

ERC - CONFAP Grant 2018

# Fake News and Artificial Intelligence

Dante A. C. Barone



European Research Council  
Established by the European Commission



# Overview

## 1. Introduction

## 2. Fake News

- a. Types and Motivations
- b. When it started
- c. How to stop Fake News

## 3. Artificial Intelligence and Fake News

- a. Challenges
- b. Research works
- c. Same solutions

## 4. ENC - CONFAP results

- a. L3S Research Center
- b. Alexandria
- c. TSE
- d. Students and Researchers at UFRGS

# Introduction

**Dante A. C. Barone**

Computer Scientist, Full Professor of  
the Informatics Institute of the  
Federal University of Rio Grande do  
Sul (UFRGS)

**ERC - CONFAP 2017 call for  
participations**

January of 2018

Hannover - Germany



# L3S Research Center

Forschungszentrum  Research Center



**Prof. Dr. tech. Wolfgang Nejdl**

## About L3S

› L3S Overview

Mentoring Guidelines

PhD Program

Graduations

Memberships

L3S in Facts

News

L3S-Team and Boards

Projects

Startup@L3S

Jobs

Events

Publications

Downloads

Contact

## About the L3S Research Center

L3S is a research center for basic and applied Web Science. L3S researchers develop innovative methods and technologies that enable intelligent and seamless access to information on the Web, link individuals and communities in all aspects of the knowledge society, and connect the Internet to the real world and its institutions. L3S is also focused on transferring its knowledge and technology to society and the economy.

Through cooperation and joint projects in the area of information retrieval, databases, semantic web, performance modeling, service computing and mobile networks, L3S has brought together numerous computer scientists with scientists from other disciplines, such as law, library science, linguistics, psychology, sociology and economics. L3S scientists cooperate in a variety of projects assigned to six research fields:

- [Intelligent Access to Information](#)
- [Next Generation Internet](#)
- [E-Science](#)
- [Web Governance](#)
- [Virtual Communities](#)
- [Cyber-Physical Systems](#)

The total research budget amounts to more than 6 million Euros per year. The long standing experience of L3S with support programs of the European Union has led to a number of collaborations with research institutes and companies across Europe. L3S is currently involved in 17 projects under the [HORIZON 2020](#) program including the projects [SoBigData](#), [iBorderCtrl](#), [AFEL](#), [Carsmand](#), [Things2Do](#) and [Desir](#). Moreover, the L3S coordinates the projects [Qualimaster](#) and the infrastructure project [e-IRSP5](#). The Starting Grant from the European Research Council (ERC) for [UniQue](#) and the ERC Advanced Grants [Alexandria](#) complete the project diversity.



Prof. Dr. Wolfgang Nejdl has been full professor of computer science at the University of Hannover since 1995. Heads the L3S Research Center as well as the Distributed Systems Institute / Knowledge Based Systems.

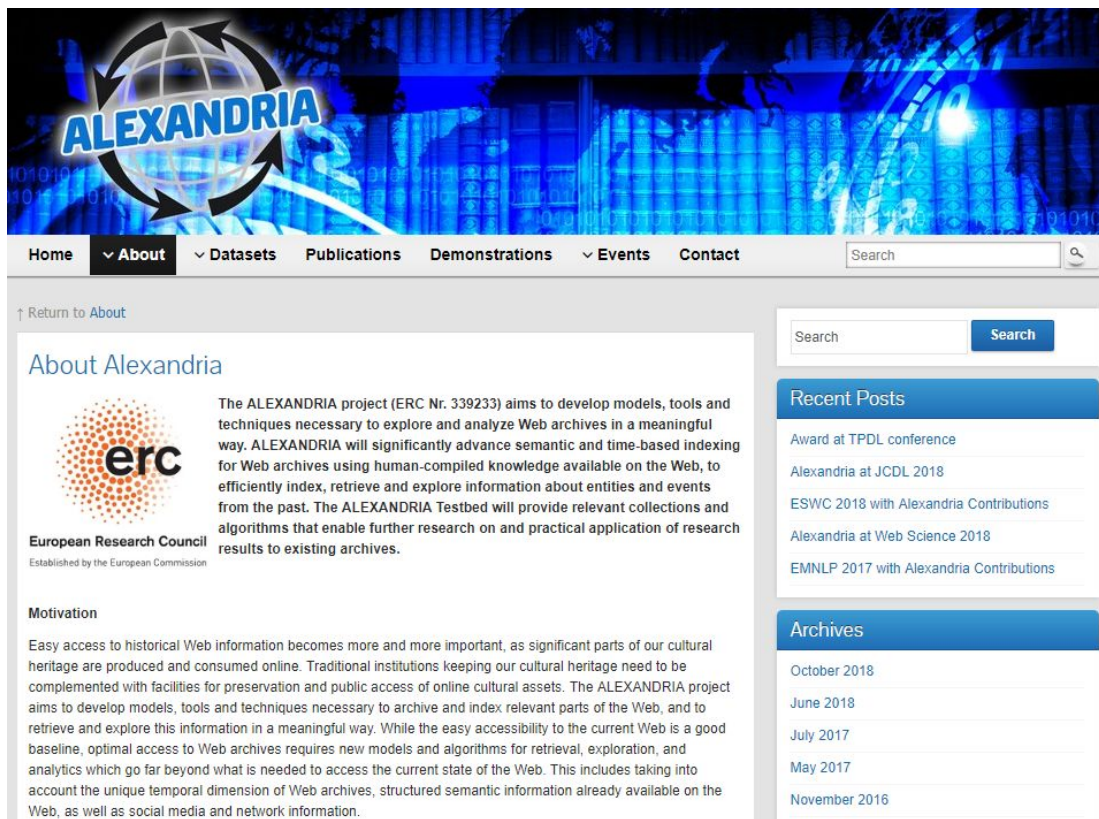
L3S downloads



L3S@WORK

# Alexandria

The ALEXANDRIA project aims to develop models, tools and techniques necessary to **explore and analyze Web archives** in a meaningful way




Home About Datasets Publications Demonstrations Events Contact

Search

↑ Return to About

## About Alexandria



**erc**  
European Research Council  
Established by the European Commission

The ALEXANDRIA project (ERC Nr. 339233) aims to develop models, tools and techniques necessary to explore and analyze Web archives in a meaningful way. ALEXANDRIA will significantly advance semantic and time-based indexing for Web archives using human-compiled knowledge available on the Web, to efficiently index, retrieve and explore information about entities and events from the past. The ALEXANDRIA Testbed will provide relevant collections and algorithms that enable further research on and practical application of research results to existing archives.

### Motivation

Easy access to historical Web information becomes more and more important, as significant parts of our cultural heritage are produced and consumed online. Traditional institutions keeping our cultural heritage need to be complemented with facilities for preservation and public access of online cultural assets. The ALEXANDRIA project aims to develop models, tools and techniques necessary to archive and index relevant parts of the Web, and to retrieve and explore this information in a meaningful way. While the easy accessibility to the current Web is a good baseline, optimal access to Web archives requires new models and algorithms for retrieval, exploration, and analytics which go far beyond what is needed to access the current state of the Web. This includes taking into account the unique temporal dimension of Web archives, structured semantic information already available on the Web, as well as social media and network information.

Search Search

### Recent Posts

- Award at TPDL conference
- Alexandria at JCDL 2018
- ESWC 2018 with Alexandria Contributions
- Alexandria at Web Science 2018
- EMNLP 2017 with Alexandria Contributions

### Archives

- October 2018
- June 2018
- July 2017
- May 2017
- November 2016

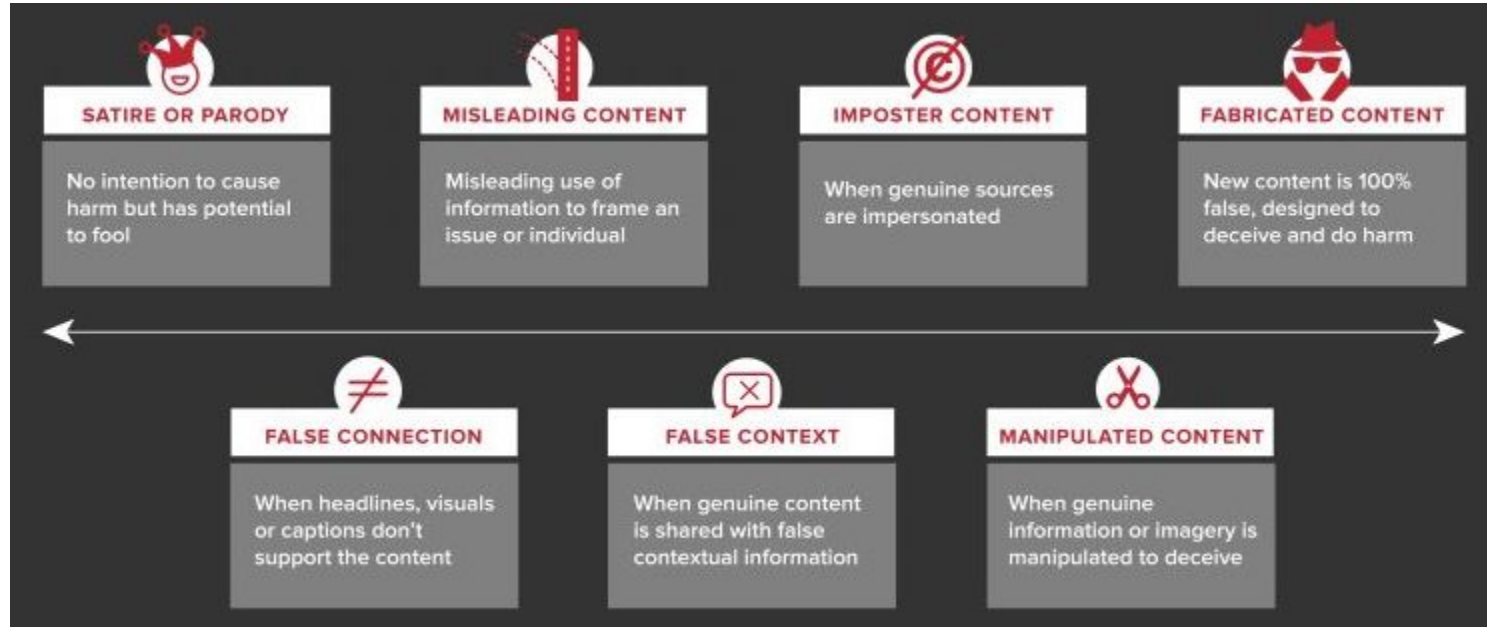
# What is Fake News?

*“ Fake news is create with the intent to **mislead** in order to **damage** an agency, entity, or person, and/or **gain** financially or politically, often using sensationalist, dishonest, or outright headlines to **increase readership**, online **sharing**, and Internet click.” - Wikipedia*

## Consequences:

- Influencing political elections
- Generate hate
- Discrimination

# Different types of Fake Content



Source: <https://firstdraftnews.org/fake-news-complicated/>

# The motivations

	 SATIRE OR PARODY	 FALSE CONNECTION	 MISLEADING CONTENT	 FALSE CONTEXT	 IMPOSTER CONTENT	 MANIPULATED CONTENT	 FABRICATED CONTENT
POOR JOURNALISM		✓	✓	✓			
TO PARODY	✓				✓		✓
TO PROVOKE OR TO 'PUNK'					✓	✓	✓
PASSION				✓			
PARTISANSHIP			✓	✓			
PROFIT		✓			✓		✓
POLITICAL INFLUENCE			✓	✓		✓	✓
PROPAGANDA			✓	✓	✓	✓	✓

Source: <https://firstdraftnews.org/fake-news-complicated/>



# When it started

**In the 13th century BC**, Rameses the Great spread lies and propaganda portraying the Battle of Kadesh as a stunning victory for the Egyptians

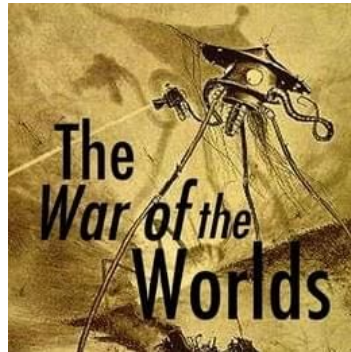
**In 1475**, a fake news story in Trent claimed that the Jewish community had murdered a two-and-a-half-year-old Christian infant named Simonino. The story resulted in all the Jews in the city being **arrested** and **tortured**.

**In 1835**, The New York Sun published articles about a real-life astronomer had observed bizarre life on the moon. The fictionalized articles successfully attracted **new subscribers**.

# When it started

**During the First World War**, was alleged a "German Corpse Factory" in which the German battlefield dead were rendered down for fats used to make nitroglycerine, candles, lubricants, human soap, and boot dubbing.

**In 1938**, an episode of the American radio drama, called "The War of the Worlds" presented as a series of simulated news bulletins, caused mass hysteria on the listeners. A group of people attacking a water tower because they falsely identified it as an alien.



# When it started

In the 21st century, the impact of Fake News became widespread. The capacity to mislead was enhanced by the widespread use of social media.

The term "**fake news**" gained importance with the electoral context in Western Europe and North America.



# Cambridge Analytica

- Private company combining **mining** and **data analysis** with strategic communication for the electoral process
- The company was able to target specific Fake News to a particular social group according to its profiles traced through social networks
- It would have played a crucial role in the **Brexit** event



# Cambridge Analytica

- Also, they had a role in **the election of Donald Trump**
- One of the objectives was to spread the idea of an Alternative Right
- **Example:** make Republicans go from the traditional " I do not like taxes " to " Obama will steal my guns with a secret army "



# Fake News Example


## WTOE 5 NEWS

YOUR LOCAL NEWS NOW

TOP STORIES COMMUNITY ENTERTAINMENT SPORTS LIFE ABOUT LATEST NEWS VID

### Pope Francis Shocks World, Endorses Donald Trump for President, Releases Statement

TOPICS: Pope Francis Endorses Donald Trump



Conservative Daily Review  
conservativedailyreview.com

Conservative Daily Review  
@ConservativeDailyReview

- Home
- Posts**
- Videos
- Photos
- About
- Likes

Like Message Share More



### EVIDENCE: Hillary Now Confirmed To Have Kuru Disease From Cannibalism (PROOF)

CLICK HERE  
CONSERVATIVEDAILYPOST.COM

Like Comment Share

Susan Munson, Connie Cook, Larry Bowling and 143 others like this. Top Comments

# How to combat Fake News sharing?



# HOW TO SPOT FAKE NEWS



## CONSIDER THE SOURCE

Click away from the story to investigate the site, its mission and its contact info.



## READ BEYOND

Headlines can be outrageous in an effort to get clicks. What's the whole story?



## CHECK THE AUTHOR

Do a quick search on the author. Are they credible? Are they real?



## SUPPORTING SOURCES?

Click on those links. Determine if the info given actually supports the story.



## CHECK THE DATE

Reposting old news stories doesn't mean they're relevant to current events.



## IS IT A JOKE?

If it is too outlandish, it might be satire. Research the site and author to be sure.



## CHECK YOUR BIASES

Consider if your own beliefs could affect your judgement.



## ASK THE EXPERTS

Ask a librarian, or consult a fact-checking site.



# How to combat Fake News using Artificial Intelligence?



# Challenges

- Already there are challenges proposed to encourage research groups to propose solutions to identify and combat Fake News.
- **Stance Detection:** is a competition where participants must provide computational models capable of asserting whether two statements of text:
  - **Agree** on a subject
  - **Disagree** on a subject
  - **Discuss** but do not express the opinion on the subject
  - **Not Related** to the subject

<b>Rank</b>	<b>Team name</b>	<b>Affiliation</b>	<b>Accuracy</b>
1	SOLAT in the SWEN	Talos Intelligence	82.02
2	Athene (UKP Lab)	TU Darmstadt	81.97
3	UCL Machine Reading	UCL	81.72

# Research works

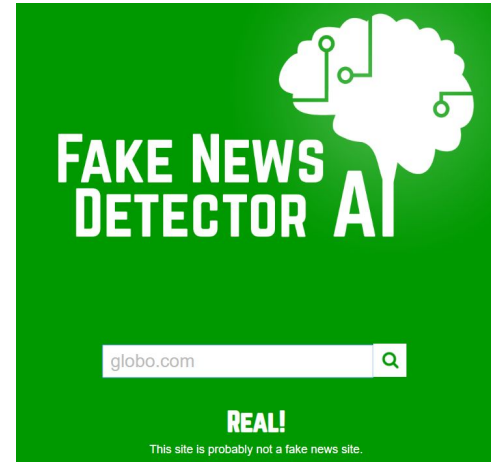
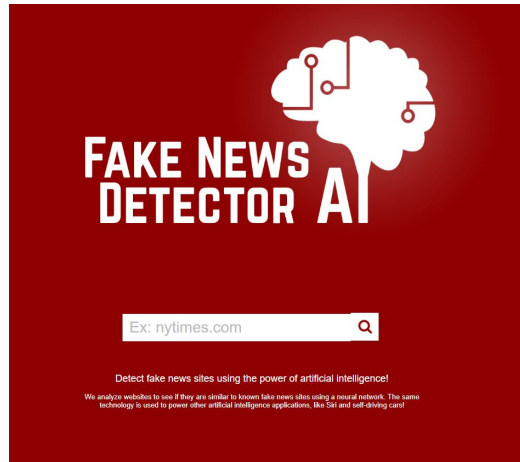
A study conducted during the campaign presidency of the United States observed the consumption of sites of 2,525 voters, concluding that:

- Fake News is more consumed among voters with **extreme political stances**
- Most texts were "overwhelmingly **pro-Trump**"
- Pro-Trump voters were **three times more likely** to consume Fake News
- The direct relationship between **Facebook** use and visits to Fake News sites
- The number suggests that consumption of **fact-checking** sites is more concentrated among voters who **do not read Fake News**

# Solutions

The **FakeNewsAI** is a website capable identify whether a given website spread Fake News or not.

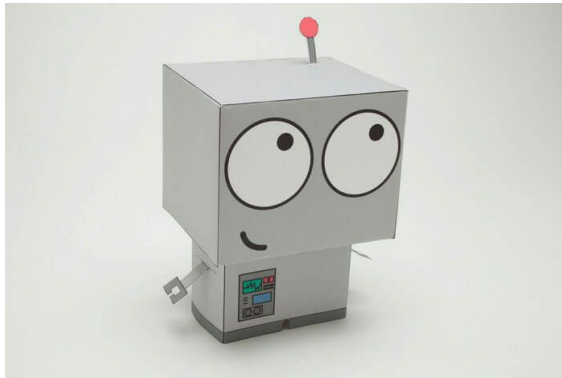
[www.fakenewsai.com](http://www.fakenewsai.com)



# Solutions

The Fake News Detector, called *Rubinho*, is an extension for Chrome and Firefox that allows you to detect and flag news directly from your Facebook into Legitimate, Fake News, Click Bait, Extremely Biased, Satire or Not news.

[www.fakenewsdetector.org](http://www.fakenewsdetector.org)



# ERC - CONFAP Results

# Students in the Fake News context

**Vinicius Woloszyn**

Ph.D. candidate in the Pos-Graduation Program in  
Computer Science at UFRGS



**Natural Language Processing, Natural Language Generation and  
Data Science**



**Scientific Researcher**

L3S Research Center  
2017 – o momento • 1 ano  
Hannover e Região, Alemanha

Fake News Detection



**Scientific Researcher**

Technische Universität Berlin  
2018 – o momento • menos de um ano  
Berlin e Região, Alemanha

Open Science

# Students in the Fake News context

## Eduardo G. Cortes

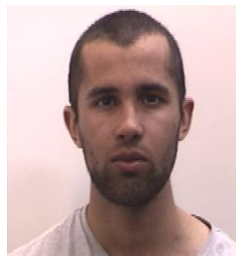


Master Degree student at UFRGS

**Natural Language Processing and  
Question Answering Systems**

Enrolled in: 2017/1

## Felipe Rosa



Undergraduate student at UFRGS  
in Computer Science

**Natural Language Processing and  
Named Entity Recognition**



# Students in the Fake News context

**Cristiano Dias**



Undergraduate student at UFRGS  
in Computer Science

**Big Data and Fake News**

# Current Research at UFRGS

## **Fake News Extractor**

This project is a collective effort to automatically extract claims from Fact-Checking websites and then consolidate a large data set for machine learning purposes. Currently, these claims are available for English, Portuguese and German.

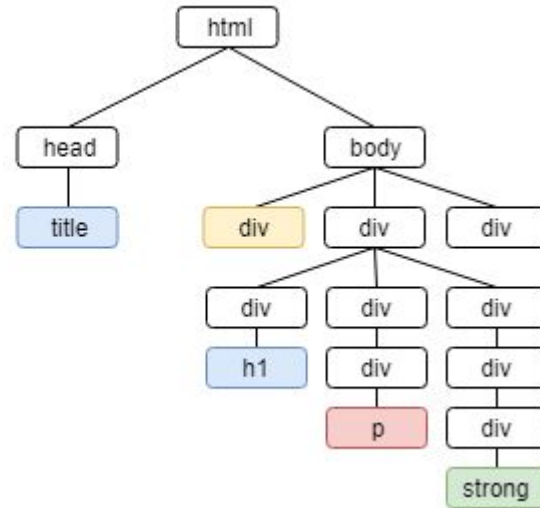
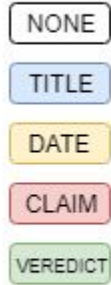
**[https://github.com/vwoloszyn/fake\\_news\\_extractor](https://github.com/vwoloszyn/fake_news_extractor)**

# Current Research at UFRGS

## Claim2Features

This project aims to give access to a smart parser to extract data full automatically from fact-checking sites without support from humans.

### Classes

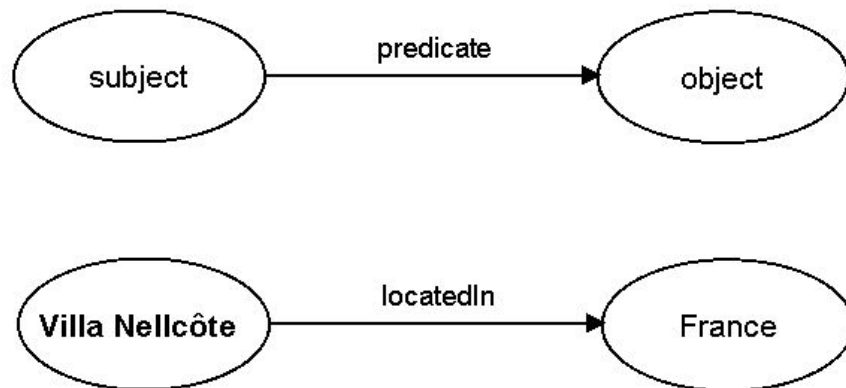


<https://github.com/vwoloszyn/claim2features>

# Current Research at UFRGS

## Fake News Knowledge Base

It's a new project that aims to create a **knowledge base** using data from **fact-checking websites**. With this resource, we intend to check if a sentence is true or false through a crossing of sentence information.



Seminário Internacional  
**Brasil - União Europeia**

**FAKE NEWS:**  
EXPERIÊNCIAS E DESAFIOS

**21/06/2018**

MINISTÉRIO DAS  
**RELAÇÕES  
EXTERIORES**



União Europeia



MINISTÉRIO DO  
**PLANEJAMENTO,  
DESENVOLVIMENTO E GESTÃO**



# Thank you !

Contacts: [barone@inf.ufrgs.br](mailto:barone@inf.ufrgs.br)



European Research Council  
Established by the European Commission

