

conta-me  
histórias



# conta-me histórias

temporal summarization and the Portuguese  
web archive

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

TPDL 2018



about me  
**arian pasquali**

Researcher

University of Porto and LIAAD - INESC TEC



# challenges

- . information overload
- . credibility crisis
- . fake news and post-truth era
- . filter bubbles
- . content based machine learning models (e.g. fact-checkers)
- . more than ever web archives are crucial resources



# importance of web archives

the importance of preserving the history of the web

~ 2%

of the websites  
disappear  
**every week \***

~ 8%

of urls disappear  
**every year \*\***

\* Frank McCown, Catherine C. Marshall, and Michael L. Nelson. 2009. Why web sites are lost. *Commun. ACM* 52, 11 (November 2009), 141-145.

\*\* Legal Information Archive, The chesapeake project report. 2008

# 2018 Arquivo.pt announces a free text search API

## Challenge Arquivo.pt Prizes

- Use public API to create innovative ways to explore the web archive

<https://sobre.arquivo.pt/en/collaborate/submissions-for-the-arquivo-pt-prizes-2018/>



troika em Portugal  Pesquisa avançada

entre:   e:

Resultados 1 a 10 de 23.776.571

[Livro Os 10 erros da Troika em Portugal - Figueiras \(Porto\) ...](#)  
**15 Abril, 2015** - [Listar versões](#)

Livro Os 10 erros da **Troika em Portugal** - Figueiras (Porto) - Livros - Revistas Livro Os 10 erros da **Troika em Portugal** olx.pt Anunciar e Vender Anunciar Homepage Pesquisar My OLX Favoritos Início Lazer Livros - Revistas Livro Os 10 erros da **Troika em Portugal** Livro Os 10 erros da **Troika em** ...

<http://figueiras.olx.pt/livro-os-10-erros-da-troika-em-portugal-iid-463457489>

[Louça quer saber para onde vai 1.ª tranche da 'troika' - .....](#)  
**31 Maio, 2011** - [Listar versões](#)

Louça quer saber para onde vai 1.ª tranche da **'troika' - Portugal** - DN Última hora Cronologia dos ... ALUGUER/CRIME Mulheres que alugam o útero **em Portugal** ELEIÇÕES/PSD Passos Coelho diz que Santana ... o DN | Classificados | Loja do Jornal **Portugal** Globo Economia Ciência Artes TV & Media Opinião ...

[http://www.dn.pt/inicio/portugal/interior.aspx?content\\_id=1866281](http://www.dn.pt/inicio/portugal/interior.aspx?content_id=1866281)

["Este programa está muito além do memorando" da 'troika' - .....](#)  
**8 Maio, 2011** - [Listar versões](#)

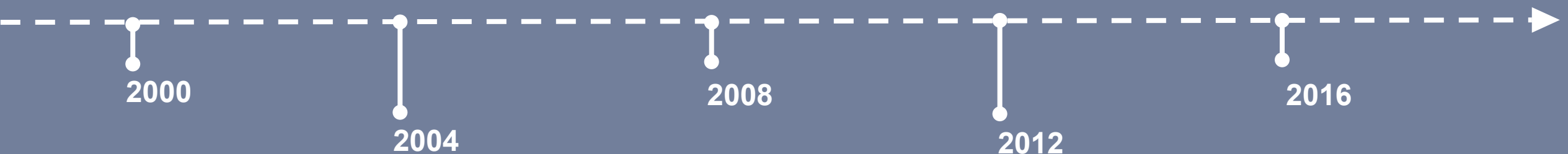
"Este programa está muito além do memorando" da **'troika' - Portugal** - DN Última hora PCP deixa **em** ... Espanha e no YouTube AJUDA EXTERNA Reino Unido põe **em** causa ajuda financeira a **Portugal** COCA COLA ... 'AJUDA EXTERNA Reino Unido põe **em** causa ajuda financeira a **Portugal** EUROVISÃO/EX-CONCORRENTES "A ...

[http://www.dn.pt/inicio/portugal/interior.aspx?content\\_id=1847261](http://www.dn.pt/inicio/portugal/interior.aspx?content_id=1847261)

# idea

what if we could generate a timeline about any given topic automatically?

can we build a left-wing versus right-wing narrative timeline?





# timelines are a natural choice to explore long term stories

## FACEBOOK

News Feed Updates: A Timeline

- AUG '14** **Click bait gets booted**  
Facebook updates the algorithm to de-prioritize headlines it sees as click bait.
- SEPT '14** **Trending factors in**  
Posts that are timely or address trending topics rise to the top of the News Feed.
- NOV '14** **Promotional posts sink**  
Posts that are deemed overly promotional on 3 criteria see dramatically lower reach.
- APR '15** **Pages appear lower**  
Facebook prioritizes updates from friends and family over posts from brands in the feed.
- JUN '15** **Video actions matter**  
Factors like turning volume on, going full screen, and more now count toward a video's popularity.
- JUL '15** **Users take control**  
Facebook gives users tools to prioritize people, and un-follow, re-follow, and discover new pages.

Keep followers engaged no matter what happens next.  
WWW.PAGEMODO.COM

## The Biggest Security Business News of 2017

## A TIMELINE OF THIS YEAR'S TOP STORIES

as reported by **Yicis**

- JANUARY**
  - 05 Europe PE buyers face higher prices in January
  - 14 European parliament urges tougher rules for plastics recycling
- FEBRUARY**
  - 21 Barcelona shifts investment opportunity focus to eastern Europe fertilizers
  - 27 Europe methanol prices firm in light of tight supply and high demand
- MARCH**
  - 05 Walk overseas demand prompts China slowdown
  - 19 PETRONAS to consider alternative growth paths
- APRIL**
  - 03 Qatar's Montajat plans 21 new offices by the end of 2014
  - 14 Petrochemicals and fertilizers growth in Qatar presents new revenue opportunities
  - 25 Green chemicals to serve richer markets amid production challenges
- MAY**
  - 05 Expanded polystyrene now increasingly compatible for recyclability

## HIGHLIGHTS OF... RED/BLUE DIALOGUE

...WITHIN THE NCDD COMMUNITY

Prepared for NCDD's May 15, 2013 Confab Call. Details at [www.ncdd.org/11337](http://www.ncdd.org/11337).

- 2004** Let's Talk America! SERIES OF REUNITING AMERICA / TRANSPARTISAN ALLIANCE CONFERENCES AND MEETINGS
- 2005** NCDD is a network of thousands of innovators who bring people together across divides to tackle today's toughest challenges.
- 2006** = NCDD's BIENNIAL EVENTS
- 2007** YOU'RE NOT AS CRAZY AS I THOUGHT
- 2008** CONSERVATIVES PANEL AT NCDD AUSTIN WITH GROVER NORQUIST
- 2009** NCCD Confab with Guests
- 2010** NO LABELS
- 2011** NCDD CATALYST AWARDS ANNOUNCED (with political bridge-building category)
- 2012** REAL DIALOGUES PROJECT WINS NCDD'S BRIDGE-BUILDING CATALYST AWARD
- 2013** 2012 ELECTORAL MAP

## Google News Timeline

Magazines | time | Add query | About Timeline | Sign in

Name of magazine, e.g. Time

Time Magazine | News | Wikipedia Events | Add More Queries

Show: Year | Size: Medium | Date: | Go

Year	1998	1999	2000	2001	2002	2003	2004	2005
Time for 1998	Time for 1999	Time for 2000	Time for 2001	Time for 2002	Time for 2003	Time for 2004	Time for 2005	
Time 1/12/98	Time 1/11/99	Time 1/11/00	Time 1/8/01	Time 1/14/02	Time 1/13/03	Time 1/12/04	Time 1/10/05	
That's All Folks	THE FUTURE OF MEDICINE	WHERE DOES HE BELONG?	HOW TO SURVIVE THE SLUMP	FLAT-OUT COOL!	THE BIGGER THREAT?	WHO IS THE REAL DEAN?	THE TSUNAMI	
7 Bullmore is 5 after his 1	January 1st, 1999 Russia begins to circulate new rubles to stem inflation and promote confidence	January 1st, 2000 YZK passes without serious, widespread computer failures, despite fears to the contrary	January 1st, 2001 A black month measuring surveillance treaty, initially signed in 1992, officially enters into force	January 1st, 2002 The Open Skies mutual surveillance treaty, initially signed in 1992, officially enters into force	January 1st, 2003 Pascal Couchepin becomes President of the Confederation in Switzerland	January 1st, 2004 Pervez Musharraf wins a vote of confidence from an electoral college consisting of the provincial assemblies, confirming him as President	January 4th, 2005 Gunmen assassinate the Governor of Baghdad, Ali Al-Hadi	January 6th, 2005 The Graniteville train disaster kills 9 and injures 250 in Graniteville, South Carolina
97 rocket 3 military 00d shortly after 1 Cape edia	January 2nd, 1999 A snowstorm leaves 14 inches (355 mm) of snow in Milwaukee, Wisconsin and 19 inches (487 mm) in Chicago, Illinois, killing 68	January 3rd, 2000 Israel and Syria hold inconclusive peace talks	January 3rd, 2001 A black month measuring surveillance treaty, initially signed in 1992, officially enters into force	January 3rd, 2002 Euro notes and coins are issued in France, Portugal, Greece, Finland, Luxembourg, Belgium, Austria, Ireland and the Netherlands	January 3rd, 2003 The Ohio State University defeats the Ohio State Buckeyes in a double-overtime in the Fiesta Bowl, 31-24, for the national Bowl Championship Series (BCS)	January 3rd, 2004 Parliament and the provincial assemblies, confirming him as President	January 4th, 2005 Gunmen assassinate the Governor of Baghdad, Ali Al-Hadi	January 6th, 2005 The Graniteville train disaster kills 9 and injures 250 in Graniteville, South Carolina
97 est Rwanda, 8a members 3 workers, 3 ously	January 2nd, 1998 A gunman shoots Antonio Teodoro Filho, Brazilian politician and radio presenter, during a broadcast	January 3rd, 2000 Israel and Syria hold inconclusive peace talks	January 3rd, 2001 A black month measuring surveillance treaty, initially signed in 1992, officially enters into force	January 3rd, 2002 Euro notes and coins are issued in France, Portugal, Greece, Finland, Luxembourg, Belgium, Austria, Ireland and the Netherlands	January 3rd, 2003 The Ohio State University defeats the Ohio State Buckeyes in a double-overtime in the Fiesta Bowl, 31-24, for the national Bowl Championship Series (BCS)	January 3rd, 2004 Parliament and the provincial assemblies, confirming him as President	January 4th, 2005 Gunmen assassinate the Governor of Baghdad, Ali Al-Hadi	January 6th, 2005 The Graniteville train disaster kills 9 and injures 250 in Graniteville, South Carolina
97 rms to e than 30 celebrations r of the last West Bank	January 4th, 1999 A gunman shoots Antonio Teodoro Filho, Brazilian politician and radio presenter, during a broadcast	January 3rd, 2000 Israel and Syria hold inconclusive peace talks	January 3rd, 2001 A black month measuring surveillance treaty, initially signed in 1992, officially enters into force	January 3rd, 2002 Euro notes and coins are issued in France, Portugal, Greece, Finland, Luxembourg, Belgium, Austria, Ireland and the Netherlands	January 3rd, 2003 The Ohio State University defeats the Ohio State Buckeyes in a double-overtime in the Fiesta Bowl, 31-24, for the national Bowl Championship Series (BCS)	January 3rd, 2004 Parliament and the provincial assemblies, confirming him as President	January 4th, 2005 Gunmen assassinate the Governor of Baghdad, Ali Al-Hadi	January 6th, 2005 The Graniteville train disaster kills 9 and injures 250 in Graniteville, South Carolina
97 ident Bill ed for his	January 4th, 1999 A gunman shoots Antonio Teodoro Filho, Brazilian politician and radio presenter, during a broadcast	January 3rd, 2000 Israel and Syria hold inconclusive peace talks	January 3rd, 2001 A black month measuring surveillance treaty, initially signed in 1992, officially enters into force	January 3rd, 2002 Euro notes and coins are issued in France, Portugal, Greece, Finland, Luxembourg, Belgium, Austria, Ireland and the Netherlands	January 3rd, 2003 The Ohio State University defeats the Ohio State Buckeyes in a double-overtime in the Fiesta Bowl, 31-24, for the national Bowl Championship Series (BCS)	January 3rd, 2004 Parliament and the provincial assemblies, confirming him as President	January 4th, 2005 Gunmen assassinate the Governor of Baghdad, Ali Al-Hadi	January 6th, 2005 The Graniteville train disaster kills 9 and injures 250 in Graniteville, South Carolina

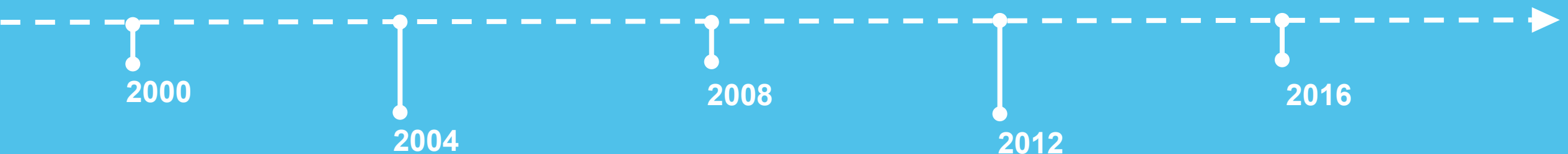


# problem definition

given a search query  
and a **timeframe**



generate a timeline  
summarization with relevant  
dates and headlines



# steps

1. query news articles using arquivo.pt API

1. identify relevant time frames

1. identify relevant headlines



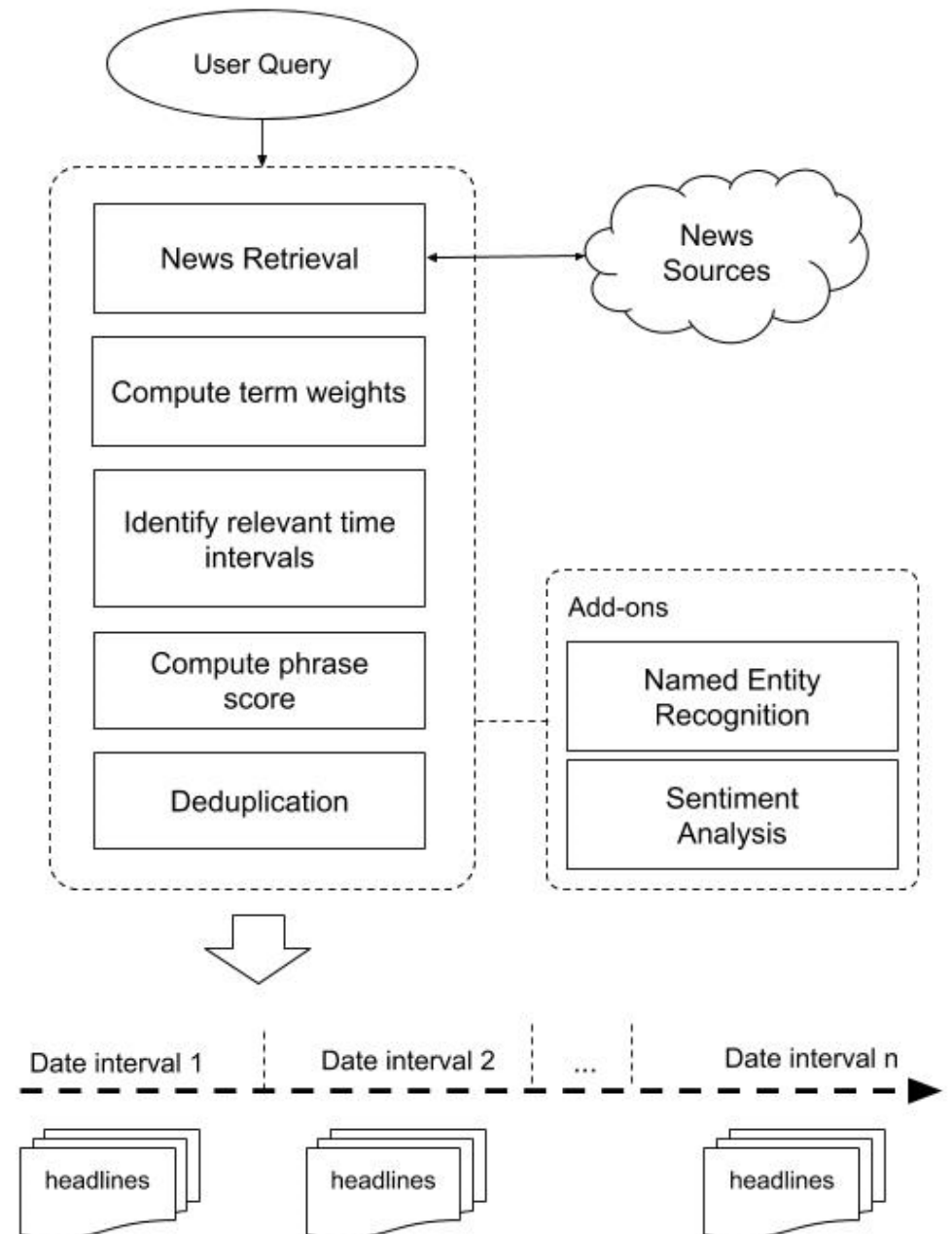
# how to calculate headline relevance?

we apply principles of keyword extraction  
algorithms to calculate headlines  
relevance from a particular time frame



# architecture

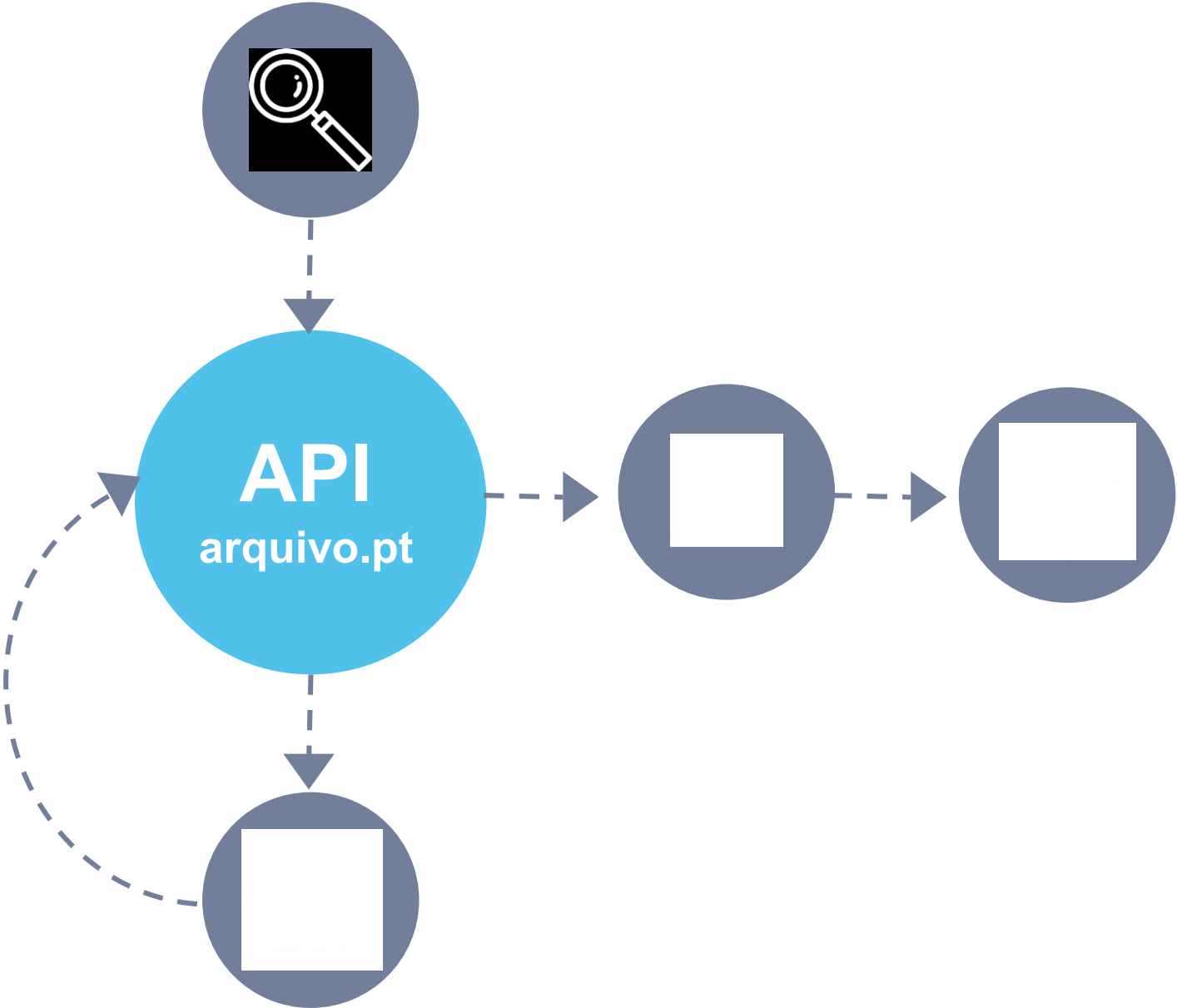
- given a search query and a timeframe
- apply principles of keyword extraction to compute relevant headlines from a particular time frame



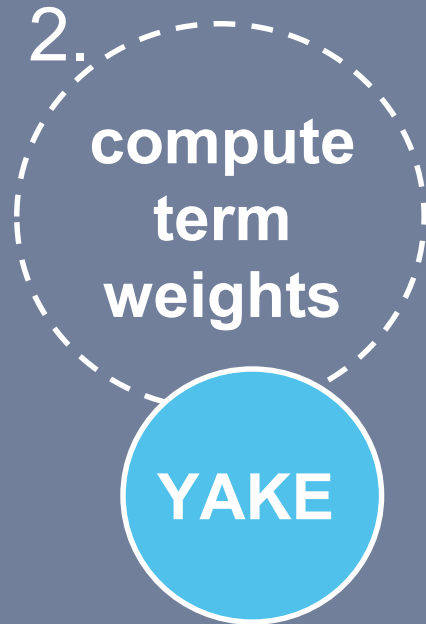
# architecture



# architecture

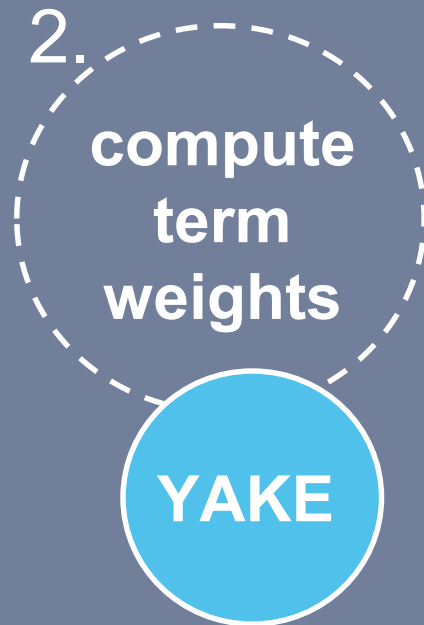


# architecture





# architecture



## More information

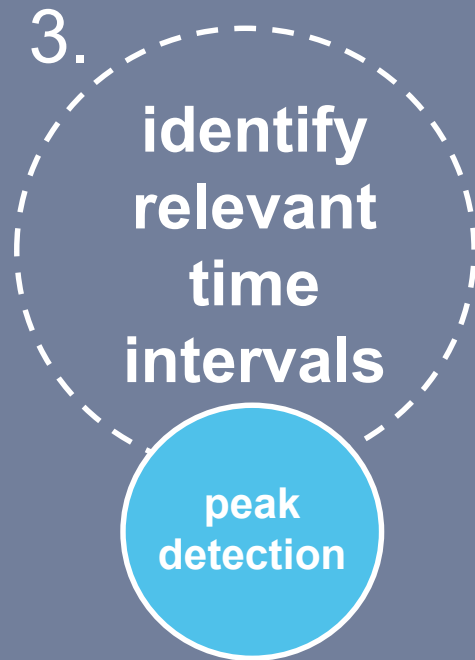
Campos, R., & Mangaravite, V., & Pasquali, A., & Jorge, A., & Nunes, C., & Jatowt, A. (2018). A Text Feature Based Automatic Keyword Extraction Method for Single Documents. Best Short Paper of ECIR'18 (<http://bit.ly/YakeDemoECIR2018>)

$$S(t) = \frac{TContext * TPos}{TCase + \frac{TFreq}{TContext} + \frac{TSent}{TContext}}$$

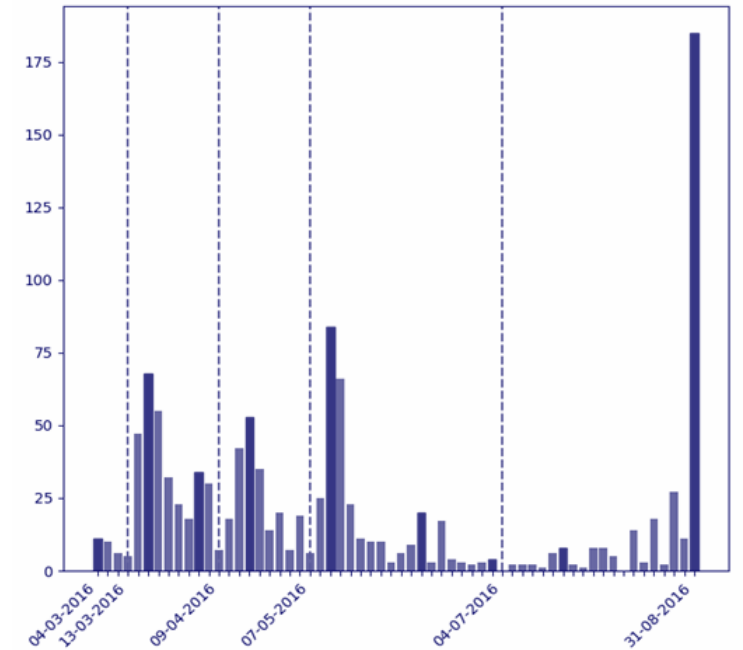
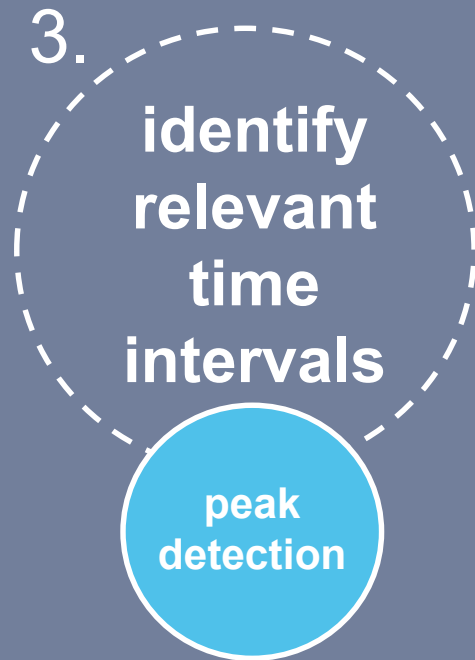
where:

- **TCase** - Casing
  - Reflects case aspect of a particular word
- **TPos** - Term position
  - Relevant terms tend to occur at the beginning of the document
- **TFreq** - Term frequency
  - Assumption that relevant words occur less
- **TSent** - term frequency on different headlines
  - How often a word appears within different headlines
- **TContent** - Term relatedness to context
  - Number of different terms that occur to the left (or right)
  - the more the number of different words around the the candidate word, more meaningless the word is likely to be

# architecture



# architecture



1. We aggregate count of news by date

2. We then detect peaks using scipy's relative extrema function

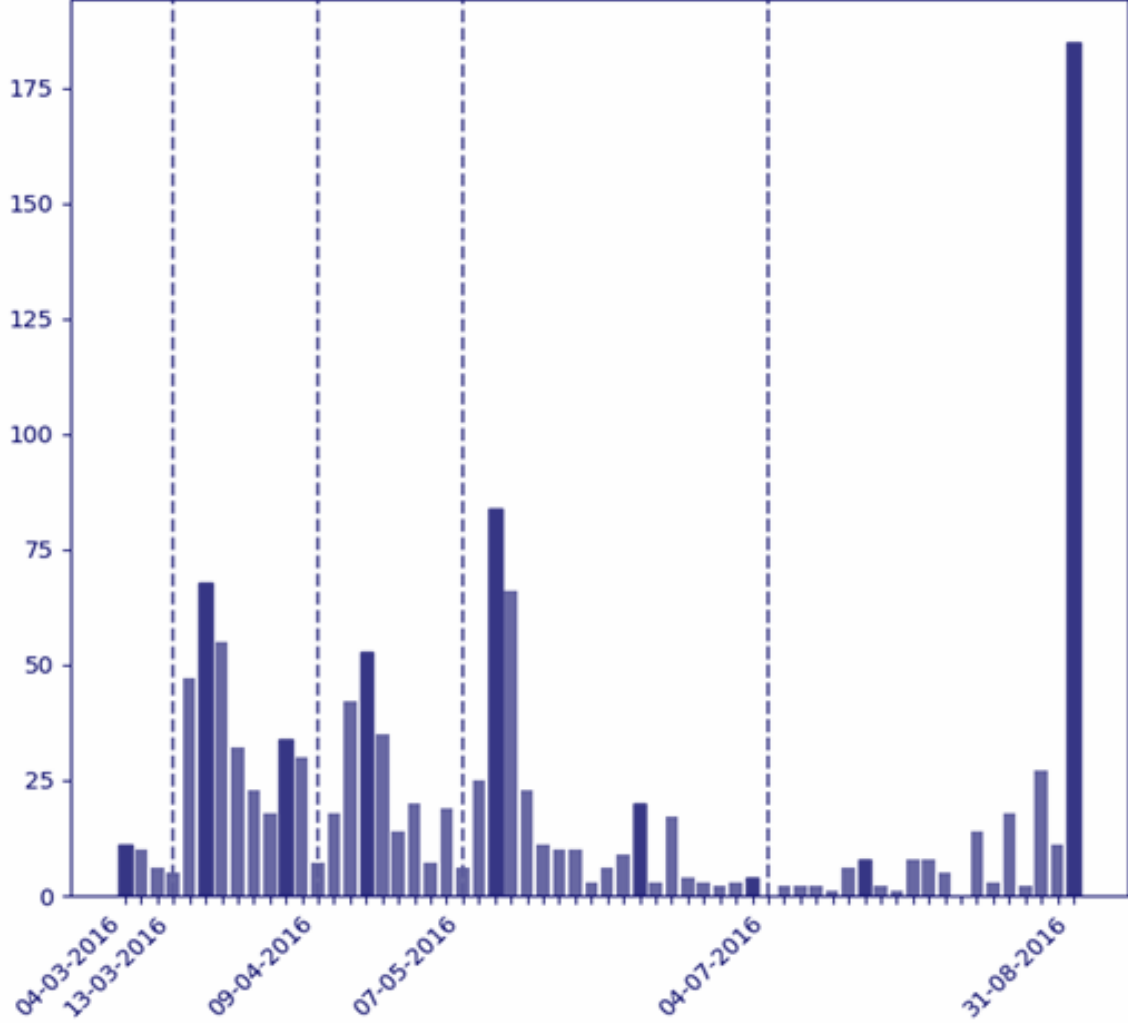
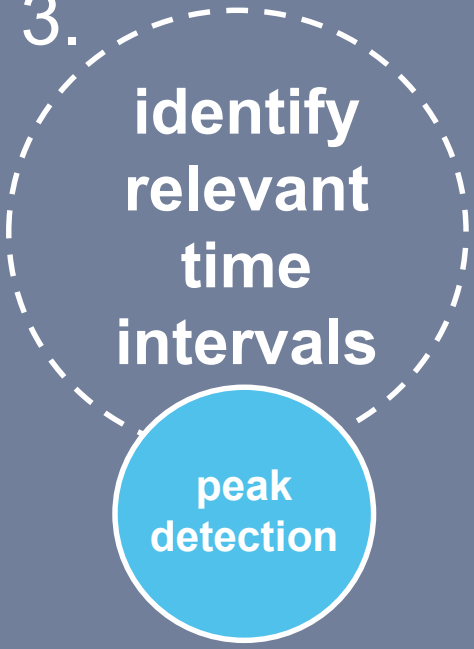
```
from scipy.signal import argrelextrema
```

```
x = np.array([2, 1, 2, 3, 2, 0, 1, 0])  
argrelextrema(x, np.greater, order=1)
```

```
>> (array([3, 6]),)
```

# architecture

3.

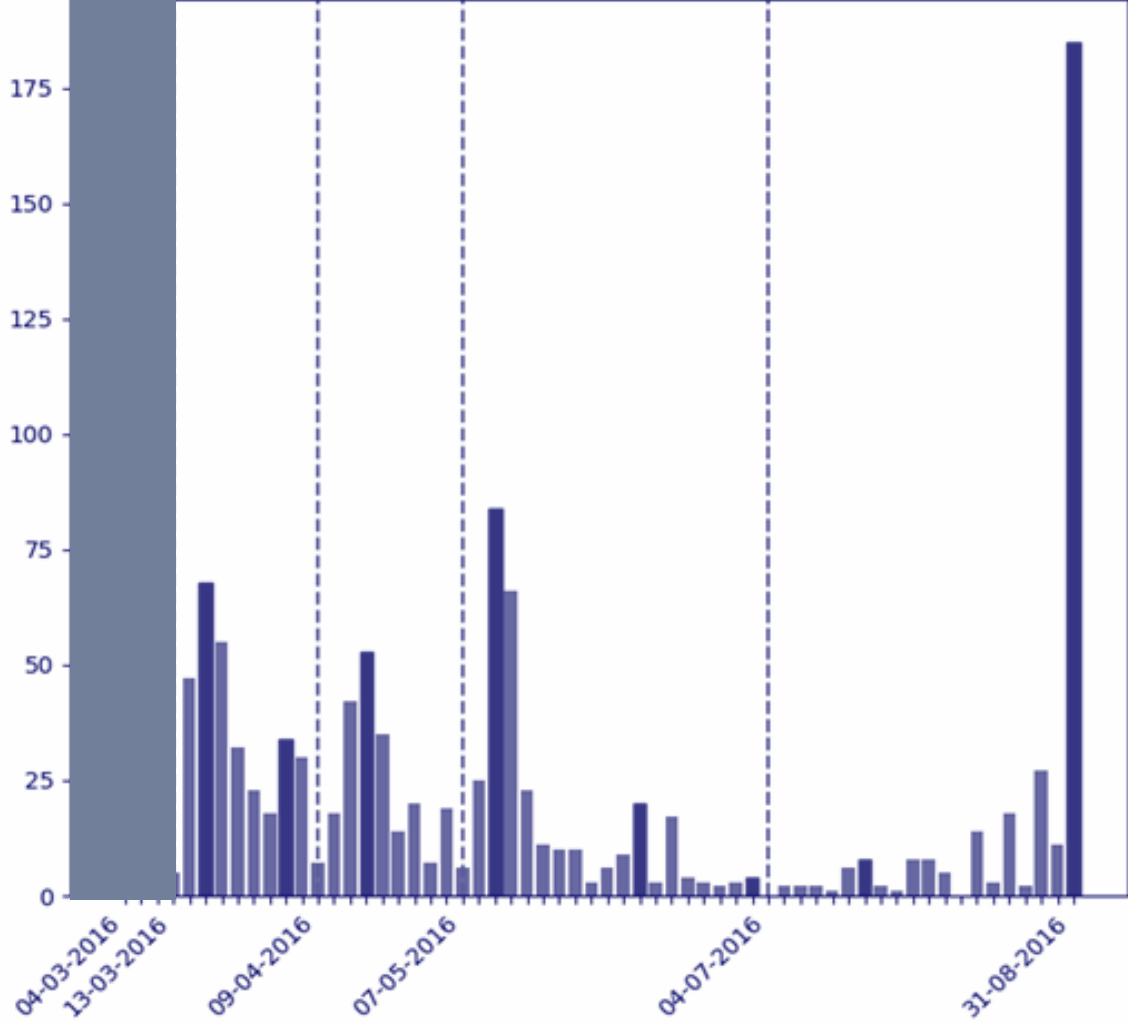


# architecture

3.

identify relevant time intervals

peak detection

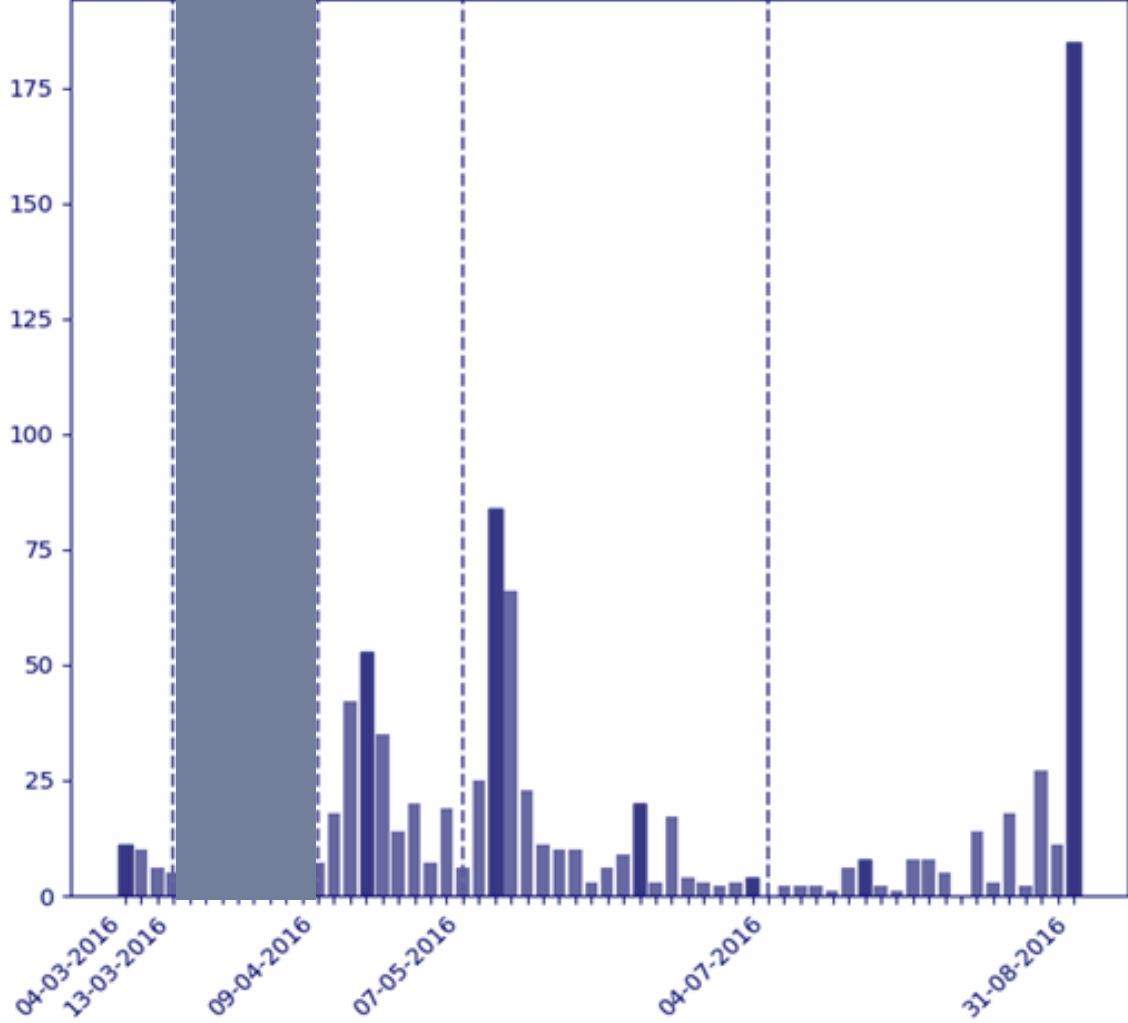


# architecture

3.

identify relevant time intervals

peak detection

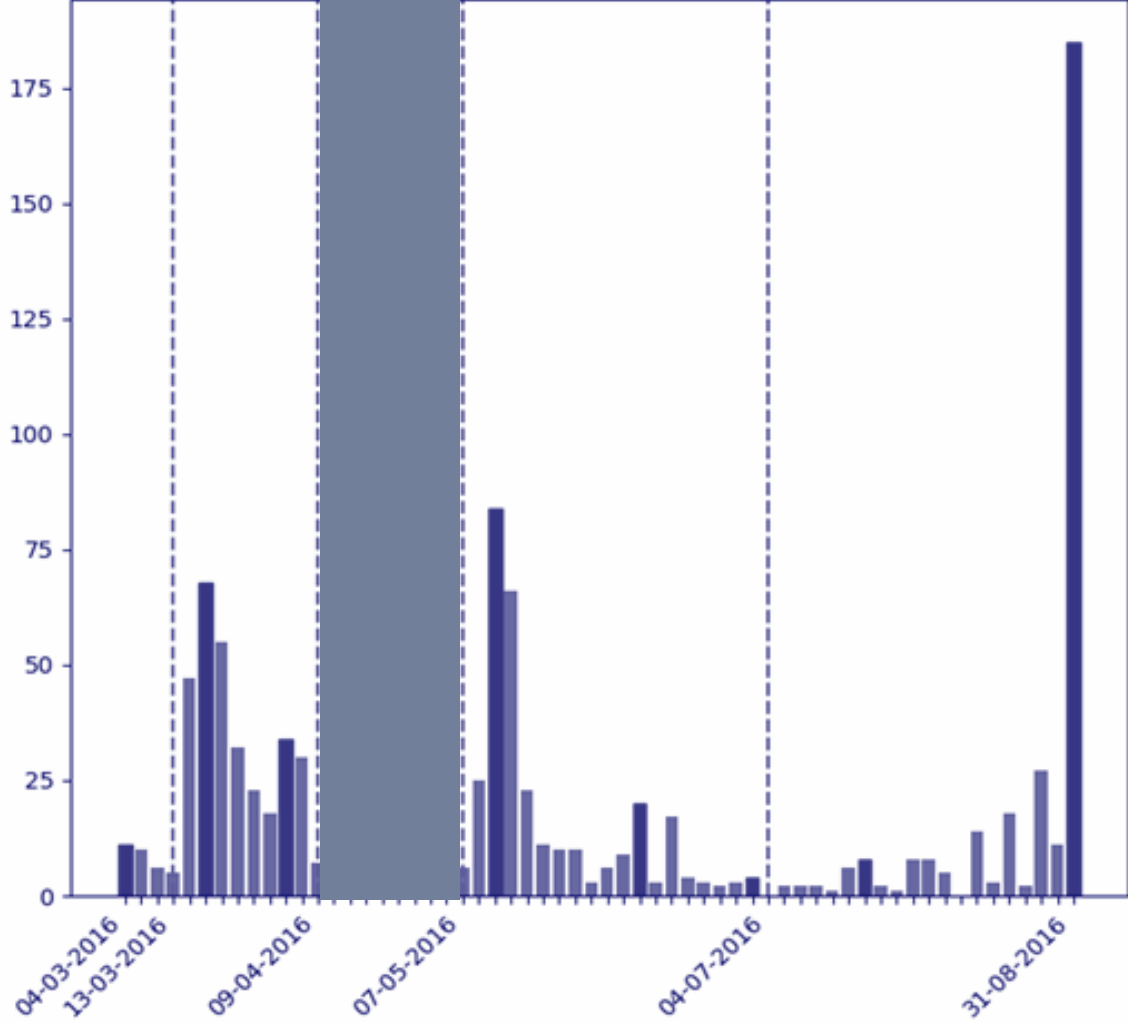


# architecture

3.

identify relevant time intervals

peak detection



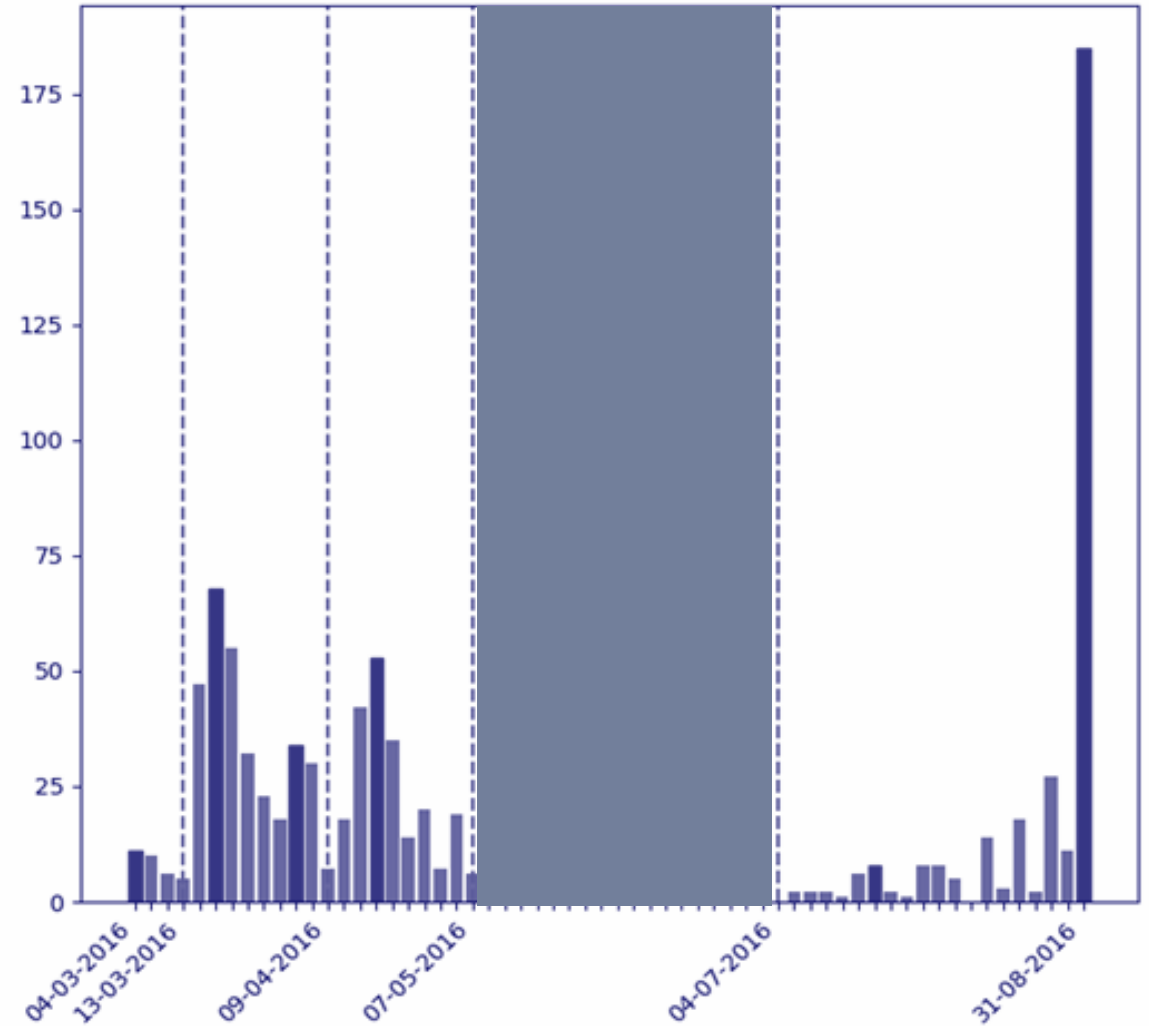


# architecture

3.

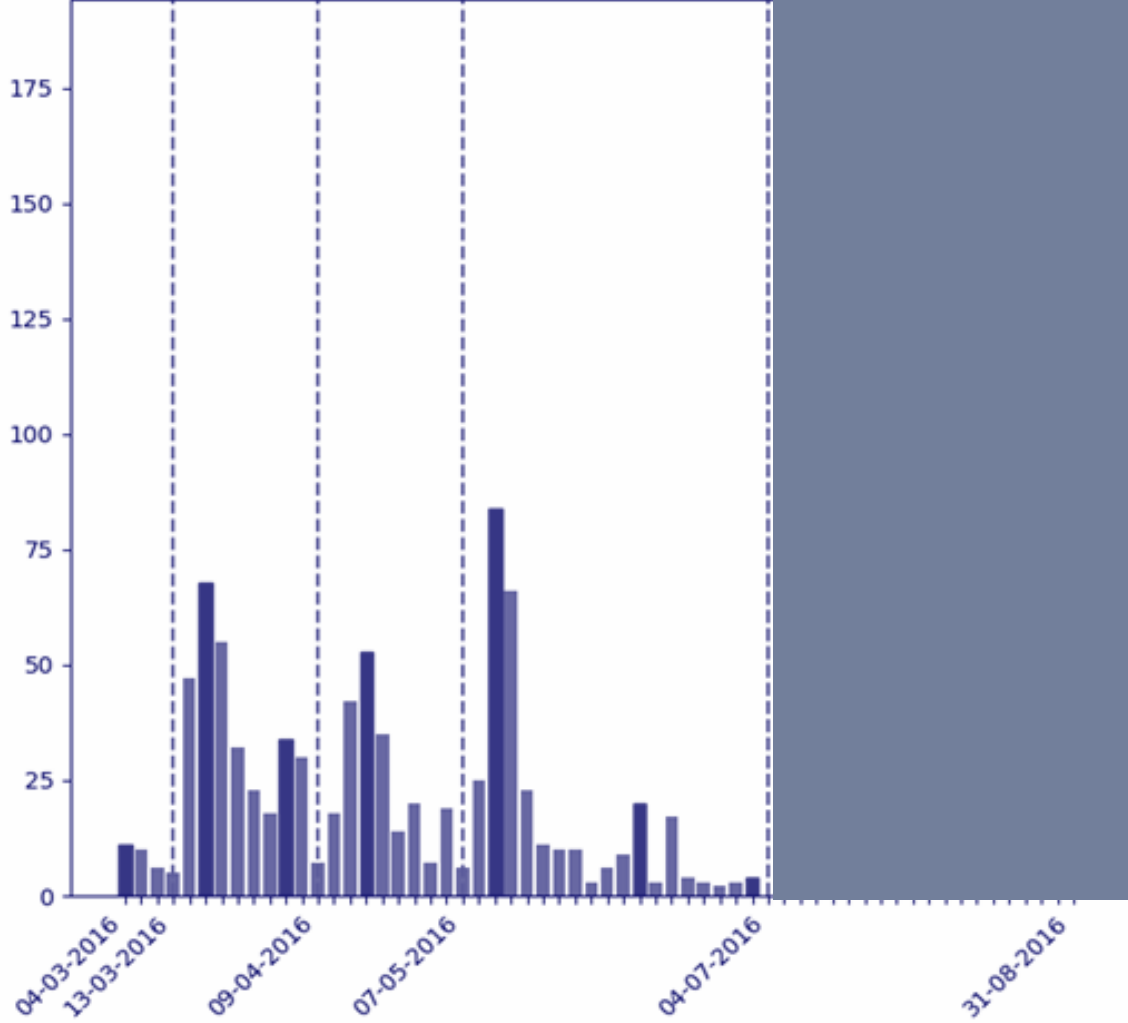
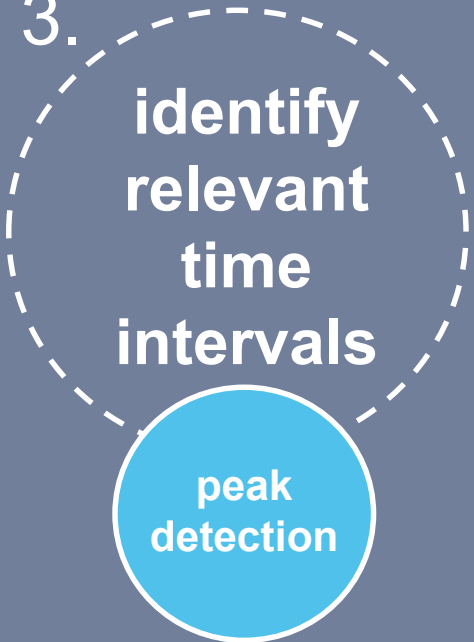
identify  
relevant  
time  
intervals

peak  
detection



# architecture

3.



# architecture



# architecture



$$S(kw) = \frac{\prod_{w \in kw} S(w)}{\text{TF}(kw) + (1 + \sum_{w \in ks} S(w))}$$

Considering headlines from each time interval where:

- **S(w)** - term weight
  - Considers weights from every term in the headline
- **TF(kw)** - Headline frequency
  - How many times this headline occurs
- **S(kw)** - Final headline score

## More information

Campos, R., & Mangaravite, V., & Pasquali, A., & Jorge, A., & Nunes, C., & Jatowt, A. (2018). A Text Feature Based Automatic Keyword Extraction Method for Single Documents. Best Short Paper of ECIR'18

(<http://bit.ly/YakeDemoECIR2018>)

# architecture

5.

user  
interface

timeline

dez 2012

nov 2013

fev 2015

## - Relevant headlines from 27, April 2014 until 20, February 2015

Jornalista e fotógrafo do The Times feridos na Síria. negative

China e Rússia vetam proposta para julgar no TPI crimes de guerra na Síria. negative

Guerra síria contamina ainda mais o Líbano. negative

Irão liga ajuda contra Estado Islâmico a negociações nucleares.

ONU acusa Síria e Estado Islâmico de crimes de guerra contra a humanidade. negative

Casa Branca confirmou morte de norte-americano jihadista na Síria. negative

Obama prepara mapa para atacar jihadistas na Síria.

ONU quer governo e Estado Islâmico julgados por crimes de guerra. negative

Obama afasta para já intervenção militar na Síria. negative

Guerra civil na Síria fez mais de três milhões de refugiados. negative

Primeiro-ministro britânico não afasta acção contra Estado Islâmico na Síria. negative

Vídeo com refém britânico compara intervenção na Síria à do Vietname.

Síria diz ter sido informada por Washington dos ataques contra o Estado Islâmico.

Presidente da síria fez rara aparição pública em mesquita de damasco.

Oposição Síria condena decapitação de cidadão britânico. negative

Turquia quer mísseis da NATO apontados para a Síria.

Turcos dão luz verde ao Exército Livre da Síria e travam curdos. positive

Criminoso de guerra nazi Brunner morreu provavelmente na Síria. negative

Terminou o ano com mais mortos na guerra civil síria. negative

# more detailed information at

## **A Text Feature Based Automatic Keyword Extraction Method for Single Documents \***

Campos, R., & Mangaravite, V., & Pasquali, A., & Jorge, A., & Nunes, C., & Jatowt, A. (2018).  
40th European Conference on Information Retrieval (ECIR'18).  
Grenoble, France. March 26 – 29. (Vol. 10772(2018), pp. 684 - 691).

## **YAKE! Collection-independent Automatic Keyword Extractor \*\***

Campos, R., & Mangaravite, V., & Pasquali, A., & Jorge, A., & Nunes, C., & Jatowt, A. (2018).  
40th European Conference on Information Retrieval (ECIR'18).  
Grenoble, France. March 26 – 29. (Vol. 10772(2018), pp. 806 - 810).

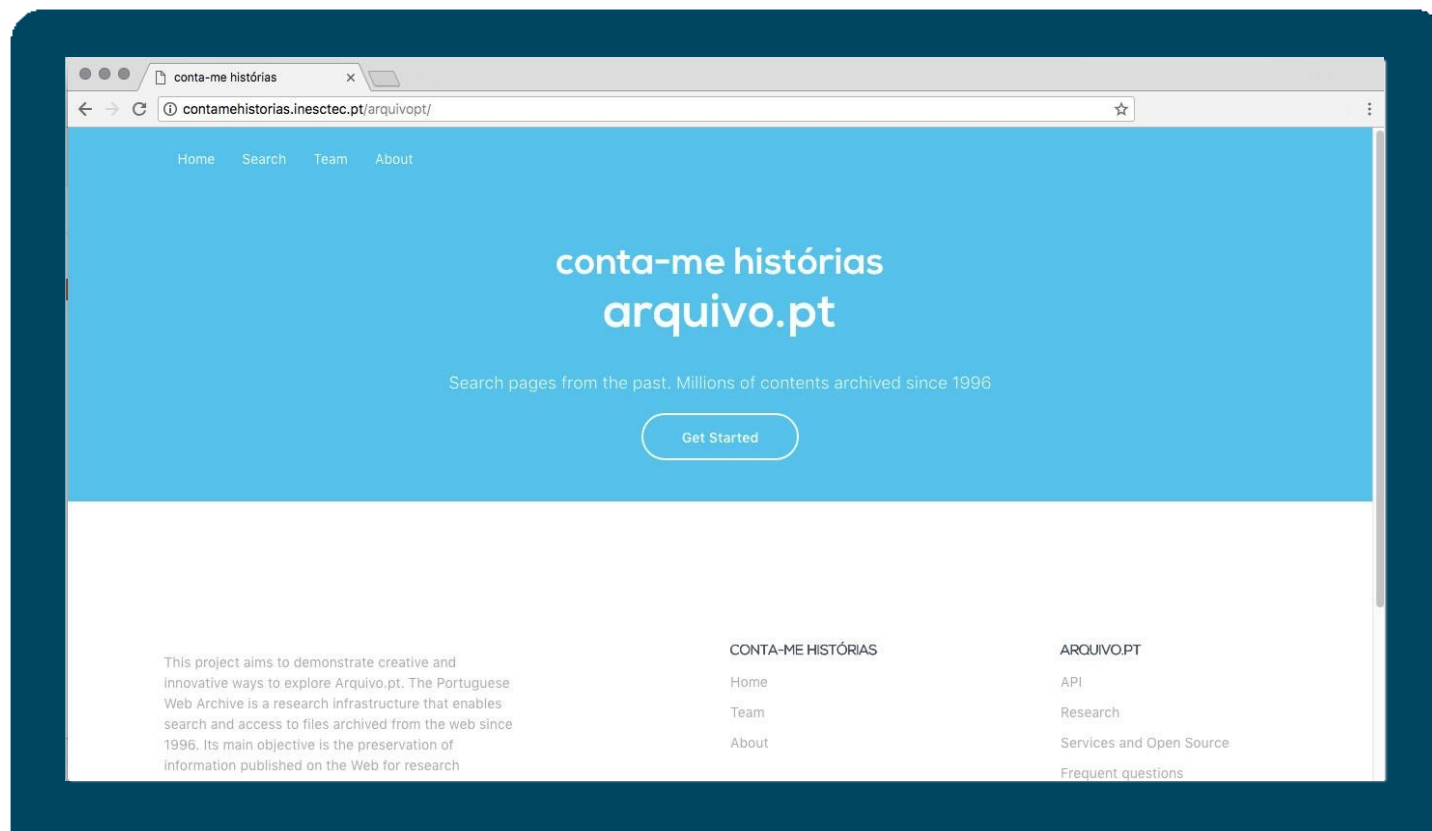


\* ECIR 2018 Best Short Paper Award

\*\* Demo

<http://bit.ly/YakeDemoECIR2018>

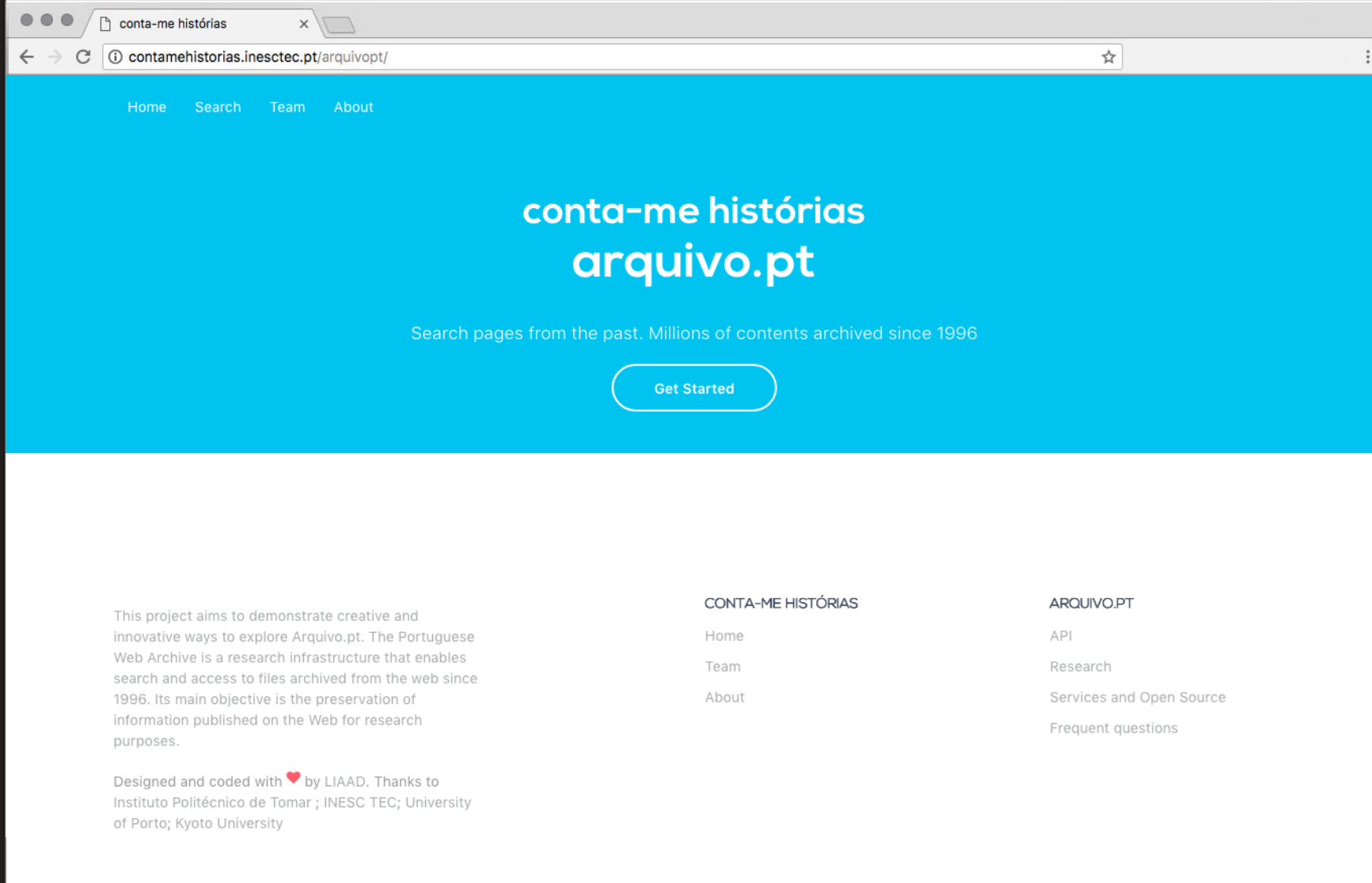
# conta-me histórias pt





<http://contamehistorias.pt>





# conta-me histórias arquivo.pt

Search pages from the past. Millions of contents archived since 1996

Get Started

This project aims to demonstrate creative and innovative ways to explore Arquivo.pt. The Portuguese Web Archive is a research infrastructure that enables search and access to files archived from the web since 1996. Its main objective is the preservation of information published on the Web for research purposes.

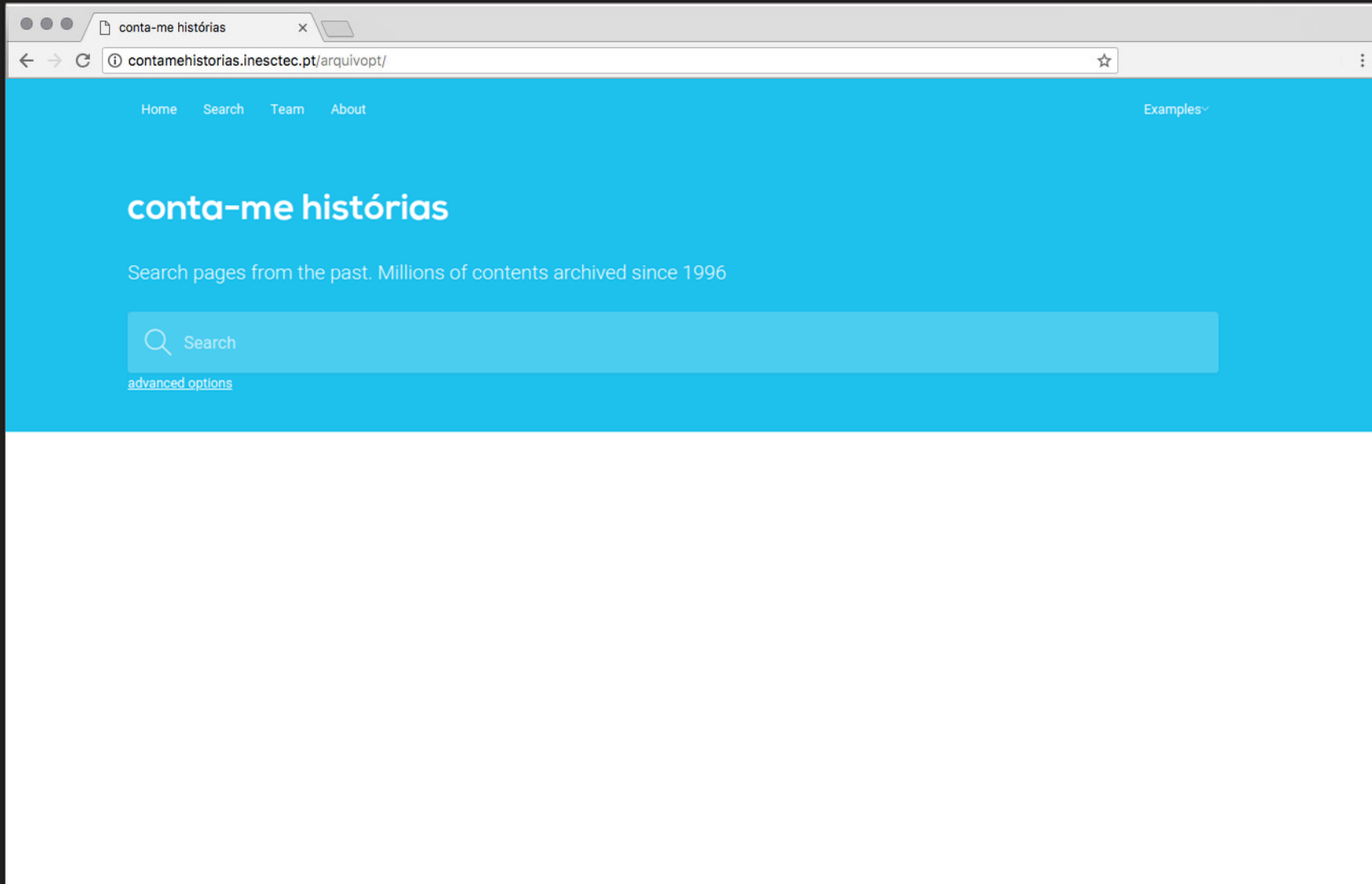
Designed and coded with ❤️ by LIAAD. Thanks to Instituto Politécnico de Tomar ; INESC TEC; University of Porto; Kyoto University

## CONTA-ME HISTÓRIAS

- [Home](#)
- [Team](#)
- [About](#)

## ARQUIVO.PT

- [API](#)
- [Research](#)
- [Services and Open Source](#)
- [Frequent questions](#)



# conta-me histórias

Search pages from the past. Millions of contents archived since 1996

🔍 **troika em portugal**

[advanced options](#)

7855 processed pages in 16.03 seconds (7.29 seconds of query time and 8.75 seconds spent analyzing content)

**7855**  
PAGES FOUND

**2290**  
PAGES ANALYZED

**17**  
UNIQUE DOMAINS

Narrative Related terms

**Narrative Timeline of Troika em Portugal** during the last 10 years

Horizontal Vertical

[Previous](#)

[Next](#)

conta-me histórias x

contamehistorias.inesctec.pt/arquivopt/

Sep 2011 Mar 2012 Mar 2013 Nov 2013 Feb 2015 Nov 2016

## - Relevant stories from **02, August 2010** until **27, September 2011**

2011-04-30 Troika vai demolir proteção dos trabalhadores portugueses **positive**  
PSD apresenta programa eleitoral em Maio  
Del Potro vence Estoril Open  
Durão Barroso diz que é indispensável que Portugal cumpra objectivos **positive**  
Desemprego em Espanha baixa em Abril **negative**  
Portugal deve ir mais além das medidas acordadas com a troika  
FIFA confirma que FC Porto é o clube português com mais títulos  
FMI e CE preocupados com hipótese de Portugal falhar metas **negative**  
Manuel Alegre e António Capucho fazem primeira análise do acordo **positive**

2011-06-12 Ministros europeus das Finanças aprovam hoje assistência financeira a Portugal **positive**  
Aprovação de pacote de austeridade é alívio para a Europa e justa para Portugal **positive**  
Portugal volta hoje ao mercado ensombrado pela Moody  
Juros de Portugal disparam em dia de leilão  
António José Seguro  
Ferreira do Amaral defende saída digna da moeda única **positive**  
Portugal recebe cinco ME em compensações europeias  
apresenta conclusões de primeira avaliação a Portugal  
Mário Soares alerta que a Grécia cair **negative**  
Eurogrupo expressa confiança na capacidade de Portugal **positive**

Sources: Ionline, Tsf, Sabado, Economico Sapo, Dinheirovivo, News Google, Lux lol, Aeiou,



Listar versões

2011

Maio

Junho

- 2
- 5
- 7
- 9
- 10
- 12 **00:06**
- 16
- 19
- 20
- 21
- 22
- 23
- 24



Actualidade | Programas | Multimédia | Blogues | Futebol | **Bolsa** | Trânsito | Totojogos | Meteorologia | Marés | Farmácias | Agenda | Jogos

PSI20 | PSI Geral | IBEX35 | CAC 40 | Euronext | FTSE | NASDAQ 100 | Outros Mercados

E.coli | Christine Lagarde | dívida | Eurogrupo | Calçado | FMI | Grécia | sindicatos

Noticiário das 00h00 | Fórum TSF | Podcast | RSS | FM | MOBILE | SMS | Fale connosco

login | registo

Economia

Siga-nos em [Facebook](#) | [Twitter](#)

Administração Pública | Orçamento | Bolsa | Empresas

Patrocínio

---

Ferramentas

Enviar | Partilhar | Imprimir | Aumentar | Diminuir

Comentar este Artigo

---

Estadísticas

515 Visualizações

0 Comentários

0 Envios

### Ministros europeus das Finanças aprovam hoje assistência financeira a Portugal

16 MAI 11 às 06:29

Os ministros das Finanças da Zona Euro e da União Europeia deverão aprovar sem dificuldade, esta segunda-feira em Bruxelas, o programa de assistência financeira a Portugal, depois de ter sido afastada a ameaça da Finlândia se opor à decisão.

O acordo político deverá ser alcançado numa reunião, esta tarde às 17h00 (16h00 de Lisboa), dos responsáveis da Zona Euro alargada aos ministros da UE (os 17 anteriores mais dez que não pertencem à Zona Euro).

A assistência financeira a Portugal será repartida em partes iguais de 26 mil milhões de euros pelo Fundo Europeu de Estabilização Financeira (FEFF).

Últimas | + Vistas | + Comentadas | + Pesquisadas

Projecto Novos Povoadores é alternativa à vida na cidade  
Ontem às 23:44

Governador do Banco de Israel é o novo candidato ao FMI  
Ontem às 23:16

Explosões no Paquistão causam 25 mortos e 100 feridos  
Ontem às 22:50

Cáritas pede criação de rede de protecção social  
Ontem às 22:39

Blue Note é a Melhor Editora de Jazz do Mundo  
Ontem às 22:24

conta-me histórias x

contamehistorias.inesctec.pt/arquivopt/

Sep 2011 Mar 2012 Mar 2013 Nov 2013 Feb 2015 Nov 2016

### - Relevant stories from **01, July 2012** until **24, March 2013**

Avaliação positiva da Troika contribui para melhorar imagem do país positive

Seleção nacional de regresso a Portugal

FMI diz que vai discutir mais austeridade na próxima avaliação da troika positive

Portugal devia mostrar relatório sobre pobreza à troika negative

Troika regressa a Portugal em semana dominada por Miguel Relvas negative

Estela Barbot acredita que troika terá bom senso na próxima avaliação positive

Deputado do CDS pede que troika corte nos juros

Berlim e EUA saudam esforços de Portugal para cumprir programa

Portugal merece ser recompensado pela troika

Passos Coelho não garante que IRS não vá aumentar

Troika obriga Portugal a pagar juros de dinheiro que não utiliza

Juros da dívida sobem em Portugal e Espanha

Juros da dívida aliviam em Portugal antes de leilão

Angela Merkel convencida de que Portugal vai cumprir memorando positive

Gaspar cita Argentina para dizer que Portugal tem de honrar compromissos positive

Portugal já merecia uma atenção especial da troika positive

Portugal desaconselhado a pedir condições iguais às da Grécia

Portugal Ways Lounge

Portugal conquista três prémios no turismo mundial positive

Presidente da AICEP defende mais Economia e menos Finanças

SOURCES: Ionline, Tsf, Sabado, Economico Sapo, Dinheirovivo, News Google, Lux lol, Aeiou,

Sentiment automatically detected based on the words that are part of the sentence



conta-me histórias x

contamehistorias.inesctec.pt/arquivopt/

Sep 2011 Mar 2012 Mar 2013 Nov 2013 Feb 2015 Nov 2016

## - Relevant stories from **06, February 2015** until **16, November 2016**

2015-02-28 Portugal reembolsará seis mil milhões de euros ao FMI em Março

Paulo Portas diz ter havido uma troika de restrições e uma troika de bem-fazer

Portugal não vai recuperar nem metade dos empregos destruídos pela troika positive

Troika culpa Governo português por falhanço do programa negative

Passos e Cavaco são a troika em Portugal

Parlamento Europeu vai investigar erros da troika em Portugal positive

Troika regressa a Portugal na próxima semana

Troika regressa a Lisboa para segunda avaliação pós-resgate

Paulo Portas numa feira em Bruxelas lembra que Portugal não é a Grécia

chega hoje a Lisboa para debater TSU e Orçamento do Estado

Troika volta a Portugal e não é para matar saudades

Costa diz que visita da troika a Portugal

Banda desenhada do vinho do Porto na Garrafeira Nacional

The Guardian e Lonely Planet escolhem as melhores dicas sobre Lisboa

Germano Silva vai ser Doutor Honoris Causa pela Universidade do Porto positive

Atletas gregos sem dinheiro lutam para competir nos Jogos do Rio de Janeiro positive

Banco de Portugal e CMVM

Troika em Portugal para analisar contas públicas

Presidente da Águas de Portugal demite-se em ruptura com o Governo

Carlos Costa diz que CMVM devia ser informada pelo BES

SOURCES: Ionline, Tsf, Sabado, Economico Sapo, Dinheirovivo, News Google, Lux lol, Aeiou,



publico.pt

entre: 01/01/1996 e: 31/12/2017

[Pesquisa avançada](#)

Ver resultados que contêm o texto: 'publico.pt'

### Lista de versões

5 368 versões de publico.pt

1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
<a href="#">13 Out</a>		<a href="#">12 Dez</a>	<a href="#">25 Jan</a>	<a href="#">11 Mai</a>	<a href="#">18 Abr</a>	<a href="#">30 Mar</a>	<a href="#">8 Fev</a>	<a href="#">2 Jan</a>	<a href="#">4 Jan</a>	<a href="#">1 Jan</a>	<a href="#">2 Fev</a>	<a href="#">1 Jan</a>	<a href="#">20 Mai</a>	<a href="#">23 Mar</a>	<a href="#">1 Jan</a>	<a href="#">1 Jan</a>	<a href="#">10 Jan</a>	<a href="#">27 Abr</a>	<a href="#">1 Jan</a>	<a href="#">1 Jan</a>	<a href="#">Disponíveis após 1 ano.</a>
<a href="#">13 Out</a>		<a href="#">12 Dez</a>	<a href="#">25 Jan</a>	<a href="#">11 Mai</a>	<a href="#">3 Mai</a>	<a href="#">25 Mai</a>	<a href="#">18 Fev</a>	<a href="#">6 Fev</a>	<a href="#">15 Jan</a>	<a href="#">4 Jan</a>	<a href="#">16 Mar</a>	<a href="#">1 Jan</a>	<a href="#">24 Mai</a>	<a href="#">1 Abr</a>	<a href="#">1 Jan</a>	<a href="#">1 Jan</a>	<a href="#">10 Jan</a>	<a href="#">27 Abr</a>	<a href="#">1 Jan</a>	<a href="#">1 Jan</a>	
			<a href="#">8 Fev</a>	<a href="#">11 Mai</a>	<a href="#">6 Mai</a>	<a href="#">28 Mai</a>	<a href="#">22 Mar</a>	<a href="#">11 Abr</a>	<a href="#">17 Jan</a>	<a href="#">4 Jan</a>	<a href="#">16 Mar</a>	<a href="#">5 Jan</a>	<a href="#">24 Jun</a>	<a href="#">5 Abr</a>	<a href="#">2 Jan</a>	<a href="#">2 Jan</a>	<a href="#">11 Jan</a>	<a href="#">27 Abr</a>	<a href="#">1 Jan</a>	<a href="#">1 Jan</a>	
			<a href="#">8 Fev</a>	<a href="#">11 Mai</a>	<a href="#">6 Mai</a>	<a href="#">1 Jun</a>	<a href="#">26 Mar</a>	<a href="#">8 Mai</a>	<a href="#">18 Jan</a>	<a href="#">4 Jan</a>	<a href="#">6 Jun</a>	<a href="#">11 Jan</a>	<a href="#">24 Jun</a>	<a href="#">8 Abr</a>	<a href="#">2 Jan</a>	<a href="#">2 Jan</a>	<a href="#">11 Jan</a>	<a href="#">27 Abr</a>	<a href="#">2 Jan</a>	<a href="#">1 Jan</a>	
			<a href="#">21 Fev</a>	<a href="#">11 Mai</a>	<a href="#">16 Mai</a>	<a href="#">4 Jun</a>	<a href="#">4 Abr</a>	<a href="#">20 Mai</a>	<a href="#">20 Jan</a>	<a href="#">6 Jan</a>	<a href="#">7 Jun</a>	<a href="#">15 Jan</a>	<a href="#">25 Set</a>	<a href="#">9 Abr</a>	<a href="#">3 Jan</a>	<a href="#">3 Jan</a>	<a href="#">12 Jan</a>	<a href="#">28 Abr</a>	<a href="#">2 Jan</a>	<a href="#">2 Jan</a>	
			<a href="#">21 Fev</a>	<a href="#">11 Mai</a>	<a href="#">22 Mai</a>	<a href="#">6 Jun</a>	<a href="#">11 Abr</a>	<a href="#">23 Mai</a>	<a href="#">22 Jan</a>	<a href="#">7 Jan</a>	<a href="#">7 Jun</a>	<a href="#">18 Jan</a>	<a href="#">26 Set</a>	<a href="#">12 Abr</a>	<a href="#">3 Jan</a>	<a href="#">3 Jan</a>	<a href="#">12 Jan</a>	<a href="#">28 Abr</a>	<a href="#">2 Jan</a>	<a href="#">2 Jan</a>	
			<a href="#">21 Abr</a>	<a href="#">11 Mai</a>	<a href="#">8 Jun</a>	<a href="#">2 Ago</a>	<a href="#">24 Abr</a>	<a href="#">26 Mai</a>	<a href="#">24 Jan</a>	<a href="#">8 Jan</a>	<a href="#">7 Jun</a>	<a href="#">20 Jan</a>	<a href="#">17 Dez</a>	<a href="#">13 Abr</a>	<a href="#">4 Jan</a>	<a href="#">4 Jan</a>	<a href="#">15 Jan</a>	<a href="#">28 Abr</a>	<a href="#">3 Jan</a>	<a href="#">2 Jan</a>	
			<a href="#">21 Abr</a>	<a href="#">11 Mai</a>	<a href="#">27 Jun</a>	<a href="#">7 Ago</a>	<a href="#">25 Mai</a>	<a href="#">6 Jun</a>	<a href="#">27 Jan</a>	<a href="#">11 Jan</a>	<a href="#">7 Jun</a>	<a href="#">23 Jan</a>	<a href="#">18 Dez</a>	<a href="#">16 Abr</a>	<a href="#">4 Jan</a>	<a href="#">4 Jan</a>	<a href="#">15 Jan</a>	<a href="#">28 Abr</a>	<a href="#">3 Jan</a>	<a href="#">2 Jan</a>	
			<a href="#">27 Abr</a>	<a href="#">11 Mai</a>	<a href="#">27 Jun</a>	<a href="#">12 Ago</a>	<a href="#">17 Jul</a>	<a href="#">7 Jun</a>	<a href="#">28 Jan</a>	<a href="#">13 Jan</a>	<a href="#">9 Jun</a>	<a href="#">26 Jan</a>		<a href="#">19 Abr</a>	<a href="#">5 Jan</a>	<a href="#">5 Jan</a>	<a href="#">17 Jan</a>	<a href="#">29 Abr</a>	<a href="#">3 Jan</a>	<a href="#">3 Jan</a>	
			<a href="#">27 Abr</a>	<a href="#">11 Mai</a>	<a href="#">28 Jun</a>	<a href="#">12 Ago</a>	<a href="#">30 Set</a>	<a href="#">10 Jun</a>	<a href="#">28 Jan</a>	<a href="#">14 Jan</a>	<a href="#">14 Jun</a>	<a href="#">1 Fev</a>		<a href="#">22 Abr</a>	<a href="#">5 Jan</a>	<a href="#">5 Jan</a>	<a href="#">17 Jan</a>	<a href="#">29 Abr</a>	<a href="#">4 Jan</a>	<a href="#">3 Jan</a>	
			<a href="#">12 Out</a>	<a href="#">11 Mai</a>	<a href="#">28 Jun</a>	<a href="#">23 Set</a>	<a href="#">10 Out</a>	<a href="#">10 Jun</a>	<a href="#">29 Jan</a>	<a href="#">14 Jan</a>	<a href="#">14 Jun</a>	<a href="#">3 Fev</a>		<a href="#">30 Abr</a>	<a href="#">6 Jan</a>	<a href="#">6 Jan</a>	<a href="#">18 Jan</a>	<a href="#">29 Abr</a>	<a href="#">4 Jan</a>	<a href="#">3 Jan</a>	
			<a href="#">12 Out</a>	<a href="#">20 Mai</a>	<a href="#">1 Jul</a>	<a href="#">28 Set</a>	<a href="#">21 Out</a>	<a href="#">11 Jun</a>	<a href="#">30 Jan</a>	<a href="#">14 Jan</a>	<a href="#">6 Ago</a>	<a href="#">9 Fev</a>		<a href="#">2 Mai</a>	<a href="#">6 Jan</a>	<a href="#">6 Jan</a>	<a href="#">18 Jan</a>	<a href="#">30 Abr</a>	<a href="#">4 Jan</a>	<a href="#">3 Jan</a>	
			<a href="#">11 Nov</a>	<a href="#">20 Mai</a>	<a href="#">9 Jul</a>	<a href="#">24 Out</a>	<a href="#">24 Nov</a>	<a href="#">12 Jun</a>	<a href="#">31 Jan</a>	<a href="#">15 Jan</a>	<a href="#">27 Set</a>	<a href="#">13 Fev</a>		<a href="#">3 Mai</a>	<a href="#">7 Jan</a>	<a href="#">7 Jan</a>	<a href="#">19 Jan</a>	<a href="#">30 Abr</a>	<a href="#">5 Jan</a>	<a href="#">4 Jan</a>	

### Narrative Timeline of Troika Em Portugal during the last 10 years

Horizontal Vertical

**16, November 2016**

- Portugal reembolsará seis mil milhões de euros ao FMI em Março
- Paulo Portas diz ter havido uma troika de restrições e uma troika de bem-fazer
- Portugal não vai recuperar nem metade dos empregos destruídos pela troika
- Troika culpa Governo português por falhanço do programa
- Passos e Cavaco são a troika em Portugal
- Parlamento Europeu vai investigar erros da troika em Portugal

2 years ago

**01, February 2015**

- Troika diz que Portugal precisa
- Última avaliação da troika arranca hoje em Lisboa
- Jerónimo de Sousa diz que saída da troika de

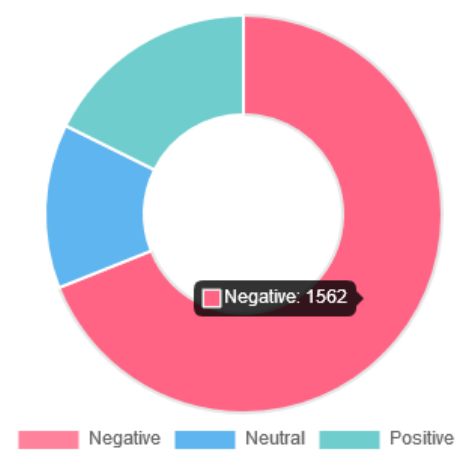
Related terms about Troika Em Portugal during the last 10 years



Sources: Ionline, Tsf, Sabado, Economico Sapo, Dinheirovivo, News Google, Lux lol, Aeiou.

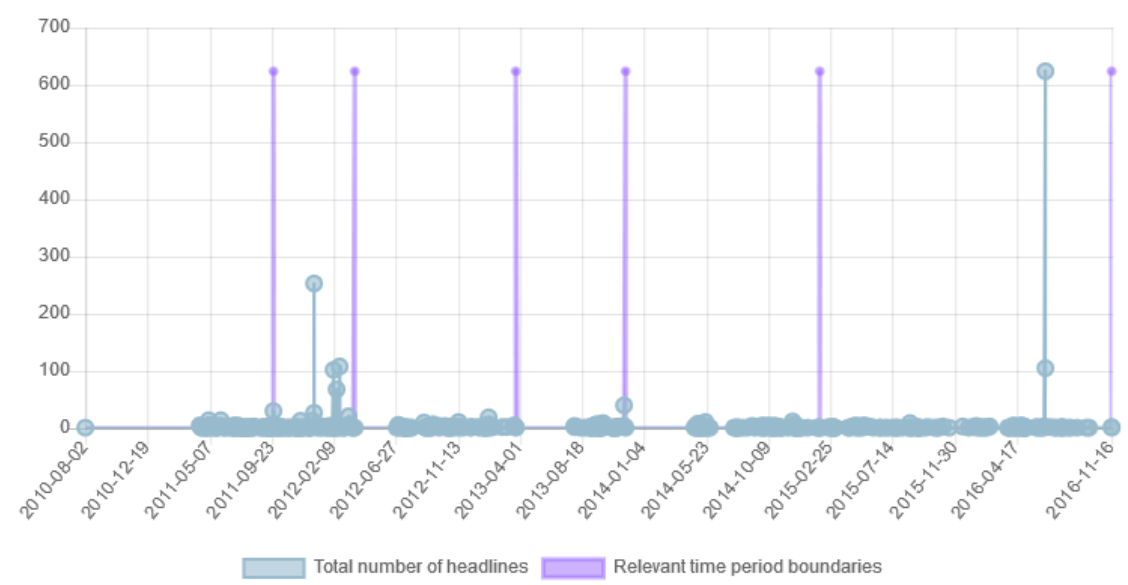
### Detailed analysis on Troika Em Portugal during the last 10 years

#### Sentiment analysis



#### Headlines by date

click at the item to open detailed search at arquivo.pt





Troika em Portugal Pesquisar

entre: 27/12/2011 e: 27/12/2011

[Pesquisa avançada](#)

Resultados 1 a 10 de 14.911

[Troika: Portugal «tem feito bons progressos em várias frente...](#)

27 Dezembro, 2011 - [Listar versões](#)

**Troika: Portugal** «tem feito bons progressos em várias frentes» - Comissão Europeia | União Europeia ...: **Portugal** «tem feito bons progressos em várias frentes» - Comissão Europeia A Comissão Europeia considerou hoje que **Portugal** «tem feito bons progressos em várias frentes» no que refere à implementação do ...  
[http://diariodigital.sapo.pt/news.asp?section\\_id=57&id\\_news=549293](http://diariodigital.sapo.pt/news.asp?section_id=57&id_news=549293)

[Troika: Portugal não deve precisar de mais dinheiro que o in...](#)

27 Dezembro, 2011 - [Listar versões](#)

**Troika: Portugal** não deve precisar de mais dinheiro que o inicialmente acordado - Bruxelas | União Europeia | Diário Digital Director: Pedro Curvelo Pesquise no DD em Todas as secções. Sociedade ... **Troika: Portugal** não deve precisar de mais dinheiro que o inicialmente acordado - Bruxelas A Comissão ...  
[http://diariodigital.sapo.pt/news.asp?section\\_id=57&id\\_news=549315](http://diariodigital.sapo.pt/news.asp?section_id=57&id_news=549315)

[Portugal na "lupa" da troika](#)

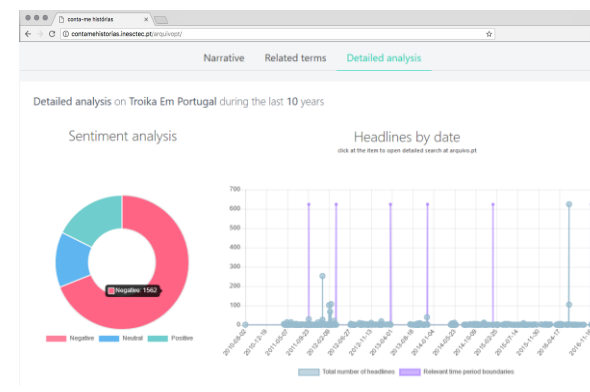
27 Dezembro, 2011 - [Listar versões](#)

**Portugal** na "lupa" da **troika** Skip to content Skip to main navigation Skip to 1st column Skip to 2nd ... Economia  
 Economia **Portugal** na "lupa" da **troika** **Portugal** na "lupa" da **troika** Terça, 08 Novembro 2011 00:00 Redacção  
 TRANSLATE english french spanish swahili **Portugal** na "lupa" da **troika** **Portugal** na "lupa" ...  
<http://opais.sapo.mz/index.php/economia/...17575-portugal-na-lupa-da-troika.html>

[Sol - Dossier - troika](#)

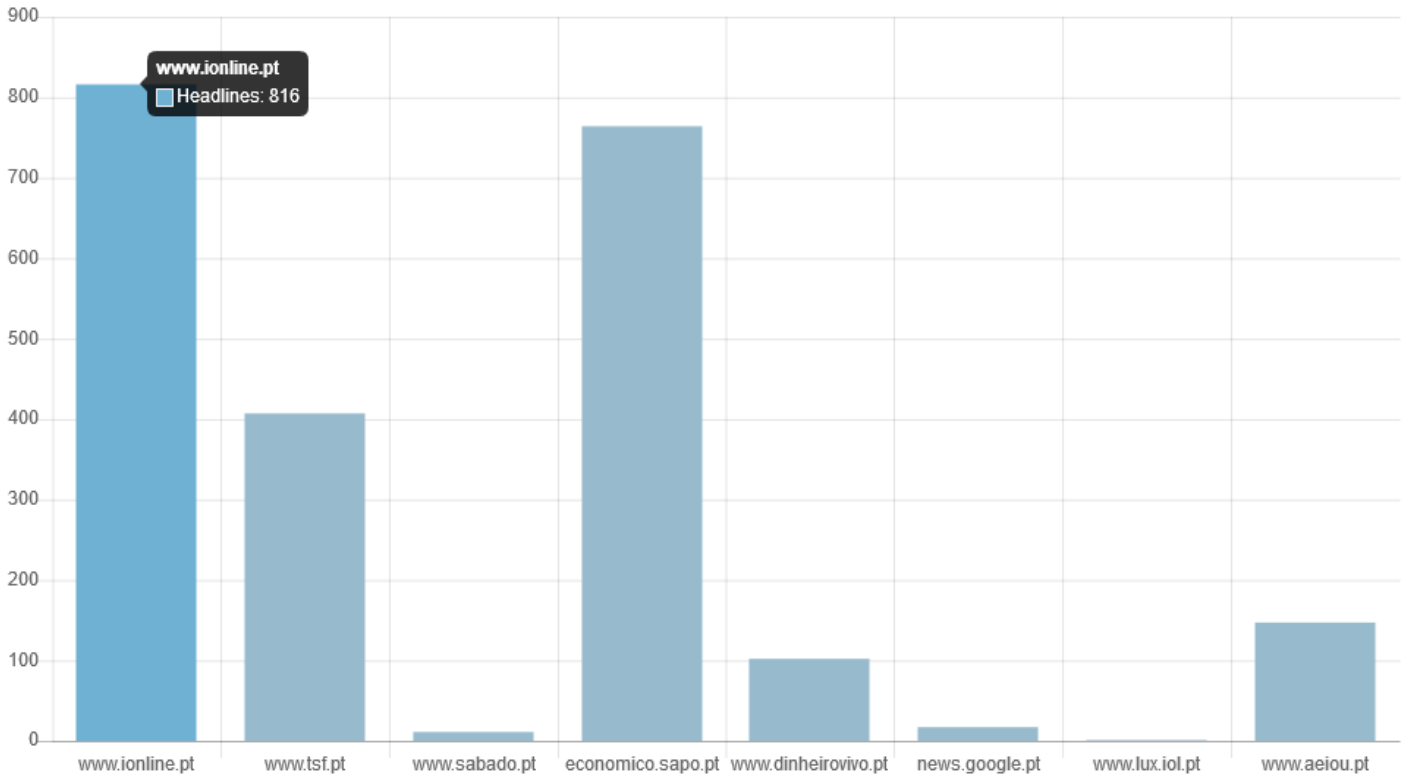
27 Dezembro, 2011 - [Listar versões](#)

Economia **Portugal** dever♦ poupar 5,2 mil mil♦es em juros durante ajuda externa A Comiss♦o Europeia ... ♦  
 ♦ comentar Economia **Troika: Portugal** n♦o deve precisar de mais dinheiro A Comiss♦o Europeia ... **'troika'** no  
 programa de ajustamento... ♦ ♦ comentar Economia Rendimento dispon♦vel cai 6% em 2012 O ...



# Headlines by news sources

click at the item to open detailed search at arquivo.pt





Troika em Portugal site:www.ionline.pt type:html

Pesquisar

entre: 02/08/2010 e: 16/11/2016

Pesquisa avançada

Resultados 1 a 10 de 346.782

Desejava ver apenas o histórico de www.ionline.pt?

Troika: Portugal à espera dos próximos episódios

24 Abril, 2011 - Listar versões

Troika: Portugal à espera dos próximos episódios Publicado em 22 de Abril de 2011 A troika começou as reuniões com os partidos e ... Impostas pela troika aos portugueses deverão ser conhecidas na próxima semana. Ainda uma viagem à ... http://www.ionline.pt/interior/index.php?p=news-print&idNota=118903

Costa. A troika em Portugal agora chama-se Portas, Passos e ...

24 Novembro, 2015 - Listar versões

Costa. A troika em Portugal agora chama-se Portas, Passos e Cavaco 24/11/15 subscrever newsletter ... Registrar Pesquisar Costa. A troika em Portugal agora chama-se Portas, Passos e Cavaco Costa. A troika em Portugal agora chama-se Portas, Passos e Cavaco ANDRÉ KOSTERS, Lusa António Costa 23/04 ... http://www.ionline.pt/388076

Troika. Portugal vai ao terceiro exame da cadeira Maior Ajus...

12 Fevereiro, 2012 - Listar versões

Troika. Portugal vai ao terceiro exame da cadeira Maior Ajustamento Económico de sempre | iOnline ... Mercados Boa Vida iModa iTestdrive iGastronomia Desporto ITV iFotogaleria Opinião iSabe Troika. Portugal ... Portugal, os bancos podem cumprir a meta já em 2013. Mas há um preço: o esmagamento do crédito empresarial ... http://www.ionline.pt/dinheiro/troika-po...ra-maior-ajustamento-economico-sempre

Troika: Portugal não deve precisar de mais dinheiro que o in...

22 Dezembro, 2011 - Listar versões

Troika: Portugal não deve precisar de mais dinheiro que o inicialmente acordado - Bruxelas ... iTestdrive





## Narrative Timeline of Donald Trump during the last 10 years

Horizontal Vertical

Previous

Next

Jan 2018 Feb 2018 Mar 2018 Apr 2018 Apr 2018 May 2018 May 2018

### - Relevant stories from **06, February 2018** until **20, March 2018**

Trump Plan to Move Drug Prevention Program to Justice Dept positive

Trump Expected to Release Democratic Memo

Trump 's Infrastructure Plan May Ignore Climate Change

American ISIS Suspect Said He Wanted to Report From Syria

Russia Wanted Trump to Win positive

Trump Edges Toward a Conventional Republican Approach

Trump to Discuss School Shootings With Survivors

Trump Condemns Florida Officer Who Did n't Act negative

Trump Blames Video Games for Mass Killings negative

Trump Embraces a Trade War negative

Trump Advisers Fervently Defend Tariffs

South Korean Official on Trump

Trade War Prospect Shakes Part of Trump Base

Trump to Push Ahead on Gun Training for School Employees

Trump Plans to View Wall Prototypes

Trump Finds Unlikely Culprit in School Shootings

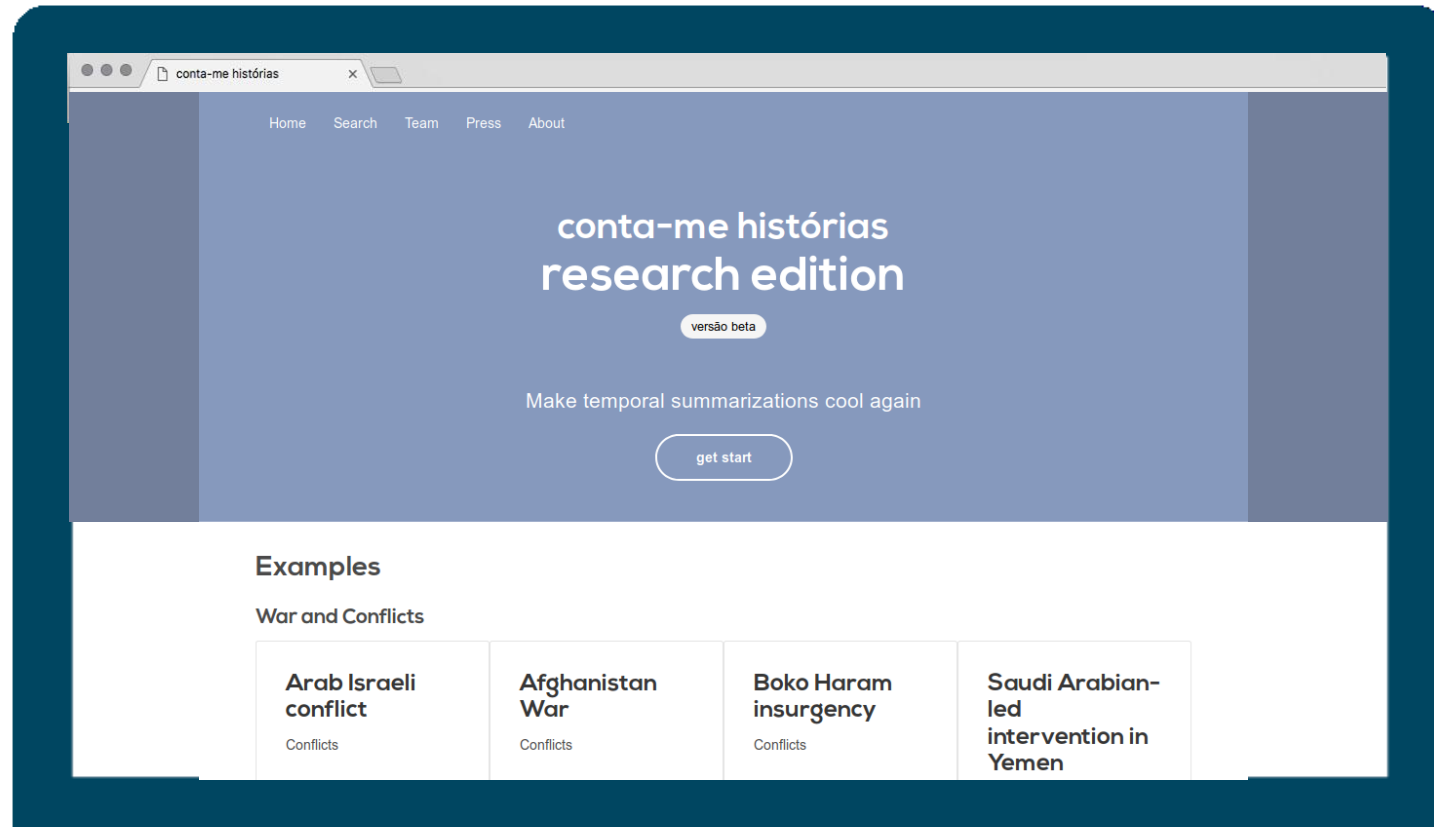
Trump Says He Made Up Trade Numbers in Talk with Trudeau

Assaults Increased When Cities Hosted Trump Rallies

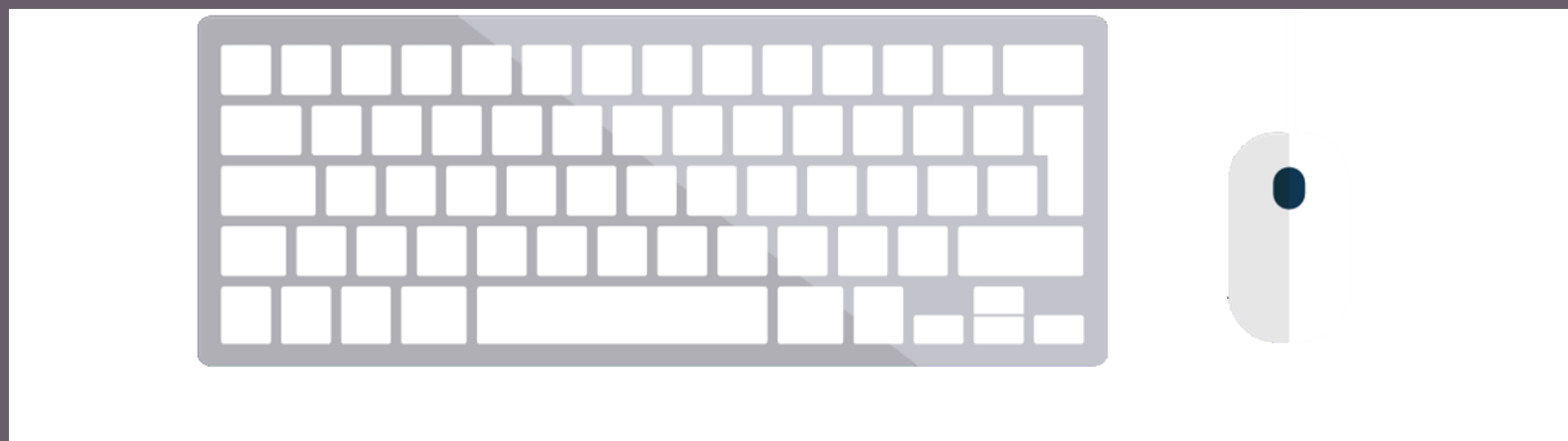
Trump to Hire Lawyer Who Has Pushed Theory That Justice Dept positive

Melania Trump Wants to End Online Bullying negative

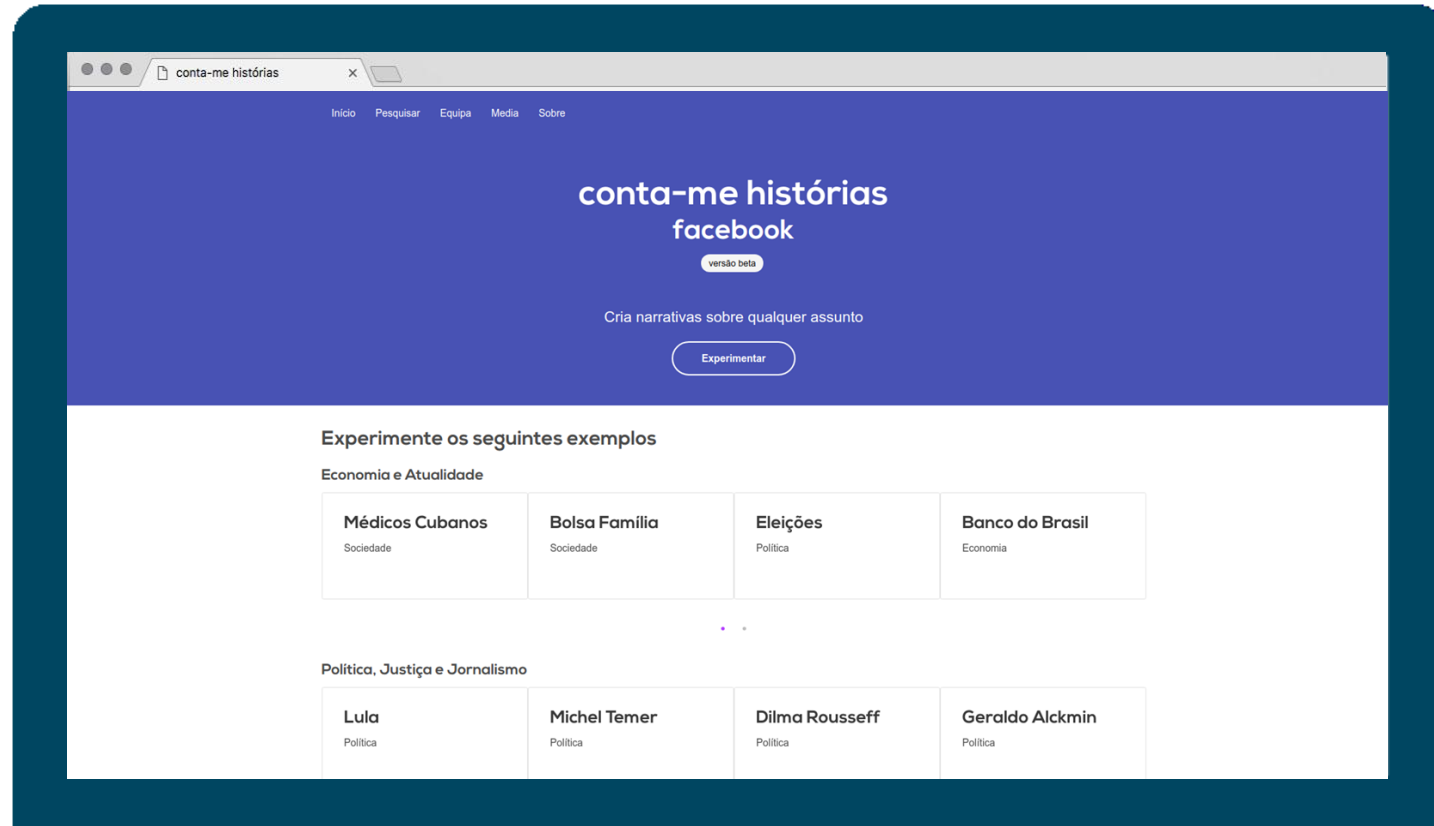
# research edition with different



<http://nlplab.inesctec.pt/contamehistorias-labs>



# facebook posts edition



<http://nlplab.inesctec.pt/contamehistorias-facebook>



# **Conta-me Histórias open source**

**open source python package**

Search or jump to... Pull requests Issues Marketplace Explore

Your repository details have been saved.

LIAAD / TemporalSummarizationFramework Unwatch 4 Star 1 Fork 0

Code Issues 1 Pull requests 0 Projects 0 Wiki Insights Settings

Conta-me Histórias, temporal summarization engine <http://contamehistorias.pt> Edit

Manage topics

45 commits 1 branch 0 releases 1 contributor

Branch: master New pull request Create new file Upload files Find file Clone or download

arianpasquali Update LICENSE.txt Latest commit a08fc6a 3 hours ago

contamehistorias	refactoring e levenshtein	a day ago
LICENSE.txt	Update LICENSE.txt	3 hours ago
MANIFEST.in	refactoring e levenshtein	a day ago
README.md	Update README.md	7 days ago
requirements.txt	refactoring e levenshtein	a day ago
setup.cfg	terminal script and pip install	17 days ago
setup.py	refactoring and new examples	15 days ago
test_conteme.py	terminal script and pip install	17 days ago

<http://github.com/LIAAD/TemporalSummarizationFramework>

# Installing

requires Python 3



```
> pip install git+https://github.com/LIAAD/TemporalSummarizationFramework.git
```

---



# Using terminal interface

## Accessing client help

```
● ● ●  
  
> contamehistorias --help  
  
Usage: contamehistorias [OPTIONS]  
  
Console script for contamehistorias.  
  
Options:  
  --query TEXT      Perform news retrieval with informed query  
  --language TEXT   Expected language in headlines  
  --start_date DATE Perform news retrieval since this date  
  --end_date DATE   Perform news retrieval until this date  
  --domains TEXT    Comma separated list of domains (www.publico.pt,www.dn.pt)  
  --verbose  
  --help            Show this message and exit.
```

# Code with Arquivo.pt as data source



```
from contamehistorias.datasources.webarchive import ArquivoPT
from contamehistorias.engine import TemporalSummarizationEngine
from datetime import datetime

domains = [ 'http://publico.pt/', 'http://www.dn.pt/', 'http://www.rtp.pt/', 'http://www.iol.pt/' ]

params = { 'domains':domains, 'from':datetime(year=2016, month=3, day=1),
          'to':datetime(year=2018, month=1, day=10) }

query = 'Dilma Rousseff'
language = "pt"

#instantiate ArquivoPT search engine
apt = ArquivoPT()
search_result = apt.getResult(query=query, **params)
#instantiate temporal summarization class
engine = TemporalSummarizationEngine()
intervals = cont.build_intervals(search_result, language)
#print results
engine.pprint(intervals)
```

# How to extend it

You can extend it to use your own collection as data source.

Extend BaseDataSource class and implement method **getResult** to return a list of object **ResultHeadLine**.

For each document you need a **title**, a **timestamp**, **source** and **url**.

```
class BaseDataSource(object):

    def __init__(self, name):
        self.name = name

    def getResult(self, query, **kwargs):
        raise NotImplementedError('getResult on ' + self.name + ' not implemented yet!')
```

# Extension example with custom dataset \*

```
from contamehistorias.engine import TemporalSummarizationEngine
from contamehistorias.datasources.signal import SignalNewsIRDataset

#collect dataset
query = "Barack Obama"
search_result = SignalNewsIRDataset().getResult(query=query)

#generate timeline
engine = TemporalSummarizationEngine()
summary_result = engine.build_intervals(search_result, "en")

#print results
engine.pprint(summary_result, verbose=True)
```

\* Signal's NewsIR'16 dataset : <http://research.signalmedia.co/newsir16/signal-dataset.html>

# Output



```
> contamehistorias_signal --query "Barack Obama"
```

```
Conta-me Historias Temporal Summarization. Signal NewsIR 16 dataset example (more info  
http://research.signalmedia.co/newsir16/signal-dataset.html  
--query "Barack Obama"
```

```
Performing search. This may take a while.  
Computing temporal summarization ...
```

## Timeline

```
2015-07-25 08:44:54 until 2015-09-07 06:53:14
```

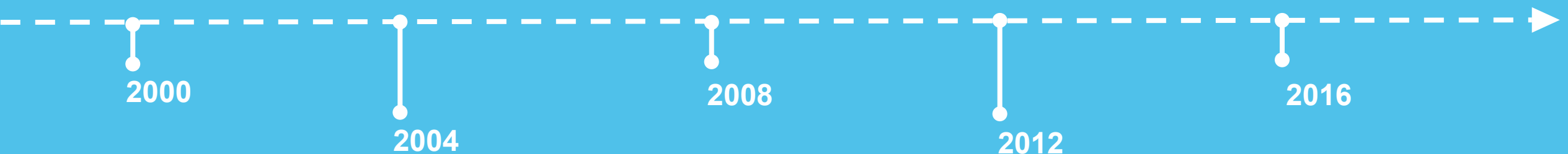
```
  obama says world must reach climate deal in paris  
  obama enthusiastic for los angeles summer olympics bid  
  democrats give obama votes to uphold iran nuke deal  
  obama scores policy win in securing votes for iran nuclear deal
```

```
2015-09-14 04:11:35 until 2015-09-21 22:27:54
```

```
  president obama honors stephen kin with the national medal of the arts  
  struggling gop presidential candidate scott walker blames obama for police deaths  
  obama unveils new college financial aid application process  
  obama will greet chinese president with handshake  
  barack obama supports ahmed mohamed
```

# next steps

- In terms of data sources support different web archives
- Different timelines for different kinds of sources (left-wing, right-wing, tabloids, etc)
- Extended formal evaluation





# thank you

Arian Pasquali  
arianpasquali@gmail.com  
INESCTEC - LIAAD  
Universidade do Porto

<http://contamehistorias.pt>  
[admin@contamehistorias.pt](mailto:admin@contamehistorias.pt)