

Prémio Arquivo.pt

Descrição Sumária do Trabalho

Identificação

- Título: Memória Política
- Área temática: Ciências Políticas, Sociais e Tecnologias de Informação
- URL: www.memoria-politica.pt
- Candidatos:
 - João Andrade, afonso2345@gmail.com
 - Maria Carneiro, mariajcarneiro56@gmail.com
 - Miguel Azevedo, lopesazevedom@gmail.com

Descrição do Trabalho

O Arquivo.pt apresenta-se como uma plataforma que tem como principal objetivo a preservação de milhões de ficheiros recolhidos da web desde a década de 1990 disponibilizando um serviço público de pesquisa sobre esta mesma informação.

Na sociedade atual, por vezes, torna-se bastante difícil identificar o que é verdade ou não, não só devido à crescente expansão dos meios digitais, que por sua vez potencia o aparecimento e divulgação de falsa informação e notícias falsas, como também devido a alguma falta de educação sobre literacia mediática, entre outras...

Assim, inspirados por esta missão do Arquivo de manter e arquivar toda esta informação da web, que se perderia inevitavelmente aos longos dos anos, e tendo também como objetivo a criação de uma plataforma que faça frente e combata a crescente desinformação, decidimos explorar o conteúdo publicado por todos os partidos políticos portugueses que têm, atualmente, presença na Assembleia de Portugal.

Nesta plataforma, não só é possível a visualização de informação relativa à Assembleia de Portugal como também informação gráfica e visual relevante a cada partido assim como um sistema de pesquisa que permite obter informação tratada presente nas páginas web arquivadas do Arquivo.pt.

A componente técnica deste projeto passou por:

- Recolher as páginas web de todos os partidos, recorrendo à API do Arquivo.pt
- Processar e limpar os dados para facilitar o processamento posterior
- Extração de informação dos dados recorrendo a vários modelos e algoritmos, como por exemplo algoritmos de Named Entity Recognition para identificar menções a

personalidades, utilização de um modelo LLM (large language model) para fazer uma análise de polaridade e subjetividade aos textos recolhidos, entre outros.

- Indexação dos dados e criação de um sistema de pesquisa dedicado às páginas dos partidos políticos.
- Desenvolvimento de um website em Nextjs que agrega toda a informação recolhida e extraída.

Objetivos

Um dos principais objetivos do Memória Política é de certa forma facilitar e melhorar o processo de chegada de informação ao utilizador, informação esta (dados políticos) que muitas vezes é de difícil acesso à população (maioria dos utilizadores não consegue facilmente obter esta informação devido à sua elevada extensão e complexidade).

Para isto o Memória Política disponibiliza uma série de análise de dados tanto visual como textual (por exemplo, sentiment analysis de cada partido por legislatura, número de personalidades mais mencionadas por ano em cada partido, etc...), colocando também ao dispor uma ferramenta de pesquisa melhorada.

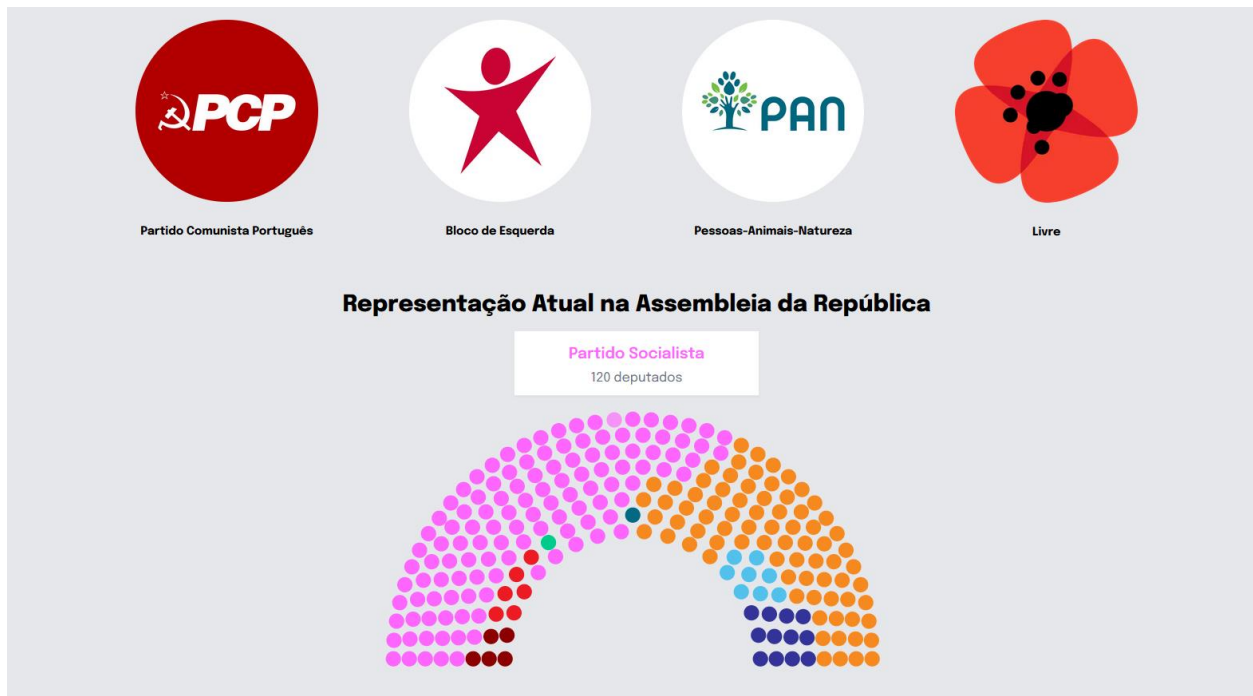
Não só temos como objetivo facilitar e melhorar a chegada como também tentamos ao máximo que a informação apresentada fosse simples e de fácil compreensão para todas as faixas etárias, permitindo assim lidar com um intervalo grande de necessidades diferentes.

Resultados Atingidos

Memória Política compila informação e dados de 8 partidos políticos, num intervalo de 27 anos, o que se transmite num total de 242 890 páginas web arquivadas com os seguintes atributos (nome partido, data, texto, título).

Maior parte da informação está organizada por partido político, sendo também possível organizar a informação por ano, tanto no motor de pesquisa como também na visualização de gráficos e outras ferramentas de análise de dados na página de cada um dos partidos.

Assim sendo, na página inicial podemos observar as palavras mais frequentes em toda a coleção, uma lista dos 8 partidos, uma representação atual da Assembleia da República e o número de deputados presente para cada um dos partidos na mesma e também o motor de pesquisa com algumas sugestões de pesquisa.



Na página de cada partido, podemos novamente ver as palavras mais frequentes para o partido, as personalidades mais mencionadas em cada ano, uma análise de sentimento por legislatura que fornece uma medida referente ao quão positivo ou negativo o discurso de cada partido é (polaridade), bem como uma classificação referente à subjetividade do mesmo e por fim podemos também visualizar uma sequência temporal de alguns dos eventos mais importantes do partido e também um 'doughnut chart' que mostra as menções a todos os outros partidos políticos.

Partido Socialista

O Partido Socialista (PS) é um partido político fundado a 19 de Abril de 1973 por militantes da Acção Socialista Portuguesa (ASP) sendo o seu espectro político de centro-esquerda.



Palavras Mais Frequentes para o Partido



Análise de Sentimento por Legislatura

Através da utilização de modelos de linguagem natural, classificamos o discurso presente nas páginas de cada partido ao longo de cada uma das legislaturas. Fornecemos uma medida referente ao quão positivo ou negativo o discurso de cada partido é (polaridade), bem como uma classificação referente à subjectividade do mesmo.

Selecione a legislatura:

XIV Legislatura



Partido(s) do Governo: PS
Maioria Absoluta: Não
Data de Início: 2019-10-25
Data de Fim: 2022-12-31

Polaridade

Com base na linguagem determinada nos textos presentes nas páginas web dos partidos, é possível classificar o seu conteúdo como positivo ou negativo.

Negativo  Positivo

Subjectividade

A subjectividade mede a quantidade de textos nas páginas web que corresponde a opiniões versus informação factual. Um valor mais alto de subjectividade, significa que durante a legislatura em questão, as páginas do partido tinham mais textos de opinião do que textos factuais.

Objectivo  Subjectivo

Na página de pesquisa podemos verificar que os resultados podem ser filtrados tanto por partido político como por ano.

Explore o Arquivo

Qual a posição dos partidos sobre o aborto? 🔍

Filtrar por Partido BLOCO

Filtrar por Ano 2009

75 Resultados Encontrados

Portal do Bloco de Esquerda - Aborto BLOCO 2009-05-04
 Portal do Bloco de Esquerda - Aborto Home Notícias Aborto Arquivo Contactos Home Eleições 2009 Notícias Autarquias Trabalho Jovens LGBT 6ª Convenção Nome e-mail Subscrição Cancelar Subscrição Aborto 04-Jul-2006 Condenação reforça urgência de mudar a lei A decisão condenatória do Tribunal de Aveiro comprova a urgência da modificação da lei que prevê a perseguição judicial das mulheres, a sua humi...

Portal do Bloco de Esquerda - Referendo à Despenalização do Aborto BLOCO 2009-05-04
 Portal do Bloco de Esquerda - Referendo à Despenalização do Aborto Home Arquivo Contactos Home Eleições 2009 Notícias Autarquias Trabalho Jovens LGBT 6ª Convenção Nome e-mail Subscrição Cancelar Subscrição Referendo à Despenalização do Aborto 02-Mar-2007 Relatório sobre o Referendo à Despenalização do Aborto , documento debatido na mesa nacional do Bloco de Esquerda (24 de Fevereiro de 2007) <...>

Portal do Bloco de Esquerda - Aborto: Sócrates quer novo adiamento BLOCO 2009-06-23
 Portal do Bloco de Esquerda - Aborto: Sócrates quer novo adiamento Home Arquivo Contactos Home Eleições 2009 Notícias Autarquias Trabalho Jovens LGBT 6ª Convenção Nome e-mail Subscrição Cancelar Subscrição Aborto: Sócrates quer novo adiamento 03-Jul-2006 Sendo impossível fazer o referendo, a deputada Helena Pinto anunciou na 6ª (28/10) que o Bloco tinha

Para terminar, temos uma página de perguntas frequentes onde é colocada alguma informação mais técnica sobre alguns aspetos da realização do trabalho, assim como alguma informação pessoal.

FAQ

Quem somos?

Somos 3 estudantes de **Engenharia Informática** na **FEUP** (Faculdade de Engenharia da Universidade do Porto) com um especial interesse na área de processamento e análise de dados e que decidiu aceitar este desafio como uma forma de melhor desenvolver as nossas capacidades, tanto técnicas como também de trabalho em equipa.

Como é constituído o corpus?

O **corpus** final utilizado está dividido em diversos **corpus** diferentes, correspondendo cada um deles a um dos partidos políticos presentes na Assembleia Portuguesa atualmente. Estes **corpus** foram criados com base nos ficheiros HTML obtidos através do Arquivo.pt que é um repositório nacional que guarda informação sobre páginas WEB Portuguesas desde 1996.

Dado que cada partido tem, não só uma data de criação diferente, como também diferentes presenças online, cada corpus abarca períodos temporais diferentes (**PSD** 1999-2020, **PS** 1996-2020, **Chega** 2019-2020, **IL** 2017-2020, **PCP** 1996-2020, **BE** 2003-2020, **Livre** 2015-2020, **PAN** 2013-2020).

Porque é que alguns partidos não estão incluídos?

A nossa escolha dos partidos a incluir baseou-se na relevância atual para a população portuguesa: é importante sabermos e conhecermos a história de quem nos representa. No entanto, futuramente, planeamos incluir o resto dos partidos existentes atualmente e, possivelmente, novos partidos que possam surgir.

A Análise de Sentimento é confiável?

A Análise de Sentimento é baseada em modelos de linguagem pré-treinados e aplicada aos textos retirados do Arquivo.pt. Para mais informação acerca da confiança do modelo, visite este link.

Originalidade e caráter inovador

O Memória Política, apesar de não ser o primeiro a tratar e analisar dados políticos, é o primeiro projeto que reúne dados de todas as páginas web arquivadas no Arquivo.pt de cada um dos partidos presentes na Assembleia Portuguesa.

Com esta abordagem conseguimos, não só permitir perceber os pontos de vista e ideais de cada um dos partidos como também conseguir perceber a relação entre os mesmos e mesmo o impacto que têm uns nos outros.

Assim, disponibilizamos ao utilizador informação que, sem o Memória Política não seria possível de obter (ou seria, mas com grandes dificuldades).

Impacto social (aplicação e utilidade social)

Um dos principais aspetos do Memória Política, como já foi mencionado antes, é a sua versatilidade e simplicidade, o que permite mesmo ao utilizador mais comum, facilmente navegar e captar informação do website. Porém, esta facilidade e simplicidade contrastam com a veracidade dos dados e o quão informativo são.

O Memória Política, para além de ser informativo num caso de uso geral, é também bastante útil em casos de uso mais específicos onde o utilizador poderá utilizar o motor de pesquisa disponível para obter informações mais representativas daquilo que está à procura. Por exemplo, e imaginando que as eleições legislativas se encontram à porta, seria do interesse do utilizador pesquisar informação relativa às propostas de cada um dos partidos para as eleições ou mesmo os seus pontos de vista e ideais.

Por último, importa também referir que este site tem como objetivo valorizar e preservar a produção jornalística e documental sobre a vida política portuguesa ao longo dos anos.

Impacto científico (aplicação e utilidade científica)

Um dos principais aspetos relevantes quanto à aplicação e utilidade científica é a fácil acessibilidade e simplicidade no acesso à informação, permitindo assim um estudo melhorado dos dados políticos nas diversas áreas de investigação que abrangem o tema.

Relevância da utilização do Arquivo.pt

Visto que este projeto tem como objetivo apresentar o crescimento, propostas e ideais dos partidos políticos desde a sua fundação até ao presente, isto não seria possível sem recorrer ao uso das APIs do Arquivo.pt que e ao Arquivo em si que desde a década de 1990 que arquiva e guarda informação de artigos e website (havendo mesmo alguns que já não se encontram disponíveis), que de outra maneira, simplesmente se perderiam no tempo.

Neste sentido, louvamos o esforço feito até aqui pelo Arquivo.pt para manter este arquivo de informação publicada na web acessível.

Recursos complementares

- Dados do Parlamento.pt, <https://www.parlamento.pt/>, recolha visual de informação para esclarecimento de dúvidas
- Dados da Wipedia.org, https://pt.wikipedia.org/wiki/Wikip%C3%A9dia:P%C3%A1gina_principal, recolha visual de informação e esclarecimento de dúvidas para cada partido