, traz imagens, vídeos e *hiperlinks*.

4 CONCLUSÃO

Cohen et al (2012) considera que poucos campos fizeram mais para auto governança e cidadania online do que o jornalismo de dados. De Souza et al (2018) aponta que a relação entre jornalistas e dados digitais é recente e ainda cresce. Assim, adequar as proficiências do jornalismo de dados é uma tarefa oportuna, quiçá essencial ao futuro da profissão. Porém, para isso é preciso expor a origem e as atuais consequências político-econômicas dessa prática.

5 CRONOGRAMA

O cronograma deve ser adequado às necessidades do trabalho.

Atividades	Projeto Final							
	Projeto 1				Projeto 2			
	1ª se m	2ª se m	3ª se m	4ª se m	1ª se m	2ª se m	3ª se m	4ª se m
Escolha do assunto do projeto	x							
Elaboração da estrutura do projeto		x						
Seleção e leitura das obras para elaboração do projeto		x						
Elaboração dos objetivos, delimitação do tema, definição do problema, etc.		x						
Elaboração da pesquisa bibliográfica e documental do projeto		x						
Coleta de dados			x					
Tratamento dos dados				x				
Revisão final do texto e elaboração da introdução e conclusão				x				
Data limite de entrega do Projeto de Estágio				x				

6 REFERÊNCIAS BIBLIOGRÁFICA

BERENDT, Bettina. Power to the agents?! In the# WebWeWant, people will critically engage with data—and data journalism can help them want to do this. In: USEWOD-PROFILES 2015. 2015, p. 29-31.

BONAQUE, Raphaël, et al. Mixed-instance querying: a lightweight integration architecture for data journalism. Proceedings of the VLDB Endowment: 2016, 9.13: 1513-1516.

BORGES-REY, Eddy. Unravelling Data Journalism: A study of data journalism practice in British newsrooms. Journalism Practice: 2016, 10.7: 833-843.

BOYLES, Jan Lauren; MEYER, Eric. Newsrooms accommodate data-based news work. Newspaper Research Journal: 2017, 38.4: 428-438.

DE SOUZA, Daniele R., et al. Interacting with Data to Create Journalistic Stories: a Systematic Review In: International Conference on Human Interface and the Management of Information. Springer, Cham, 2018. p. 685-704.

DIAKOPOULOS, Nicholas; GOLDENBERG, Sergio; ESSA, Irfan. Videolyzer: quality analysis of online informational video for bloggers and journalists In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM, 2009. p. 799-808.

FELLE, Tom. Digital watchdogs? Data reporting and the news media's traditional 'fourth estate' function. Journalism, 2016, 17.1: 85-96.

FRANCISCO-REVILLA, Luis. Digital libraries for computational journalism In: Proceedings of the 12th ACM/IEEE-CS joint conference on Digital Libraries. ACM, 2012. p. 365-366.

GARCÍA, Xosé López; RODRÍGUEZ, Ana Isabel Vázquez; PEREIRA, Xosé Fariña. Technological Skills and New Professional Profiles: Present Challenges for Journalism. 2017.

GARRISON, Bruce. Tools daily newspapers use in computer-assisted reporting. Newspaper Research Journal, 1996, 17.1-2: 113-126.

GRAY, Jonathan; CHAMBERS, Lucy; BOUNEGRU, Liliana. The data journalism handbook: How journalists can use data to improve the news. " O'Reilly Media, Inc." (2012)

LANGER, Julia; ZEILLER, Michael. Evaluation of the User Experience of Interactive Infographics in Online Newspapers.

LEWIS, Seth C.; USHER, Nikki. Code, collaboration, and the future of journalism: A case study of the Hacks/Hackers global network. Digital Journalism, 2014, 2.3: 383-393.

LEWIS, Seth C.; WESTLUND, Oscar. Big data and journalism: Epistemology, expertise, economics, and ethics. Digital journalism, 2015, 3.3: 447-466.

LIMA, Soraia. Jornalismo de Dados no Brasil: Tendências e Desafios In: SILVA, Tarcízio; STABILE, Max. Monitoramento e Pesquisa em Mídias Sociais –Metodologias, aplicações e inovações. Análise de redes em Mídias Sociais. São Paulo: Uva Limão, 2016, p.279-299.

MAYO, Justin; LESHNER, Glenn. Assessing the credibility of computer-assisted reporting. *Newspaper Research Journal*, 2000, 21.4: 68-82.

SPLENDORE, Sergio, et al. Educational strategies in data journalism: A comparative study of six European countries. *Journalism*, 2016, 17.1: 138-152.

WU, You, et al. On one of the few objects In: *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2012. p. 1487-1495.