

VISUALIZAÇÃO DE AFINIDADES ENTRE PARLAMENTARES MEDIANTE DADOS DE VOTAÇÕES NO SENADO BRASILEIRO

Ricardo César Gonçalves Sant'Ana*
Fernando de Assis Rodrigues**

artigo de revisão

RESUMO:

A ampliação da participação cidadã na esfera pública depende diretamente do acesso as informações relativas a atuação dos representantes eleitos, principalmente no que diz respeito a suas decisões nas votações durante seus mandatos. A Ciência da Informação pode colaborar neste processo, propondo e avaliando modelos de acesso a estas informações que podem ser obtidos através de dados a serem disponibilizados pelos sítios oficiais do Poder Legislativo nas esferas Federal, Estadual e Municipal. Propõe-se neste artigo a análise do processo de coleta e uso de dados sobre votações de senadores com vistas a apropriação deste modelo para coleta e uso de dados nas demais esferas. A partir dos dados é analisada a elaboração de uma matriz de afinidades que permita identificar a relação entre cada um dos parlamentares com os demais, em função das similaridades das decisões tomadas no conjunto das votações abertas. É analisa também a elaboração de visualizações iniciais e a ampliação do escopo da pesquisa através da aplicação dos dados obtidos em todas as afinidades entre os parlamentares e na obtenção de uma afinidade média entre partidos, permitindo novas dimensões de análise aos dados coletados. A elaboração das matrizes completas das relações de afinidades entre os parlamentares, pode propiciar um horizonte rico de possibilidades para elaboração de novas formas de visualização e análise, ampliando a visibilidade das ações parlamentares junto a sociedade.

PALAVRAS-CHAVE: Dados Abertos. Visualização de dados. Transparência Pública.

* Doutor em Ciência da Informação pela Universidade Estadual Paulista Júlio de Mesquita Filho, Brasil. Professor do Programa de Pós-Graduação em Ciência da Informação da Universidade Estadual Paulista Júlio de Mesquita Filho, Campus de Marília.
E-mail: ricardosantana@marilia.unesp.br.

** Bacharel em Sistemas de Informação pela Universidade do Sagrado Coração, Brasil. Doutorando no Programa de Pós-Graduação em Ciência da Informação da Universidade Estadual Paulista Júlio de Mesquita Filho, Campus de Marília.
E-mail: fernando@elleth.org.

I INTRODUÇÃO

Mesmo com os recursos proporcionados pelas Tecnologias da Informação e Comunicação, são grandes as dificuldades na ampliação da participação popular nas administrações públicas, mas pode-se buscar novas formas de acesso a dados que favoreçam a participação cidadã, por exemplo, no acompanhamento de votos de parlamentares. Apresenta-se neste artigo uma proposta de construção de recurso de visualização destas

votações, que propiciem análises de suas inter-relações e ainda uma breve descrição do quão disponíveis estão os dados sobre as votações dos parlamentares no Senado Brasileiro.

Adotou-se um modelo de grafo social para exibir as afinidades entre os senadores, conforme trabalho realizado por Chris Wilson sobre votações realizadas no Senado norte-americano em que foram utilizados dados sobre votações dos senadores para gerar uma relação de afinidade, exibindo coerência entre os votos do senador Arlen Specter em relação as votações de

membros dos partidos Democrata e Republicano (WILSON, 2009).

O conceito de construção de grafos sociais baseia-se na definição matemática da teoria dos grafos em que são estudadas as relações entre elementos de um conjunto. Estes elementos são denominados vértices e cada uma das relações é composta por um par de vértices e é denominada aresta. O estudo destas ligações entre elementos de um conjunto possibilita a sua aplicação em diversas áreas do conhecimento (DIESTEL, 2000; WILSON, 1996). Assim, obtêm-se regras para exibição de relações entre dois ou mais objetos em um determinado contexto, sendo necessário escolher o conjunto de dados a ser analisado e a regra ou fórmula que irá definir as relações entre os elementos do conjunto, formando as arestas. Um exemplo de uso desta teoria seria um gráfico representando parte da World Wide Web com páginas representadas como vértices e os links como as arestas (LOCKE, 2000)

Odehahn (2010) utilizou-se do modelo de visualização proposto por Wilson (2009), considerando a premissa de que o senado norte-americano é costumeiramente entendido pela sociedade como sendo composto por duas coalisões bastante estáveis, em função do conceito bipartidário. Assim, Odehahn (2010) buscou em sua proposta criar um grafo que pudesse exibir um quadro mais amplo, revelando que existe um fator de variabilidade na formação das relações dos senadores de ambos partidos ao longo do tempo quando toma-se por base as votações durante a legislatura.

Na construção de seu quadro, Odehahn (2010) teve acesso a dados primários (*raw data*) sobre as votações, ou seja, utilizou dados sobre os votos dos senadores como fonte de informação para criar uma representação gráfica das afinidades oriundas de suas decisões.

No trabalho realizado por Odehahn (2010), para a construção de uma visualização a partir destes dados, foi necessário o desenvolvimento de um grafo (a partir dos dados abertos disponibilizados) que explorasse estas questões visualmente para a observação da antítese com as premissas ensinadas nos colégios estudantis americanos, que definem o senado como um corpo inerente e conservativo, no sentido oposto e literal ao de mudanças de opiniões entre seus membros. De um lado os republicanos e de outro (oposto) os democratas (ODEWAHN, 2010).

Assim foi possível exibir, tratar e visualizar os dados já existentes, mas de uma forma em que se conseguiu extrair e analisar “algo” oculto, não-visível em uma outra abordagem ou em outra visualização dos mesmos dados – demonstrando estas conexões entre senadores (a priori ocultas nos dados primários), ao invés de visualizar apenas as votações singularmente em uma lista ordenada.

O conceito de dado primário identifica dados que não tenham sido obtidos a partir de transformações ou manipulações de tal forma que representem valores o mais próximo possível de suas origens, ou seja, não são resultantes de somas, agrupamentos, análises ou qualquer outro processamento. Assim, os dados primários não podem permitir divisões em que partes destes dados mantenham referência semântica no contexto de sua aplicação. Portanto, ao analisar o faturamento de um determinado empreendimento, por exemplo, o valor total de vendas de um dia é o resultado da soma das notas emitidas e não é um dado primário, ao contrário do valor de uma nota de venda que representa um dado primário para este contexto de análise.

Considera-se o acesso a dados primários como um dos fatores-chave de sucesso na construção de alternativas para acesso à dados evitando-se eventuais falhas de interpretação e ainda permitindo variados caminhos de análise em função do alto grau de liberdade de combinações e possíveis processamentos que um dado primário oferece.

Embora os conjuntos de dados primários [...] requeiram agregação e sínteses, ao mesmo tempo estes dados apresentam uma importante vantagem, que é precisamente o fato de que eles podem ser analisados, sintetizados e reagrupados de muitas formas. Em outras palavras, dados primários são flexíveis, e ao tornar estes dados disponíveis ao público as agências tem explicitamente convidado a sociedade a propor novos a estes dados, de acordo com as necessidades e desejos dos usuários destas informações. (EUA, 2011)

Outro fator importante no acesso à dados e a possibilidade destes dados serem tratados diretamente por máquinas, ou seja, eles deve estar disponíveis de tal forma que sejam passíveis de

utilização sem a necessidade de redigitação ou de tratamentos adicionais para sejam passíveis obtenção direta. Esta característica propicia a interoperabilidade de aplicações e o uso destes dados por pesquisas externas com independência do formato inicial e de especificidades ligadas a plataforma tecnológica (BERNERS-LEE, 2009). Assim, por exemplo, um dado que está disponível em um arquivo que tem as características de apresentação, tal como um arquivo em formato PDF (*Portable Document Format*) pode não estar acessível para tratamento direto pela máquina, já que exige interpretação sobre as posições em que os valores são apresentados para que a semântica seja obtida.

No ano de 2007, na cidade de Sebastopol nos Estados Unidos, foi reunido um grupo de trabalho multidisciplinar para o desenvolvimento de princípios que possam conceituar um dado como dado aberto (*open data*). Esta iniciativa foi nomeada como "*Open Government Working Group*". A análise deste trabalho foi iniciada com base nos registros de órgãos governamentais norte-americanos e resultou em oito princípios para avaliar se um determinado registro pode ser considerado um dado aberto. Estes princípios atribuem que um dado aberto deve ser completo, primário, temporal (deve ser disponibilizado a tempo de ser ainda útil), acessível, processável por máquina, não-discriminatório (o usuário não tem necessidade de identificar-se), não-proprietário e livre de licença (OGD, 2007). Portanto um dado aberto deve ser primário, mas nem todo dado primário é um dado aberto. Considerou-se neste artigo as duas características: dados primários e dados tratáveis por máquinas como sendo as principais na definição de um dado como dado aberto (OGD, 2007).

A proposta deste artigo é o aporte do modelo norte-americano na criação de um modelo de visualização para o cenário brasileiro, mas com o objetivo de analisar se os dados primários de votações abertas (no período de 2007 a 2010) e as Tecnologias de Comunicação e Informação do senado brasileiro dão sustentação para afirmar a existência de uma estrutura divisória das coalisões entre os partidos que compõem a base aliada do governo, a oposição e os independentes.

2 OBJETIVO E PROCEDIMENTOS

Este artigo tem o objetivo de apresentar uma análise sobre o processo de coleta e

elaboração de estrutura visual dos dados das matérias votadas pelo Senado brasileiro através de um grafo, neste contexto, significando uma coleção de vértices (*nodes*) e arestas (*edges*) e não um gráfico de coordenadas (ODEWAHN, 2010, p. 123) de relações sociais, construído com base em uma matriz de afinidades, adaptada do modelo conceitual norte-americano de Andrew Odewahn (2010), exposto no livro "*Beautiful Visualization*", de Julie Steele e Noah Iliinsky.

Para a obtenção dos dados primários foi utilizado apenas os canais oferecidos pelas Tecnologias de Comunicação e Informação (TIC) oficiais do Senado Federal brasileiro, mais precisamente os metadados dos senadores e os dados primários de votações abertas, no período de 2007 a 2010. Com isto, puderam ser analisadas, de forma visual questões mais profundas, como a possibilidade que estes resultados sintetizados indiquem a existência ou não de afinidades entre partidos (e seus membros) que compõem oposição e governo.

2.1 Construção do grafo social

No artigo publicado por Chris Wilson (2009), os vértices representam os senadores e cada vértice recebe um rótulo oculto (*hidden label*) que ao ser disparado o comportamento *mouseover* do navegador (ação de deixar o ponteiro do mouse em cima de um objeto) apresenta o nome do respectivo senador. Os vértices são coloridos e cada cor representa um partido (vermelho no caso dos membros republicanos, azul para os democratas, verde para senadores os independentes e amarelo na falta desta informação). As arestas representam as ligações entre dois senadores quando os mesmos votaram igualmente mais de 65% no período determinado.

No estudo de Odewahn (2010) foi acrescentada em sua abordagem a separação dos membros em função de seus partidos. O grafo apresentou os democratas à esquerda e os republicanos à direita para um melhor entendimento das ligações entre os membros dos partidos com o grupo de sua oposição. Também foi definido que cada sessão bial legislativa é a unidade básica de tempo dos grafos. Exemplificando, a centésima quinta sessão corresponde ao período iniciado no dia três de janeiro de 1997 e finalizado no dia três de janeiro de 1999. Sua principal justificativa para a escolha

desta unidade de tempo é que o corpo do senado norte-americano passa por mudanças a cada período, visto as eleições de novos senadores e o desligamento de outros (ODEWAHN, 2010, p. 125).

Para Odewahn (2010) são necessários três passos para a criação deste modelo de visualização: a aquisição dos dados primários; o cálculo e a geração da matriz de afinidade e; a construção do grafo a partir destes dados.

2.2 A aquisição dos dados primários

Para a construção da visualização é necessário a obtenção de dois conjuntos de dados: o primeiro é o conjunto de metadados contendo informações individuais sobre cada senador, com elementos como o seu nome e partido. O segundo conjunto é composto pelos dados primários de votações feitas por estes senadores em uma determinada matéria, no período desejado (ODEWAHN, 2010). Os metadados são fator chave para o funcionamento e interoperabilidade de sistemas de informação, principalmente quando há a necessidade de aquisição de dados primários de uma base de dados externa e-ou não-visível a um software, já que são os próprios metadados que garantem a correta representação do conteúdo informacional de um determinado recurso (SANTOS; ALVES, 2009). Após a definição dos metadados que irão compor uma representação de um conjunto de dados, é necessário escolher qual o formato de metadados possui condições de representá-lo com eficiência; e assim criar base para a correta condução de uma conectividade com as mais variadas TIC e aplicações ferramentais existentes. Assim, os formatos ou padrões de metadados são estruturas padronizadas, que através do conjunto de dados-atributos, permitem a correta representação do conteúdo informacional (SANTOS; ALVES, 2009).

O principal obstáculo nesta fase foi a ausência destes dados primários em web sites governamentais norte-americanos de maneira estruturada e legível por máquina (ODEWAHN, 2010, p. 125). Contudo, projetos civis como o GovTrack (<http://www.govtrack.us>) organizam e padronizam estes dados primários (sobre senadores e seus votos em matérias) em metadados que são legíveis por máquina – e ainda os disponibilizam em conformidade com

o conceito de dados abertos – permitindo a obtenção dos mesmos dados oficiais dos web sites do governo americano, mas agregando um padrão de metadado. Esta modelagem permite que outros softwares utilizem estes recursos através de tecnologias que auxiliam a interoperabilidade, como o XML (*eXtensive Markup Language*), o JSON (*JavaScript Object Notation*), entre outros. O GovTrack possui, em sua base, dados primários estruturados e completos de votações do congresso retroativas até 1991; e estruturados, porém incompletos, pré-1991. Esta iniciativa civil está em acordo com os oito princípios necessários para a denominação “dados abertos” (OGD, 2007).

No sítio do Senado do Brasil (<http://www.senado.gov.br>), os dados sobre os senadores do exercício vigente são encontradas em um documento (<http://www.senado.gov.br/senadores>), em formato HTML (*Hypertext Markup Language*), com um metadado de tipo simples, em nível básico, para indexação apenas da página como um todo por motores de busca, através do uso de *crawlers* (SENSO, 2003, p. 103), que são algoritmos de motores de busca que coletam informações de um conteúdo ou arquivo disponível na Internet e armazenam em uma base de dados, os endereços originais em que o arquivo/página está armazenado.

Portanto, os dados ali contidos sobre os senadores não são estruturados para um uso facilitado por aplicativos, gerando a necessidade de tratamento através de uma camada de abstração, ou seja, de um algoritmo de conversão para um determinado padrão estruturado, legível por máquina, como, por exemplo, o formato XML; e com isto não atende a um dos oito princípios de dado aberto: que estes dados devem ser processáveis por máquina (OGD, 2007). Esta característica dificulta a utilização dos dados por ferramentas tecnológicas sem a intervenção humana.

Com relação à atividade legislativa, ou seja, os votos do colegiado, por estarem apenas disponíveis em documentos digitais no formato PDF (*Portable Document Format*), não é possível um acesso direto aos dados ali contidos, também necessitando uma camada abstração para uso automatizado e o *download* manual dos resultados das matérias para tratamento e uso. A obtenção destes resultados é feita através de uma consulta (<http://www.senado.gov.br/>

atividade/votacao/default.asp) pré-formatada e diretamente no navegador, não existindo um acesso direto aos seus dados primários sem intervenção humana.

Para a criação do grafo de visualização das afinidades entre os votos dos senadores, foi necessário identificar quais senadores fizeram parte do colegiado em um exercício anterior: de 2007 a 2010. Estes dados foram encontrados em formato similar a página do colegiado em exercício vigente, ou seja, em um documento no formato HTML e também disponível para impressão (<http://www.senado.gov.br/senadores/periodos/LegisAnter53.shtm>). Neste documento foi possível encontrar três elementos: o nome do senador, sua unidade federativa e seu partido na época em que foi eleito para aquele exercício.

Para a extração destes dados, novamente foi preciso o desenvolvimento de um algoritmo que capturou este documento como um todo e que transformou do formato dos dados ali contidos, extraindo-os e convertendo-os em uma estrutura que permitiu inserir estas informações em um banco de dados externo ao sítio do senado, construído especificamente para a visualização proposta.

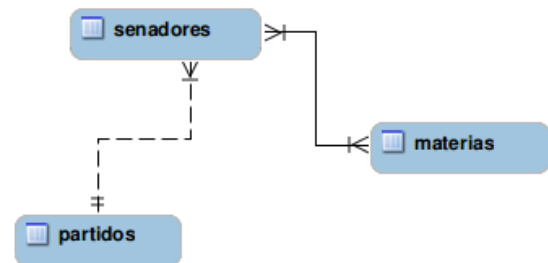
Esta obtenção dos dados sobre as atividades legislativas do Senado, mais especificamente sobre as votações abertas foi realizada utilizando apenas os dados disponibilizados pelo sítio oficial do senado, sem o uso de nenhum projeto mantido fora do escopo oficial.

Neste trabalho foi um recurso de busca, disponibilizado no sítio do Senado, específico para pesquisa das atividades legislativas sobre matérias já votadas. O tipo de proposição das matérias foi ignorado, bem como qualquer outro tipo de filtro, exceto a unidade de tempo proposta: matérias votadas entre 2007 e 2010. O volume de matérias encontradas, com votações abertas para este período, foi: 34 em 2007, 50 em 2008, 64 em 2009 e 28 em 2010 – totalizando 176 matérias.

Este período foi escolhido por possuir dados mais recentes e em um período legislativo fechado no momento da elaboração desta pesquisa: primeiro semestre do ano de 2011 (RODRIGUES; SANTANA, 2011). Todavia, este modelo de construção pode ser aplicado em qualquer período e exercício legislativo,

desde que se consiga acesso aos metadados dos parlamentares e os dados do movimento das atividades legislativas.

Figura 1 - Representação lógica dos dados primários e relacionamentos entre senadores, partidos e votações



Fonte: Elaboração dos autores

Uma alternativa viável para o uso dos dados obtidos das fontes oficiais para a geração dos grafos e posterior visualização, é transformá-los em registros de uma base de dados relacional (ODEWAHN, 2010), preferencialmente instanciada em uma banco de dados de código aberto (*open source*), como por exemplo os bancos *SQLite Database* ou o *MySQL*. Contudo, a modelagem entidade-relacionamento bipartidária proposta por Odewahn (2010) não se encaixa no cenário brasileiro, principalmente devido a característica brasileira de representação multipartidária. A Figura 1 apresenta o modelo lógico proposto para os dados primários necessários para a construção específica do grafo sobre as afinidades entre os parlamentares com base nas votações em plenário.

2.3 Geração da matriz de afinidade

Após a obtenção dos dados primários e transformação para uma base relacional, o segundo passo é a extração, a partir destes dados primários, das informações sobre as afinidades entre os senadores no período de tempo escolhido.

Pode-se construir estas conexões calculando uma matriz de afinidades, alimentada com os dados (votos) destes senadores. A matriz de afinidade (Figura 2) possui uma estrutura

similar ao conceito de linhas, colunas e células de uma planilha. Suas colunas e linhas representam todas as matérias que tiveram votação aberta em um determinado período, sendo que as matérias da coluna representam os votos de um senador (ex.: senador “a”) e as das linhas representam de outro parlamentar (ex.: senador “b”).

Se caso o encontro destas linhas tenha valores comuns para ambos, ou seja, se no encontro de uma mesma matéria, os dois

senadores votaram igualmente, marca-se um ponto de afinidade. Posteriormente, se a soma dos pontos de afinidades, dividido pelo número total de matérias, for maior que o fator “0,65”, fator proposto por Odewahn (2010) a partir do artigo de Wilson (2009), cria-se uma aresta entre os senadores no grafo. A matriz deve ser executada de dois em dois senadores até que todos sejam confrontados entre si (ODEWAHN, 2010, p. 124).

Figura 2 – Matriz de afinidade dos votos entre dois senadores

		Senador a					
		Matéria 1	Matéria 2	Matéria 3	Matéria 4	Matéria n	Matéria 6
Senador b	Matéria 1	-	Σ	Σ	Σ	Σ	Σ
	Matéria 2	-	-	Σ	Σ	Σ	Σ
	Matéria 3	-					
	Matéria 4	-					
	Matéria n	-					
	Matéria 6	-					

Cria-se uma aresta quando:

$$\frac{\sum (\text{senador a, senador b})}{\text{Total de Votações}} > 0,65$$

Fonte: Odewahn, 2010, tradução nossa

Na geração das afinidades das votações abertas do Senado brasileiro foi utilizado o mesmo algoritmo proposto por Odewahn (2010). O tratamento destes dados pode ser feito com qualquer linguagem de programação, já que é irrelevante a tecnologia para criar a matriz de afinidades contando que o algoritmo respeite a regra proposta. A matriz de afinidades pode gerar milhares de arestas em um período de tempo, dependendo diretamente do volume de matérias computadas e como o volume de dados destas afinidades é relativamente extenso para ser calculado *on-demand* por um computador pessoal, os resultados obtidos devem ser gravados em uma nova instância do banco de dados.

2.4 Construção do grafo de vértices e arestas

O terceiro passo para a construção da visualização é a transformação dos dados gravados pela matriz de afinidades, através de uma

ferramenta tecnológica, em um grafo. Dentre as várias ferramentas tecnológicas disponíveis para a construção de gráficos, utilizou-se o software *open source* GraphViz (<http://www.graphviz.org/>) que já possui a construção nativa de grafos de vértices e arestas – chamado “neato”. O modelo de vértices e arestas segue modelo similar ao proposto por Wilson (2009), em que os vértices representam os senadores e as arestas representam as afinidades entre eles. Assim, cada vértice recebe um rótulo (*label*) numérico que representa um senador em uma lista (externa ao grafo) ordenada alfabeticamente.

Na construção do grafo para o modelo brasileiro, foi necessário obter um elemento adicional dos dados primários: a representação partidária. O Brasil possui um sistema político multipartidário e, portanto, as cores de cada vértice representam um partido que possuiu ao menos uma representação no período 2007-2010. Os partidos identificados foram: Democratas;

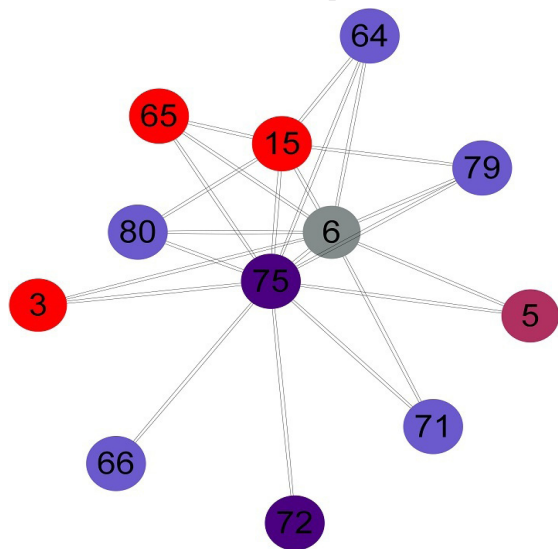
PC do B; PDT; PMDB; PP; PR; PRB; PSB; PSC; PSDB; PSOL; PT; PTB; PV; e ainda senadores sem afiliação partidária.

3 RESULTADOS

Concluída a sequência de três passos proposta por Odewahn (2010), o resultado imagético das afinidades das votações abertas entre os senadores, no período de 2007 a 2010, é demonstrado na Figura 3. A imagem exhibe apenas os membros que possuíram uma afinidade maior ou igual ao fator utilizado (0,65). Os senadores que não possuem afinidade maior que o valor 0,65 com nenhum outro membro da casa não são exibidos.

Este *cluster* é formado principalmente pelos partidos PMDB (5 membros), PT (3 membros), PTB (2 membros), DEM (1 membro), PSB (1 membro) em um total de 12 membros com grandes afinidades em suas votações, principalmente conduzida pelos senadores do PTB (75) e do PSB (6) que apresentam alto índice de afinidade com os senadores do PT (vértices vermelhos), com a oposição (senadores na parte inferior do grafo) e com o PMDB (vértices azuis-ardósia, que possuem senadores da base do governo e centro-oposição).

Figura 3 - Grafo de vértices e arestas representando senadores que possuíram afinidades entre seus votos abertos no período de 2007 a 2010

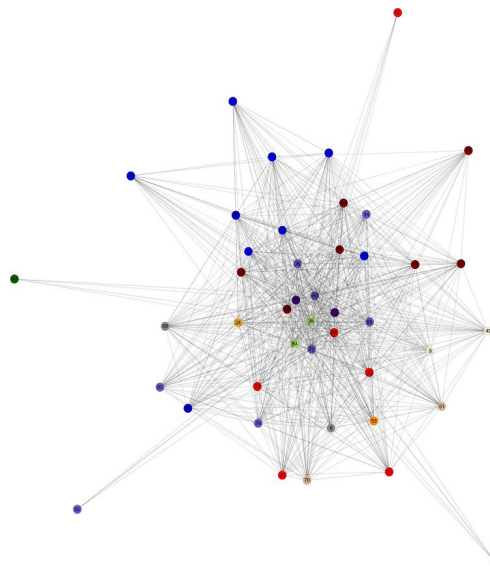


Fonte: Elaborado pelos autores.

Em seguida, é apresentada uma análise de unidades de tempo menores e isoladas dentro do mesmo mandato deste colegiado de senadores, utilizando a mesma modelagem de construção do grafo, porém com matérias votadas em apenas um ano, com o objetivo principal de isolar fatos importantes no período de tempo proposto.

A Figura 4 demonstra o grafo de afinidades entre os senadores baseados em matérias votadas apenas no ano de 2010. São encontradas afinidades entre membros dos partidos: PMDB (9 membros), PSDB (9 membros), DEM (7 membros), PT (6 membros), PRB (2 membros), PTB (2 membros), PDT (2 membros), PSB (2 membros), PR (2 membros), PC DO B (1 membro), PP (1 membro), PSOL (1 membro). No entanto, nesta figura aparecem membros dos partidos PSDB e DEM em uma forte afinidade na parte superior do grafo e quatro membros do PT permeados por membros de siglas com menor número de representantes no colegiado, como o PR, PSB e o PRB; mas estes possuem fortes ligações com a base aliada do governo. Partidos como o PSOL (senador de número 43) mantém uma longa distância da base aliada e da oposição, apesar deste partido de estar representado naquele mandato e neste grafo por apenas um único senador. Este grafo torna perceptível um núcleo de membros articulados em torno das votações do ano de 2010.

Figura 4 - Grafo de vértices e arestas representando os senadores que possuíram afinidades entre seus votos abertos no ano de 2010



Fonte: Elaborado pelos autores.

Acrescenta-se aos resultados obtidos o conjunto de todas as votações abertas dos senadores e das respectivas afinidades calculadas entre todos eles, o que nos permite identificar um quadro geral de posicionamento em relação a afinidade média dos senadores de cada partido em relação aos demais partidos.

Esta afinidade média foi obtida pela soma entre todas as votações entre cada um dos parlamentares em relação aos demais parlamentares e dividindo-se esta soma pelo número de elementos calculados. Estes dados, coletados em planilhas contendo um código do parlamentar, o partido que está filiado e seguido de outras três colunas

contendo o código de cada um dos demais parlamentares, seus respectivos partidos e o índice calculado entre cada um deles foi tratada pelo aplicativo de consolidação de afinidades entre partidos (v.exe), disponibilizado no site: <http://dadosabertos.info>, e que permite a elaboração de quadros de correlação de afinidades entre as votações dos representantes de cada partido, conforme apresentado nas figuras 5, 6, 7 e 8. Situações em que o partido conta apenas com um filiado, no cálculo de seu índice de afinidade dentro do próprio partido resultaram em índice 0,00.

Figura 5 - Índice de afinidade entre partidos no Senado nas votações abertas de 2007

PARTIDO	DEM	PMDB	PT	PSDB	PR	PDT	PTB	PP	PC DO B	PSOL	PRB	PFL	PSB
DEM	0,1194	0,1378	0,1551	0,1396	0,1329	0,1138	0,1051	0,2059	0,1557	0,0606	0,1202	0,0000	0,1488
PMDB	0,1378	0,3269	0,3717	0,2608	0,3101	0,2658	0,2428	0,4468	0,3964	0,1737	0,2675	0,0930	0,3739
PT	0,1551	0,3717	0,4545	0,2484	0,3667	0,3157	0,2849	0,5049	0,4436	0,1520	0,3272	0,0771	0,4330
PSDB	0,1396	0,2608	0,2484	0,3174	0,2261	0,1971	0,1807	0,3592	0,2878	0,2038	0,1670	0,1174	0,2815
PR	0,1329	0,3101	0,3667	0,2261	0,2735	0,2565	0,2353	0,4353	0,3765	0,1235	0,2618	0,0792	0,3490
PDT	0,1138	0,2658	0,3157	0,1971	0,2565	0,2382	0,1904	0,3529	0,3235	0,1176	0,2559	0,0678	0,2745
PTB	0,1051	0,2428	0,2849	0,1807	0,2353	0,1904	0,1607	0,3419	0,2941	0,1140	0,1893	0,0627	0,2745
PP	0,2059	0,4468	0,5049	0,3592	0,4353	0,3529	0,3419	0,0000	0,5294	0,2353	0,3529	0,1176	0,4902
PC DO B	0,1557	0,3964	0,4436	0,2878	0,3765	0,3235	0,2941	0,5294	0,0000	0,1765	0,3088	0,1228	0,4412
PSOL	0,0606	0,1737	0,1520	0,2038	0,1235	0,1176	0,1140	0,2353	0,1765	0,0000	0,0882	0,0969	0,1863
PRB	0,1202	0,2675	0,3272	0,1670	0,2618	0,2559	0,1893	0,3529	0,3088	0,0882	0,2353	0,0199	0,2892
PFL	0,0000	0,0930	0,0771	0,1174	0,0792	0,0678	0,0627	0,1176	0,1228	0,0969	0,0199	0,0984	0,0969
PSB	0,1488	0,3739	0,4330	0,2815	0,3490	0,2745	0,2745	0,4902	0,4412	0,1863	0,2892	0,0969	0,4020

Fonte: Elaborado pelos autores.

Figura 6 - Índice de afinidade entre partidos no Senado nas votações abertas de 2008

PARTIDO	DEM	PMDB	PT	PSDB	PR	PDT	PTB	PP	PC DO B	PSOL	PRB	PSB	PSC	S/PART
DEM	0,2122	0,1913	0,2404	0,2385	0,2434	0,1515	0,1929	0,1838	0,3150	0,2650	0,1588	0,2819	0,2463	0,0300
PMDB	0,1913	0,2418	0,3247	0,2187	0,3266	0,1873	0,2561	0,2664	0,3927	0,3127	0,2145	0,3909	0,1955	0,0300
PT	0,2404	0,3247	0,4244	0,2800	0,4342	0,2441	0,3410	0,3369	0,5108	0,4015	0,2877	0,5138	0,2569	0,0462
PSDB	0,2385	0,2187	0,2800	0,2669	0,2808	0,1733	0,2164	0,1862	0,3677	0,3169	0,1692	0,3200	0,2569	0,0262
PR	0,2434	0,3266	0,4342	0,2808	0,4367	0,2575	0,3444	0,3750	0,5050	0,4050	0,2800	0,5200	0,2850	0,0500
PDT	0,1515	0,1873	0,2441	0,1733	0,2575	0,1280	0,1956	0,2133	0,3000	0,2567	0,1633	0,2950	0,1700	0,0200
PTB	0,1929	0,2561	0,3410	0,2164	0,3444	0,1956	0,2589	0,2756	0,4000	0,2956	0,2200	0,4144	0,2111	0,0422
PP	0,1838	0,2664	0,3369	0,1862	0,3750	0,2133	0,2756	0,0000	0,4200	0,3200	0,3600	0,4500	0,2400	0,0200
PC DO B	0,3150	0,3927	0,5108	0,3677	0,5050	0,3000	0,4000	0,4200	0,0000	0,5200	0,3400	0,6100	0,3400	0,0400
PSOL	0,2650	0,3127	0,4015	0,3169	0,4050	0,2567	0,2956	0,3200	0,5200	0,0000	0,2800	0,4600	0,2800	0,0000
PRB	0,1588	0,2145	0,2877	0,1692	0,2800	0,1633	0,2200	0,3600	0,3400	0,2800	0,0000	0,3700	0,2000	0,0200
PSB	0,2819	0,3909	0,5138	0,3200	0,5200	0,2950	0,4144	0,4500	0,6100	0,4600	0,3700	0,6200	0,3400	0,0600
PSC	0,2463	0,1955	0,2569	0,2569	0,2850	0,1700	0,2111	0,2400	0,3400	0,2800	0,2000	0,3400	0,0000	0,0400
S/PART	0,0300	0,0300	0,0462	0,0262	0,0500	0,0200	0,0422	0,0200	0,0400	0,0000	0,0200	0,0600	0,0400	0,0000

Fonte: Elaborado pelos autores.

Figura 7 - Índice de afinidade entre partidos no Senado nas votações abertas de 2009

PARTIDO	DEM	PMDB	PT	PSDB	PR	PDT	PTB	PP	PC DO B	PSOL	PRB	PSB	PSC	PV	S/PART
DEM	0,3116	0,2698	0,2988	0,3087	0,3069	0,2156	0,3503	0,3192	0,2690	0,2031	0,2885	0,4057	0,1953	0,0703	0,0273
PMDB	0,2698	0,2287	0,2658	0,2620	0,2783	0,1826	0,3089	0,2633	0,2352	0,1594	0,2645	0,3539	0,1563	0,0555	0,0219
PT	0,2988	0,2658	0,2951	0,2928	0,3014	0,2193	0,3499	0,3041	0,3017	0,1791	0,3053	0,4153	0,1875	0,0709	0,0331
PSDB	0,3087	0,2620	0,2928	0,3046	0,2958	0,1982	0,3423	0,2896	0,2563	0,1563	0,2823	0,3828	0,1885	0,0458	0,0250
PR	0,3069	0,2783	0,3014	0,2958	0,3021	0,1908	0,3354	0,2461	0,2227	0,1719	0,2969	0,3848	0,1172	0,0508	0,0234
PDT	0,2156	0,1826	0,2193	0,1982	0,1908	0,1548	0,2575	0,2545	0,2522	0,1518	0,2176	0,3080	0,1741	0,0759	0,0212
PTB	0,3503	0,3089	0,3499	0,3423	0,3354	0,2575	0,4107	0,3828	0,3535	0,2129	0,3848	0,4775	0,2402	0,0742	0,0273
PP	0,3192	0,2633	0,3041	0,2896	0,2461	0,2545	0,3828	0,0000	0,3750	0,2813	0,2891	0,4531	0,2969	0,0938	0,0313
PC DO B	0,2690	0,2352	0,3017	0,2563	0,2227	0,2522	0,3535	0,3750	0,0000	0,1719	0,3047	0,4219	0,2813	0,0781	0,0391
PSOL	0,2031	0,1594	0,1791	0,1563	0,1719	0,1518	0,2129	0,2813	0,1719	0,0000	0,1563	0,2734	0,1406	0,0938	0,0313
PRB	0,2885	0,2645	0,3053	0,2823	0,2969	0,2176	0,3848	0,2891	0,3047	0,1563	0,2969	0,4102	0,1328	0,0391	0,0352
PSB	0,4057	0,3539	0,4153	0,3828	0,3848	0,3080	0,4775	0,4531	0,4219	0,2734	0,4102	0,5000	0,2969	0,1250	0,0469
PSC	0,1953	0,1563	0,1875	0,1885	0,1172	0,1741	0,2402	0,2969	0,2813	0,1406	0,1328	0,2969	0,0000	0,0938	0,0000
PV	0,0703	0,0555	0,0709	0,0458	0,0508	0,0759	0,0742	0,0938	0,0781	0,0938	0,0391	0,1250	0,0938	0,0000	0,0156
S/PART	0,0273	0,0219	0,0331	0,0250	0,0234	0,0212	0,0273	0,0313	0,0391	0,0313	0,0352	0,0469	0,0000	0,0156	0,0156

Fonte: Elaborado pelos autores.

Figura 8 - Índice de afinidade entre partidos no Senado nas votações abertas de 2010

PARTIDO	DEM	PMDB	PT	PSDB	PR	PDT	PTB	PP	PC DO B	PSOL	PRB	PSB	PSC	PV
DEM	0,3565	0,3209	0,3855	0,3763	0,4023	0,3300	0,3170	0,5312	0,4933	0,3996	0,4687	0,4888	0,3437	0,1562
PMDB	0,3209	0,3045	0,3873	0,3339	0,3929	0,3152	0,3048	0,5018	0,4929	0,3875	0,4688	0,4759	0,3000	0,1518
PT	0,3855	0,3873	0,4659	0,4074	0,4786	0,3857	0,3719	0,6179	0,6036	0,4571	0,5875	0,5982	0,3393	0,2071
PSDB	0,3763	0,3339	0,4074	0,4054	0,4169	0,3490	0,3281	0,5670	0,5156	0,3750	0,5033	0,5190	0,3147	0,1964
PR	0,4023	0,3929	0,4786	0,4169	0,4583	0,3854	0,3763	0,6161	0,6161	0,5089	0,5759	0,5938	0,3929	0,1875
PDT	0,3300	0,3152	0,3857	0,3490	0,3854	0,2833	0,3053	0,5060	0,4762	0,3750	0,4673	0,4732	0,3095	0,1667
PTB	0,3170	0,3048	0,3719	0,3281	0,3763	0,3053	0,2670	0,4949	0,4745	0,3827	0,4515	0,4566	0,3061	0,1327
PP	0,5312	0,5018	0,6179	0,5670	0,6161	0,5060	0,4949	0,0000	0,7857	0,6071	0,7679	0,7857	0,5000	0,2500
PC DO B	0,4933	0,4929	0,6036	0,5156	0,6161	0,4762	0,4745	0,7857	0,0000	0,6071	0,7500	0,7500	0,4286	0,2143
PSOL	0,3996	0,3875	0,4571	0,3750	0,5089	0,3750	0,3827	0,6071	0,6071	0,0000	0,5536	0,5536	0,4643	0,0714
PRB	0,4687	0,4688	0,5875	0,5033	0,5759	0,4673	0,4515	0,7679	0,7500	0,5536	0,6786	0,7321	0,4286	0,2500
PSB	0,4888	0,4759	0,5982	0,5190	0,5938	0,4732	0,4566	0,7857	0,7500	0,5536	0,7321	0,7143	0,4286	0,2500
PSC	0,3437	0,3000	0,3393	0,3147	0,3929	0,3095	0,3061	0,5000	0,4286	0,4643	0,4286	0,4286	0,0000	0,0714
PV	0,1562	0,1518	0,2071	0,1964	0,1875	0,1667	0,1327	0,2500	0,2143	0,0714	0,2500	0,2500	0,0714	0,0000

Fonte: Elaborado pelos autores.

As ocorrências de índice de afinidade igual ou superior a 0,65 no caso das relações entre os partidos é muito menor, já que leva em conta a participação de todos os seus parlamentares filiados. Na coleta realizada, no ano de 2007, o maior índice de afinidade encontrado foi de 0,5294, entre os membros do PC do B e do PP. No ano de 2008, a maior afinidade foi encontrada na relação entre os próprios membros do PSB, e entre partidos a maior afinidade foi entre o PCdoB e o PSB. Referente ao ano de 2009, mais uma vez a afinidade nas votações entre os membros do PSB foi a maior com 0,5, e entre partidos, foi de 0,4775, entre o PTB e o PSB. Somente na matriz referente ao ano de 2010 foram calculados índices de afinidade maiores que 0,65, sendo que estas ocorrências encontram-se destacadas na figura 8.

A identificação da maior afinidade entre os partidos percebida no ano de 2010 pode ser destacada também pela obtenção da média das afinidades calculadas em cada um dos anos, sendo que no ano de 2007 a média foi de 0,2331, em 2008 foi 0,2584, em 2009 o cálculo indicou o menor índice, 0,2141 e em 2010 o maior índice, 0,4175.

Outra análise que pode ser proposta é a relação de um partido com os demais, sendo que neste caso, optou-se pelo Partido dos Trabalhadores por ser o principal partido do governo executivo, para que se possa estabelecer uma matriz das afinidades entre os parlamentares de cada partido ao longo dos anos, permitindo, assim, percepção da flutuação das afinidades de cada partido e, também a cada ano, conforme apresentado no figura 9.

Figura 9 - Índice de afinidade entre partidos em relação ao Partido dos Trabalhadores (PT) nas votações abertas de 2007 a 2010

ANO\PART	DEM	PMDB	PT	PSDB	PR	PDT	PTB	PP	PC DO B	PSOL	PRB	PFL	PSB	PSC	PV	S/PART	Média
2007	0,1551	0,3717	0,4545	0,2484	0,3667	0,3157	0,2849	0,5049	0,4436	0,1520	0,3272	0,0771	0,4330	0,0000	0,0000	0,0000	0,2584
2007	0,2404	0,3247	0,4244	0,2800	0,4342	0,2441	0,3410	0,3369	0,5108	0,4015	0,2877	0,0000	0,5138	0,2569	0,0000	0,0462	0,2902
2009	0,2988	0,2658	0,2951	0,2928	0,3014	0,2193	0,3499	0,3041	0,3017	0,1791	0,3053	0,0000	0,4153	0,1875	0,0709	0,0331	0,2388
2010	0,3855	0,3873	0,4659	0,4074	0,4786	0,3857	0,3719	0,6179	0,6036	0,4571	0,5875	0,0000	0,5982	0,3393	0,2071	0,0000	0,3933
Média	0,2700	0,3374	0,4100	0,3072	0,3952	0,2912	0,3369	0,4410	0,4649	0,2974	0,3769	0,0193	0,4901	0,1959	0,0695	0,0198	0,2952

Fonte: Elaborado pelos autores.

Percebe-se que quando analisados em seu conjunto, pode-se ter uma percepção mais clara de aspectos como coesão partidária, ou seja, o índice de afinidade entre as decisões tomadas pelos membros de um mesmo partido e também a proximidade que cada um dos partidos apresenta entre si.

A partir de dados como estes, pode-se elaborar diferentes visualizações, e ainda, acompanhar a coerência entre o que é apresentado nos discursos como linha de conduta para tomada de decisão e a realidade das votações de cada um dos parlamentares e de seus respectivos partidos.

4 CONCLUSÕES

A descrição do processo de obtenção dos dados sobre as votações abertas demonstram que a escassez na disponibilização de dados primários das votações do senado brasileiro reduz as alternativas para construção de visualizações das votações dos parlamentares; bem como a falta de estrutura e formato destes dados no sítio oficial dificulta a recuperação e o tratamento da informação. Fica claro a não consideração dos oito princípios de dado aberto na apresentação destes dados pelo Senado até o momento, o que facilitaria uso diversificado destas informações possibilitando a construção de aplicações para estes dados pela própria sociedade.

A partir do acesso a dados e da elaboração de visualizações destes dados como os obtidos sobre cada uma das votações dos parlamentares, podem ser construídas análises, como a identificação de *clusters* partidários, da real posição de cada representante em relação aos demais e aos partidos, bem como a coerência de cada um de nossos representantes nas votações durante seus mandatos.

Não podemos associar nenhum *cluster* partidário no período de 2007-2010 a partir dos dados coletados, nem entre os partidos aliados

e os opositores. Também não é possível afirmar a existência de grupos de senadores dentro de partidos. Não há um padrão a partir dos dados coletados que identifique a existência de dois grupos em oposição (de senadores da oposição e governo) no senado brasileiro. Considera-se que os demais senadores, que participaram do colegiado entre 2007-2010 e estão ausentes no grafo por falta de afinidade com todos os colegiados em mais de 65%, tiveram uma postura dispersa nos votos abertos. ou seja, em nenhum momento são parte de um grupo fixo de senadores.

No grafo do ano de 2010 (Figura 4), a quantidade de associações aumenta devido à articulação da oposição e governo em prol da aprovação (ou veto) de matérias. Entretanto, é importante a ressalva que, no ano de 2010, o número de matérias em votação capturadas no sítio do Senado Federal foi o menos expressivo em comparação aos outros anos no período escolhido – 28 matérias. Todavia, este período também possui votações abertas de matérias importantes para a sociedade como a inelegibilidade de candidatos visando à proteção da probidade administrativa pública; a criação do Fundo Social (FS); a cessão onerosa ao exercício de atividades da PETROBRAS na pesquisa e lavra de petróleo, de gás natural e hidrocarbonetos fluidos; a prorrogação de benefícios para a Zona Franca de Manaus; a permanência atemporal do Fundo de Combate e Erradicação da Pobreza; a nova redação que impede a utilização da aposentadoria dos magistrados como medida disciplinar e a permissão à perda de cargo; e o aumento para cento e oitenta dias na duração do período da licença à gestante, muitas delas com forte apelo popular.

O acesso a dados e as muitas possibilidades de construção de visualização para estes dados, é elemento chave para construção de uma participação cidadã e para a concretização de uma relação mais transparente entre o legislativo e a sociedade.

Artigo recebido em 26/06/2012 e aceito para publicação em 17/03/2013

VIEWING BY PARLIAMENTARY AFFINITIES BETWEEN DATA VOTES IN THE SENATE BRAZILIAN

ABSTRACT: *The expansion of citizen participation in the public sphere depends directly of access information on the performance of elected representatives, especially with regard to their voting decisions during their terms. The Information Science can assist in this process, proposing and evaluating models of access to such information that may be obtained from data to be provided by the official websites from legislative, in levels federal, state and municipal. It is proposed in this paper to analyze the process of collecting and using of data about votes of senators in order to promote the use of this model in other spheres. From analyzed data a affinity matrix was developed to identifying the relationship between each of the parliamentary with others, based on the similarities of the decisions taken in all the open votes. It was also analyzed the development of initial visualizations and the extension of the scope of the search through the application of data obtained in all the affinities between parliamentarians and to obtain an average affinity between parties, enabling new dimensions of analysis to the data collected. The preparation of the full matrix of relations of affinity among parliamentarians can provide a new horizon of possibilities for developing new forms of visualization and analysis, increasing the visibility of parliamentary actions with society.*

Keywords: *Open Data. Data Visualization. Public Transparency.*

REFERÊNCIAS

BERNERS-LEE, Tim. On the next web. TED Conference 2009. Disponível em: <http://www.ted.com/index.php/talks/tim_berniers_lee_on_the_next_web.html>. Acesso em: 28 set. 2011.

DIESTEL, R. **Graph theory**. New York: Springer-Verlag. 2000.

EUA. The Obama Administration's Commitment to Open Government: status report. Disponível em: <http://www.whitehouse.gov/sites/default/files/opengov_report.pdf>. Acessado em: 29 jan. 2011.

LOCKE, Stephen. **Graph Theory**. Department of Mathematical Sciences at Florida Atlantic University: Boca Raton, FL. 2000. Disponível em: <<http://math.fau.edu/locke/GRAPHTHE.HTM>>. Acesso em: 25 set. 2011.

ODEWAHN, Andrew. Visualizing the U.S. Senate Social Graph (1991-2009). In: STEELE, Julie; ILIINSK Noah. **Beautiful Visualization: looking at data through the eyes of experts**. O'Reilly: Sebastopol, CA, 2010.

OGD - Open Government Data. 8 Principles of Open Government Data. 2007. Disponível em: <<http://www.opengovdata.org/home/8principles>>. Acesso em: 22 set. 2011.

RODRIGUES, F. A. ; SANTANA, R. C. G. . Visualizando grafos sociais do Senado brasileiro (2007-2010) através de ferramentas de Tecnologia da Informação e Comunicação. In: ENCONTRO INTERNACIONAL DE INFORMAÇÃO, CONHECIMENTO E AÇÃO (EIICA), 7., 2011, Marília,SP. **Resumos...** Marília,SP : UNESP, 2011. p. 27-28.

SANTOS, Plácida Leopoldina Ventura Amorim da Costa; ALVES, Rachel Cristina Vesú. Metadados e web semântica para estruturação da web 2.0 e web 3.0. **DataGramZero**, v. 10, n. 6, nov./dez. 2009. Disponível em: <http://www.dgz.org.br/dez09/Art_04.htm>. Acesso em: 28 jun. 2011.

SENSO, J. A.; ROSA PIÑERO, A. de la. El concepto de metadato. Algo más que descripción de recursos electrónicos. **Ciência da Informação**, Brasília, v. 32, n. 2, p. 95-106, maio/ago. 2003. Disponível em: <<http://revista.ibict.br/ciinf/index.php/ciinf/article/view/126/107>>. Acesso em: 29 jun. 2011.

WILSON, R. **Introduction to graph theory**. 4. ed. Harlow, UK: Longman. 1996.

WILSON, Chris. The Senate Social Network. **Slate Magazine**, Washington, abr. 2009. Disponível em: <<http://www.slate.com/id/2217204/>>. Acesso em: 01 jun. 2011

