



*Empoderando vidas.
Fortalecendo nações.*

**Relatório referente ao Projeto BRA/07/004 –
Democratização de Informações no Processo de
Elaboração Normativa**

**Produto 2 - Relatório contendo proposta de solução para
gerar relatório agregado que consolide e organize as
contribuições recebidas através ferramentas e aplicações
de consulta pública**

Diego Rabatone Oliveira

Brasil, Abril/2015, v1

Produto 2 - Relatório contendo proposta de solução para gerar relatório agregado que consolide e organize as contribuições recebidas através ferramentas e aplicações de consulta pública

Contrato n. 2015/000042-00

Valor do produto: R\$6.200,00 (seis mil e duzentos reais)

Data da entrega: 08/03/2015

Nome do Consultor: Diego Rabatone Oliveira

Supervisora: Sabrina Durigon Marques

Abril/2015, v1

Diego Rabatone Oliveira

Produto 2 - Relatório contendo proposta de solução para gerar relatório agregado que consolide e organize as contribuições recebidas através ferramentas e aplicações de consulta pública/ Diego Rabatone Oliveira. – Brasil, Abril/2015, v1-

23 p. : il. (algumas color.) ; 30 cm.

Supervisora: Sabrina Durigon Marques

Relatório técnico – Secretaria de Assuntos Legislativos do
Ministério da Justiça (SAL/MJ), Abril/2015, v1.

1. Consolidação. 2. Relatório. 3. Organização. 4. Métodos Ágeis. I. Sabrina Durigon Marques. II. Secretaria de Assuntos Legislativos. III. Ministério da Justiça. IV. Produto 2 - Relatório contendo proposta de solução para gerar relatório agregado que consolide e organize as contribuições recebidas através ferramentas e aplicações de consulta pública.



Este trabalho está licenciado com uma Licença [Creative Commons - Atribuição 4.0 Internaci-](https://creativecommons.org/licenses/by/4.0/)

onal.



Resumo

Este produto objetiva primeiramente a proposição de soluções para geração de relatório agregado que organize e consolide as contribuições recebidas pelas ferramentas e aplicações de consulta pública. Além disso, também será complementada a proposta de organização das atividades de desenvolvimento, considerando a utilização de métodos ágeis.

Palavras-chaves: consolidação, relatório, organização, Métodos Ágeis.



Lista de ilustrações

Figura 1 – Diagrama do fluxo de um processo de análise de sentimentos	10
Figura 2 – Diagrama do Fluxo de Desenvolvimento	14
Figura 3 – Opções da ferramenta Hypotes.is ao selecionar um texto	16
Figura 4 – Destaques sobrepostos nos textos	17
Figura 5 – Lista de trechos de texto destacados	18
Figura 6 – Lista de trechos de texto destacados	19

Glossário

branch é uma ramificação na árvore de desenvolvimento do projeto no sistema de versionamento git. Para mais informações ver: <<http://git-scm.com/book/pt-br/v1/Ramificaç~ao-Branching-no-Git-O-que-é-um-Branch>>. 11, 13

deploy ação de implementar um serviço, ou atualização de um serviço, para que o mesmo possa ser utilizado pelos usuários. 11

fork é uma cópia do projeto que tem seu desenvolvimento acontecendo de forma independente da árvore de desenvolvimento do projeto original. 11

hotfix é o desenvolvimento emergencial para a resolução de algum problema do sistema que está em produção. HotFixes possuem prioridade máxima e ocorrem fora do ciclo padrão de gestão do projeto. 11

issue Uma tarefa a ser realizada. No escopo deste projeto deve-se buscar a formulação de tarefas minimalistas. 9–11, 13

merge é a ação de agregar propostas de modificação de diferentes *branches* ou *forks*. 13

milestone Conjunto de tarefas (*issues*) com prazo de finalização, neste projeto está sendo utilizado como organizador dos *Sprints*. 10

sprint é a unidade básica de desenvolvimento, no caso deste projeto é o trabalho (a ser) realizado durante uma semana. 10, 13

synset é um conjunto de sinônimos cognitivos, ou seja, um conjunto de verbos que expressam um mesmo conceito. 18

toolkit é um conjunto de ferramentas para realizar diferentes tarefas dentro de um mesmo escopo de atividade ou trabalho. 19

Natural Language Processing termo em inglês para Processamento de Linguagem Natural. 19

Processamento de Linguagem Natual é uma subárea da inteligência artificial e da linguística que estuda os problemas da geração e compreensão automática de linguas humanas naturais. 19



pull-request requisição realizada com propostas de modificação no código fonte do repositório principal do projeto. [10](#)



Siglas

NLP *Natural Language Processing*. 19, *Glossário: Natural Language Processing*

PLN Processamento de Linguagem Natural. 19, *Glossário: Processamento de Linguagem Natural*

PR *pull-request*. 10, 11, 13, *Glossário: pull-request*

SAL/MJ Secretaria de Assuntos Legislativos do Ministério da Justiça. 7

Sumário

	Lista de ilustrações	4
	Glossário	5
	Siglas	7
1	INTRODUÇÃO	9
1.1	Objetivo	9
1.2	Análise de Sentimentos	10
2	DESENVOLVIMENTO	11
2.1	Metodologia de Trabalho	11
2.1.1	Ferramentas de planejamento	11
2.1.2	Rotina de Planejamento e Acompanhamento	12
2.1.3	Processo e fluxo de desenvolvimento	13
2.1.4	Procedimento de integração de um <i>Pull-Request</i>	15
2.2	Propostas para Sistematização de Contribuições	16
2.2.1	Anotações	16
2.2.2	Análise de Sentimentos	20
2.2.2.1	WordNet.Br 1.0 e SentiLex-PT	20
2.2.2.2	Apache OpenNLP	21
2.2.2.3	Criando seu classificador com TextBlob	21
3	CONCLUSÃO	22
	Referências	23

1 Introdução

As manifestações ocorridas em Junho de 2013 trouxeram definitivamente para o debate público uma crise de representação política do atual sistema, concretizada principalmente pelo desagrado da população para com os serviços públicos, sejam eles transporte, educação, saúde, etc (SANTANA; SILVA, 2013). Conforme pontua Svab (2014):

A crescente adoção de mecanismos de participação pela Internet trouxe consigo esperança sobre a possibilidade de superação da dicotomia *participação versus representação*, além de também significar incremento no processo de transparência e abertura do governo.

Em 2007 a Secretaria de Assuntos Legislativos do Ministério da Justiça (SAL/MJ) iniciou o projeto Pensando o Direito objetivando promover uma democratização do processo de elaboração legislativa no Brasil, mas é agora, com as consultas públicas do Marco Civil da Internet¹ e de Proteção de Dados Pessoais² que esta iniciativa começa de fato a atingir um público muito mais amplo.

Com essa nova abrangência, identificam-se novos desafios, dentre eles “a transparência e clareza na sistematização das discussões promovidas;” (SVAB, 2014). Este, assim como outros, é um desafio para o qual ainda não se tem uma solução consolidada, principalmente se considerado o fato de que existem diversas metodologias de participação, e cada uma delas demanda um ou mais processos de sistematização.

1.1 Objetivo

Considerando o contexto exposto anteriormente, o presente relatório objetiva propor soluções para permitir uma sistematização mais automatizada, pública e transparente das consultas públicas em vigor coordenadas pela SAL/MJ, a saber Regulamentação do Marco Civil da Internet e Anteprojeto de Lei de Proteção de Dados Pessoais.

Além disso, também faz parte dos objetivos do presente trabalho propor uma metodologia de organização do desenvolvimento das ferramentas de participação com a finalidade de tornar o processo mais organizado, transparente, público e que permita a

¹ Consulta Pública da Regulamentação do Marco Civil da Internet - <<http://participacao.mj.gov.br/marcocivil/>> - Acessado em 08/04/2015

² Consulta Pública do Anteprojeto de Lei de Proteção de Dados Pessoais - <<http://participacao.mj.gov.br/dadospessoais/>> - Acessado em 08/04/2015

colaboração de novos agentes, sejam eles funcionários da instituição, consultores externos ou mesmo integrantes da sociedade civil que desejem contribuir.

1.2 Análise de Sentimentos

Com a necessidade de se avaliar massivas quantidades de informações, advindas de um grande número de usuários e representando suas opiniões, surgiu uma nova área de pesquisa denominada mineração de opinião, ou análise de sentimentos, cuja finalidade é extrair de forma resumida e sintética a opinião (sentimento) dos usuários com sobre a entidade à qual o usuário está se referindo (SILVA, 2005 apud CRUVINEL, 2014).

As tecnologias para mineração de opiniões são muito mais desenvolvidas, diversas e disseminadas para conteúdos de língua inglesa, mas já existem diversos recursos disponíveis para a realização de tais análises em língua portuguesa.

Existem diversas técnicas diferentes para se realizar uma “análise de sentimento” de uma opinião, podendo ser o resultado da análise uma simples classificação binária - *positivo* x *negativo*, por exemplo; uma classificação em mais de uma categoria - *positivo* x *neutro* x *negativo*; ou mesmo um resultado numa escala contínua. A mais comum das análises é a que oferece uma saída binária - *positivo* x *negativo*; *a favor* x *contra* (PANG; LEE, 2008).

De forma geral, a Figura 1 resume o processo de análise de sentimentos, e na subseção 2.2.2 estão expostas algumas ferramentas que podem ser utilizadas em algumas dessas etapas para análise de texto em português.

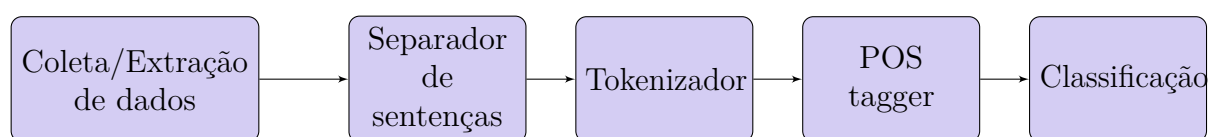


Figura 1: Diagrama do fluxo de um processo de análise de sentimentos

Fonte: Autoria Própria

De forma resumida, o diagrama acima mostra que primeiro temos a coleta dos dados, em seguida o há a separação das sentenças a serem analisadas. Cada sentença passa pelo *Tokenizador*, que separa as palavras em tokens individuais, para que na etapa seguinte a classe gramatical de cada token seja analisada, considerando seu contexto no texto. Por fim há o Classificador, que define o “sentimento” das sentenças.

2 Desenvolvimento

Neste capítulo estão as propostas referentes a este produto. Primeiramente referentes à organização do processo de desenvolvimento e posteriormente as propostas referentes à sistematização das contribuições dos usuários nas consultas públicas.

2.1 Metodologia de Trabalho

Nesta seção serão expostos os procedimentos relativos à metodologia de trabalho da equipe de desenvolvimento.

2.1.1 Ferramentas de planejamento

O planejamento das atividades da equipe de desenvolvimento será baseado em tarefas (*issues*), que serão criadas na área de gestão de projeto do repositório “**participacao-sitebase**”¹. Cada *issue* deve ser minimalista, ou seja, deve-se buscar a criação de tarefas que sejam as mais simples possíveis.

A visualização e organização dessas tarefas se dará por meio da ferramenta chamada HuBoard², que é uma interface no estilo KanBan³ (SUGIMORI et al., 1977), para gerenciamento de *issues* de repositórios no GitHub⁴.

No *HuBoard* tem-se 4 estados para as tarefas:

Backlog - tarefas criadas e ainda não alocadas;

Pronto para Trabalhar - tarefas que devem ser trabalhadas no presente ciclo;

Trabalhando - tarefas que estão sendo desenvolvidas no momento; e

Feito - tarefas que já foram finalizadas mas ainda não avaliadas pela equipe.

¹ **participacao-sitebase** é o repositório principal do projeto e está hospedado no endereço: <<https://github.com/pensandoodireito/participacao-sitebase/issues>> - Acessado em 02/04/2015

² HuBoard - <<https://huboard.com>> - Acessado em 02/04/2015

³ O “sistema KanBan” é a materialização do sistema *just-in-time* e foi desenvolvido por Taiichi Ohno, enquanto Vice-Presidente da Toyota Motors na década de 1950. Ele é um sistema de controle de fluxo de atividades que se utiliza de cartões. - <<http://en.wikipedia.org/wiki/Kanban>> - Acessado em: 02/04/2015

⁴ GitHub - repositório de projetos - <<http://github.com>> - Acessado em 02/04/2015

Além do *GitHub* e do *HuBoard*, a equipe também utilizará o serviço chamado *slack*⁵ como ferramenta de comunicação interna. Para melhorar a comunicação o *HuBoard* também será integrado ao *slack*, de maneira que qualquer mudança ocorrida no *HuBoard* seja registrada no *slack*.

2.1.2 Rotina de Planejamento e Acompanhamento

O primeiro pressuposto da presente metodologia é que serão estabelecidos ciclos de desenvolvimento de uma semana. Para tanto, foi definido que às segundas-feiras serão realizadas as reuniões semanais de alinhamento e planejamento.

Nesta reunião, a primeira tarefa a ser realizada é a avaliação das atividades desenvolvidas desde a reunião de planejamento anterior. Essa avaliação se dá inicialmente pela verificação de cada uma das *issues* presentes no estado **Feito** do *HuBoard*. Para verificar a implementação de uma *issue*, a equipe acompanhará a pessoa no papel de **integrador** implementando o *pull-request* (PR) relativo àquela *issue*, validando o resultado final e em seguida arquivando a mesma.

Após a verificação de todas as *issues* no estado **Feito**, a equipe verificará se alguma *issue* prevista para aquele *sprint* não foi implementada. Esta checagem ocorrerá por meio da verificação da *milestone*, utilizando os filtros do *HuBoard*. Caso todas as *issues* tenham sido implementadas, o *milestone* pode ser encerrado, caso contrário, deve-se avaliar o motivo de não se ter atingido o objetivo planejado e as *issues* não implementadas voltam para a etapa de planejamento.

A seguir procede-se para a etapa de planejamento, que compreende a criação de novas *issues* em decorrência do trabalho que fora desenvolvido e também de demandas “externas” e de “clientes”. Após a criação e avaliação de todas as *issues*, a equipe deve decidir quais serão trabalhadas no próximo *sprint* semanal, alocando-as ao *milestone* respectivo.

Por fim, espera-se que a equipe faça um rápido acompanhamento diário das atividades junto ao gestor da equipe. Este acompanhamento tem a finalidade maior de identificar alguma atividade que está consumindo um tempo excessivo e, eventualmente, reduzir sua prioridade em detrimento de outras tarefas a serem realizadas durante o *sprint*.

⁵ Slack - <<http://slack.com>> - Acessado em 04/04/2015

2.1.3 Processo e fluxo de desenvolvimento

Para o processo e fluxo de desenvolvimento resume-se aqui, na Figura 2 o proposto em Oliveira (2015).

Do diagrama, destacam-se as seguintes regras:

1. Os desenvolvedores trabalharão em seus próprios *forks*;
2. todo desenvolvimento deve ocorrer em *branches* criada para cada *issue*;
3. a criação de uma nova *branch* deve ser realizada a partir da *branch master*.

```
1 git checkout -b issueXX master
```

Código 2.1: Sugestão de comando para criar nova branch baseada na master

4. ao encerrar o desenvolvimento em uma *branch* o desenvolvedor deve fazer um **PR** para a *branch master* do repositório principal.
5. os desenvolvedores devem sempre manter sua *branch master* local sincronizada com a *branch master* do repositório principal.

Além disso, quando for necessário realizar algum desenvolvimento emergencial para resolver problemas no sistema que está em produção, será utilizado o expediente de *hotfix*. Para tanto, o desenvolvedor responsável por implementar a solução deverá criar uma nova *branch* baseada na *branch production* do repositório principal, desenvolver a solução nesta *branch* e, em seguida, realizar um **PR** para a *branch production* do repositório principal e outro **PR** para a *branch master* do repositório principal. Em seguida o responsável pelo projeto deve ser informado para que possa ocorrer a integração dos **PRs** e o *deploy* no site em produção.

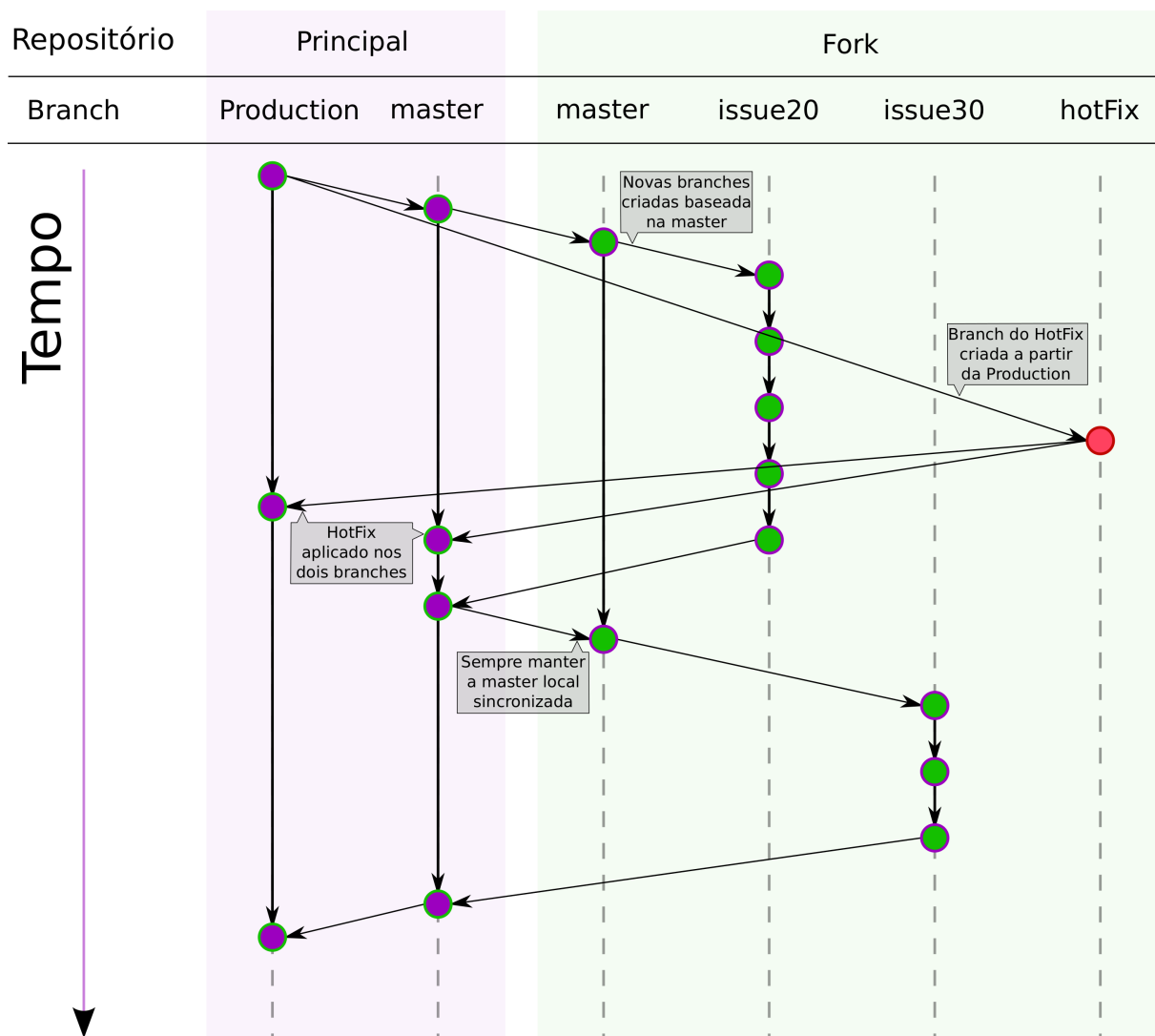


Figura 2: Diagrama do Fluxo de Desenvolvimento

Fonte: Autoria Própria

2.1.4 Procedimento de integração de um *Pull-Request*

A integração de códigos desenvolvidos durante o *sprint*, deve ser realizada no *branch master* do projeto, para posteriormente ser transferida para o *branch production*, que é o que estará em produção. É importante que o procedimento de integração não seja realizado utilizando a ferramenta de *merge* da interface do GitHub, mas sim que seja realizado, de acordo com o que será apresentado abaixo, numa cópia local no computador da pessoa responsável pela integração.

O primeiro passo para integração de um *PR* é criar um novo *branch* localmente, recomenda-se que o nome do *branch* seja o nome da *issue* a ser resolvida.

```
1 git checkout -b issueXX master
```

Código 2.2: Cria uma nova branch chamada issueXX baseada na branch master

O passo seguinte é incorporar as mudanças propostas no *PR*. As mudanças estão no repositório do(a) desenvolvedor(a) que originou o *PR*, na *branch* de nome *issueXX*.

```
1 git pull https://github.com/<user>/<nome-do-repositorio>.git issueXX
```

Código 2.3: Recebendo as modificações do branch a ser incorporado

Pode ser necessário realizar, manualmente, o trabalho de *merge*, caso seja, é importante que a pessoa responsável pelo desenvolvimento do código que originou o *PR* esteja junto a quem está integrando os códigos para auxiliar na resolução do conflito.

Neste momento, o projeto local está com a modificação proposta implementada. Deve-se observar no *PR* que está no GitHub se existem observações relativas a mudanças que devem ser feitas no painel de administração do wordpress para completar a resolução da *issue*. Caso existam, as recomendações devem ser seguidas e posteriormente registradas no wiki de documentação (<<https://github.com/pensandoodireito/participacao-sitebase/wiki>>) do projeto, na página relativa ao *sprint* sendo avaliado.

Em seguida a equipe deve validar as modificações, verificando o resultado final. Caso decida-se que a *issue* está resolvida, agora iremos integrar as mudanças na *branch master* do projeto.

```
1 git checkout master  
git merge --no-ff issue37  
3 git push origin master
```

Código 2.4: Enviando as modificações aceitas para o branch master

2.2 Propostas para Sistematização de Contribuições

Nesta seção será feita uma proposta inicial para auxiliar no processo de sistematização dos comentários e contribuições feitas nas consultas públicas, e também serão apresentados alguns recursos que futuramente podem ser trabalhados para a criação de ferramentas que também auxiliem, de outros pontos de vista, o processo de sistematização.

2.2.1 Anotações

Conforme acordado, a primeira solução a ser implementada no sentido de auxiliar no processo de sistematização das contribuições será baseado na ferramenta *Hypothes.is*⁶.

De forma geral, o que esta ferramenta propõe é adicionar uma nova camada a um website que permita a realização de anotações em quaisquer trechos de texto. Uma analogia possível seria a utilização de um papel vegetal sobre um documento para que se possa realizar anotações no documento sem modificá-lo.

Ao selecionar um trecho de texto, a ferramenta apresenta duas opções, conforme apresentado na Figura 3. O usuário pode adicionar um comentário (*note*) relativo ao trecho selecionado ou apenas destacar o trecho de texto (*highlight*).

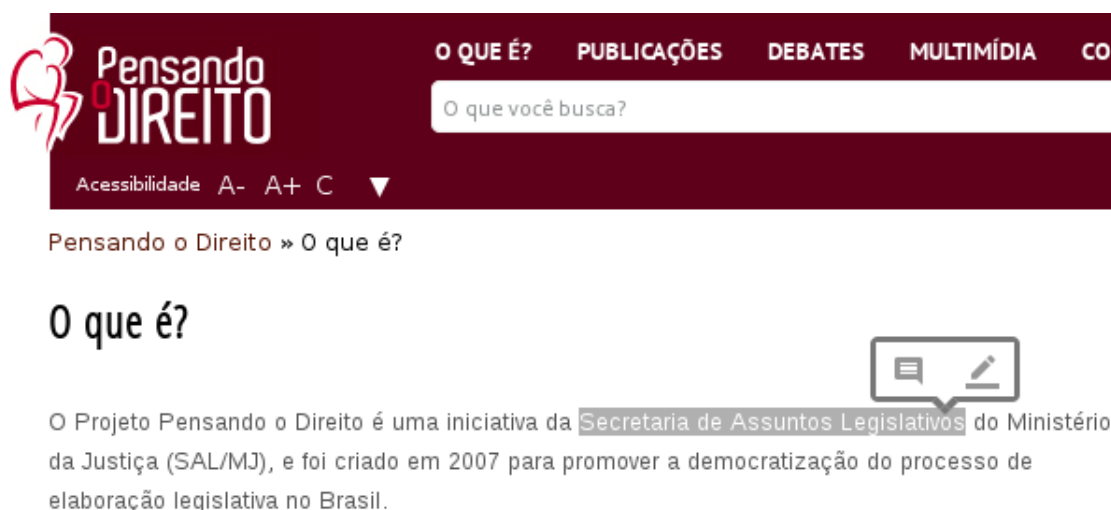


Figura 3: Opções da ferramenta Hypotes.is ao selecionar um texto

Fonte: Autoria Própria

⁶ *Hypothes.is* - <<http://hypothes.is>> - “Nossa missão é trazer uma nova camada para a web. Utilize o *Hypothes.is* para debater, colaborar, organizar sua pesquisa ou fazer anotações pessoais.” (tradução própria) - Acessado em 10/04/2015

Na Figura 4, podemos observar um parágrafo no qual foi utilizado o recurso de destaque (*highlight*). Nele pode-se perceber alguns trechos que receberam mais de um destaque e, por conta disto, apresentam-se grifados numa coloração mais escura decorrente, desta sobreposição de grifos. O mesmo valeria caso os trechos em questão tivessem sido comentados, eles também receberiam o grifo de destaque, que pode ser sobreposto escuracendo o trecho. Assim, trechos com grifos mais escuros representam trechos mais comentados, ou seja, trechos que receberam mais atenção.

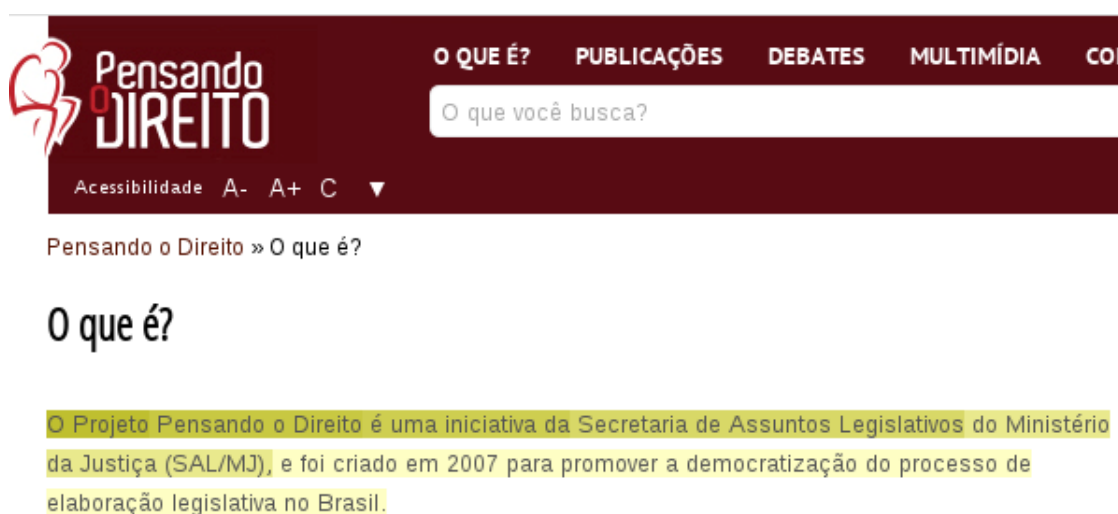


Figura 4: Destaques sobrepostos nos textos

Fonte: Autoria Própria

Na Figura 5 podemos observar a barra lateral da ferramenta. Nela estão os trechos que foram selecionados e destacados, assim como também estão trechos que foram comentados. Ainda existe a opção para buscar comentários que tenham determinados termos e a opção de ordenar os comentários e destaques por ordem de criação - do mais antigo para o mais novo e do mais novo para o mais antigo; além da ordenação por “localização”, na qual os comentários são mostrados ordenados de acordo com a ordem de ocorrência do trecho ao qual o comentário se refere.

BRASIL
Acesso à informação
Participar
Search...
Sign in

O QUE É?
PUBLICAÇÕES
DEBATES
MULTIMÍDIA
CONTATO

O que você busca?
Buscar

Acessibilidade
A-
A+
C

Pensando o Direito » O que é?

O que é?

O Projeto Pensando o Direito é uma iniciativa da Secretaria de Assuntos Legislativos do Ministério da Justiça (SAL/MJ), e foi criado em 2007 para promover a democratização do processo de elaboração legislativa no Brasil.

A partir do lançamento de editais para a contratação de equipes de pesquisa, o Projeto mobiliza setores importantes da sociedade – Academia, instituições de pesquisa, ONG's entre outros – para a realização de estudos em temas de interesse da Secretaria.

Para atingir o propósito de conferir com maior efetividade às normas perante a realidade social, são privilegiadas pesquisas aplicadas, de caráter empírico, com o emprego de estratégias qualitativas e quantitativas que informem e fortaleçam o debate no processo de produção de leis e demais atos normativos.

Além de qualificar a atuação da SAL/MJ, órgão que tem por objetivo institucional a preservação da ordem jurídica, essa iniciativa busca fortalecer as instituições democráticas voltadas à construção de um ordenamento normativo mais próximo da realidade da população brasileira. A aproximação entre Academia e Parlamento permite que a discussão política seja associada a argumentos, dados e informações embasados em pesquisas com comprovação empírica, propiciando a construção de

Showing 1 selected annotation. Clear selection.
Sorted by location

O Projeto

O Projeto Pensando o Direito é uma iniciativa da Secretaria de Assuntos Legislativos do Ministério da Justiça (SAL/MJ), e foi criado em 2007 para promover a democratização do processo de elaboração legislativa no Brasil.

O Projeto Pensando o Direito é uma iniciativa da Secretaria de Assuntos Legislativos

O Projeto Pensando o Direito é uma iniciativa da

O Projeto Pensando o Direito

O Projeto Pensando o Direito é uma iniciativa da Secretaria de Assuntos Legislativos do Ministério da Justiça (SAL/MJ),

Tweets de @pr

NOTÍCIA

22/04/2015
Volume 53 – A 1 da sociedade ci

10/04/2015
Volume 52 – Vi Práticas Instituc

16/03/2015
Publicado result consultores

15/01/2015
Abertos editais consultores

Empoderando vidas.
Fortalecendo nações.

PND
UD

Página 18

Figura 5: Lista de trechos de texto destacados

Fonte: Autoria Própria

Além da possibilidade de simples destaque, a ferramenta também permite comentários atrelados a trechos de texto selecionados. Estes comentários podem ainda receber *tags*, e sobre eles podem haver outros comentários, permitindo assim algum tipo de debate/discussão. Na Figura 6 podemos observar um comentário relativo a um destaque no texto e as tags marcadas no comentário. Além disso, pode-se perceber também que elementos, como um vídeo, também podem receber comentários.

A partir do lançamento de editais para a contratação de equipes de pesquisa, o Projeto mobiliza setores importantes da sociedade – Academia, instituições de pesquisa, ONG's entre outros – para a realização de estudos em temas de interesse da Secretaria.

Para atingir o propósito de conferir com maior efetividade às normas perante a realidade social, são privilegiadas pesquisas aplicadas, de caráter empírico, com o emprego de estratégias qualitativas e quantitativas que informem e fortaleçam o debate no processo de produção de leis e demais atos normativos.

Além de qualificar a atuação da SAL/MJ, órgão que tem por objetivo institucional a preservação da ordem jurídica, essa iniciativa busca fortalecer as instituições democráticas voltadas à construção de um ordenamento normativo mais próximo da realidade da população brasileira. A aproximação entre Academia e Parlamento permite que a discussão política seja associada a argumentos, dados e informações embasados em pesquisas com comprovação empírica, propiciando a construção de normas mais democráticas e efetivas.

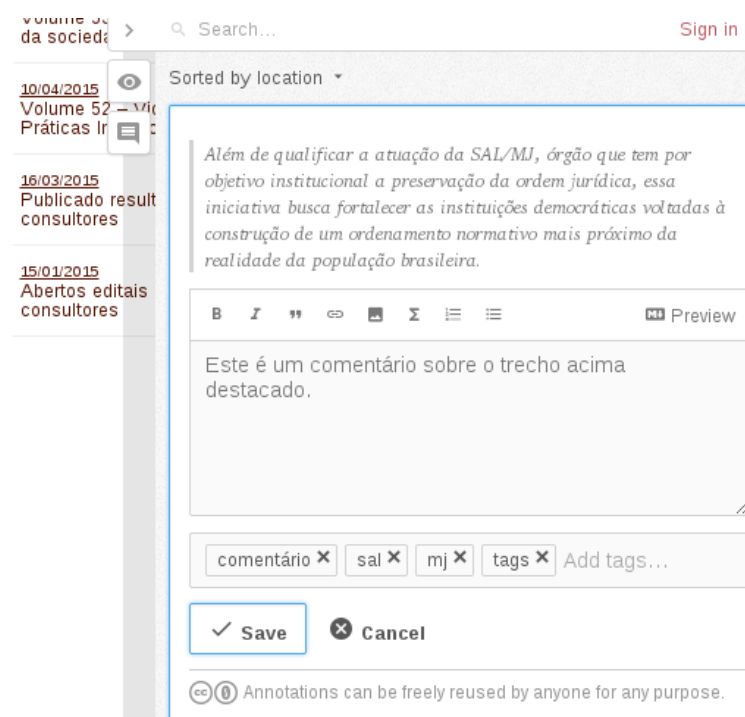
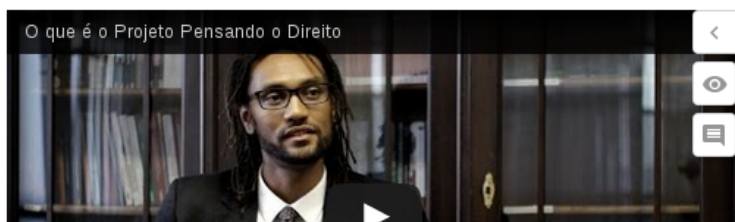


Figura 6: Lista de trechos de texto destacados

Fonte: Autoria Própria

2.2.2 Análise de Sentimentos

Para além da proposta de utilização da ferramenta *Hypotes.is* como auxiliar do processo de sistematização das contribuições, exposta na Seção 2.2.1, iremos agora indicar alguns recursos a serem melhor estudados e avaliados para o desenvolvimento de novas ferramentas que auxiliem no processo de sistematização das contribuições.

2.2.2.1 WordNet.Br 1.0 e SentiLex-PT

A **WordNet.Br 1.0** (SILVA, 2005) é uma base de dados de verbos da língua portuguesa construída de forma alinhada com a versão 2.0 da WordNet de Princeton (UNIVERSITY, 2010; MILLER, 1995), base de verbos na língua inglesa.

Esta versão conta com 5860 verbos em 3713 *synsets*. Um *synset* é um conjunto elementos que são considerados semanticamente equivalentes para o propósito de extração de informações.

Além do WordNet.Br existe ainda o **SentiLex-PT**(CARVALHO; SILVA, 2010; CARVALHO; SILVA, 2015), um léxico de sentimento para o português construído por 7.014 lemas e 82.347 formas flexionadas. Mais especificamente este léxico descreve - entre parênteses número de formas flexionadas:

- 4.799 (16.863) adjetivos;
- 1.081 (1.280) substantivos;
- 489 (29.504) verbos; e
- 666 (34.700) expressões idiomáticas.

Para cada entrada do SentiLex-PT temos como atributos de sentimentos o alvo do sentimento, a polaridade do predicador e o método de atribuição de polaridade.

O SentiLex-PT está disponível para download em seu [site oficial](http://dmir.inesc-id.pt/project/SentiLex-PT_02) (<http://dmir.inesc-id.pt/project/SentiLex-PT_02>), enquanto o WordNet.Br apenas consta como um serviço web para consulta pontual. Seria de grande relevância buscar um diálogo com os responsáveis pelo WordNet.Br no sentido de torná-lo também uma base pública que possa ser incorporada a aplicações de análise de ferramentas, assim como torná-lo um webservice que permita automatizações das consultas.

2.2.2.2 Apache OpenNLP

O Apache OpenNLP⁷ é um *toolkit* livre (*Apache License, Version 2.0*) desenvolvido pela Apache Foundation que utiliza recursos de aprendizado de máquina (*machine learning*) para o processamento de textos em linguagem natural (*Processamento de Linguagem Natual (PLN)* ou *Natural Language Processing (NLP)*). Este *toolkit* realiza a maioria das tarefas tradicionais de sistemas de PLN, tarefas essas que servem de base para a construção de sistemas mais complexos. Vale destacar que nem todas as tarefas estão disponíveis para a língua portuguesa. As tarefas presentes no OpenNLP são:

- Tokenização;
- segmentação de sentenças;
- tagueamento de trechos de discursos;
- extração de entidade nomeada;
- *chunking*;
- parseamento; e
- resolução de co-referência.

2.2.2.3 Criando seu classificador com TextBlob

Uma das técnicas mais comuns para classificação de textos é a chamada Análise Bayesiana. Já existem diversas ferramentas que facilitam sua aplicação, como é o caso da biblioteca *TextBlob*⁸. Neste documento não entraremos em detalhes relativos à sua aplicação, apenas deixamos aqui a indicação de sua utilização, em especial seguindo o tutorial da própria biblioteca disponível em: <http://textblob.readthedocs.org/en/dev/classifiers.html>.

⁷ Apache OpenNLP - <https://opennlp.apache.org/> - Acesso em 15/04/2015

⁸ *TextBlob* - Biblioteca Python para PLN - <http://textblob.readthedocs.org/> - Acessado em 16/04/2015

3 Conclusão

Este documento apresenta a necessidade e a importância do processo de sistematização de contribuições realizadas em ambientes de colaboração, em especial quando se trata da relação entre sociedade civil e governos. Esta ainda é uma área completamente em aberto e que demanda muitas experimentações e estudos.

Existem diversas formas de a sociedade civil trabalhar em parceria com o governo, sendo uma delas a colaboração na construção das ferramentas de participação social. Para tanto, é fundamental que o processo de desenvolvimento de ferramentas esteja bem documentado e seja orientado a colaborações de terceiros. Assim, apresentamos na seção [2.1](#) uma proposta de Metodologia que visa contemplar colaboradores externos e pontuais.

Além disso, como parte do processo de aprendizado e experimentação em participação social, propõe-se aqui a implementação da ferramenta *Hypotes.is* para ser utilizada na sistematização das contribuições realizadas na consulta pública do Marco Civil da Internet e do Proteção de Dados Pessoais. Posteriormente se faz necessária uma avaliação de como se deu a utilização desta ferramenta neste processo.

Por fim, são apresentadas também algumas alternativas de ferramentas e tecnologias que poderiam ser utilizadas em conjunto com o *Hypotes.is* para automatizar o processo de avaliação e sistematização das contribuições.

Referências

- CARVALHO, P.; SILVA, M. J. *SentiLex-PT 02 Website*. 2010. Disponível em: <http://dmir.inesc-id.pt/project/SentiLex-PT_02>. Citado na página 20.
- CARVALHO, P.; SILVA, M. J. Sentilex-pt: Principais características e potencialidades. *Oslo Studies in Language*, v. 7, n. 1, 2015. Citado na página 20.
- CRUVINEL, G. W. F. Análise de sentimentos em redes sociais digitais: Uma aplicação no contexto político. 2014. Disponível em: <https://www.academia.edu/7882935/An%C3%A1lise_de_Sentimentos_em_Redes_Sociais_Digitais_-_Uma_aplica%C3%A7%C3%A3o_no_contexto_pol%C3%ADtico>. Citado na página 10.
- MILLER, G. A. Wordnet: A lexical database for english. *Commun. ACM*, ACM, New York, NY, USA, v. 38, n. 11, p. 39–41, nov. 1995. ISSN 0001-0782. Disponível em: <<http://doi.acm.org/10.1145/219717.219748>>. Citado na página 20.
- OLIVEIRA, D. R. *Produto 1 - Relatório contendo avaliação e propostas de melhorias técnicas para ferramentas e aplicações de consulta pública, incluindo plugins*. [S.l.], 2015. <https://github.com/diraol/PNUD-2015-000042-00/blob/master/produto01.pdf>. Citado na página 13.
- PANG, B.; LEE, L. Opinion mining and sentiment analysis. *Found. Trends Inf. Retr.*, Now Publishers Inc., Hanover, MA, USA, v. 2, n. 1-2, p. 1–135, jan. 2008. ISSN 1554-0669. Disponível em: <<http://dx.doi.org/10.1561/1500000011>>. Citado na página 10.
- SANTANA, B.; SILVA, D. B. Brasil: «No es por 0, 20. Es por los derechos». *Nueva Sociedad*, v. 247, p. 4–15, 2013. Disponível em: <http://www.nuso.org/upload/articulos/3973_1.pdf>. Citado na página 9.
- SILVA, C. S. Bento Carlos Dias da. *WordNet.Br*. 2005. Disponível em: <<http://143.107.183.175:21380/wordnetbr/>>. Citado 2 vezes nas páginas 10 e 20.
- SUGIMORI, Y. et al. Toyota production system and kanban system materialization of just-in-time and respect-for-human system. *The International Journal of Production Research*, Taylor & Francis, v. 15, n. 6, p. 553–564, 1977. Citado na página 11.
- SVAB, H. *Espaço Digital e Participação Real*. Dissertação (Monografia) — Faculdade de Filosofia e Ciências Humanas da Universidade Federal de Minas Gerais, 2014. Programa de Pós-Graduação em Democracia Participativa, República e Movimentos Sociais. Citado na página 9.
- UNIVERSITY, P. *About WordNet*. 2010. Disponível em: <<http://wordnet.princeton.edu/>>. Citado na página 20.