



# Conta-me Histórias

## Venha recordar as notícias do passado

Ricardo Campos, Arian Pasquali, Vítor Mangaravite, Alípio Jorge, Adam Jatowt

Instituto Politécnico de Tomar, Universidade do Porto, INESC TEC, Universidade de Kyoto



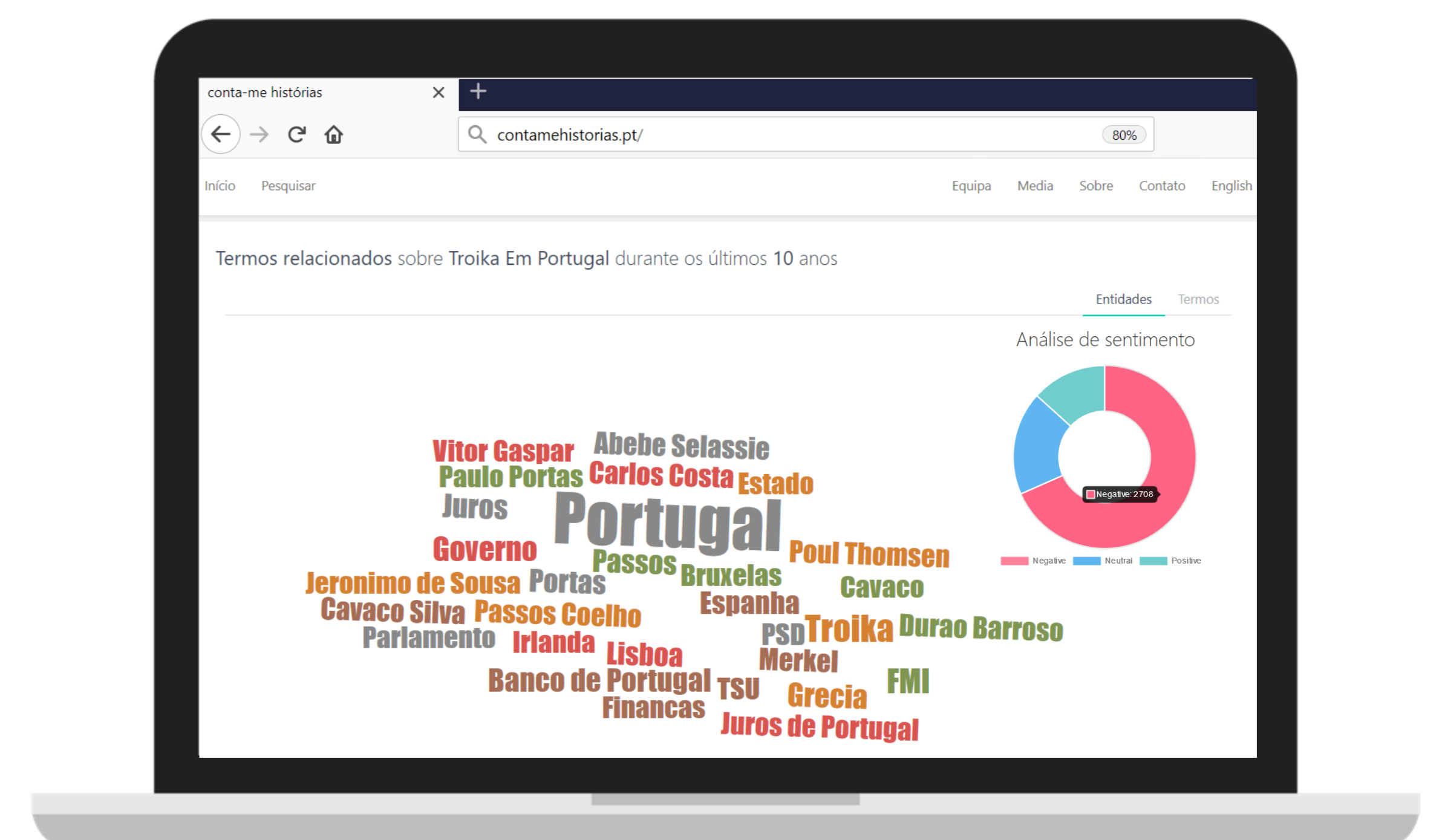
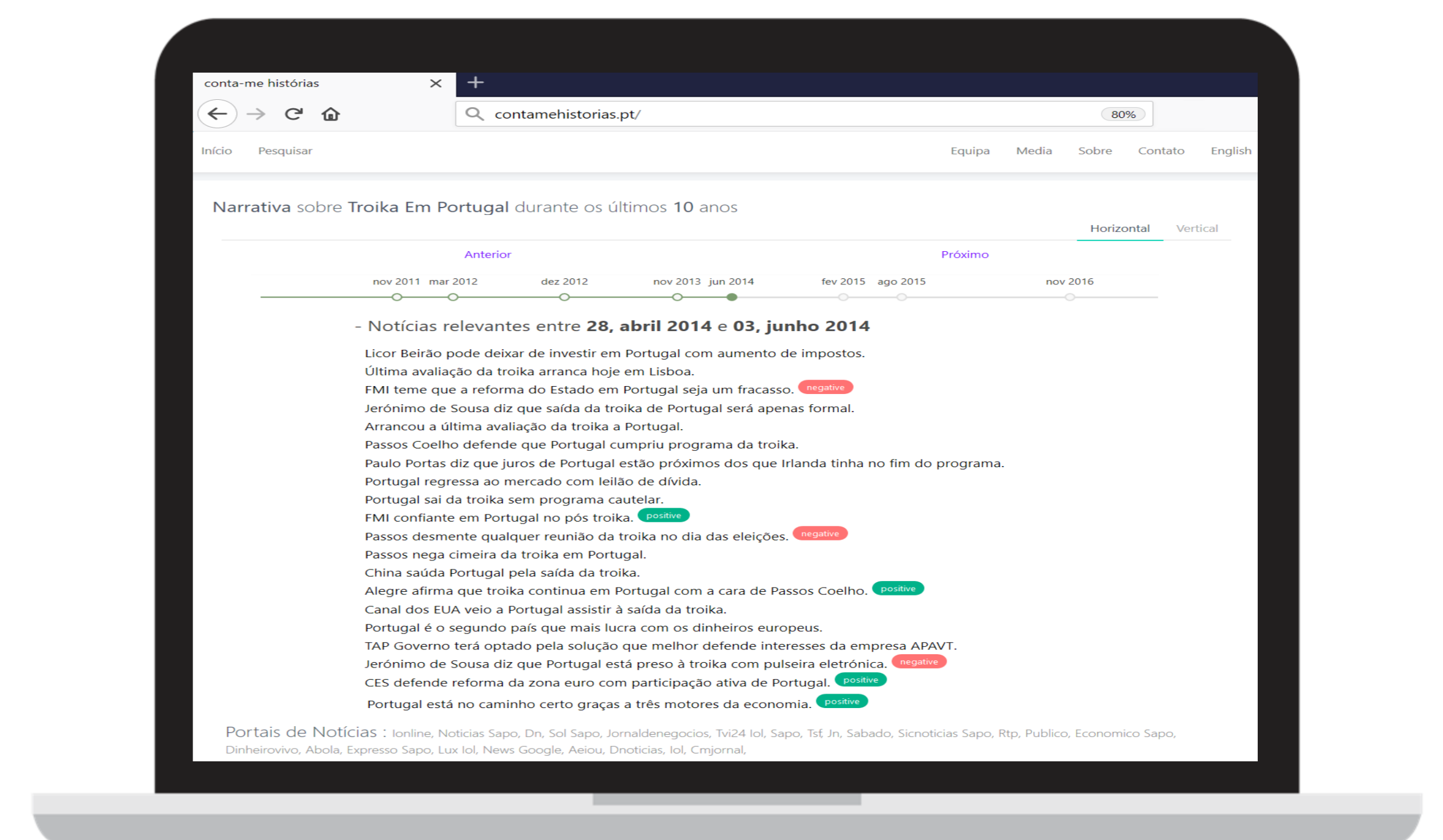
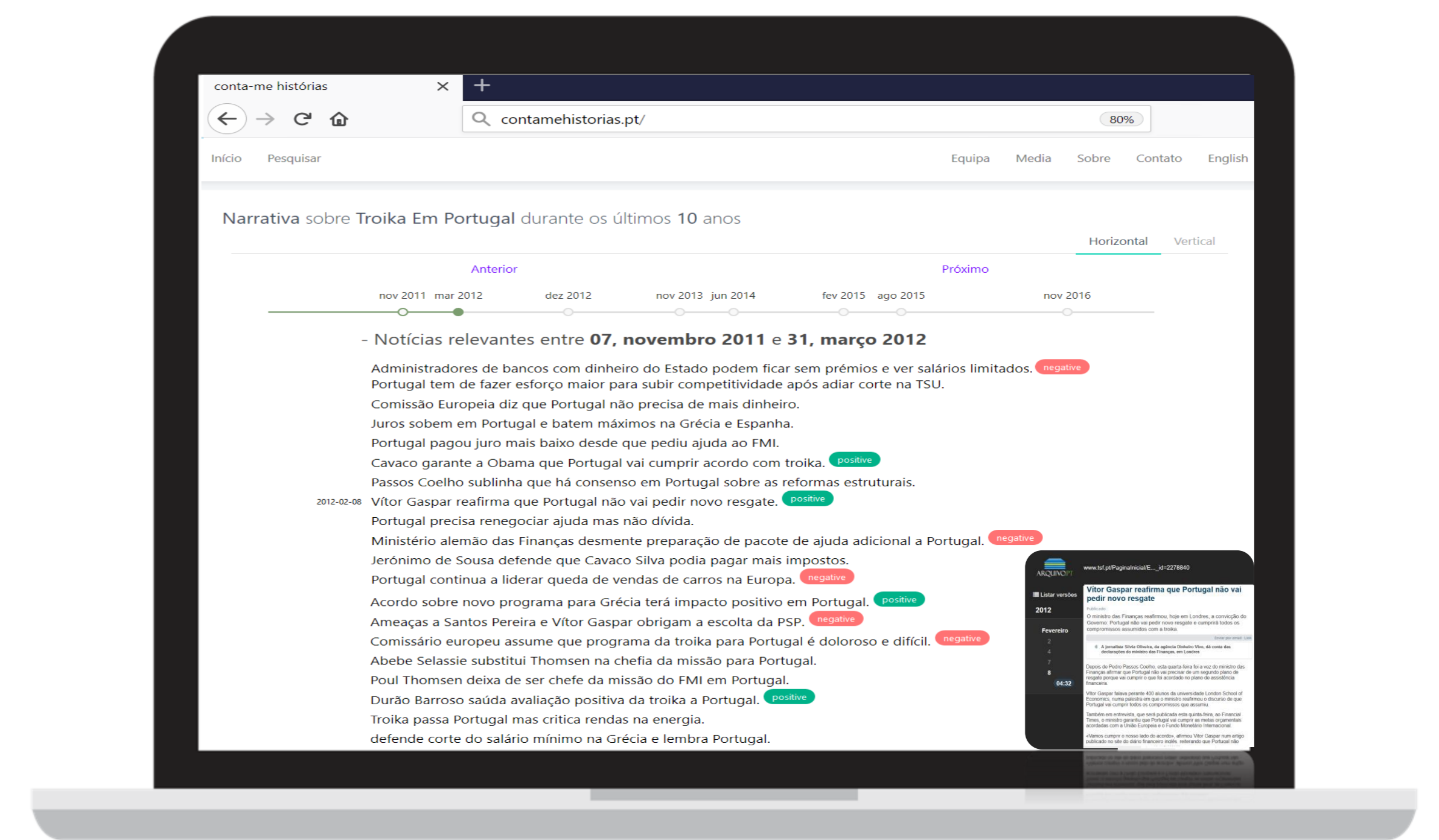
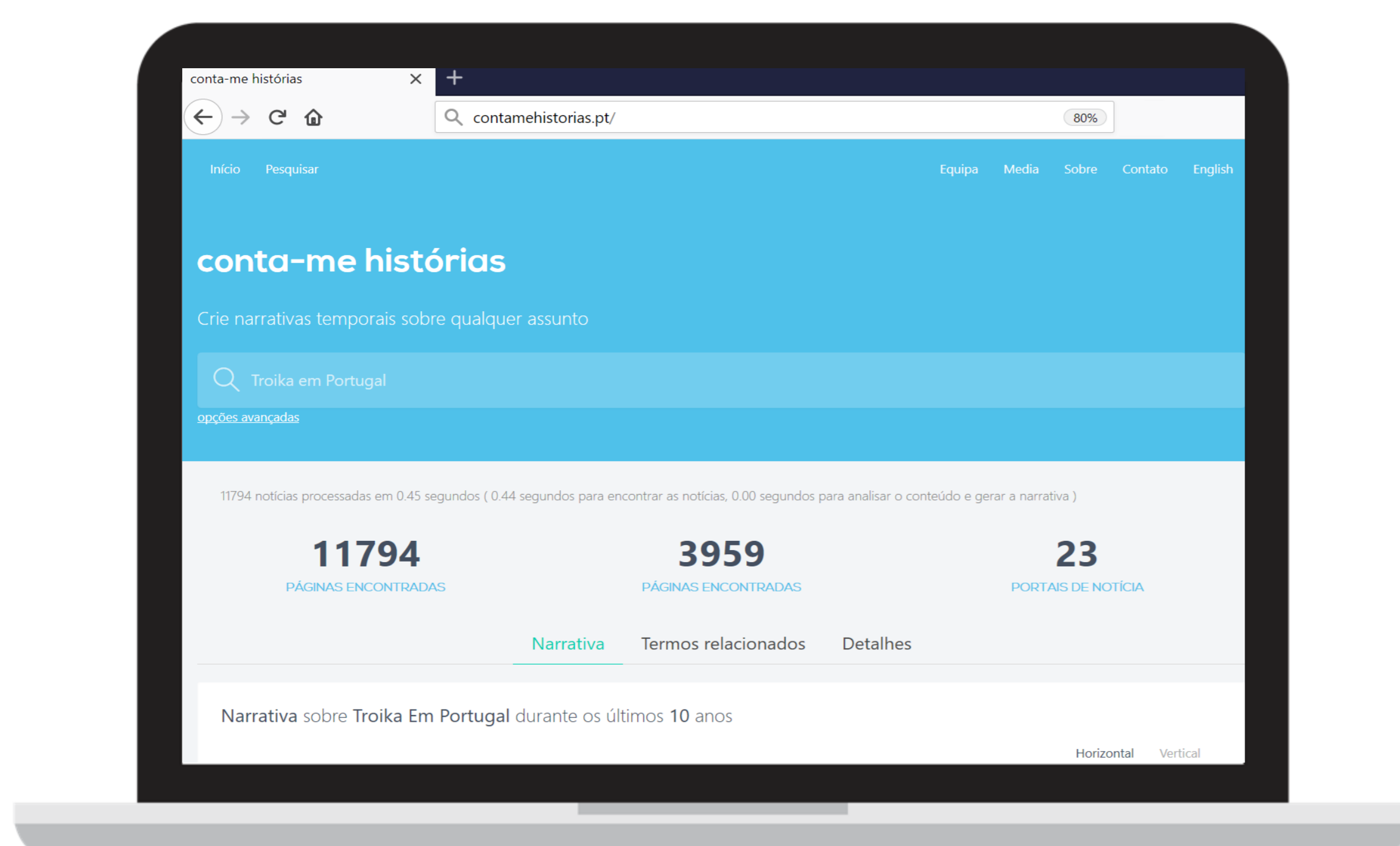
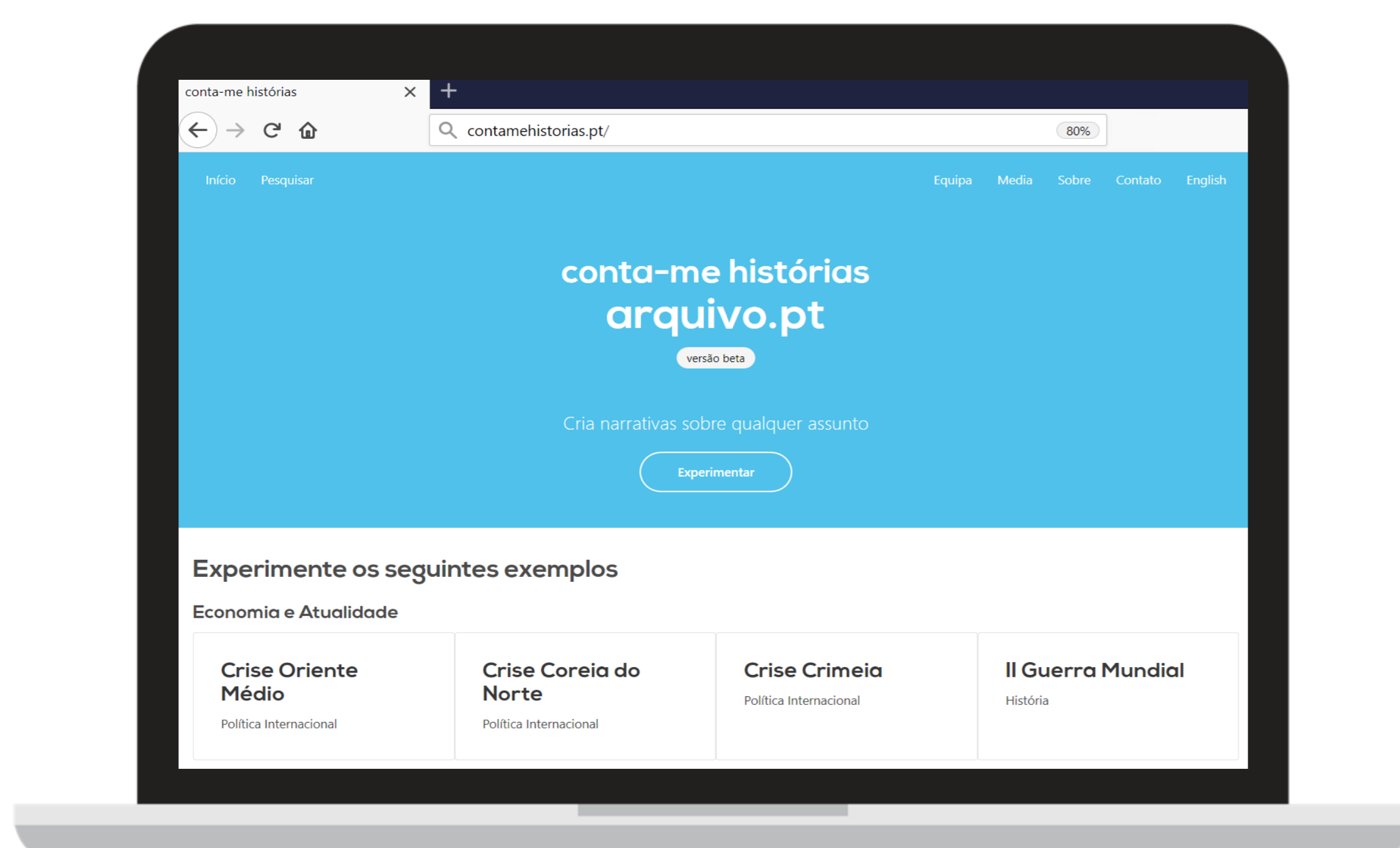
Vencedor dos Prémios Arquivo.pt 2018

<http://contamehistorias.pt>

Conta-me Histórias é um projeto científico que permite gerar automaticamente um resumo temporal interativo de notícias preservadas pelo Arquivo.pt (Arquivo da Web Portuguesa), oferecendo aos utilizadores a possibilidade de navegar no passado e sumarizar a informação disponível online.

Já imaginou poder revisitar o processo **Casa Pia** com um único clique? Recordar os tempos da **Troika em Portugal**? A polémica em torno do novo **Acordo Ortográfico**?

Venha conhecer o Conta-me Histórias!



## Contribuições

A nossa abordagem (adaptável a outros tipos de fontes de dados (e.g., patentes, redes sociais, etc) incluindo diferentes línguas) é baseada no YAKE [1], um algoritmo de extração de palavras que ajuda à deteção dos títulos mais relevantes, e num método de deteção de picos que ajuda à seleção dos períodos de tempo mais importantes para a pesquisa efetuada.



1. Código open-source. Qualquer pessoa pode extendê-lo e usá-lo com diferentes datasets



2. Extrator de palavras relevantes

Conta-me Histórias API  
<http://bit.ly/ContameCode>

YAKE  
<http://yake.inesctec.pt>

## Referências

- Campos, R. and Mangaravite, V. and Pasquali, A. and Jorge, A.M. and Nunes, C. and Jatowt, A.: A Text Feature Based Automatic Keyword Extraction Method for Single Documents. ECIR, 684–691 (2018). **Best short paper ECIR 2018.**
- Rocha, C. and Jorge, A. and Sionara, R. and Brito, P. and Pimenta, C. and Rezende, S. PAMPO: using pattern matching and pos-tagging for effective Named Entities recognition in Portuguese, (2016)
- Silva, M. J. and Carvalho, P. and Costa, C and Sarmento L.: Automatic Expansion of a Social Judgment Lexicon for Sentiment Analysis. Technical Report TR 10-08. University of Lisbon, Faculty of Sciences, LASIGE, (2010)

## Agradecimentos

Este trabalho é parcialmente financiado pela ERDF através do programa COMPETE 2020 no âmbito do projeto POCI-01-0145-FEDER-006961, e por fundos nacionais através da FCT como parte do projeto UID/EEA/50014/2013. Foi também financiado pelo projeto “TEC4Growth - Pervasive Intelligence, Enhancers and Proofs of Concept with Industrial Impact/NORTE-01-0145-FEDER-000020” o qual foi financiado pelo NORTE 2020, no âmbito do Portugal 2020 e da ERDF.

