



*Empoderando vidas.
Fortalecendo nações.*

Projeto BRA/12/018 - Desenvolvimento de Metodologias de Articulação e Gestão de Políticas Públicas para Promoção da Democracia Participativa

Produto 04 - Proposta de adaptações e incrementos para a interface do portal federal de participação social e suas ferramentas

com elementos visuais e de usabilidade para mecanismos de priorização de conteúdos e
autorregulação

Renato Fabbri



Secretaria-Geral da Presidência da República

Produto 04 - Proposta de adaptações e incrementos para a interface do portal federal de participação social e suas ferramentas

Contrato n. 2013/000566

Objeto da contratação: Aporte de conhecimentos e tecnologias para especificação de vocabulário e ferramentas assistidas que utilizam processamento de linguagem natural e análise de redes complexas para o conteúdo do portal da participação social.

Valor do produto: R\$ 21,600 (vinte e um mil e seiscentos reais)

Data de entrega: 28 de Agosto de 2014

Nome do consultor: Renato Fabbri

Nome da supervisora: Gabriella Vieira Oliveira Gonçalves



Secretaria-Geral da Presidência da República

Fabbri, Renato

Proposta de adaptações e incrementos para a interface do portal federal de participação social e suas ferramentas: com elementos visuais e de usabilidade para mecanismos de priorização de conteúdos e autorregulação / 2014.

Total de folhas: 17

Supervisora: Gabriella Vieira Oliveira Gonçalves

Secretaria: SNAS

Secretaria-Geral da Presidência da República

Palavras-chave: reconhecimento de padrões, redes complexas, processamento de linguagem natural, participação social.



Esta obra é licenciada sob uma licença Creative Commons - Atribuição-NãoComercial. 4.0 Internacional.



Empoderando vidas.
Fortalecendo nações.

Sumário

1	Introdução	7
1.1	Contexto e importância da consultoria	7
1.2	Contexto e importância do Produto	7
1.2.1	Objetivos	7
1.2.2	Resultados esperados	8
1.2.3	Caráter inovador	8
1.2.4	Aparato em hardware	9
2	Desenvolvimento	9
2.1	Etapas de desenvolvimento anteriores a este produto	10
2.1.1	Sistematização ontológica da participação online	10
2.1.2	Triplificação dos dados do participa.br	10
2.1.3	Levantamento do endpoint SparQL	10
2.1.4	Análises iniciais, modelos	10
2.2	Etapas de desenvolvimento deste produto	10
2.2.1	Reuniões com equipe do participa.br	10
2.2.2	Estudos de aprofundamento e amadurecimento	10
2.2.3	Escrita deste documento	11
2.2.4	Especificação da API de recomendação	11
2.2.5	Implementação do serviço para aquisição dos dados, processamento e entrega em JSON	11
2.2.6	Disponibilização das rotinas no IPython Notebook	11
2.2.7	Proposta de implementações no portal federal de participação social	11
2.3	Justificativa do método	11
2.4	Justificativa das fontes	11
2.5	Confronto entre os resultados esperados e os alcançados	12
3	Usos dos resultados	12
4	Conclusão	13
4.1	Comentários, sugestões, recomendações	13
4.2	Impacto do Produto para a elaboração, gestão e/ou avaliação de políticas públicas de participação social	13
4.3	Como o Produto deverá impactar o público-alvo das políticas públicas a que se refere	14
5	Agradecimentos	14



*Empoderando vidas.
Fortalecendo nações.*

A	Especificação da API de recomendação de recursos do participa.br	17
B	Rotinas de acesso e processamento de dados do participa.br para as recomendações	17
C	Infraestrutura do sistema de recomendações	17
D	Instalação e modificação do sistema de recomendações	17
E	Propostas de implementações na interface do portal federal de participação social	17



*Empoderando vidas.
Fortalecendo nações.*

Resumo

Este documento descreve rotinas de priorização de conteúdo e de autorregulação para o portal federal de participação social.

Palavras-chave: reconhecimento de padrões, redes complexas, processamento de linguagem natural, participação social.



Empoderando vidas.
Fortalecendo nações.

1 Introdução

1.1 Contexto e importância da consultoria

Em confluência com o portal federal de participação social (Participa.br) e o Plano Nacional de Participação Social (PNPS), esta consultoria propõe métodos de classificação e priorização de conteúdo e formas de autorregulação para o portal. O presente produto apresenta “formas de priorização de conteúdo e de autorregulação” implementadas na forma de um sistema de recomendação de recursos para usuários.

1.2 Contexto e importância do Produto

1.2.1 Objetivos

Este produto tem por objetivo principal a disponibilização de um sistema de recomendação de recursos do participa.br para usuários, tanto via critérios personalizados quanto considerando comunidades e linha editorial. Através deste sistema de recomendação, ficam facilitados, até mesmo prontamente disponíveis, diversos processos de autorregulação, de geração de resumos, de geração de relatórios e análises informativas. Objetivos secundários são:

- Disponibilização de uma API HTTP, para uso no participa.br, conforme requisitado pela equipe do Participa.br em diversos itens dos produtos [?].
- A exposição destes algoritmos de recomendação aos visitantes em páginas HTML comuns, para edição execução dos trechos de código utilizados pela plataforma federal de participação social. Feito no IPython Notebook levantado para este trabalho [?].
- Aproveitamento do endpoint SparQL com os dados do participa.br, fortalecendo as tecnologias de dados linkados e web 3.0.
- Em reunião com a consultora Daniela Feitosa, foi delineada a pertinência de um sistema de recomendação de perfis para um ActionItem em andamento para o participa.br [?]. Este produto visa suprir esta necessidade através da API disponibilizada.
- A entrega das tecnologias livres com simplicidade e boa documentação, favorecendo ao máximo o aproveitamento deste trabalho para melhoras, geração de derivados e novas e independentes tecnologias. Isso pode ser observado no repositório git público deste produto [?].
- Compatibilizar ao máximo a entrega deste produto às demandas da equipe do participa.br.
- Cumprir de forma precisa e pertinente a descrição deste quarto produto no Termo de Referência.



Empoderando vidas.
Fortalecendo nações.

1.2.2 Resultados esperados

Em um aspecto mais amplo, o resultado do produto é iniciar um processo aberto de apropriação das análises e mecanismos de autorregulação para o portal federal de participação social.

Como resultados diretos destes resultados, constam:

- a habilitação para uso da API HTTP para recomendação de recursos do participa.br.
- A interface para apreensão e inovação dos algoritmos, não somente exemplificada ou projetada, mas operante e disponível.
- Um plano de implementação para o participa.br, que utiliza a API de recomendação para priorização de conteúdo e autorregulação.
- Transparência absoluta no trabalho, com toda a documentação e código computacional online em um repositório git que contem o histórico de implementação.
- Algoritmos implementados em código de simples leitura, para facilitar a implementação em Ruby ou Javascript, quando houverem recursos maduros para estas linguagens. (No momento, Python possui mais recursos e mais maduros tanto para redes complexas quanto para processamento de linguagem natural, o que justifica o servidor.)
- Habilitação de implementações em andamento, como o plugin de recomendação de perfis [?], para recomendar amigos para visitantes ou participantes.

1.2.3 Caráter inovador

Centralmente, este trabalho é inovador na aplicação de recursos de análise de redes sociais de forma comunitária, entregando as tecnologias e priorizando a reutilização. Este recurso é de vital interesse para a democracia participativa no contexto atual, com as revoluções da internet e redes sociais. Permite, em última instância, que haja uma inteligência para aproveitamento das estruturas sociais, e que esta inteligência seja pública, transparente, minimizando vetores vigilatistas ou turvos.

Há a inovação na arquitetura em software, apresentando traços de web 3.0 como os dados linkados como base de conhecimento e os múltiplos recursos online acessados no funcionamento usual (ao menos endpoint sparql para acesso aos dados, API HTTP Flask para tratar os dados e gerar estatísticas e estruturas de interesse, participa.br para interface e contexto pertinente).

Há inovação na difusão científica. Dados Linkados, Processamento de Linguagem Natural (PLN) e Redes Complexas (RC) são três termos empregados em textos científicos, consistindo de áreas que recebem pesquisas, revistas dedicadas e até carreiras científicas inteiras. Neste trabalho este conteúdo está em português e se presta a orientar contribuições de outras partes interessadas e de curiosos.



Empoderando vidas.
Fortalecendo nações.

Os métodos de recomendação em si não foram confrontados exaustivamente com a literatura, mas possivelmente possui também inovações nos procedimentos e certamente no contexto de implementação, tanto de relevância social quanto de aparato tecnológico.

1.2.4 Aparato em hardware

O sistema de recomendações precisa ser mantido online, e isso implica na manutenção de uma estrutura em hardware e software que extrapola o objeto deste produto. As especificações desta arquitetura de software e hardware estão em um documento escrito pelo consultor a pedido da PR para a SNAS e DITEC [?]. Para fins de pesquisa e de entrega deste produto, o esquema da figura ?? é realizado em máquinas de pesquisa da USP, concedidas para a pesquisa de doutorado do consultor. Tal configuração é razoável pois a pesquisa conflui e é potencializada por esta consultoria, mas está prevista uma infraestrutura própria da PR para estes serviços. Além disso, os usos atuais ainda são moderados, sem causar sobrecarga ao serviços de computação em nuvem da USP.

[Figura/Diagrama do endpoint+serv flask+frontEnd noosfero]

2 Desenvolvimento

O produto é descrito no Termo de Referência desta consultoria assim: “Documento com proposta de adaptações e incrementos para a interface do portal e suas ferramentas que inclua elementos visuais e de usabilidade para mecanismos de priorização de conteúdos e auto-regulação com base nos metadados gerados nativamente pela plataforma do portal e nas análises de PLN e RC produzidas pelas ferramentas assistidas”.

Dada a dimensão do participa.br, tanto da estrutura em software e das práticas participativas, quanto de mobilização humana, as propostas de adaptações para o portal são muitas e estão em diversos produtos deste e de outros consultores. Uma forma especialmente pertinente de realizar este produto, confluyente com o Termo de Referência desta consultoria, com o trabalho dos gestores e comunidade e com os produtos deste e de outros consultores [?], é a entrega de um sistema de recomendação de recursos para o participa.br.

Conforme discutido em reunião com gestores e comunidade do participa.br [1], em termos de priorização de conteúdo, quase todos os casos de interesse para o atual participa.br podem ser considerados sistemas de recomendação. Além disso, os mecanismos de autorregulação são convenientemente calcados na adição de conteúdo e na espontaneidade dos participantes, o que pode novamente ser considerado através de sistemas de recomendação.

Desta forma, foi idealizada a entrega de um sistema de recomendação de recursos do participa.br para seus usuários, comunidades e linha editorial. A subseção a seguir detalha componentes da



Empoderando vidas.
Fortalecendo nações.

versão operante deste sistema de recomendação de recursos. A seção 3 compreende previsões e propostas de integração com o participa.br.

2.1 Etapas de desenvolvimento anteriores a este produto

2.1.1 Sistematização ontológica da participação online

Através de estudos e reuniões presenciais e online, Ontologia de Participação Social (OPS) foi revisada [2] e a Ontologia do Participa.br (OPA) foi feita [3].

2.1.2 Triplificação dos dados do participa.br

Feito um script para triplificar os dados do Participa.br, ou seja, enriquecimento semântico e escrita em RDF dos dados em Postgresql da instância Noosfero do Participa.br [4].

2.1.3 Levantamento do endpoint SparQL

Para uso dos dados triplificados, pode-se recorrer a diversos métodos de leitura e disponibilização. Um método-chave é a disponibilização dos dados rdf (*triple store*) em um *endpoint sparql*. Para os fins de testes e usos leves (avisar consultor), está disponibilizado um endpoint SparQL em servidores da USP para pesquisa do consultor [5].

2.1.4 Análises iniciais, modelos

Análises dos dados do participa.br foram abertas no IPython Notebook, com ênfase no texto produzido e nas redes formadas [6].

2.2 Etapas de desenvolvimento deste produto

2.2.1 Reuniões com equipe do participa.br

Especial agradecimentos ao Ricardo Poppi, Ronald Costa, Enaile Ladanza, Joênio Costa, Daniela Feitosa e Fernando Cruz.

2.2.2 Estudos de aprofundamento e amadurecimento

Em especial, foram lidos todos os produtos entregues pelos consultores, visitados cursos no coursera e literatura científica [?].



Empoderando vidas.
Fortalecendo nações.

2.2.3 Escrita deste documento

Através da escrita deste documento, várias informações foram organizadas e sistematizadas, facilitando compreensão do contexto para especificação da API e entrega do produto.

2.2.4 Especificação da API de recomendação

Basicamente: quais URLs para recomendar quais recursos para quem, conforme explicitado no Apêndice A.

2.2.5 Implementação do serviço para aquisição dos dados, processamento e entrega em JSON

Para a API da etapa acima, foi necessário desenvolver um servidor Flask cujas rotinas estão expostas conforme a etapa abaixo.

2.2.6 Disponibilização das rotinas no IPython Notebook

Para expor os critérios de recomendação, e auxiliar análises e versões diferentes, as rotinas estão em um IPython Notebook. Veja o Apêndice B para uma exposição e detalhamento destas rotinas.

2.2.7 Proposta de implementações no portal federal de participação social

Com base no sistema de recomendações, nas demandas levantadas em reuniões e nos produtos já feitos por este e outros consultores, foi sistematizado um conjunto de propostas de acréscimos para o portal federal, detalhado no Apêndice E.

2.3 Justificativa do método

São usados diversos métodos neste produto. As justificativas se concentram nas tecnologias de web 3.0, de dados linkados, e nas demandas da equipe do participa.br, detectadas em reuniões e documentações produzidas [?].

2.4 Justificativa das fontes

A equipe do participa.br é a equipe da Presidência da República voltada para participação social. As fontes externas utilizadas, como artigos, livros e videos, são parte da pesquisa acadêmica do consultor, que possui produção bibliográfica própria.



Empoderando vidas.
Fortalecendo nações.

2.5 Confronto entre os resultados esperados e os alcançados

No Termo de Referência desta consultoria, este produto é descrito como “Documento com proposta de adaptações e incrementos para a interface do portal”. Este produto entrega estas propostas na forma de modificações a serem implementadas no portal federal de participação social. Além disso, para facilitar estas implementações e fazer a conexão com os desenvolvimentos ontológicos e de dados linkados, foi implementado e disponibilizado um sistema de recomendação de recursos para o participa.br. Um terceiro expoente deste terceiro produto é a visita aos produtos dos outros consultores que, aliada às reuniões e pesquisas do consultor, fortalece a integração e pertinência do produto.

3 Usos dos resultados

Há usos previstos do sistema de recomendação, pois são demandas da equipe. Em especial, há o uso para habilitar o plugin de recomendação de perfis em desenvolvimento (ActionItem [7]).

Há usos próprios dos participantes, como consultas aos recursos recomendados para si e para outros, e usos para a linha editorial do portal. Estes usos podem ser feitos no próprio servidor que disponibiliza o serviço de recomendação. Podem ser adaptados para a interface do participa.br, com plugins ou temas apropriados. Algumas sugestões de implementações estão no Apêndice E.

Outro uso previsto, e que idealmente contará com incentivos da comunidade, é a evolução destes métodos de recomendação de perfis para melhor atender aos usos do portal e das comunidades. O consultor sugere que as comunidades sejam convidadas a apresentar uma ou mais pessoas com conhecimentos ou disposição para algoritmos e sistematizações. Esta pessoa passaria por uma reunião para apreensão dos métodos de recomendação e análise, com vistas a proposição de melhoras e ampliação das funcionalidades.

As rotinas de recomendação estarão disponíveis online, e editáveis e executáveis no browser no IPython Notebook, o que deve facilitar dinâmicas de compartilhamento e amadurecimento destes processos. Este uso dos resultados deste produto é ainda mais precioso do que a API em si ou o uso dela dentro do participa.br. Mesmo assim, dada a complexidade das técnicas usadas, das tecnologias envolvidas e dos propósitos, o uso da API dentro do participa.br trará menos desafios que a apreensão, aproveitamento e melhora destes métodos pela comunidade de democracia participativa.

Um uso previsto é contemplar recursos de uma biblioteca digital assim como contemplar buscas e itens da biblioteca digital do participa.br como referência para indicar recursos ao participante. Análises dos dados do participa.br podem ser feitas diariamente e constar como documentos da biblioteca digital. Outras análises especiais podem entrar como itens na biblioteca digital, facilitadas com as rotinas já disponibilizadas, e facilitando outras, pois não exige instalação alguma, apenas visitas para páginas html via browsers usuais.



Empoderando vidas.
Fortalecendo nações.

4 Conclusão

4.1 Comentários, sugestões, recomendações

O consultor recomenda que haja amadurecimentos em encontros, tanto para melhora deste legado do participa.br, quanto para apreensão dos métodos pela comunidades interessadas.

As estruturas básicas para as análises e recomendações, que são as redes (de amizade e de interação) e os histogramas de palavras, podem ser convenientemente incluídos na triplificação dos dados, de forma que as recomendações e análises fiquem mais leves. Por hora, para viabilizar o uso, o sistema de recomendações refaz estas estruturas caso seja visitado o caminho **recomenda/atualiza**. Estas estruturas são:

- Rede de amizades.
- Rede de interação.
- Histograma de palavras de todos os textos e comentários do participa.br.
- Seleção das X (200-400) palavras mais ocorrentes para caracterizar o domínio).
- Histograma de palavras de cada usuário.
- Contagem das ocorrências das palavras selecionadas para caracterizar o domínio (portal participativo).

Os métodos de recomendação implementados utilizam separadamente critérios linguísticos ou de interação. Os casos e resultados já são interessantes e até complexos para uma primeira versão. Há planos para implementações de métodos que utilizam ambos recursos linguísticos e de relacionamento (topológicos). Por hora, ambas as características podem ser utilizadas operando as pontuações das recomendações com base no texto com pontuações das recomendações com base em relacionamento.

4.2 Impacto do Produto para a elaboração, gestão e/ou avaliação de políticas públicas de participação social

Este trabalho torna disponível uma porção de avaliações automáticas dos processos participativos que ocorreram ou ocorrerem no participa.br. Estas mesmas avaliações podem ser usadas continuamente, facilitando a gestão. A elaboração pode se beneficiar da análise das experiências passadas, facilitada pelos métodos aqui presentes.

Outro impacto pra a elaboração e gestão é basear o método de participação na utilização de métodos de recomendação de recursos. Por exemplo: pode-se recomendar que participantes



Empoderando vidas.
Fortalecendo nações.

com características X faça provocações. Estas provocações são enviadas para outros participantes como recomendações. Ainda a outros participantes pode ser recomendada a sistematização destes resultados. Um exemplo mais próximo da democracia participativa atual é uma etapa em que os participantes escrevem para recomendados participantes para fazer contato e iniciar discussões.

4.3 Como o Produto deverá impactar o público-alvo das políticas públicas a que se refere

As comunidades afeitas à participação social poderão aproveitar estas análises para melhor assimilação dos processos participativos, para relatórios, para gerar novos métodos de recomendação que melhor atendam aos interesses específicos das comunidades ou aos interesses da democracia participativa.

Um impacto imediato é a transparência reforçada dos processos participativos, pois não somente os dados, mas as análises e as recomendações dos recursos do participa.br estão online e publicamente disponíveis.

Outro impacto imediato é a valorização da participação. Sempre que um participante, por exemplo, fizer uma amizade ou comentar uma postagem ou outro comentário, ele estará constando nas redes envolvidas, modificará as análises atuais, e poderá ver seu nome e de outros em ranqueamentos diversos.

5 Agradecimentos

O consultor Renato Fabbri agradece ao Joenio Costa pelo template em \LaTeX para os produtos. Agradece à Daniela Feitosa pela reunião para demanda de recomendação de perfis. Agradece aos supervisores do trabalho realizado em torno do participa.br: Ricardo Poppi e Ronald Costa. Agradece ao labMacambira.sf.net e todas as comunidades de software e cultura livre que compõe esta contribuição.



Empoderando vidas.
Fortalecendo nações.

Referências

- [1] *Pad de proposta do produto 4: documento e scripts.* <https://etherpad.mozilla.org/pnud4>.
- [2] “Ontologia de participação social,” <http://tinyurl.com/p2doueu>.
- [3] “Ontologia do participa,” <http://tinyurl.com/lcccwop>.
- [4] “Especificação da triplificação dos dados do participa.br,” <http://tinyurl.com/k74z3yl>.
- [5] *Endpoint SparQL com dados do Participa.br.* <http://200.144.255.210:8082/>.
- [6] *Repositório Git do produto 3: documento e scripts.* <https://github.com/ttm/pnud3>.
- [7] *ActionItem do plugin de recomendação de perfis do participa.br.* <http://noosfero.org/Development/ActionItem3234>.



Empoderando vidas.
Fortalecendo nações.

Abreviações e jargão

RC: Redes Complexas

PLN: Processamento de Linguagem Natural

OPS: Ontologia de participação Social

OPA: Ontologia do Participa.br

MMISSA: Monitoramento Massivo e Interativo da Sociedade pela Sociedade para Aproveitamento

AARS: A Análise de Redes Sociais

MyNSA: Monitoring yields Natural Streaming and Analysis

PNPS: Plano Nacional de Participação Social

RDF: Resource Description Framework

HTTP: Hypertext Transfer Protocol

SPARQL: Simple Protocol and RDF Query Language

endpoint SPARQL: ponto de acesso, geralmente HTTP, a dados em RDF via buscas em SPARQL.

Participa.br: Portal federal de participação social.

IPython Notebook: instância online para rodar scripts Python

Mateor: arcabouço para páginas reativas e com funcionamento distribuído.

D3js: biblioteca de visualização de dados.



Empoderando vidas.
Fortalecendo nações.

A Especificação da API de recomendação de recursos do participa.br

Embora as rotinas no Apêndice [?] possam ser adaptados para os mais diversos fins, na API disponibilizada há três campos principais e dois auxiliares:

- Recurso: o recurso a ser recomendado: participantes, comunidades, trilhas, artigos ou comentários.
- Destinatário: para quem está sendo feita a recomendação: participante, comunidade ou linha editorial. Campo auxiliar “idd” para id do destinatário (comunidade ou participante).
- Método: método para a recomendação: top(ológico), tex(tual) ou hib(rido). Campo auxiliar de polaridade sim(ilar), dis(similar) ou mis(ta).

A url usada para consulta HTTP incorpora os parâmetros da forma usual: `http://<urlDoServidor>/recomenda?recurso=participante&destinatario=comunidade&idd=mirosc&metodo=top&polaridade=mis`

B Rotinas de acesso e processamento de dados do participa.br para as recomendações

C Infraestrutura do sistema de recomendações

D Instalação e modificação do sistema de recomendações

E Propostas de implementações na interface do portal federal de participação social