

Democracia Aumentada:

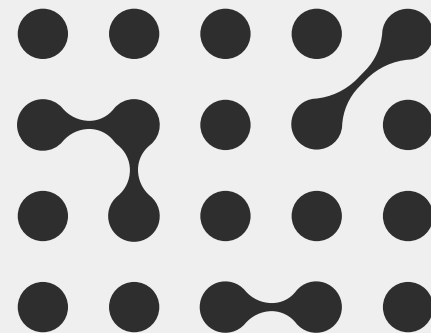
Um Sistema para a Democracia Digital

Enzo Bustos Da Silva

Orientadora: Anna Helena Reali Costa

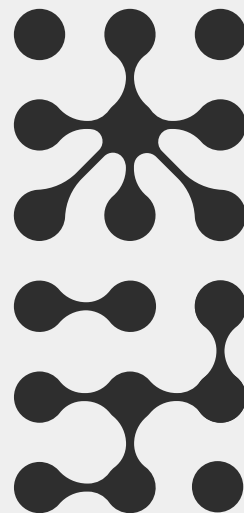
Colaboradores: Thomas Palmeira Ferraz, André Seidel Oliveira

■ SIICUSP - 2021



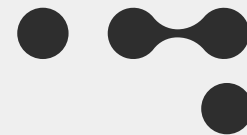
01

RELEVÂNCIA





MOTIVAÇÃO



Aumento na descrenças nas instituições democráticas

01

Política atual

- ☐ Internet
- ☐ Espaço de debate
- ☐ Redes Sociais

02

Órgãos públicos ultrapassados

- Longas reuniões
- Documentos complexos
- Descrença nas Instituições



Fonte: gpsalvador.org

RELEVÂNCIA

Seleção de informações na Democracia Digital

03

Democracia Digital

- ☐ Uso político das novas tecnologias
- ☐ Informação, Participação, Discussão
- ☐ Excesso de conteúdo e Fake News

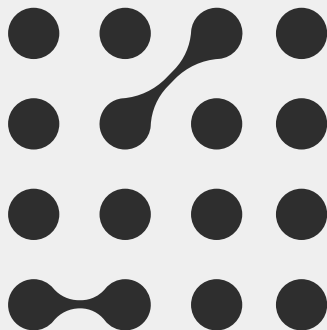
04

Democracia Aumentada

- Inteligência Artificial e PLN
- Processar e estruturar
- Acessibilidade na interpretação



Fonte: biblioo.info



METODOLOGIA E RESULTADOS

Coleta de Dados



Web Scraping

Construção de um robô para coletar os dados diretamente do site do DAR

17 DE SETEMBRO DE 2020

37

Não foi sempre assim! É o Conselho Superior de Obras Públicas que diz que falta capacidade de planeamento e organização, faltam programas, falta capacidade de gestão. Esta é uma incapacidade quase afiliva, repito, quase afiliva!

O Sr. **Bruno Dias** (PCP): — Chame cá o Dr. Vitor que ele explica!

A Sr.ª **Filipa Roseta** (PSD): — O Prof. António Costa Silva disse hoje de manhã que o problema era a falta de capacidade da Administração Pública, o que rejeito. Não são os funcionários públicos que têm falta de qualificações. Não são, porque eu conheço-os, não é por aí, não é esse o problema. Sabem qual é o problema? O problema é uma cultura burocrática, que está a aumentar cada vez mais, que cria dificuldades para vender facilidades, abrindo portas à corrupção...

A Sr.ª **Mónica Quintela** (PSD): — Muito bem!

A Sr.ª **Filipa Roseta** (PSD): — ... e que de cada vez que há um problema e há um programa que se quer executar abre-se uma «via verde», que é o que se está a preparar para se fazer agora com os fundos, para andar mais rapidamente do que as pessoas normais! Isto, sim, é a pedra angular que temos de destruir se queremos verdadeiramente acabar com a corrupção.

Portanto, Sr. Deputado, gostava de saber a sua opinião e gostaria que todos declarássemos, como inimigo número um, a cultura burocrática que está cada vez mais a aumentar, para todos, destruindo as pessoas, as empresas, a economia e a justiça no nosso País.

Aplausos do PSD.

O Sr. **Presidente** (Fernando Negrão): — Para responder, tem a palavra o Sr. Deputado João Gonçalves Pereira.
Faça favor, Sr. Deputado.

O Sr. **João Gonçalves Pereira** (CDS-PP): — Sr. Presidente, Sr.ªs e Srs. Deputados, agradeço as questões que me foram colocadas.

Começo por responder ao Sr. Deputado da Iniciativa Liberal, João Cotrim de Figueiredo, dizendo-lhe que o País ouviu o Prof. António Costa Silva — aliás, ele desdobrou-se em entrevistas — dizer que tem uma visão, que tem um plano. Só não sabemos qual é o plano e a visão de António Costa! O que importa saber e discutir é o que é que desse plano e dessa visão vai andar para a frente e o que é que vai ficar para trás. Isso é que importa saber.

Há muita propaganda, já estamos habituados a isso, mas há uma diferença entre a apresentação inicial e a apresentação de ontem, e nisso dou razão ao Sr. Deputado: agora aparece muito mais Estado, Estado e mais Estado e falta empreendedorismo, economia aberta. Portanto, parecia que o plano tinha um determinado rumo e, de repente, fez uma inversão de marcha.

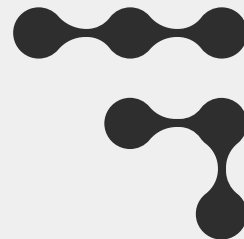
Em relação ao Sr. Deputado do PCP, Bruno Dias, que parece ter metido uma espécie de cassette «Avante»...

O Sr. **Bruno Dias** (PCP): — Finíssima!

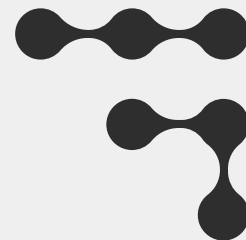
O Sr. **João Gonçalves Pereira** (CDS-PP): — O Sr. Deputado fez aqui uma declaração, não colocou perguntas, portanto, não tenho nenhuma resposta a dar-lhe, porque não me fez nenhuma pergunta! Entendeu fazê-lo e é legítimo fazer aqui uma declaração.

O Sr. **Bruno Dias** (PCP): — Se calhar, não entendeu, esse é que é o problema!

O Sr. **João Gonçalves Pereira** (CDS-PP): — Quanto às questões colocadas pela Sr.ª Deputada Filipa Roseta, do PSD, bem, a incapacidade de execução deste Governo é conhecida e eu diria até que é uma marca do Partido Socialista: anuncia muito, faz muita propaganda, muitas reflexões, muitos debates, mas quando



Fonte:
debates.parlamento.pt



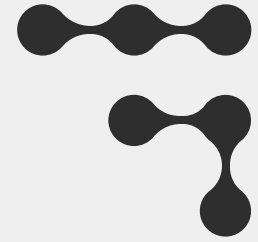
Estruturação das Atas



Tabularização

As atas foram então estruturadas em planilhas divididas por cada fala dos deputados e segmentada por iniciativa legislativa

	Transcript	Date	Person	Party	Text	Initiatives
498	DAR-002	18 DE SETEMBRO DE 2020	André Ventura	CH	Sr. Presidente, Srs. Deputados: Temos de começ...	projeto de resolução n.º 471, projeto de lei n...
499	DAR-002	18 DE SETEMBRO DE 2020	Presidente	Fernando Negrão	Sr. Deputado, chamo a sua atenção para o tempo...	projeto de resolução n.º 471, projeto de lei n...
500	DAR-002	18 DE SETEMBRO DE 2020	André Ventura	CH	Vou terminar, Sr. Presidente. Portanto, não há...	projeto de resolução n.º 471, projeto de lei n...
501	DAR-002	18 DE SETEMBRO DE 2020	Duarte Alves	PCP	Quando pagamos!... I SÉRIE — NÚMERO 2	projeto de resolução n.º 471, projeto de lei n...
502	DAR-002	18 DE SETEMBRO DE 2020	André Ventura	CH	Se pesquisarmos na internet por PCP e nacional...	projeto de resolução n.º 471, projeto de lei n...
503	DAR-002	18 DE SETEMBRO DE 2020	Presidente	Fernando Negrão	Tem de terminar, Sr. Deputado.	projeto de resolução n.º 471, projeto de lei n...

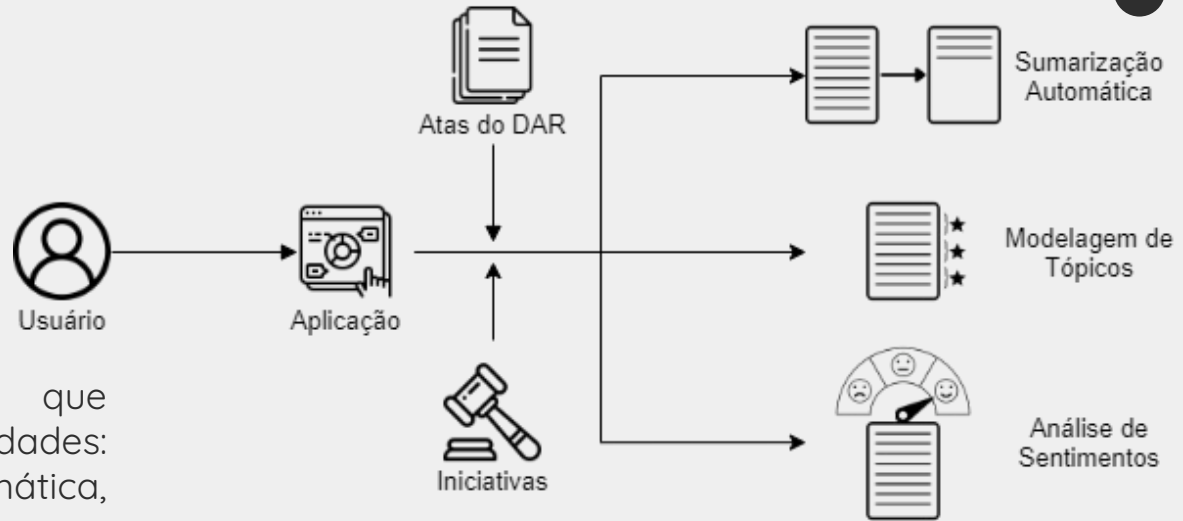


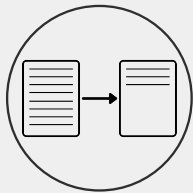
Sistema de Democracia Aumentada



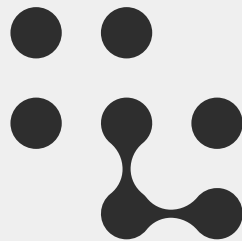
Arquitetura

Esquema do sistema que constitui 3 funcionalidades: Sumarização Automática, Modelagem de Tópicos e Análise de Sentimentos. Todas utilizam Inteligência Artificial

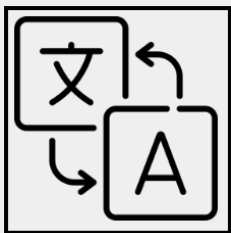




Sumarização Automática



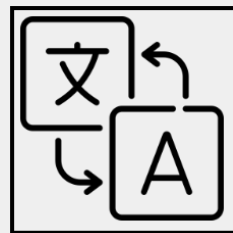
Fala dos
Deputados



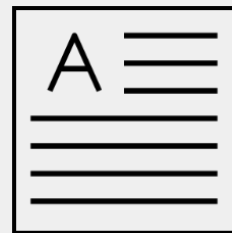
M2M100:
pt \rightarrow en



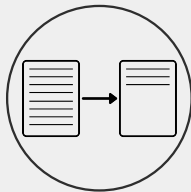
PEGASUS:
Sumarização
Abstrativa



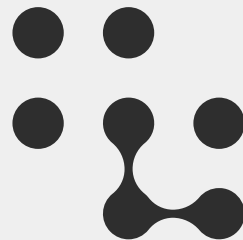
M2M100:
en \rightarrow pt



Resumo



Sumarização Automática



22

I SÉRIE — NÚMERO 7

Fonte: debates.parlamento.pt

O Sr. **Presidente**: — Para formular um pedido de esclarecimento, tem a palavra o Sr. Deputado André Ventura, do Chega.

O Sr. **André Ventura** (CH): — Sr. Presidente, ouvimos o Sr. Deputado António Filipe falar de política de justiça e dizer que não podemos ter uma justiça em que os cidadãos não podem confiar.

É curioso ouvir isto do Partido Comunista, ouvir dizer que há vários arguidos e condenados em cargos públicos, quando, sempre que instado a comentar, diz que não se mistura justiça com política e que não confunde arguidos com condenados nem condenados com o exercício de cargos políticos. Por isso, seria bom hoje ficasse claro o que querem aqui dizer!

Protestos do Deputado do PCP Bruno Dias.

Mas seria também importante perguntar ao Partido Comunista se está ou não disponível para conseguir superar o bloqueio do Tribunal Constitucional em matéria de enriquecimento ilícito e se está ou não disponível para uma proposta que, finalmente, consiga levar quem enriquece ilícitamente a ser condenado pela nossa justiça ou se é só de boca e se, na verdade, não vamos fazer nada.

Mas hoje, Sr. Deputado, não posso deixar de lhe fazer uma pergunta. Sobre a prisão perpétua, ouvi o Sr. Deputado dizer que era um choque e que violava flagrantemente as nossas normas. Por isso, com esta oportunidade, tendo em conta que, na grande maioria dos países comunistas do mundo, há prisão perpétua — até há pior, mas há prisão perpétua, pelo menos —, gostaria que nos dissesse se o Partido Comunista Português vai ou não estar de acordo e é capaz de apoiar uma medida, como a que o Chega já introduziu no sistema parlamentar português, de prisão perpétua para casos de violação de menores, homicídio, terrorismo e casos graves de corrupção, uma vez que em quase todos os países comunistas do mundo isso acontece.

Gostaria ainda de lhe fazer outra pergunta, Sr. Deputado. O Sr. Deputado falou em incompatibilidade de exercício entre a vida pública e o setor privado, como se os privados fossem todos uns criminosos e uns corruptos. É tempo de o Partido Comunista dizer também se vai ou não apoiar a proposta do Chega que prevê que quem faça negócios com o Estado e em nome do Estado nunca mais possa exercer cargos na empresa com a qual negociou em nome do Estado. É que estamos fartos de ministros que fazem concessões e depois vão trabalhar para as pontes onde fizeram essas concessões ou paraministros que propuseram obras e que acabaram a trabalhar nas empresas a que adjudicaram essas mesmas obras.

Era isto que o Partido Comunista podia fazer e era isto que, para lá da conversa sobre a política de justiça e sobre melhorar as celas individuais, como o Bloco de Esquerda quer fazer, os portugueses gostariam, efetivamente, de ver resolvido em Portugal.

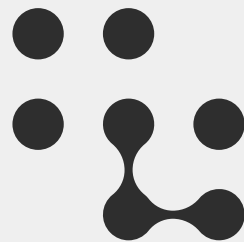
“Portanto, com esta oportunidade, tendo em conta o fato de que na grande maioria dos países comunistas há prisão perpétua, eu gostaria de dizer-nos se o Partido Comunista Português vai ou não concordar e é capaz de apoiar uma medida, como que a Chega já introduziu no sistema parlamentar português, de prisão perpétua por casos de menores violações, assassinatos, terrorismo e sérios casos de corrupção, já que em quase todos os países comunistas do mundo isso acontece.”

Discurso do deputado **André Ventura (CH)** na sessão do dia 30/09/2020 resumido, em resposta à declaração política do PCP sobre medidas de combate à corrupção.

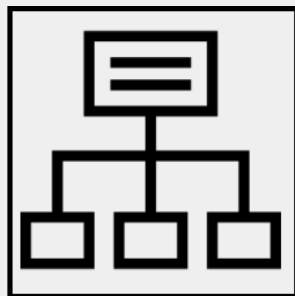
Taxa de compressão: 1:5



Modelagem de Tópicos



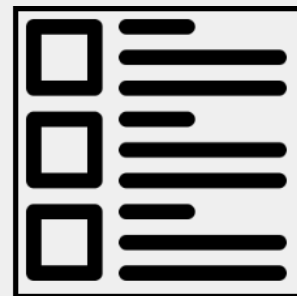
Fala dos
Deputados



BERTopic:
agrupamento
de palavras



Zero-Shot:
Convergência
em Temas

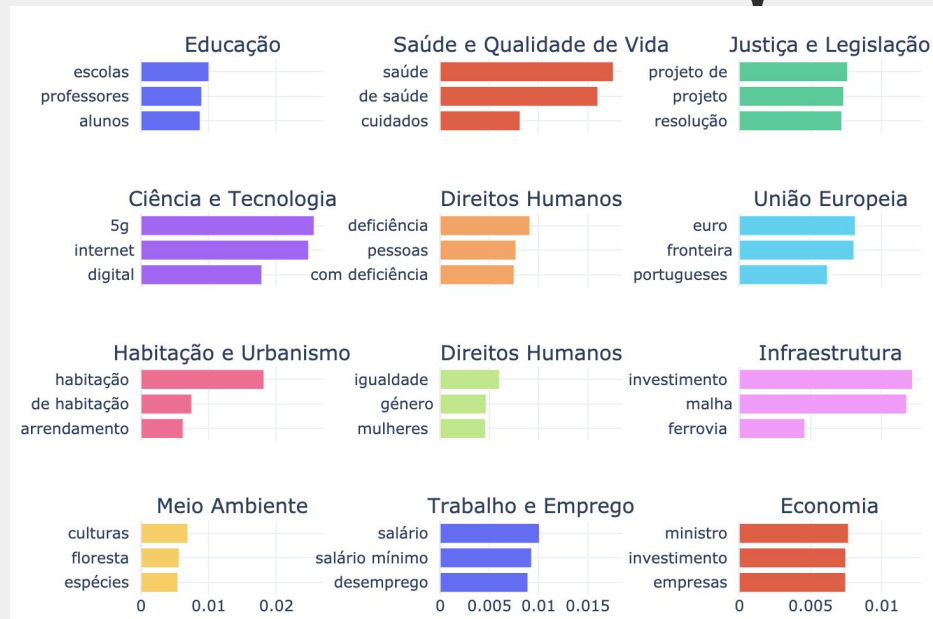


Temas

Modelagem de Tópicos

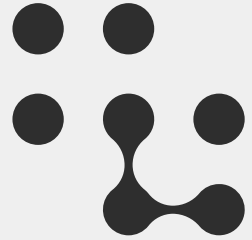


Agrupa palavras que estão relacionadas a um mesmo assunto. De modo a buscar quais são as discussões que mais interessam o usuário baseado nestes tópicos.





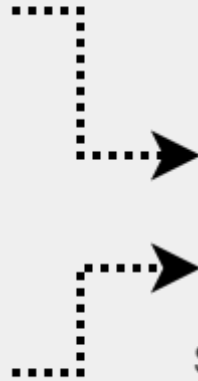
Análise de Sentimentos



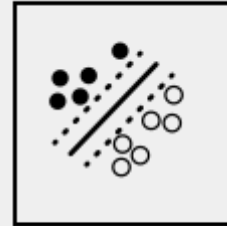
Fala dos
Deputados



Iniciativas
Legislativas



Separação entre
propostas de
aliados ou
oposição do
governo



SVM:
Classificação
da polaridade

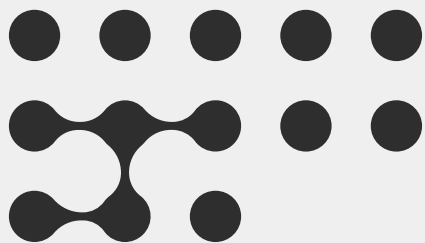
Análise de Sentimentos



Determina qual a polaridade do sentimento expresso em um texto. Usada para gerar uma métrica de confiabilidade em partidos e deputados.



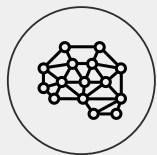
Fonte: pt.dreamstime.com



03

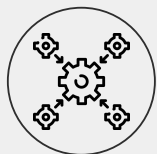
CONCLUSÃO

Conclusão



IA e PLN

Associação benéfica entre técnicas de ML dentro da política



Integração

Interface amigável para unir todas as funcionalidades



Qualidade

Melhora geral na qualidade da informação



Base

Criação para base para outros projetos que abordem política





AGRADECIMENTOS



Programa de
Bolsas Itaú



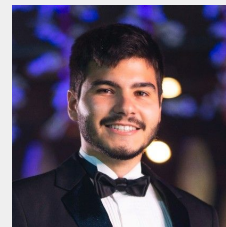
Centro de Ciência
de Dados



Anna Helena Reali Costa
(orientadora)



Thomas Palmeiras
(colaborador)

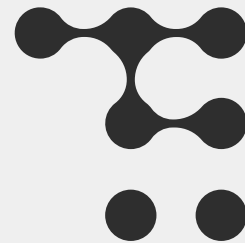


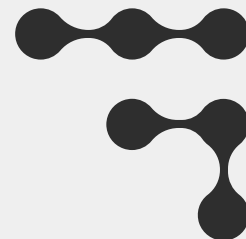
André Seidel Oliveira
(colaborador)





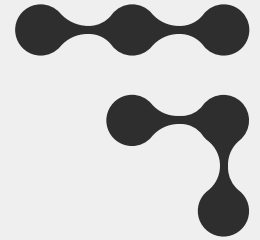
Dúvidas





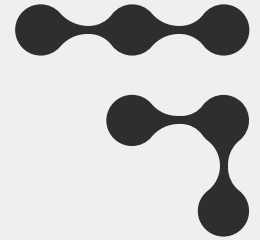
Sobre o Projeto

- Surgiu pelo contato da professora Anna Reali com uma rede de cientistas políticos e sociais de Portugal
- Democracia Aumentada: Extrapolar o fazer político atual
- Ter uma relação interdisciplinar entre exatas e humanas
- Mudança de planos...



Trabalho Futuro

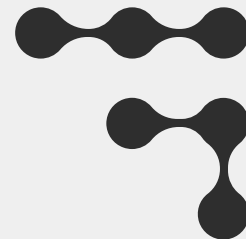
- Análise de Sentimentos
- Interface na Web
- Expansão de Funcionalidades
- Internacionalização
- Novas tecnologias



Impactos

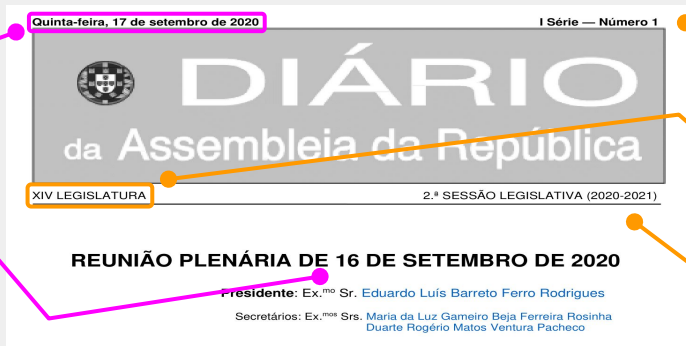
- Dia a Dia
- Dinâmica política
- O que está em pauta?
- Discussões mais embasadas
- Segurança na votação

Construção da Base



Data

Dia da Reunião
Plenária



Transcrição
Identificação
da Ata

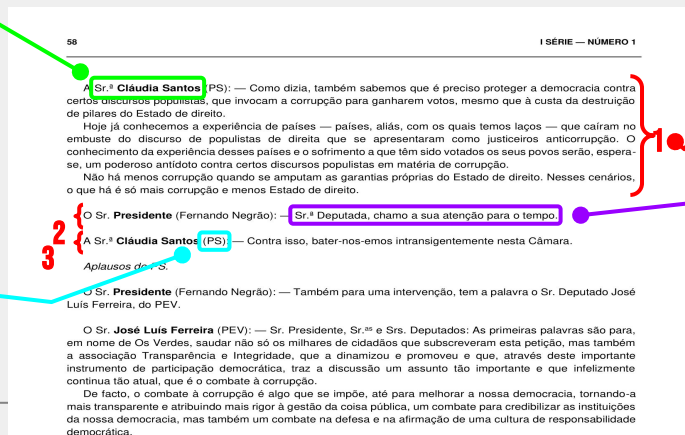
- Uso de Selenium
- Uso de RegEx
- Uso de Pandas

Pessoa

Nome do
Deputado que
está falando

Partido

Filiação política da
pessoa que está
falando



Posição

Ordem das
linhas de fala
em um diálogo

Texto

Fala de um
deputado



Mais a fundo: Sumarização

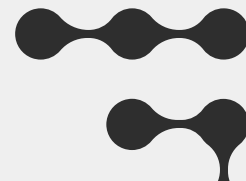
- Vários modelos foram testados:
 - Extrativos: BertExtSum, BART,
 - Abstrativos: BertExt, BigBird, Pegasus X-Sum, Pegasus Large, Distilbart, Bart-Large
- Por conta da qualidade de encadeamento foi preferível modelos abstrativos
- Embora hajam métricas para sumários como a Rouge, a maior parte dos resultados foi validado por humanos do nosso grupo
- Rouge funciona através do overlapping entre n-grams do ground-truth (sumários/traduições feitas por humanos) e o gerado pelo modelo
- Geralmente funciona melhor para extrativos, já que a paráfrase pode diminuir o valor em modelos abstrativos
- Para a tradução foi utilizado apenas o M2M100 por não ter apresentado nenhum problema nas traduções



Mais a fundo: Sumarização

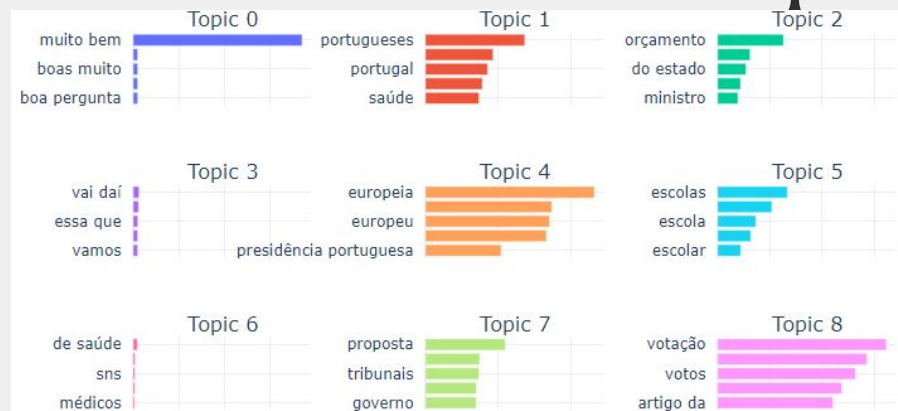
Resumo em Inglês	Resumo em Português
Therefore, with this opportunity, having regard to the fact that in the vast majority of the communist countries of the world there is perpetual imprisonment — even worse, but there is perpetual imprisonment, at least — I would like you to tell us whether the Portuguese Communist Party will or will not agree and is able to support a measure, such as that Chega has already introduced to the Portuguese parliamentary system, of perpetual imprisonment for cases of minor violations, murder, terrorism and serious cases of corruption, since in almost all communist countries of the world this happens.	Portanto, com esta oportunidade, tendo em conta o fato de que na grande maioria dos países comunistas do mundo há prisão perpétua – ainda pior, mas há prisão perpétua, pelo menos – eu gostaria de dizer-nos se o Partido Comunista Português vai ou não concordar e é capaz de apoiar uma medida, como que a Chega já introduziu no sistema parlamentar português, de prisão perpétua por casos de menores violações, assassinatos, terrorismo e sérios casos de corrupção, já que em quase todos os países comunistas do mundo isso acontece.

- O principal erro da etapa de sumarização é relacionado à referência, por exemplo: “O sr. Deputado André Ventura (CH) argumentou contra a proposta de Lei que o PCP colocou em pauta”
- Pode virar “Proposta de Lei for recusada pelo sr. Deputado” → tendo um erro de referência
- Outro erro é gerado quando o limite de tokens do modelo é atingido gerando repetição
- Erros do tipo Alucinação não foram percebidos para o Pegasus-Large, embora mais frequentes para o XSum (erros em fatos apontados no texto)



Mais a fundo: Topic Modeling

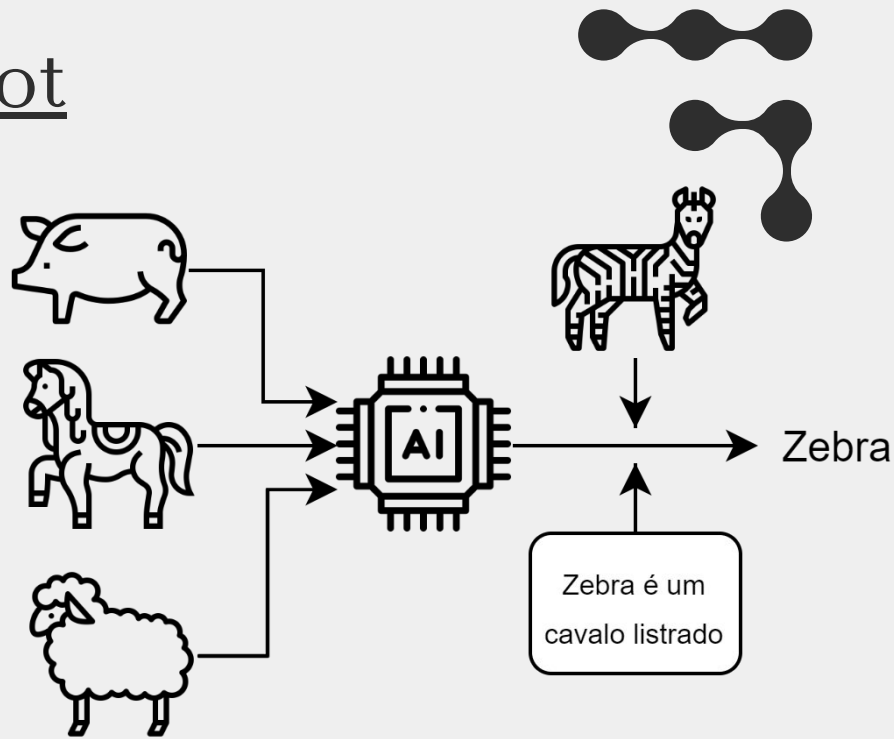
- BERTopic funciona de modo não supervisionado e olhando para todo o conjunto de atas consegue saber que tipo de palavras estão próximos semanticamente falando
- Isso cria tópicos genéricos nomeado Topic 1, 2, ... Alguns são facilmente identificados como algum tema por um humano, outros são aleatórios, como tópicos para palavras de saudação
- O zero-shot é responsável por dar nome aos bois, e classificar quais conjuntos de palavras são de um dado tema

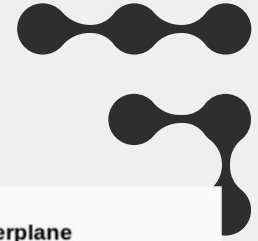


- Esse conjunto pode chega a ter mais de 300 tópicos, que são forçados aos mais relevantes por Hierarchical Clustering e depois passam para o zero-shot afunilar em nos Temas, que foram escolhidos.

Mais a fundo: Zero-Shot

- Zero-Shot Classification é quando um modelo é submetido a exemplos de teste que não foram aprendidos durante a fase de treino do modelo
- O roBERTa é treinado como a maioria dos Transformers através de masks de texto . Para ilustrar:
- No nosso caso, essa descrição adicional são justamente o tema, e usamos uma template.
- Basicamente perguntamos pro modelo “Esse conjunto de palavras está mais próximo de qual desses temas?” e dados cerca de 20 temas





Mais a fundo: Sentiment Analysis

- Como é algo que ainda está em construção, planejamos utilizar algo que já foi testado em paper como o: 'Aye' or 'No'? Speech-level Sentiment Analysis
- Que utiliza justamente essa repartição entre falas pró/contra governo antes de aplicar o SVM para determinar a polaridade do sentimento
- A ideia do SVM é basicamente o modelo aprender o melhor hiperplano que divide o conjunto de dados

