
Fake news as we feel it:

Perception and conceptualization of the term "fake news" in the media

UF *m* G

UNIVERSIDADE FEDERAL
DE MINAS GERAIS



Universiteit
Leiden



**BERKMAN
KLEIN CENTER**
FOR INTERNET & SOCIETY
AT HARVARD UNIVERSITY

Evandro Cunha, Gabriel Magno, Josemar Caetano,
Douglas Teixeira & Virgilio Almeida

THE COLLINS WORD OF THE YEAR 2017 IS...

FAKE NEWS

**FAKE
NEWS**

n (,feɪk 'njuːz)

False, often sensational,
information disseminated
under the guise of news.

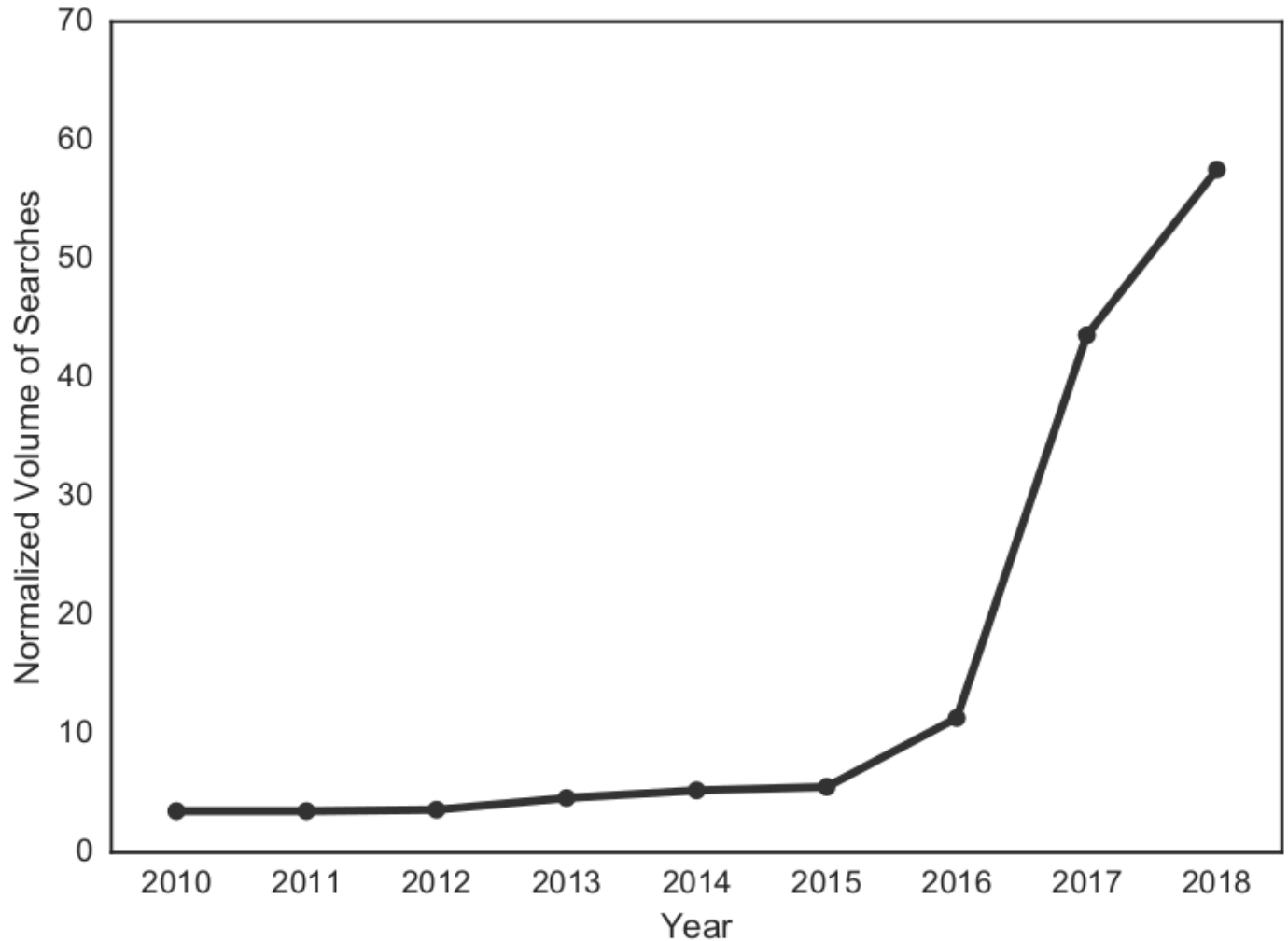


#collinsWOTY

Web search behavior

Data from
Google Trends

Increase in the
public interest
in the term
“fake news” after
2016 US
presidential
elections



Was the rise of the public interest in the term “fake news” accompanied by changes in its conceptualization and in the perception about it?

Sociolexicological theories (e.g. Matoré, 1953; Cambraia, 2013) defend the existence of a considerable **relationship between linguistic and extralinguistic factors** with regards to the vocabulary

Our hypothesis: the sudden increase in interest in the **phenomenon fake news** might have altered the general usage of the expression “fake news”

Was the rise of the public interest in the term “fake news” accompanied by changes in its conceptualization and in the perception about it?

Sociolexicological theories (e.g. Matoré, 1953; Cambraia, 2013) defend the existence of a considerable **relationship between linguistic and extralinguistic factors** with regards to the vocabulary

Our hypothesis: the sudden increase in interest in the **phenomenon fake news** might have altered the general usage of the expression “fake news”

Data source

NOW Corpus (Corpus of News on the Web)
corpus.byu.edu/now

Articles from **online newspapers and magazines**
written in English and based in **20 different**
countries from **2010 to the present time**

41,124 occurrences of the term "fake news"

Concordance lines / contexts

she has to phone her and assure her that she's still alive before the **fake news** gets to her. # Jaden Smith, pictured here at the 2016 MTV

IN THE ROAD. Facebook CEO Mark Zuckerberg will need to address the amount of **fake news** that gets passed around in his site. Photo by Lluís Gene/AFP # MANILA

are too blurry. At a time when misinformation, hoaxes, trolls, and **fake news** have already polluted social networks, there's no better editor than Facebook to

a left-wing bias by keeping conservative stories from its Trending Topics section and allowing a **fake news** story in the same section. # The debate involves whether Facebook should act

replacing them with an automated system, the algorithm seemed to be flawed, placing **fake news** stories, celebrity gossip and misidentified key words in lieu of more serious news

with algorithms that quickly demonstrated the difficulty of automating news editorial judgment by promoting a **fake news** story. # In his open letter, Hansen points out that the types

n't spread the rumor, " said Vanke on Wechat yesterday. " It is **fake news**, the crowd of lurkers can leave. " That came a day after

those unaware, this particular comment gained momentum after American John Cain filmed several **fake news** casts in 2014 that featured a man yelling this phrase to a female reporter

us -- not even the Geeky Gods of Silicon Valley -- is immune to the **fake news** hoax. # Media illiteracy is not, it turns out, a trait

Stallone has key role in ' Guardians of the Galaxy 2' # A **fake news** story -- allegedly from US news site CNN -- was being shared on Facebook

a rumour circulated on social media and appeared to confirm his death. # A **fake news** story was being shared on Facebook, with fans clicking on the spam article

written e-mail and blog had metastasized into hundreds of professionally written blogs, postings and **fake news** items. There were thousands of pages of accusations of professional wrongdoing

of themselves. Their attempt to divide ipob through propaganda can't fly. Their **fake news** items are so childish. Buhari should quickly change the leadership of the DSS

them. # David Kravets The senior editor for Ars Technica. Founder of TYDN **fake news** site. Technologist. Political scientist. Humorist. Dad of two boys.

Related News # Solly has been the subject of several rumours online after a **fake news** site released a sex song apparently recorded by the artist. # The song

replaced them with an algorithm on Friday, only to find the algorithm promoting completely **fake news** on Sunday. # Rarely in recent tech history has a downsizing decision come

this past weekend. # Over the weekend, Facebook was seen posting a **fake news** story about Fox News (the western world's Times Now) that the

. (AFP / LEON NEAL) mb.com.ph # Facebook is now facing problems after **fake news** stories ended up in its now fully automated Trending news section. # Over

Concordance lines / contexts

she has to phone her and assure her that she's still alive before the **fake news** gets to her. # Jaden Smith, pictured here at the 2016 MTV

IN THE ROAD. Facebook CEO Mark Zuckerberg will need to address the amount of **fake news** that gets passed around in his site. Photo by Lluís Gene/AFP # MANILA

are too blurry. At a time when misinformation, hoaxes, trolls, and **fake news** have already polluted social networks, there's no better editor than Facebook to

a left-wing bias by keeping conservative stories from its Trending Topics section and allowing a **fake news** story in the same section. # The debate involves whether Facebook should act

replacing them with an automated system, the algorithm seemed to be flawed, placing **fake news** stories, celebrity gossip and misidentified key words in lieu of more serious news

with algorithms that quickly demonstrated the difficulty of automating news editorial judgment by promoting a **fake news** story. # In his open letter, Hansen points out that the types

Words immediately surrounding a key term are more relevant to the conceptualization of this term than words further away from it

Stanhope has key role in 'Guardians of the Galaxy 2' # A **fake news** story — allegedly from US news site CNN — was being shared on Facebook

a rumour circulated on social media and appeared to confirm his death. # A **fake news** story was being shared on Facebook, with fans clicking on the spam article

written e-mail and blog had metastasized into hundreds of professionally written blogs, postings and **fake news** items. There were thousands of pages of accusations of professional wrongdoing

of themselves. Their attempt to divide ipob through propaganda can't fly. Their **fake news** items are so childish. Buhari should quickly change the leadership of the DSS

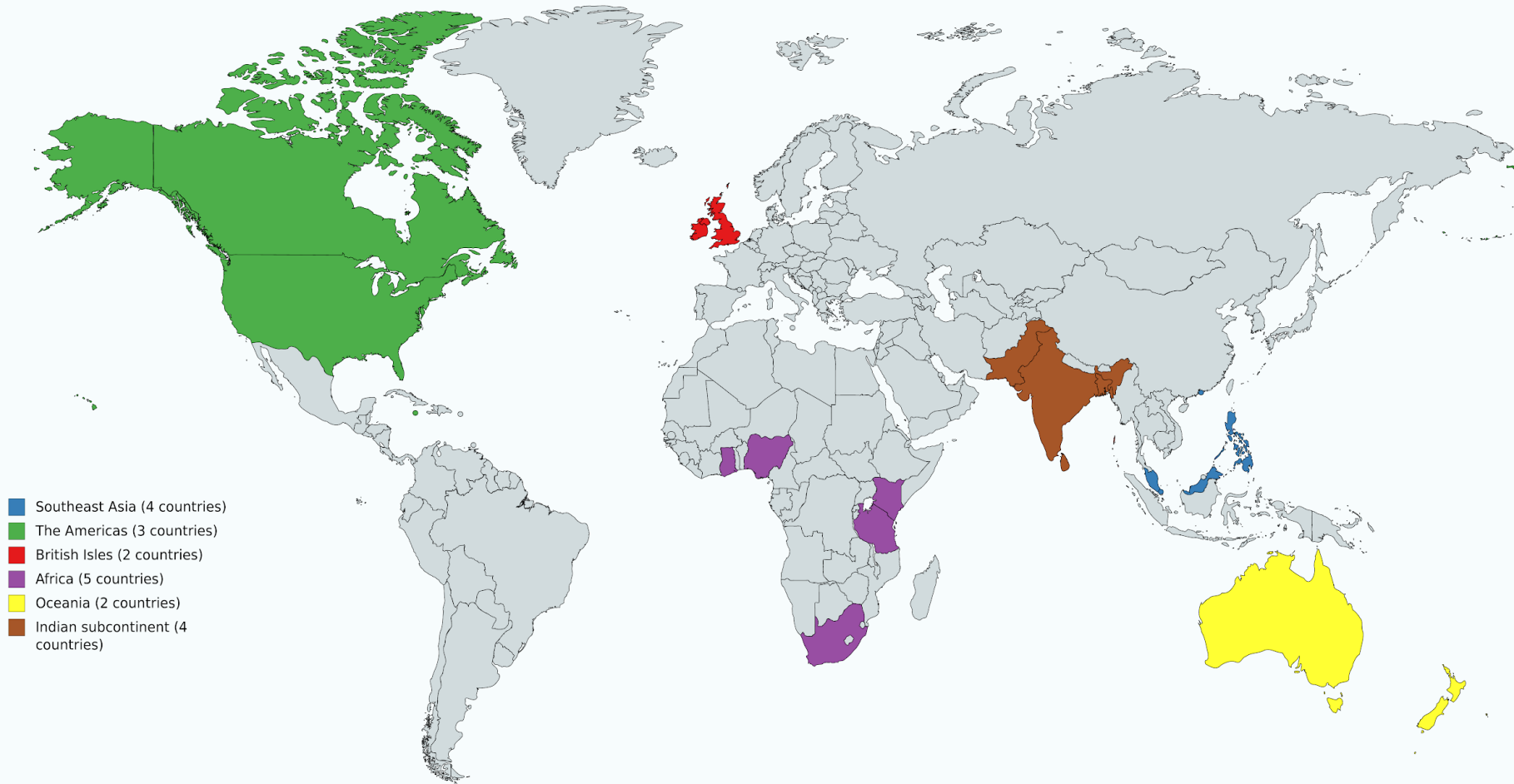
them. # David Kravets The senior editor for Ars Technica. Founder of TYDN **fake news** site. Technologist. Political scientist. Humorist. Dad of two boys.

Related News # Solly has been the subject of several rumours online after a **fake news** site released a sex song apparently recorded by the artist. # The song

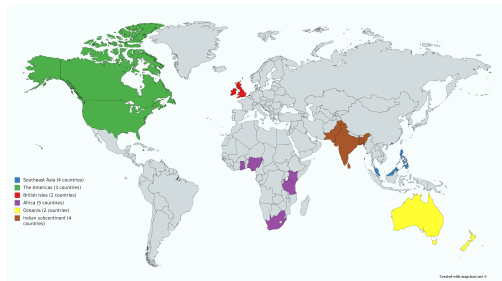
replaced them with an algorithm on Friday, only to find the algorithm promoting completely **fake news** on Sunday. # Rarely in recent tech history has a downsizing decision come

this past weekend. # Over the weekend, Facebook was seen posting a **fake news** story about Fox News (the western world's Times Now) that the

. (AFP / LEON NEAL) mb.com.ph # Facebook is now facing problems after **fake news** stories ended up in its now fully automated Trending news section. # Over



Six regions



Two periods

Year	2010	2011	2012	2013	2014	2015	2016	2017	2018
Occurrences	24	43	57	64	89	95	4,766	25,293	10,693
Period	before US election						after US election		
Occurrences	494						40,630		

Five analyses

Co-occurring **named entities**

Semantic fields of the surrounding vocabulary

Co-occurrence **networks**

Topics addressed

Polarity

1 → Co-occurring named entities

Region	Period	Entities
Africa	before after	PDP, Ekiti, Nigeria Donald Trump, Facebook, US
British Isles	before after	Facebook, The Daily Show, Stephen Colbert Donald Trump, Facebook, US
Indian subcontinent	before after	Shahid Afridi, King Salman of Saudi Arabia, BJP Facebook, Donald Trump, US
Oceania	before after	Twitter, The Daily Show, NBC Donald Trump, Facebook, US
Southeast Asia	before after	Korina Sanchez, US, China Facebook, Donald Trump, US
The Americas	before after	The Daily Show, Jon Stewart, Onion News Network Donald Trump, Facebook, CNN

1 → Co-occurring named entities

Region	Period	Entities
Africa	before after	PDP, Ekiti, Nigeria ◀ Donald Trump, Facebook, US
British Isles	before after	Facebook, The Daily Show, Stephen Colbert Donald Trump, Facebook, US
Indian subcontinent	before after	Shahid Afridi, King Salman of Saudi Arabia, BJP ◀ Facebook, Donald Trump, US
Oceania	before after	Twitter, The Daily Show, NBC Donald Trump, Facebook, US
Southeast Asia	before after	Korina Sanchez, US, China ◀ Facebook, Donald Trump, US
The Americas	before after	The Daily Show, Jon Stewart, Onion News Network Donald Trump, Facebook, CNN

1 → Co-occurring named entities

Region	Period	Entities
Africa	before	PDP, Ekiti, Nigeria
	after	Donald Trump, Facebook, US
	before	Facebook, The Daily Show, Stephen Colbert

Global standardization of the entities related to fake news

	after	Facebook, Donald Trump, US
Oceania	before	Twitter, The Daily Show, NBC
	after	Donald Trump, Facebook, US
Southeast Asia	before	Korina Sanchez, US, China
	after	Facebook, Donald Trump, US
The Americas	before	The Daily Show, Jon Stewart, Onion News Network
	after	Donald Trump, Facebook, CNN

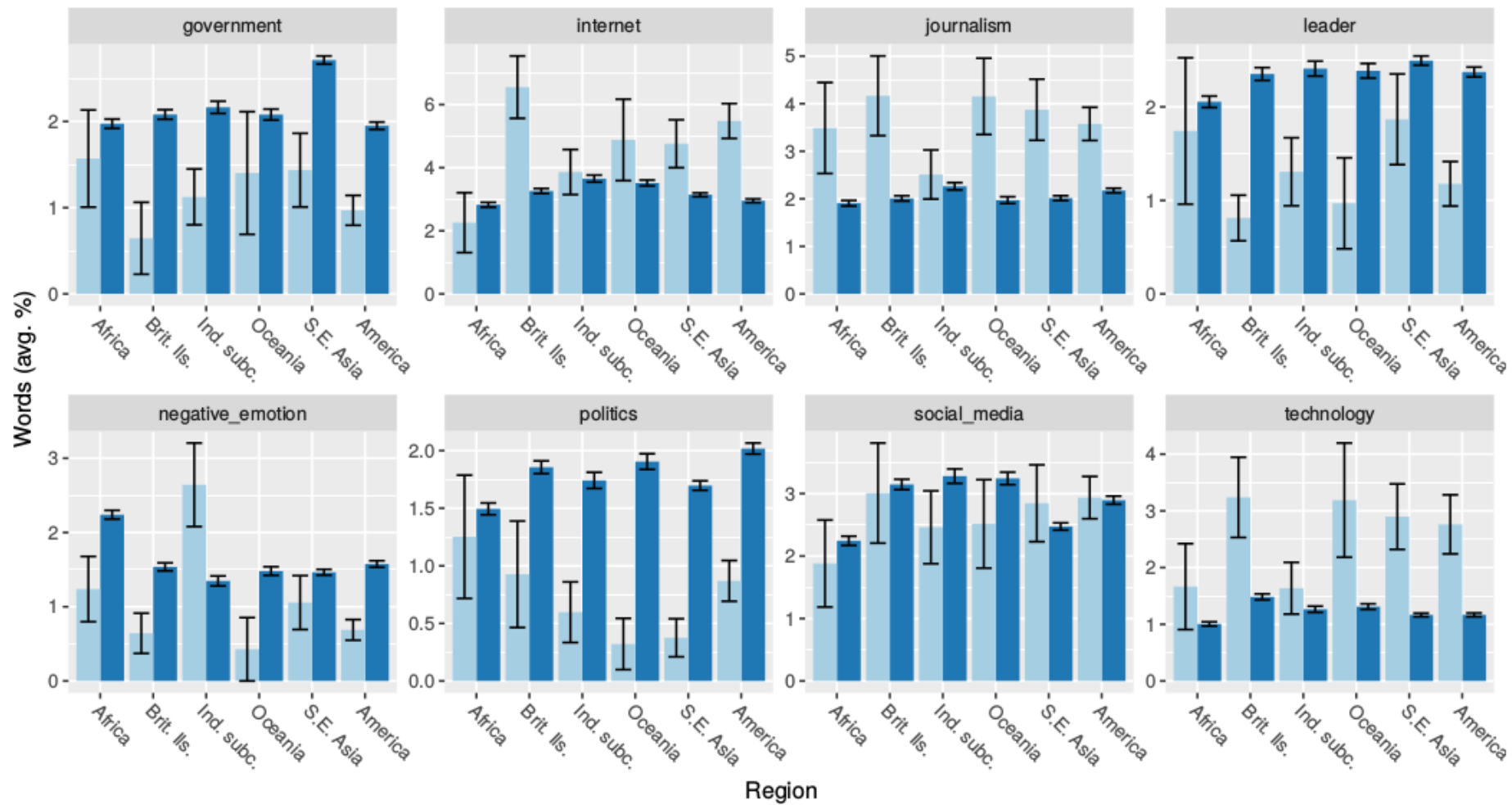
2 → Semantic fields of the surrounding vocabulary

We classified the surrounding vocabulary according to categories (from the tool Empath) that represent the following semantic fields:

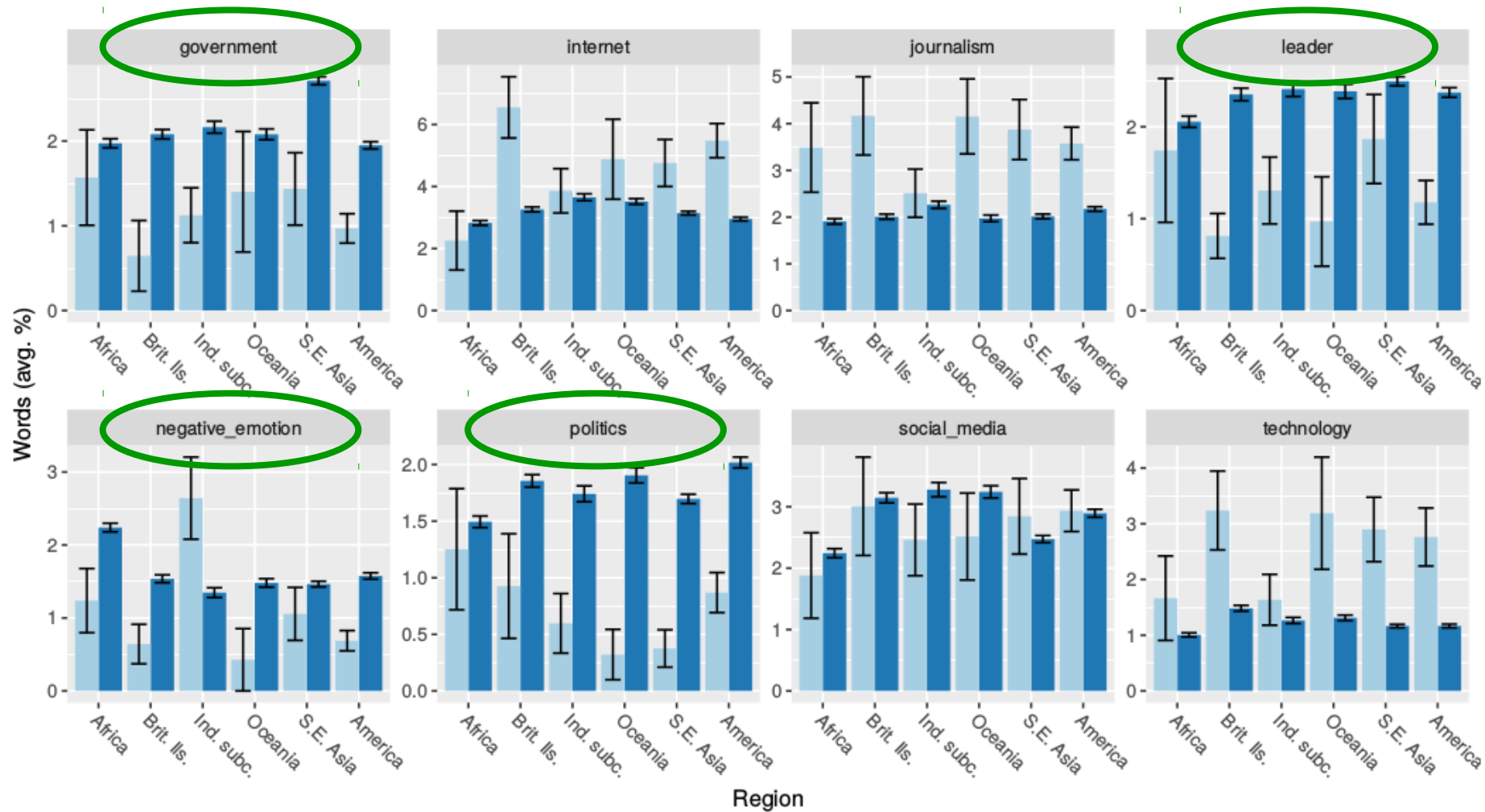
*government, internet, journalism, leader,
negative emotion, politics, social media, technology*

E.g.: *internet* → *homepage, download, hacker* + 76 words
journalism → *report, article, newspaper* + 66 words

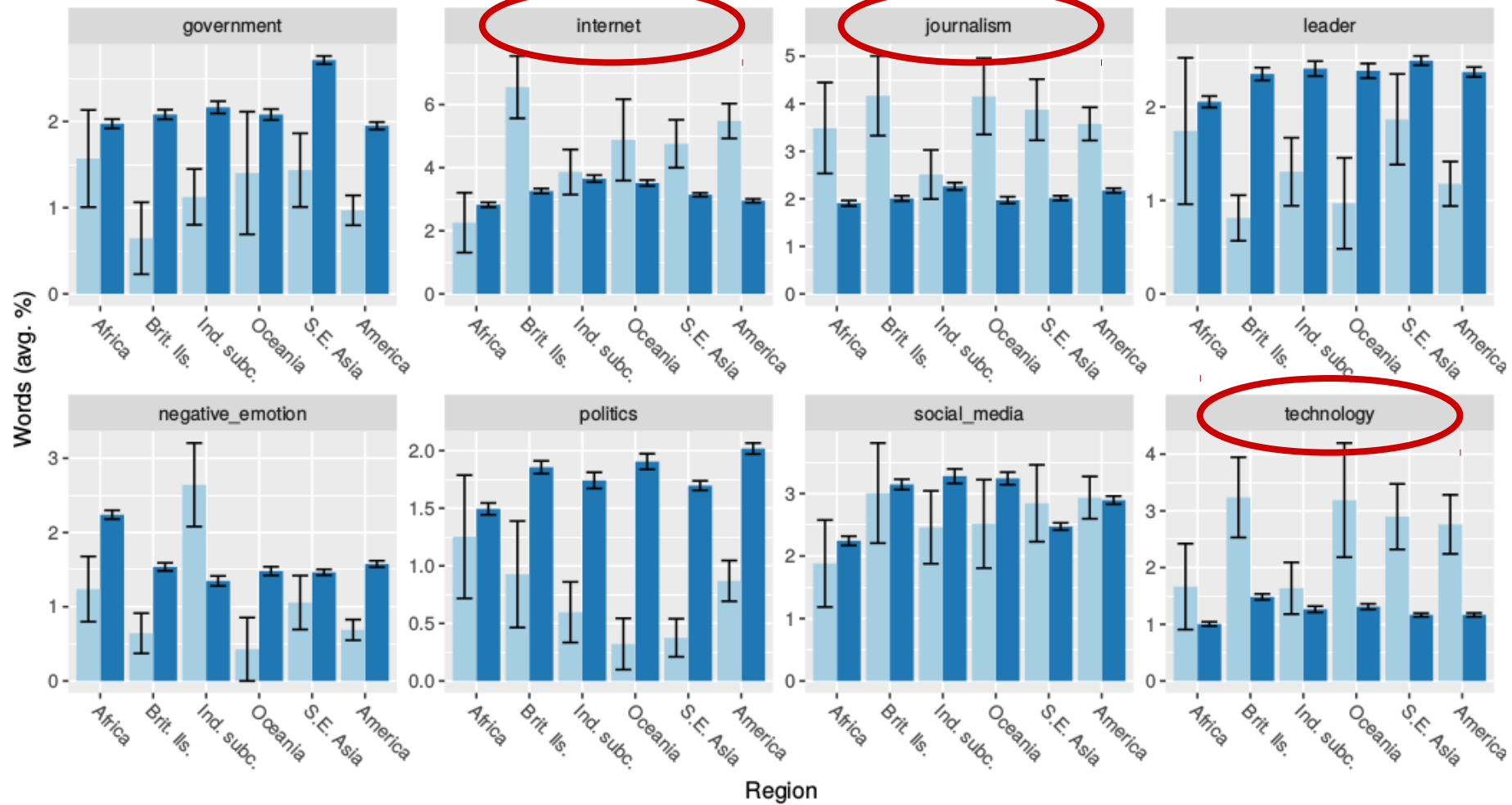
period before US election after US election

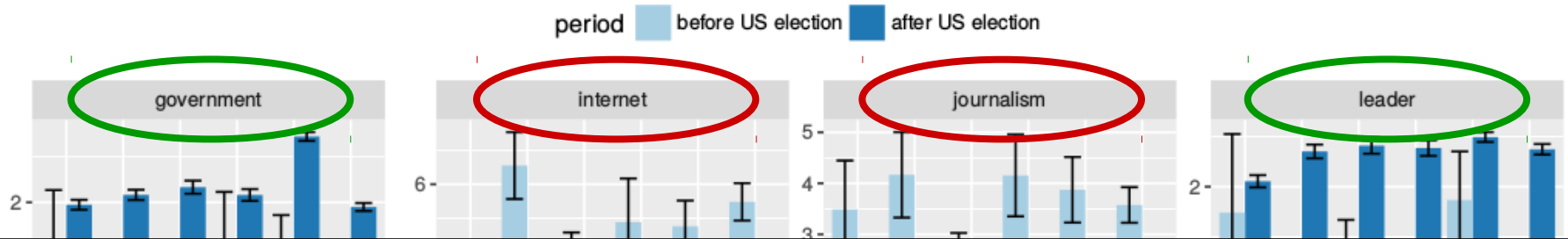


period before US election after US election



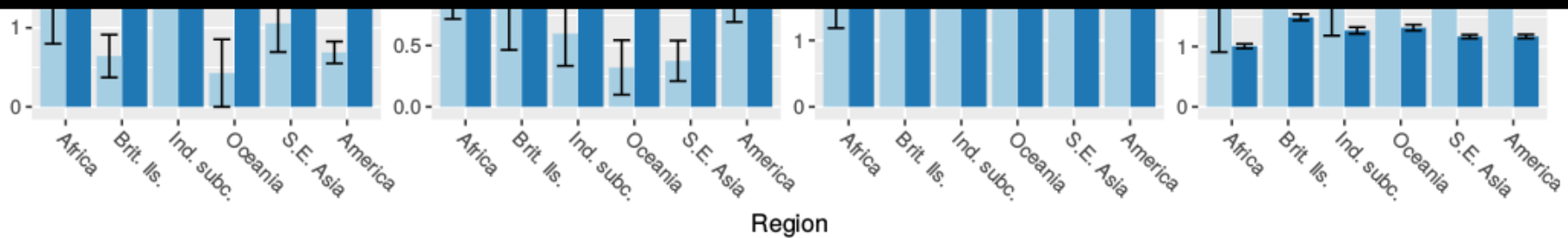
period before US election after US election





Before: “fake news” more mentioned with terms related to the **environment where they occur** (Internet, newspapers etc.)

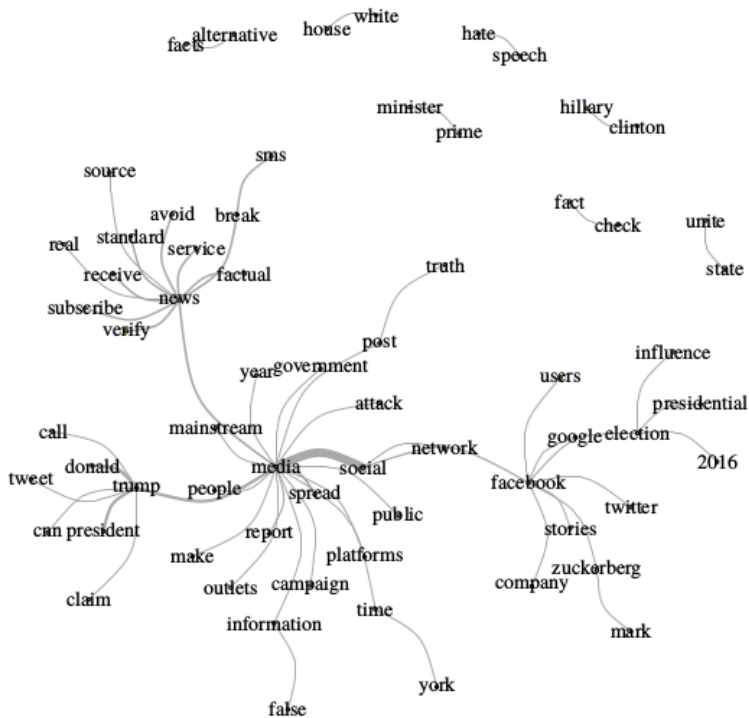
After: “fake news” more mentioned with terms more close to the **content of the news themselves** (politics, elections etc.).



3 → Co-occurrence networks



(a) Before US election



(b) After US election

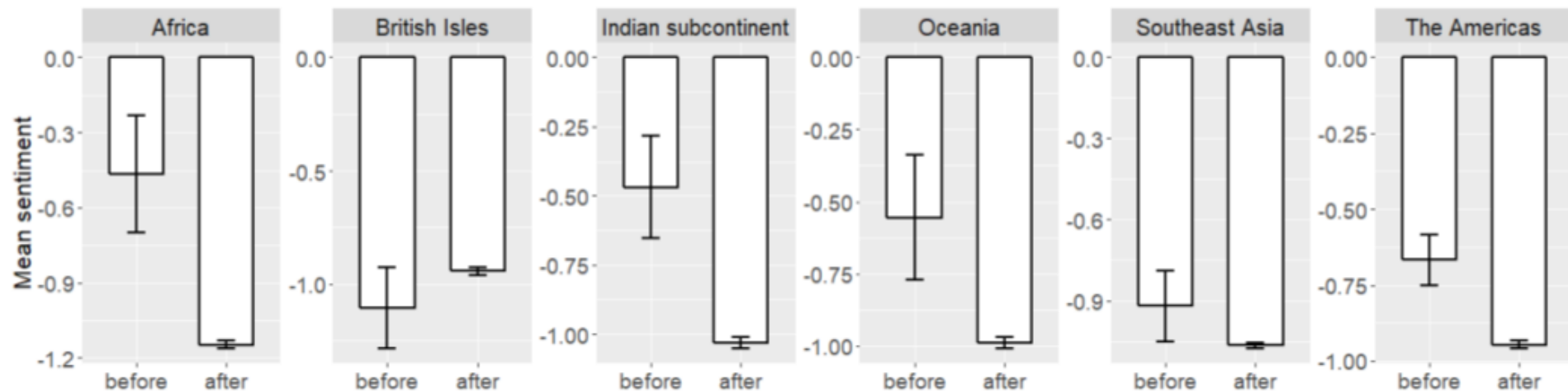
4 → Topics addressed

Topics obtained with latent Dirichlet allocation (LDA)

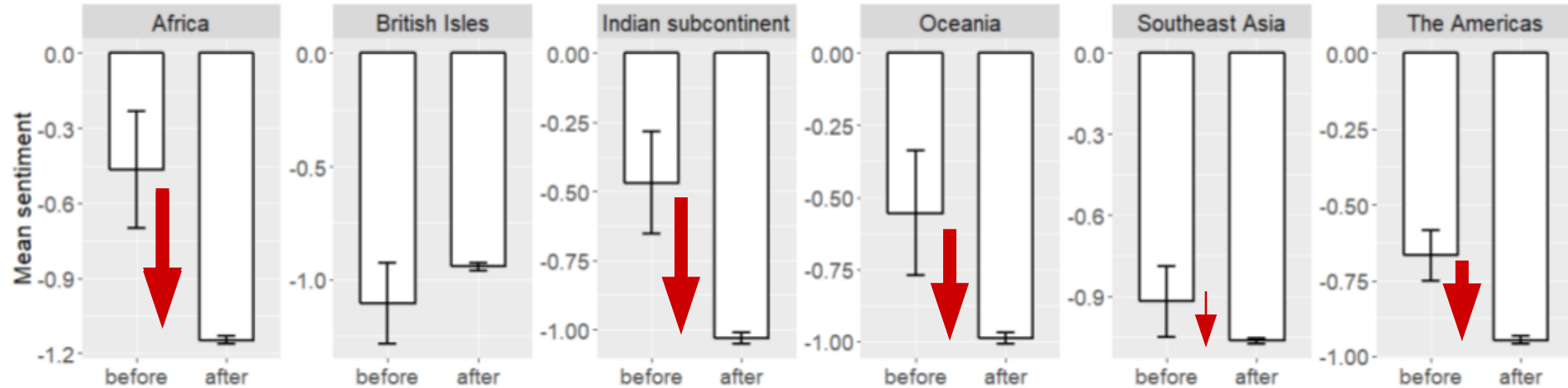
Before the 2016 US election: less cohesive topics, but most are related to **Internet** and to the **spread of misinformation, in all regions**

After the 2016 US election: a relevant frequency of words related to **politics and online social networks** in the main topics

5 → Polarity



5 → Polarity



Concept of fake news increasingly associated with **negative concepts**, like misinformation, manipulation and extremism

Summary of main results

Interest for the term “fake news” **increased** → rise of **news about it** and of **Google Search queries** for this expression

Growth accompanied by **change of framing** → from the **media industry** itself to **political affairs**

Entities not only **changed towards political topics**, but also **suffered from global standardization** after the US election

Negativity of the contexts **increased** after 2016

Summary of main results

Interest for the term “fake news” **increased** → rise of **news about it** and of **Google Search queries** for this expression

Growth accompanied by **change of framing** → from the **media industry** itself to **political affairs**

Entities not only **changed towards political topics**, but also **suffered from global standardization** after the US election

Negativity of the contexts **increased** after 2016

Summary of main results

Interest for the term “fake news” **increased** → rise of **news about it** and of **Google Search queries** for this expression

Growth accompanied by **change of framing** → from the **media industry** itself to **political affairs**

Entities not only **changed towards political topics**, but also **suffered from global standardization** after the US election

Negativity of the contexts **increased** after 2016

Summary of main results

Interest for the term “fake news” **increased** → rise of **news about it** and of **Google Search queries** for this expression

Growth accompanied by **change of framing** → from the **media industry** itself to **political affairs**

Entities not only **changed towards political topics**, but also **suffered from global standardization** after the US election

Negativity of the contexts **increased** after 2016

Concluding remarks

Our goal: a **quantitative** analysis on the **perception and conceptualization** of the term “**fake news**” in a corpus of **news articles** published from **2010 to 2018** in **20 countries**

Rise of interest in the term “fake news” was accompanied by **changes in its conceptualization and in the perception about it**, as hypothesized

Possibilities of future work: consider a **larger spectrum of periods** and add analyses using **data from other sources**, such as **edits of the Wikipedia article on fake news**

Concluding remarks

Our goal: a **quantitative** analysis on the **perception and conceptualization** of the term “**fake news**” in a corpus of **news articles** published from **2010 to 2018** in **20 countries**

Rise of interest in the term “fake news” was accompanied by **changes in its conceptualization and in the perception about it**, as hypothesized

Possibilities of future work: consider a **larger spectrum of periods** and add analyses using **data from other sources**, such as **edits of the Wikipedia article on fake news**

Concluding remarks

Our goal: a **quantitative** analysis on the **perception and conceptualization** of the term “**fake news**” in a corpus of **news articles** published from **2010 to 2018** in **20 countries**

Rise of interest in the term “fake news” was accompanied by **changes in its conceptualization and in the perception about it**, as hypothesized

Possibilities of future work: consider a **larger spectrum of periods** and add analyses using **data from other sources**, such as **edits of the Wikipedia article on fake news**

Fake news as we feel it:

Perception and conceptualization of the term
"fake news" in the media

Thank you!

Obrigado!

Спасибо!

UF *m* G

UNIVERSIDADE FEDERAL
DE MINAS GERAIS



Universiteit
Leiden



**BERKMAN
KLEIN CENTER**
FOR INTERNET & SOCIETY
AT HARVARD UNIVERSITY

Evandro L. T. Paradela Cunha

www.dcc.ufmg.br/~evandrocunha

evandrocunha@dcc.ufmg.br

twitter.com/Cunha_et_al
