

# FANDANGO META-FORUM

## 09.10.2019

FRANCESCO SAVERIO NUCCI



# FANDANGO

# FANDANGO Identity Card

- ▶ **Topic: ICT-14-2017:** Big Data PPP: cross-sectorial and cross-lingual data integration and experimentation **Topic: a) Data integration activities**
- ▶ **Project Acronym: FANDANGO:** FAke News discovery and propagation from big Data ANalysis and artificial inteliGence Operations
- ▶ **Consortium: 8 organisations** from **5 countries** (Italy, Belgium, Greece, Ireland, Spain)
- ▶ **Application Areas:** Climate (VRT), Immigration (CIVIO), The European Policies (ANSA, VRT)
- ▶ **Duration:** 36 months
- ▶ **Funding:** € 2.879.250

# Fandango Multidisciplinary team

- ▶ 1. **ICT - large industry and Small Medium Enterprises (ENG, LiveTech and Siren)**: to implement, integrate and customise the **technology platform and its components** and **exploitation of the final results** in the European and Global Big Data markets.
- ▶ 2. **Media and Information - Large Industries (VRT, ANSA)** to provide data and specific tools, run the pilots, test FANDANGO solutions and disseminate best practices to other Media and Information Industries.
- ▶ 3. **Research organisations (CERTH, UPM)**: to provide high-quality research results on several disciplines and in several FANDANGO components
- ▶ 4. **No profit organisation (Civio)** to provide different expertise on specific data transparency and data journalism.

# FAKE NEWS DAILY

Misleading  
headline

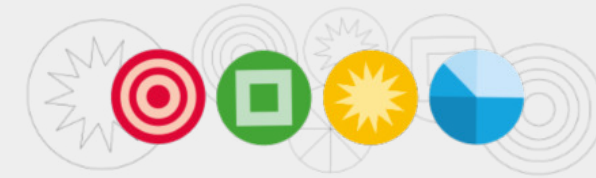


Alternative facts

## BIG DATA AND AI AGAINST FAKE NEWS

# Fake News Challenge

- ▶ Fake News is a complex problem with many different perspectives:
- ▶ **Educational** and cultural one:
  - ▶ Improve the people attention of the problem, see Finland example,
  - ▶ polarization of the actors in social network, see studies by Quattrociochi and results from European HLEG
- ▶ **Political**
  - ▶ Cambridge Analytica affair
  - ▶ Results from HLEG on misinformation
  - ▶ EU is going to support a network of independent fact checkers, with infrastructure and technical tools
- ▶ But, naturally it is also a **technical** problem



# BEYOND 'FAKE NEWS'

## 10 TYPES OF MISLEADING NEWS

| Type                         | Description  | Type                         | Description  | Impact  | Motivation            |
|------------------------------|--|------------------------------|--|---------|-----------------------|
| <b>propaganda</b><br>        | <ul style="list-style-type: none"> <li>adopted by governments, corporations and non-profits to manage attitudes, values and knowledge</li> <li>appeals to emotions</li> <li>can be beneficial or harmful</li> </ul>  | <b>partisan</b><br>          | <ul style="list-style-type: none"> <li>ideological and includes interpretation of facts but may claim to be impartial</li> <li>privileges facts that conform to the narrative whilst forgoing others</li> <li>emotional and passionate language</li> </ul>                             | neutral | politics/power        |
| <b>clickbait</b><br>         | <ul style="list-style-type: none"> <li>eye catching, sensational headlines designed to distract</li> <li>often misleading and content may not reflect headline</li> <li>drives ad revenue</li> </ul>   | <b>conspiracy theory</b><br> | <ul style="list-style-type: none"> <li>tries to explain simply complex realities as response to fear or uncertainty</li> <li>not falsifiable and evidence that refutes the conspiracy is regarded as further proof of the conspiracy</li> <li>rejects experts and authority</li> </ul> | low     | humour/fun            |
| <b>sponsored content</b><br> | <ul style="list-style-type: none"> <li>advertising made to look like editorial</li> <li>potential conflict of interest for genuine news organisations</li> <li>consumers might not identify content as advertising if it is not clearly labeled</li> </ul> | <b>pseudoscience</b><br>     | <ul style="list-style-type: none"> <li>purveyors of greenwashing, miracle cures, anti-vaccination and climate change denial</li> <li>misrepresents real scientific studies with exaggerated or false claims</li> <li>often contradicts experts</li> </ul>                              | medium  | money                 |
| <b>satire and hoax</b><br>   | <ul style="list-style-type: none"> <li>social commentary or humour</li> <li>varies widely in quality and intended meaning may not be apparent</li> <li>can embarrass people who confuse the content as true</li> </ul>                                     | <b>misinformation</b><br>    | <ul style="list-style-type: none"> <li>includes a mix of factual, false or partly-false content</li> <li>intention may be to inform but author may not be aware that the content is false</li> <li>false attributions, doctored content and misleading headlines</li> </ul>            | high    | (mis)inform           |
| <b>error</b><br>             | <ul style="list-style-type: none"> <li>established news organisations sometimes make mistakes</li> <li>mistakes can hurt the brand, offend or result in litigation</li> <li>reputable orgs publish apologies</li> </ul>                                    | <b>bogus</b><br>             | <ul style="list-style-type: none"> <li>entirely fabricated content spread intentionally to disinform</li> <li>guerrilla marketing tactics; bots, comments and counterfeit branding</li> <li>motivated by ad revenue, political influence or both</li> </ul>                            | high    | money, politics/power |

### DIG DEEPER...

|                          |  |                         |  |
|--------------------------|--|-------------------------|--|
| <b>false attribution</b> | Authentic images, video or quotes are attributed to the wrong events or person | <b>misleading</b>       | Content does not represent what the headline and captions suggest                    |
| <b>counterfeit</b>       | Websites and Twitter accounts that pose as a well-known brand or person        | <b>doctored content</b> | Content, such as statistics, graphs, photos and video have been modified or doctored |



www.eavi.eu

- Different typologies of Fake News
- Classification example, demonstrating the topics complexity and the different possible approaches



N.B. The impact and motivation assignments are not definitive and should be used as a guide for discussion



# How to spot fake-news?

ACCORDING TO EPRS

European  
Parliamentary  
Research  
Service

## 8. JOIN THE MYTH-BUSTERS<sup>(2)</sup>

Keep on top of the latest tricks and narratives used by those spreading disinformation.

Report fake stories.  
Tell your friends.

## 1. CHECK THE CONTENT

Are the facts and figures accurate? Is the article biased? A credible media outlet keeps one-sided opinions where they belong – in op-eds, not in news articles.

## 2. CHECK THE OUTLET

Do you know it? Does the URL look strange? Check the 'about' section. Who is behind it? Who is funding it? Double-check what other (trustworthy) sources say.

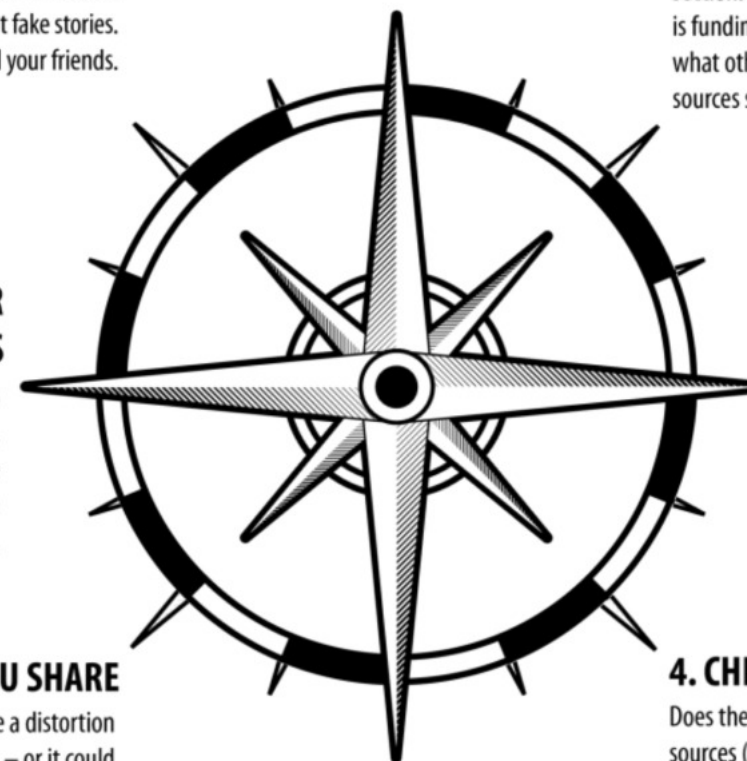
## 7. QUESTION YOUR OWN BIASES

Sometimes a story is just too good or entertaining to be true.

Take a deep breath, compare with reliable sources and keep a cool head.

## 3. CHECK THE AUTHOR

Does this person even exist? A well-respected journalist always has a track record. If the author has made up his or her name (or does not mention it), the rest is also likely to be fake.



## 6. THINK BEFORE YOU SHARE

The story could be a distortion of real or old events – or it could be satire. The headline could be designed to spark strong emotions. If an event is real, reliable media will cover it.

## 4. CHECK THE SOURCES

Does the author use reliable sources (for example, well-established and respected media outlets)? Are the quoted experts real specialists? If the story uses anonymous (or no) sources, it could be fake.

## 5. CHECK THE PICTURES

Images are powerful, and it is easy to manipulate them. An image search can show if it has been used before in a different context. The InVID plugin<sup>(1)</sup> can help you detect manipulation of videos or pictures.

# Fandango in a nutshell: what we do

In order to address the **aggressive emergence of fake news**, post-truths, and misinformation, Fandango provides online web app and services **to support media professionals** with the following high-level features:

- ▶ **Misinformation detection and trustworthiness scoring**, based on Big Data analysis techniques (ML models and Graph Analysis)
- ▶ Support **user data investigation**, through an interactive exploration of news, open data and verified claims databases

**Note:** We are not going to develop autonomous system, but supporting system! **Human** MUST be in the loop, to have the **final decision**



# Disinformation scoring features

Fandango provides a set of disinformation scoring features by **analyzing the different components of news:**

- ▶ **Text** (headline, body)
- ▶ **Authors & Source**
- ▶ **Media** (images, videos)

# Disinformation scoring features: Text analysis

**Machine learning model** will be trained to recognize features in the headline and body of a news: test a completely “context agnostic” classifier, it means that words are not considered as single feature in the model, lot of models trained, from simple to complex ones, and take the one with best performs in terms of accuracy and precision.

Features applied to headline and body of article: **Simple frequency features**

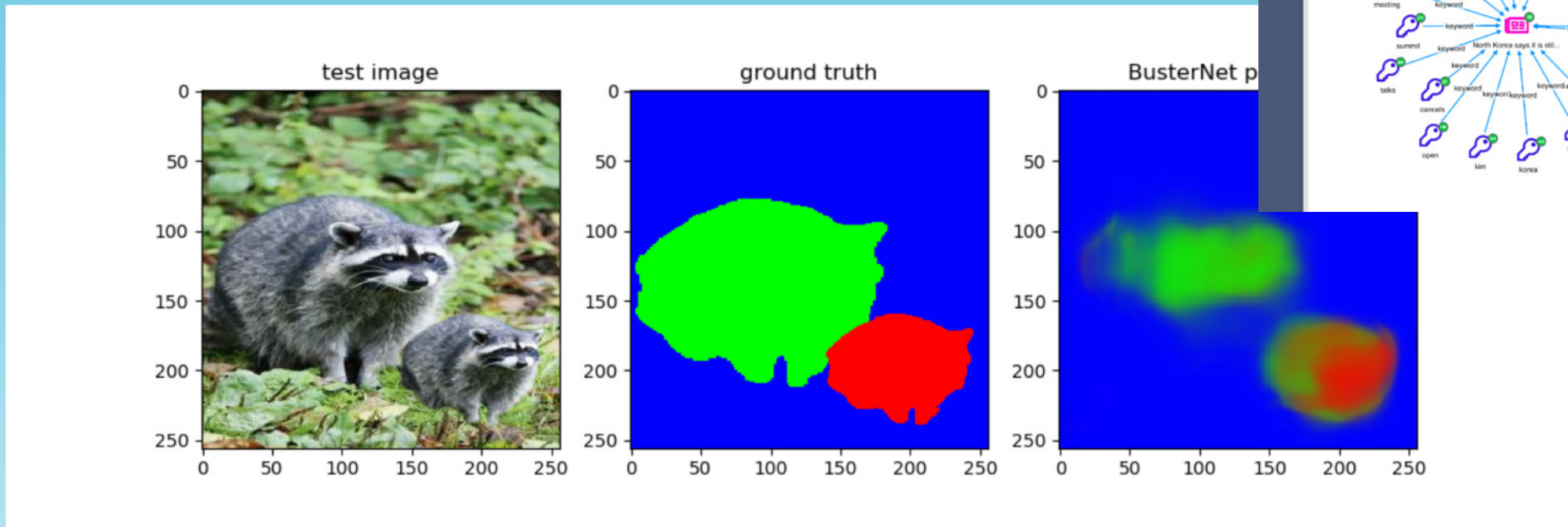
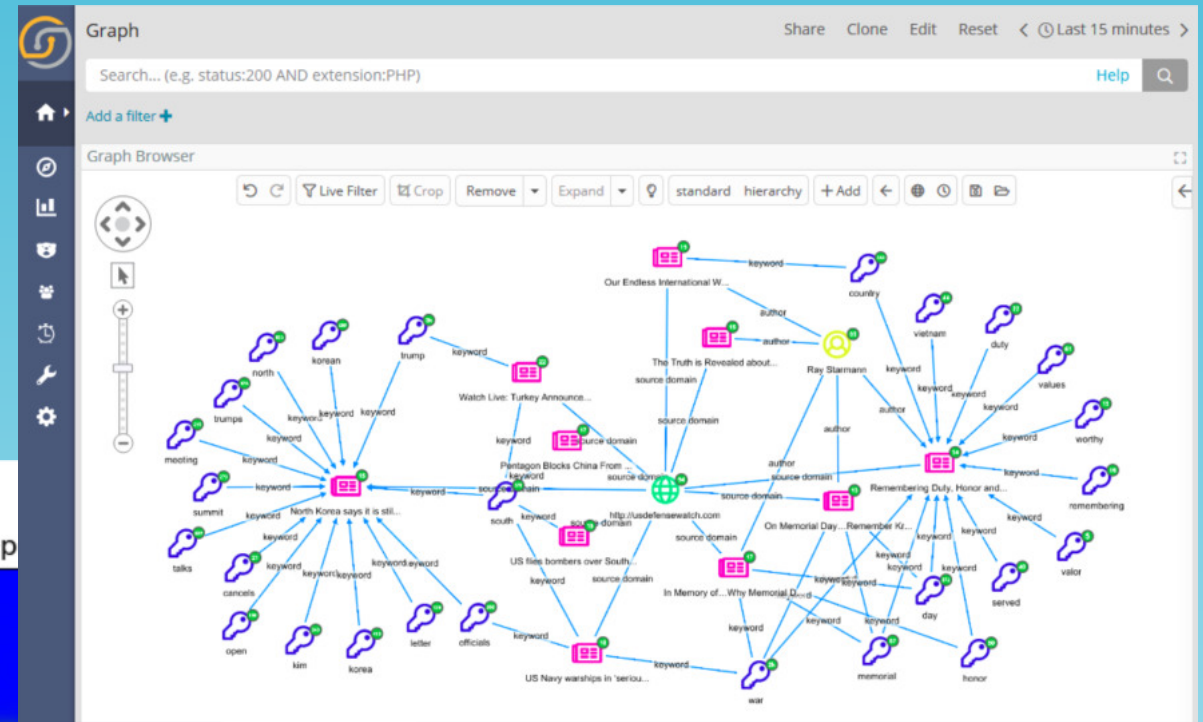
- ▶ Counting Stopwords - the stopwords are the most frequent word in a given language (e.s the, and, or, with.. ciao, il, la, ... ola etc.), Characters Counter , Punctuation Counter

**Part of The Speech (POS) features:** Counting POS (adjectives, adverbs, verbs, conjunctions)

**Advance frequency features:** Word average per paragraph, Lexical Diversity

**Readability indices**

# Other project approaches: Image, Video, DeepFake, Knowledge Graph



# Thank you !!

Francesco Saverio Nucci  
francesco.nucci@eng.it