FANDANGO

FANDANGO Project: Advanced analytics services to detect disinformation

ZAGABRIA, NEM SUMMIT 2019 , 23TH MAY 2019

European Commission | This project has received funding from the European Union's Horizon 2020 research and innovation programme under **grant agreement No 780355**

ENGINEERING   LIVETech LIVING TECHNOLOGY   CERTH CENTRE FOR RESEARCH & TECHNOLOGY HELLAS   SIREN DATA INTELLIGENCE   vrt   CIVIO   POLITÉCNICA   ANSA

BIG DATA AND AI AGAINST FAKE NEWS

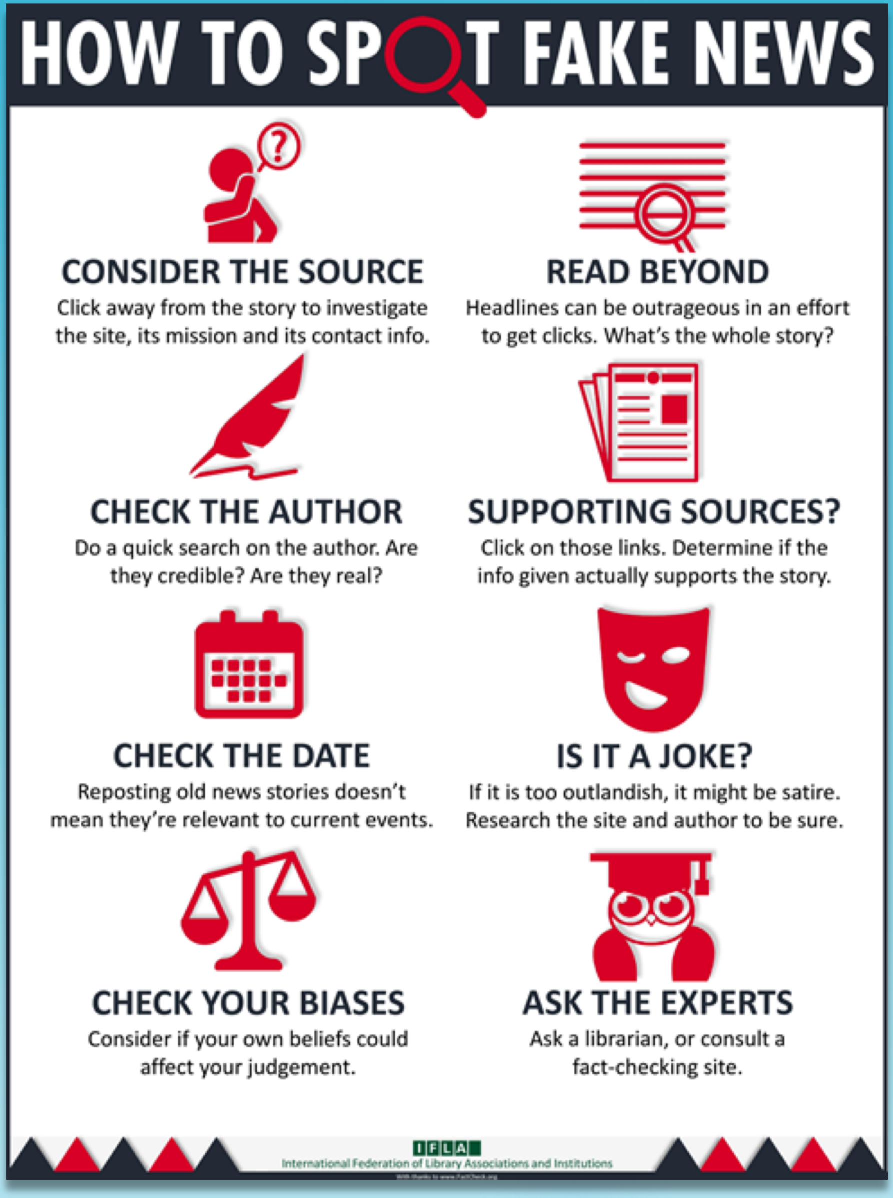
European Commission | This project has received funding from the European Union's Horizon 2020 research and innovation programme under **grant agreement No 780355**

ENGINEERING
LIVE Tech LIVING TECHNOLOGY
CERTH CENTRE FOR RESEARCH & TECHNOLOGY HELLAS
SIREN DATA INTELLIGENCE
vrt
CIVIO
POLITÉCNICA
ANSA

How to spot fake-news? According to IFLA

THE INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTION

(based on FactCheck.org's 2016 article How to Spot Fake News)



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.

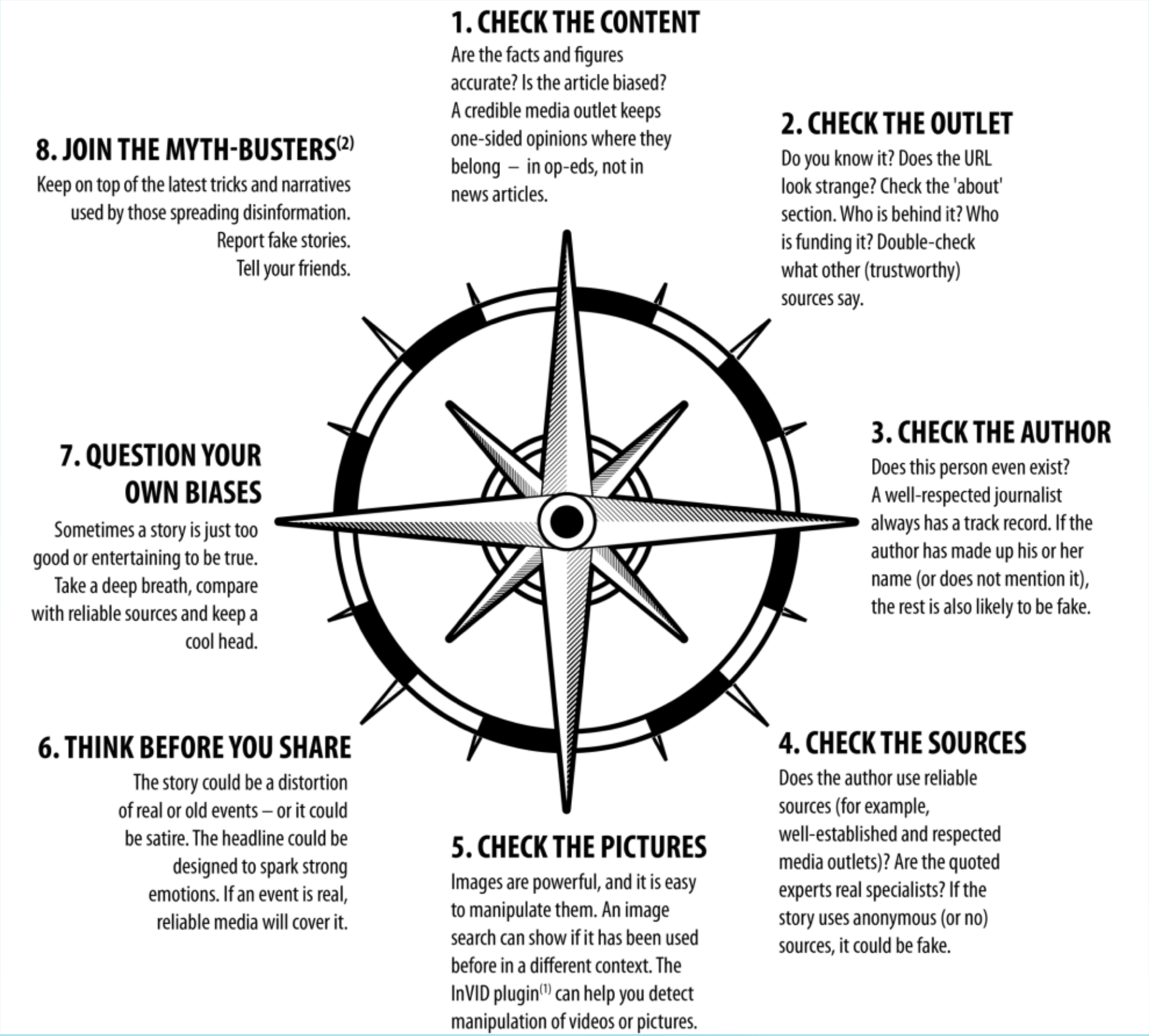
IFLA
International Federation of Library Associations and Institutions
With thanks to www.FactCheck.org

European Commission
This project has received funding from the European Union's Horizon 2020 research and innovation programme under **grant agreement No 780355**

FANDANGO

How to spot fake-news? According to EPRS

EUROPEAN PARLIAMENTARY RESEARCH SERVICE



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.

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.

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[1] can help you detect manipulation of videos or pictures.

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.

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.

8. JOIN THE MYTH-BUSTERS[2]

Keep on top of the latest tricks and narratives used by those spreading disinformation. Report fake stories. Tell your friends.

European Commission | This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 780355

How Fandango aims to tackle disinformation?

providing an online service that will support professionals with the following features:

- News disinformation detection and scoring, based on Big Data analysis techniques (ML models and Graph Analysis)
- data investigation, through an interactive exploration of news, open data and verified claims databases.

European Commission | This project has received funding from the European Union's Horizon 2020 research and innovation programme under **grant agreement No 780355**

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)

European Commission | This project has received funding from the European Union's Horizon 2020 research and innovation programme under **grant agreement No 780355**

Disinformation scoring features
Text analysis: our approach (ML)

- machine learning model will be trained to recognize features in the new's headline and body
- We are testing a completely "**context agnostic**" classifier, it means that words are not considered as single feature in the model.
- We are testing different machine learning algorithms in a greedy way, with an accurate features selection. It means that we train lot of models, from simple to complex ones, an take the one with best performs in terms of accuracy and precision.

European Commission | This project has received funding from the European Union's Horizon 2020 research and innovation programme under **grant agreement No 780355**

Disinformation scoring features
Text analysis: our ML features (1/2)

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)

European Commission | This project has received funding from the European Union's Horizon 2020 research and innovation programme under **grant agreement No 780355**

Disinformation scoring features
Text analysis: our ML features (2/2)

Features applied to title and text of an article are:

Advance frequency features

- Word average per paragraph
- Lexical Diversity

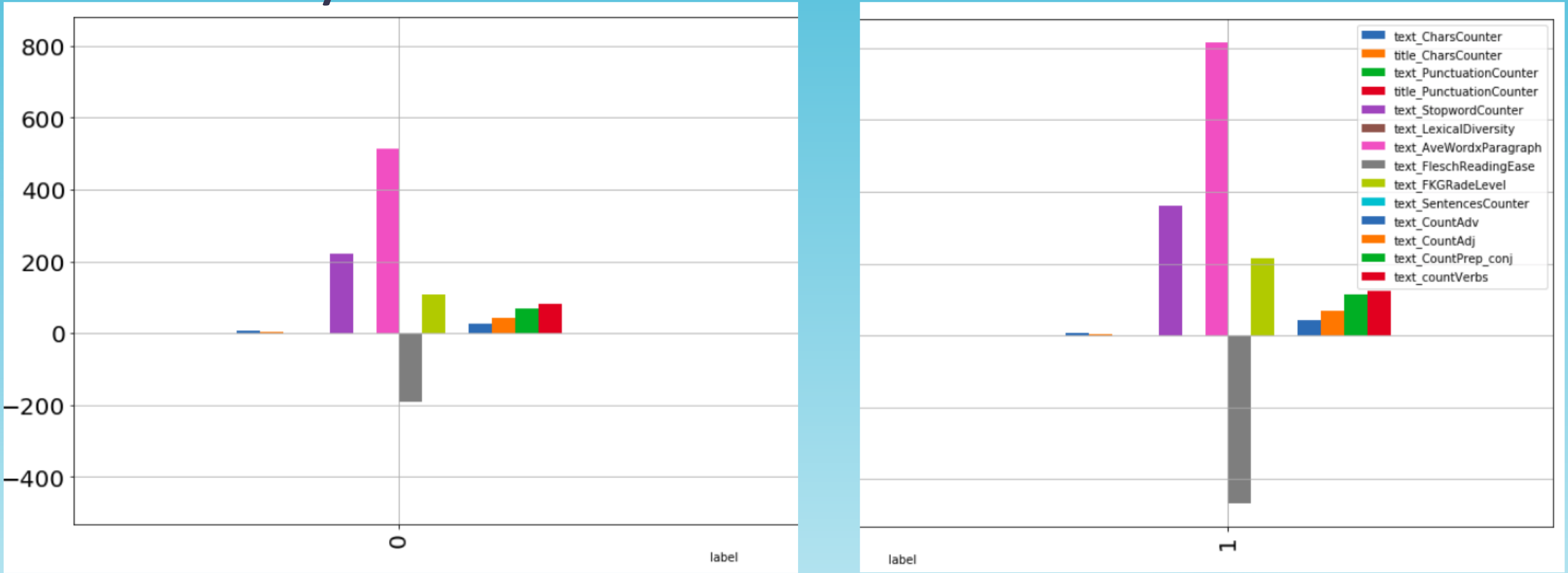
readability indices: readability tests designed to assess the understandability of a text.

- Flesch Index of readability
- FKG Read Level
- Automated Readability Index(ARI)

European Commission | This project has received funding from the European Union's Horizon 2020 research and innovation programme under **grant agreement No 780355**

Disinformation scoring features

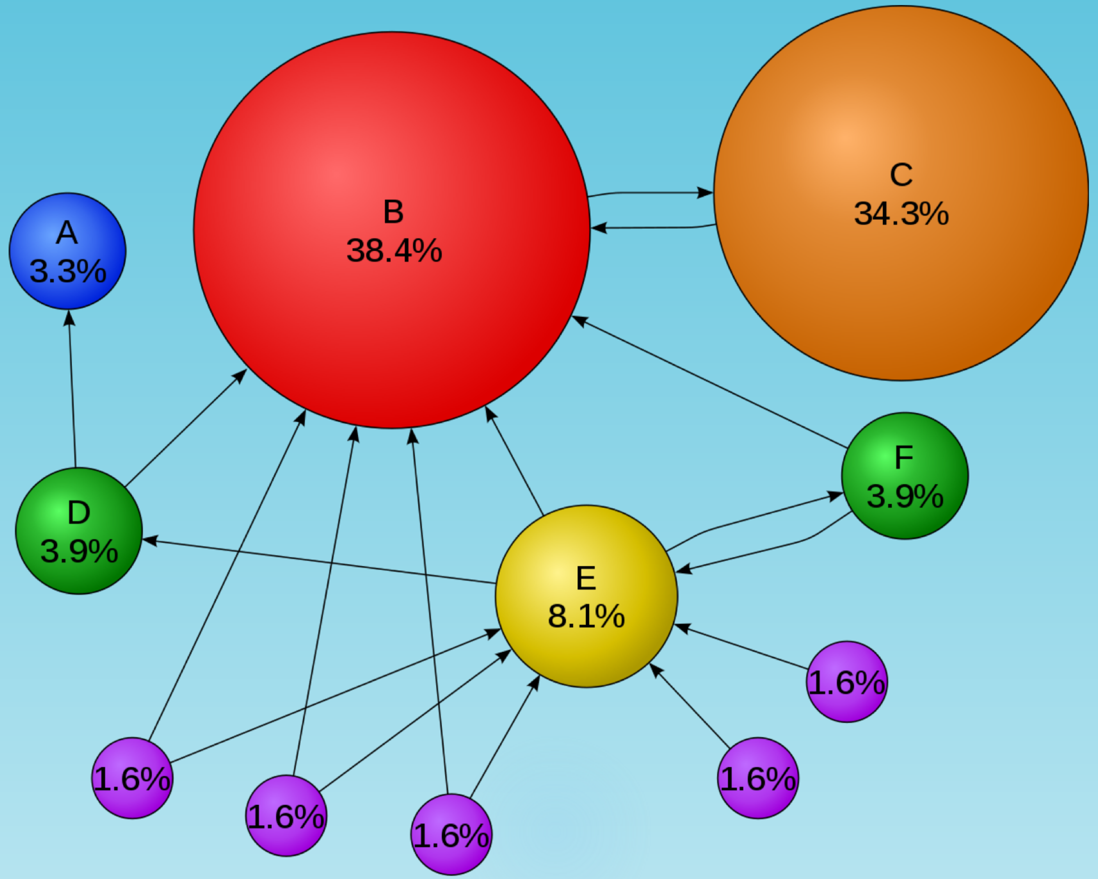
Text analysis: differences in term of features



European Commission | This project has received funding from the European Union's Horizon 2020 research and innovation programme under **grant agreement No 780355**

Disinformation scoring features Authors & Source analysis: our approach (GA)

- We apply importance algorithms to get the impact of each entity (e), including authors and organizations, in FANDANGO’s network.
- Weighted credibility indicator using the disinformation score of the articles provided by the text analyzer.
- Scoring result: will be between 0 and 1 if we have information enough for a particular entity, otherwise it will yield -99 to remark the lack of information to compute the analysis.



European Commission | This project has received funding from the European Union’s Horizon 2020 research and innovation programme under **grant agreement No 780355**

Disinformation scoring features
Authors & Source analysis: our approach (GA)

- In one hand, supervised learning has the most impressing results in the deep learning methodologies
- On the other hand, the problem with big data is that it's impossible to have fully annotated datasets to train the DL models in that manner
- a semi-supervised approach is the most realistic in terms of annotation effort / time compared to the final outcomes
- Graph based techniques are very effective in representing linearly non-separable data, but there is a limitation in working in a full adjacency matrix of the data, which translates to some millions of elements in a large dataset
- to overcome this issue we examined a sampling / batch learning approach that works quite effectively both in small scale and (most importantly) in large scale graphs

European Commission | This project has received funding from the European Union's Horizon 2020 research and innovation programme under **grant agreement No 780355**

Disinformation scoring features

Image & Video analysis

Spatiotemporal Analytics and out of context

Our goal is the detection of out of context content.
Our approach is based on comparing topics and entities extracted from the body of an article with the topic extracted from an image or video.

we are focused to detect 3 main types of out-of-context:

- two news about different topics but containing the same image or a manipulated version (i.e. using the same guerrilla images in two different war scenarios)
- two different news share the same image (or a manipulated version) but the publication dates are far from each other
- two different news share the same image (or a manipulated version) but the news are about two different location entities.

European Commission | This project has received funding from the European Union's Horizon 2020 research and innovation programme under **grant agreement No 780355**

Disinformation scoring features

Image & Video analysis

Spatiotemporal Analytics and out of context

The technique is to apply LDA NLP (Latent Dirichlet allocation - natural language processing) methods. This is the most common topic model currently in use, based on the intuition that documents cover a small number of topics and that topics often use a small number of words.

Input: News articles (**DATASET**) gathered from the web (as json files) and stored in 'Articles DB'.

Output: Entities, leading words, topics (for each article)

European Commission | This project has received funding from the European Union's Horizon 2020 research and innovation programme under **grant agreement No 780355**

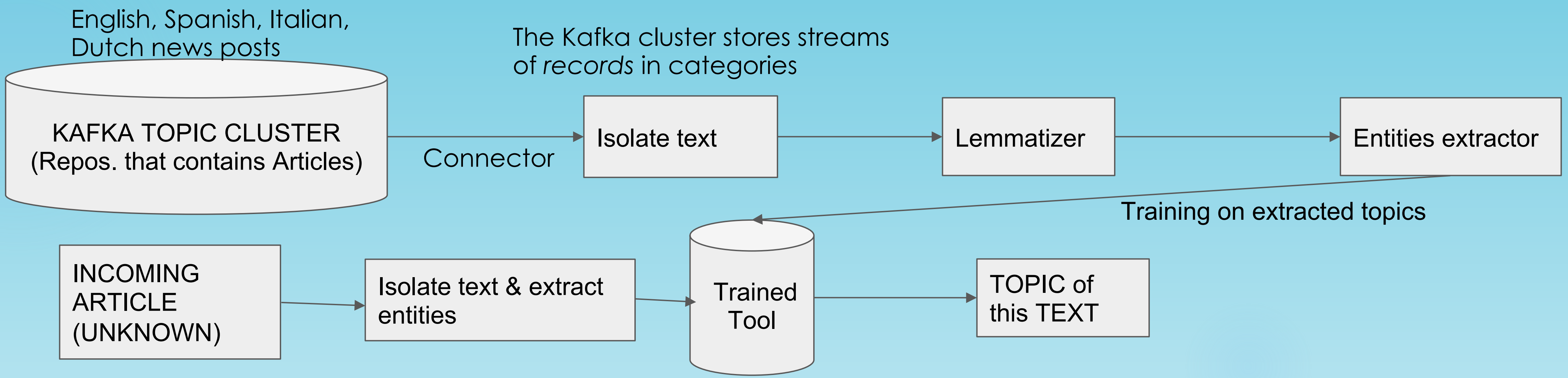
FANDANGO

Disinformation scoring features

Image & Video analysis

Spatiotemporal Analytics and out of context

Data Flow


English, Spanish, Italian,
Dutch news posts
The Kafka cluster stores streams of *records* in categories
KAFKA TOPIC CLUSTER
(Repos. that contains Articles)
Connector
Isolate text
Lemmatizer
Entities extractor
Training on extracted topics
INCOMING ARTICLE (UNKNOWN)
Isolate text & extract entities
Trained Tool
TOPIC of this TEXT

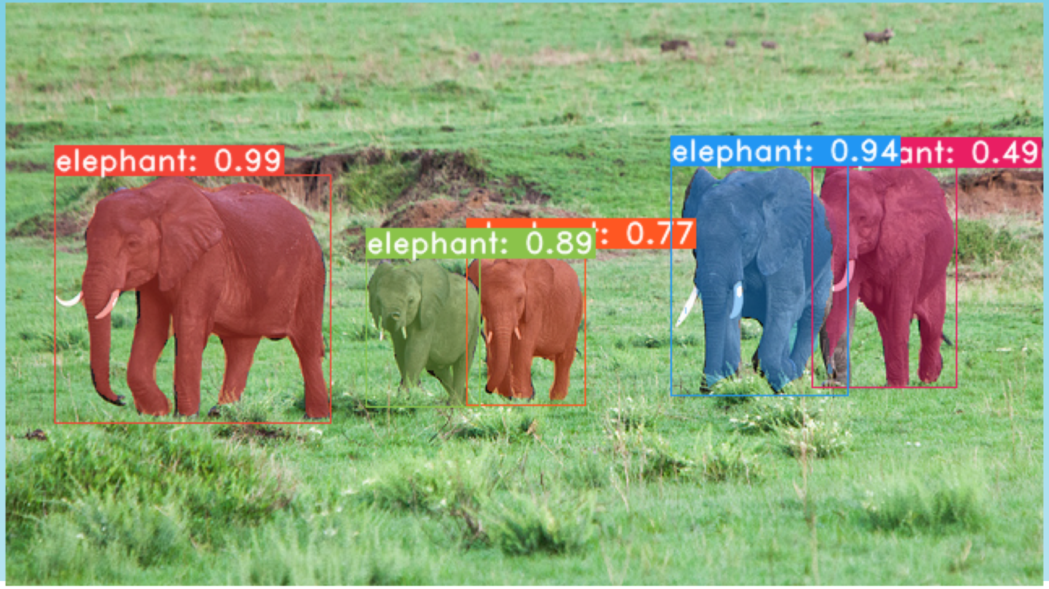
European Commission | This project has received funding from the European Union's Horizon 2020 research and innovation programme under **grant agreement No 780355**

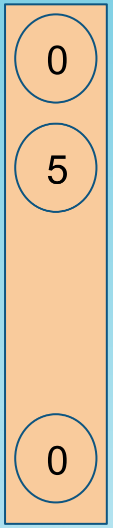
Disinformation scoring features

Image & Video analysis

Spatiotemporal Analytics and out of context: Object Detection

- Use pre-trained model to detect objects
  - Create a vector with all detected objects (probability greater than 30%)
  - This vector describes the content of the image





European Commission | This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 780355

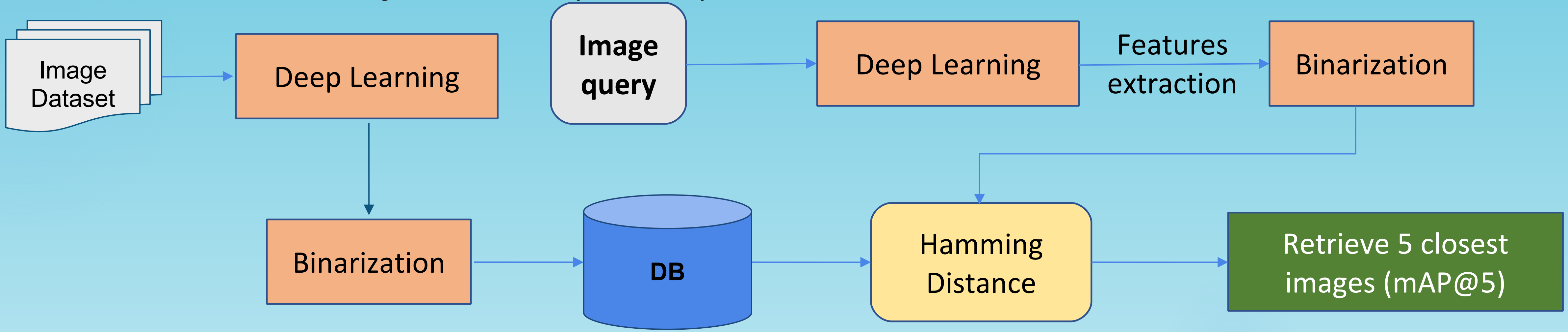
FANDANGO

Disinformation scoring features

Image & Video analysis

Spatiotemporal Analytics and out of context: Image Similarity

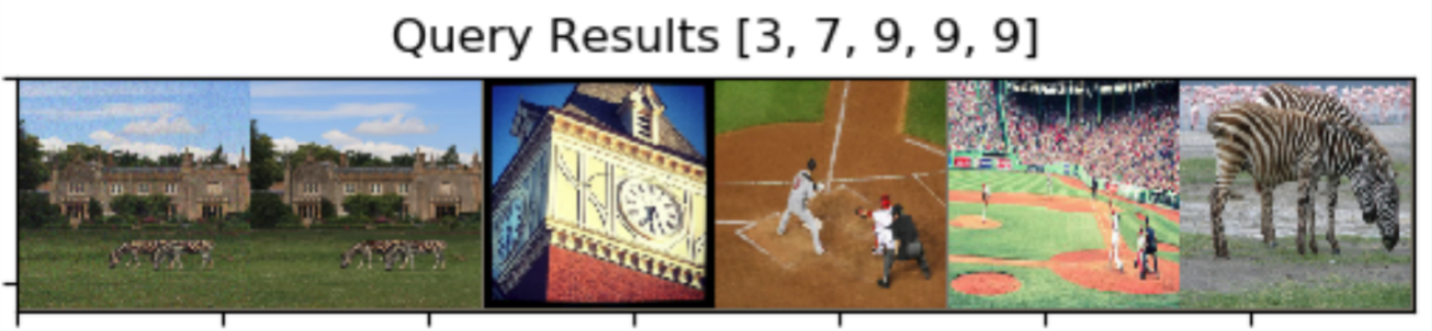
Test Datasets = synthetic, ukbench, copydays, mscoco
Queries =150
Image database = 10k
Mean average precision (mAP@5)



European Commission | This project has received funding from the European Union's Horizon 2020 research and innovation programme under **grant agreement No 780355**

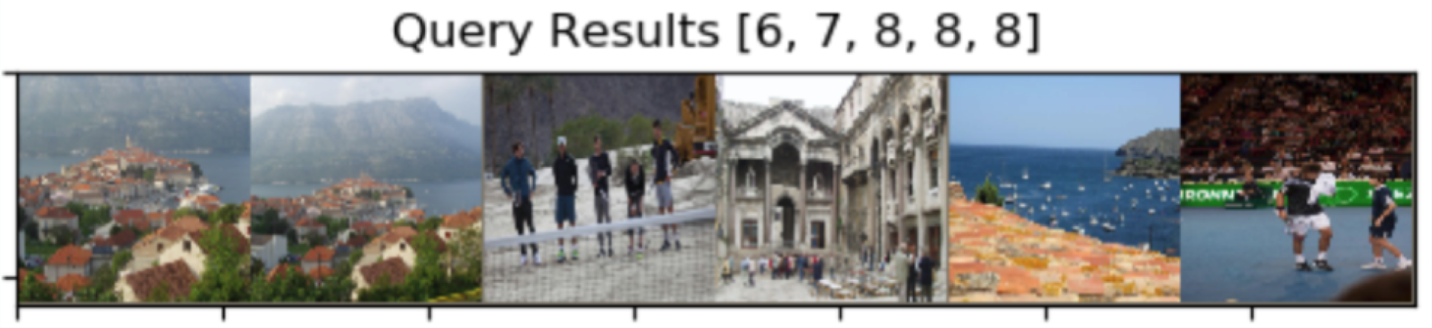
Results Visualization 1/2

- **MSCOCO gaussian noise**



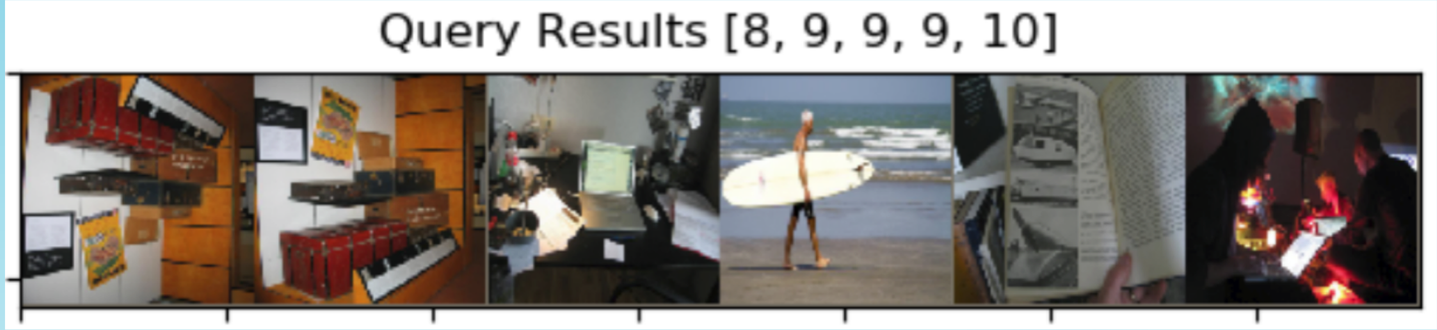
Query Results [3, 7, 9, 9, 9]

- **COPYDAYS 50% crop**



Query Results [6, 7, 8, 8, 8]

- **MSCOCO vertical flip**



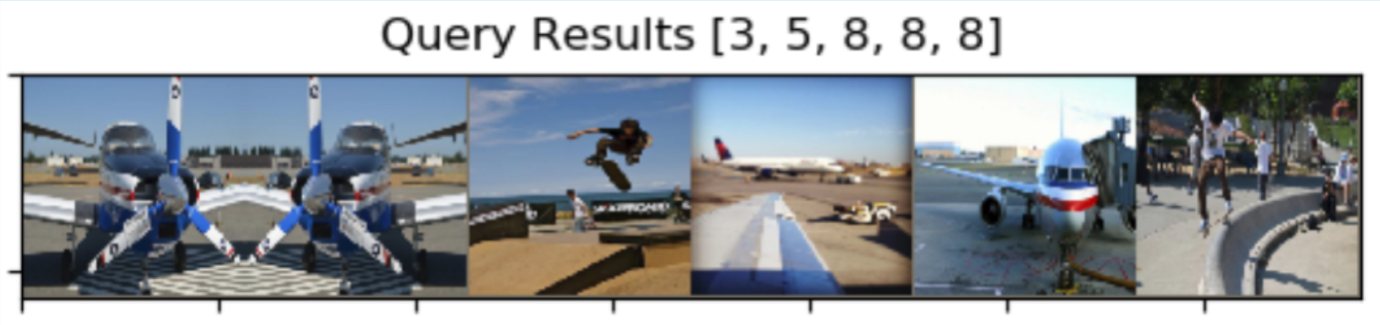
Query Results [8, 9, 9, 9, 10]

European Commission | This project has received funding from the European Union's Horizon 2020 research and innovation programme under **grant agreement No 780355**

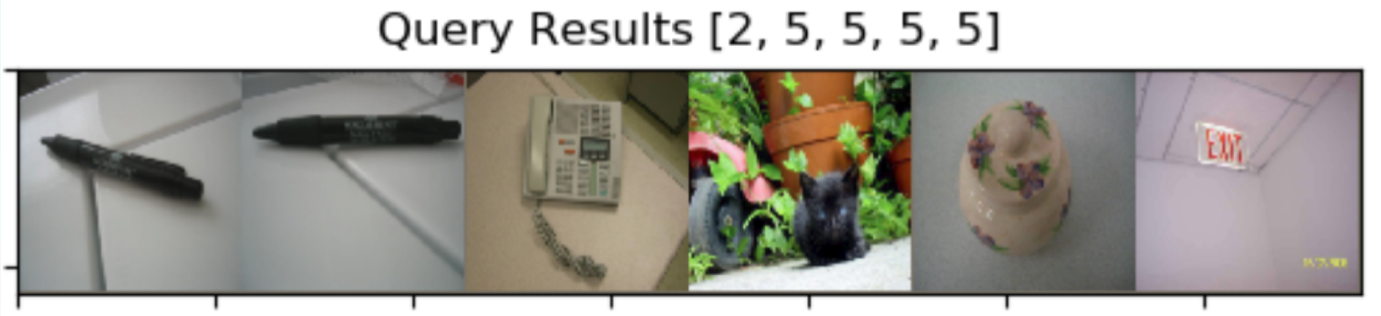
FANDANGO

Results Visualization 2/2

- **MSCOCO horizontal flip**



- **UKBENCH Dataset**



European Commission | This project has received funding from the European Union's Horizon 2020 research and innovation programme under **grant agreement No 780355**

ENGINEERING LIVETech LIVING TECHNOLOGY CERTH CENTRE FOR RESEARCH & TECHNOLOGY HELLAS SIREN DATA INTELLIGENCE vrt CIVIO POLITÉCNICA ANSA

Disinformation scoring features

Image & Video analysis

Copy-Move Detection

Goal. Build a model to classify images as fake or pristine. Provide visualizations showing the fake objects

Model. CNN Binary Classifier

Datasets. CERTH synthetic (~2K), CASIA small (~2K), CoMoFoD (200), 1st Image Forensics Challenge (~2K)

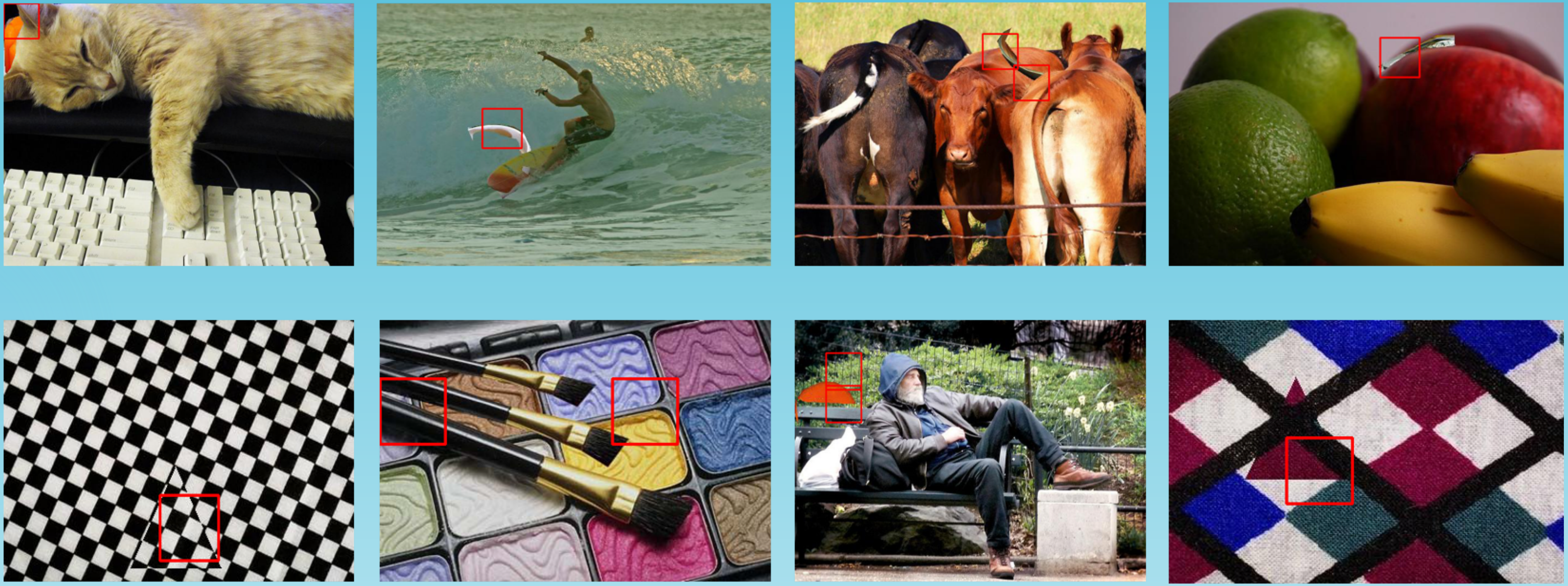
Preprocessing. Patch extraction using binary masks for random sampling on pristine, sampling on the boundaries of fake objects for fake images.

Evaluation. Perform the exact same preprocessing to test images to obtain fake and pristine patches. Evaluate images as fake even if a single patch is predicted as fake with probability > threshold. Pristine if no patch is predicted as fake.

European Commission | This project has received funding from the European Union's Horizon 2020 research and innovation programme under **grant agreement No 780355**

Copy-Move Detection:Visualizations

(Fakeness) probability threshold: 0.8. Patches that got high probability are marked with a red square



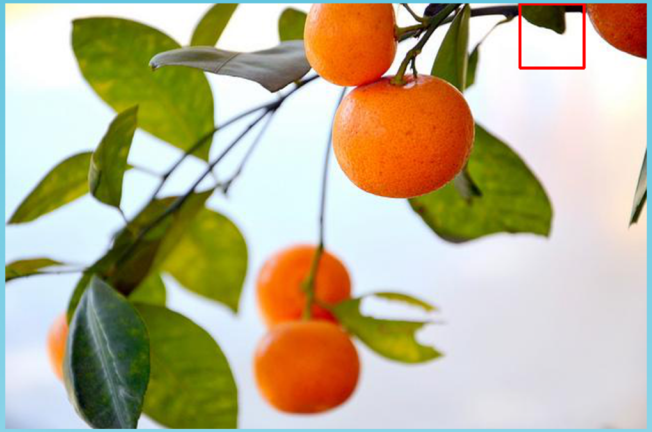
Evaluation on Fake Images

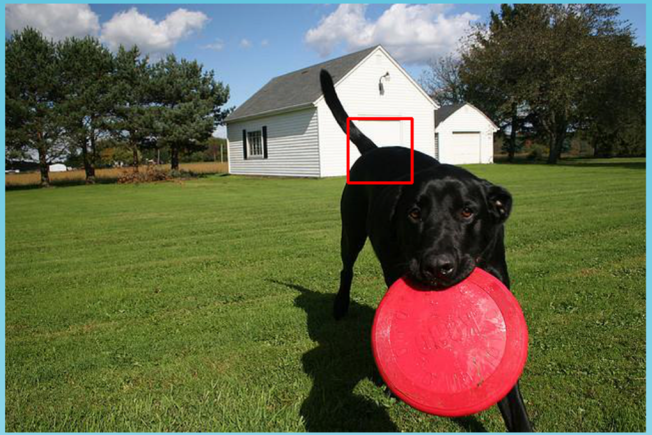
European Commission | This project has received funding from the European Union's Horizon 2020 research and innovation programme under **grant agreement No 780355**

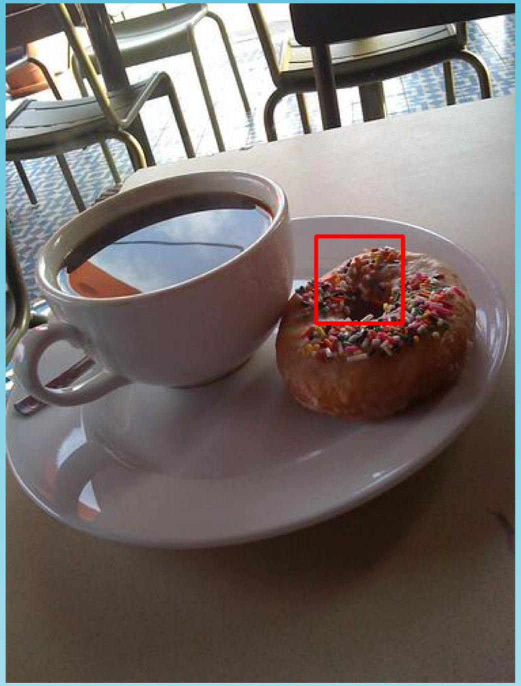
FANDANGO

Copy-Move Detection:Visualizations

(Fakeness) probability threshold: 0.8. Patches that got high probability are marked with a red square





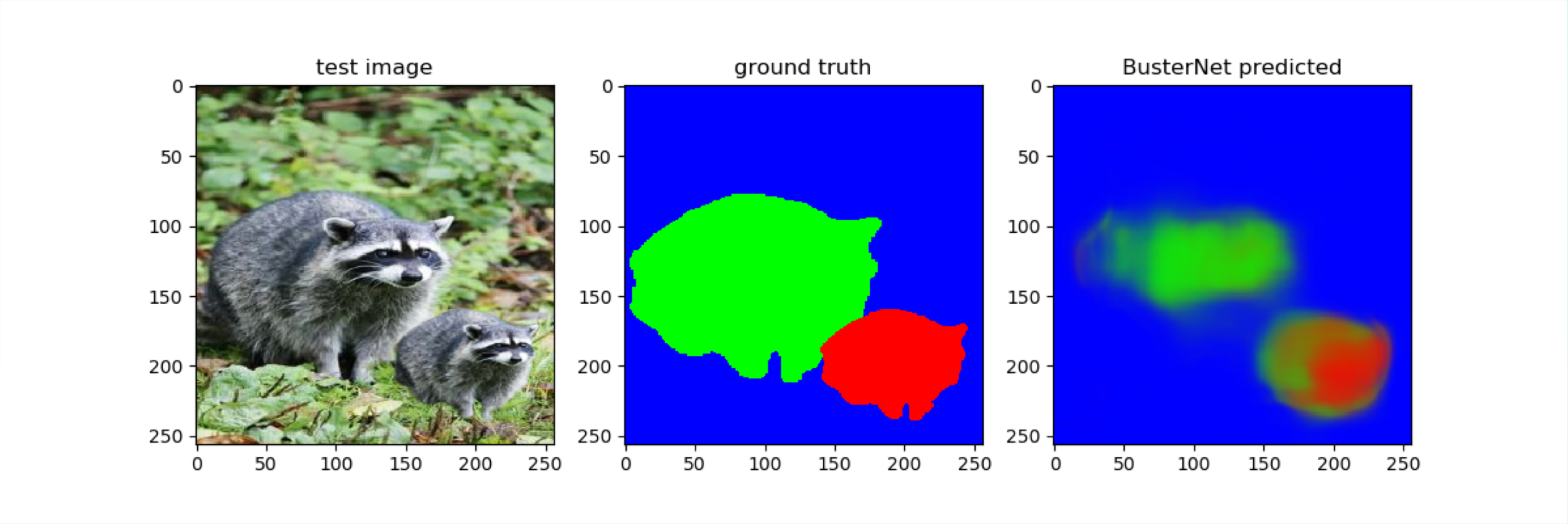


Evaluation on Pristine Images

European Commission | This project has received funding from the European Union's Horizon 2020 research and innovation programme under **grant agreement No 780355**

Copy-Move Detection: Visualizations

The final tool that will be obtained, is a multi – class classifier, able to predict three – class masks for the given fake, or pristine images



European Commission | This project has received funding from the European Union's Horizon 2020 research and innovation programme under **grant agreement No 780355**

Data Investigation features
Claim analysis: claim reviews

GOAL: support data investigation, providing similar claim reviews of a examined claim

FANDANGO will search its internal Claim database, collected from trustworthy sources that provide Claim Reviews, and display the most similar Claims and its associated Claim Reviews

It performs text comparison and tf/idf similarity analysis, which provides solid results on western languages, combined with a custom weighted pipeline to provide an overall similarity score of the Claim.

European Commission | This project has received funding from the European Union's Horizon 2020 research and innovation programme under **grant agreement No 780355**

Data Investigation features
Claim analysis: Open Data references

GOAL: support data investigation, providing direct links to Open Data

A list of links to Open Data references, with its titles, will be provided to the user based on a Claim selection or Article topic.

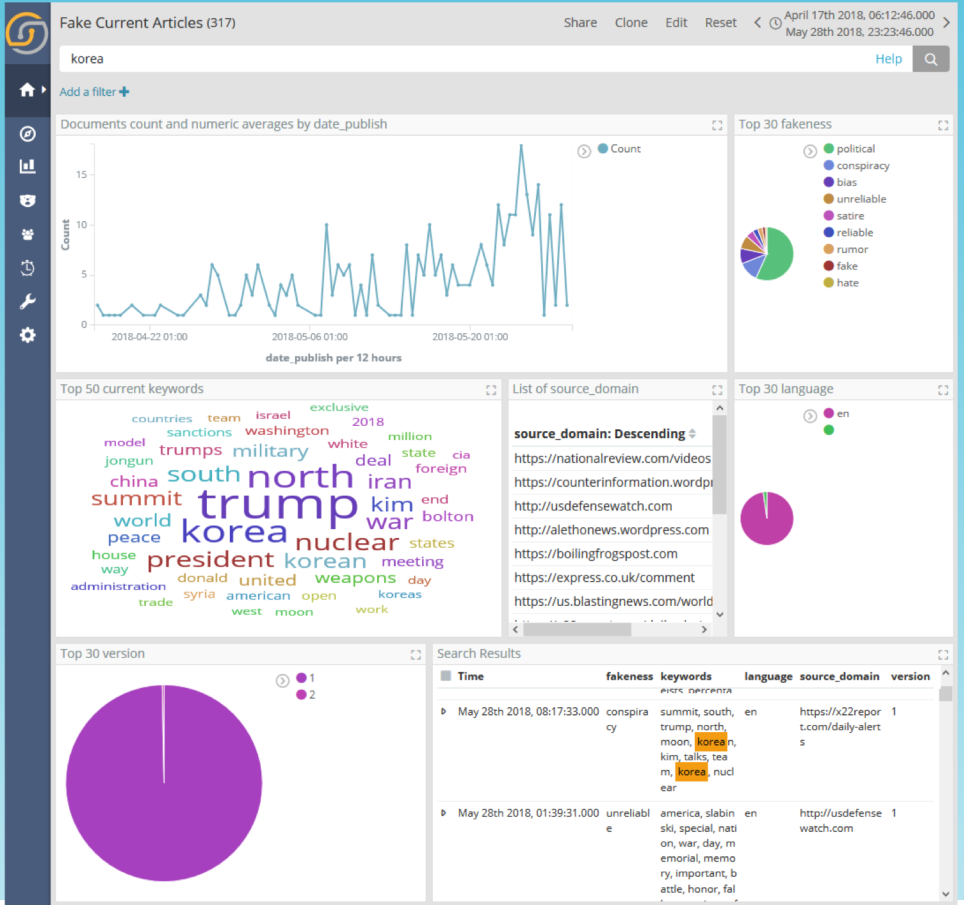
It performs a text comparison and tf/idf similarity analysis combined with a custom weighted pipeline to provide an overall similarity score of the Claim and topics.

European Commission | This project has received funding from the European Union's Horizon 2020 research and innovation programme under **grant agreement No 780355**

FANDANGO

Data Investigation features

Dashboards & Knowledge Graph


Fake Current Articles (317)

Share Clone Edit Reset

April 17th 2018, 06:12:46.000 → May 28th 2018, 23:23:46.000

korea

Help

Add a filter +

Documents count and numeric averages by date_publish

Top 30 fakeness

Top 50 current keywords

List of source_domain

source_domain: Descending

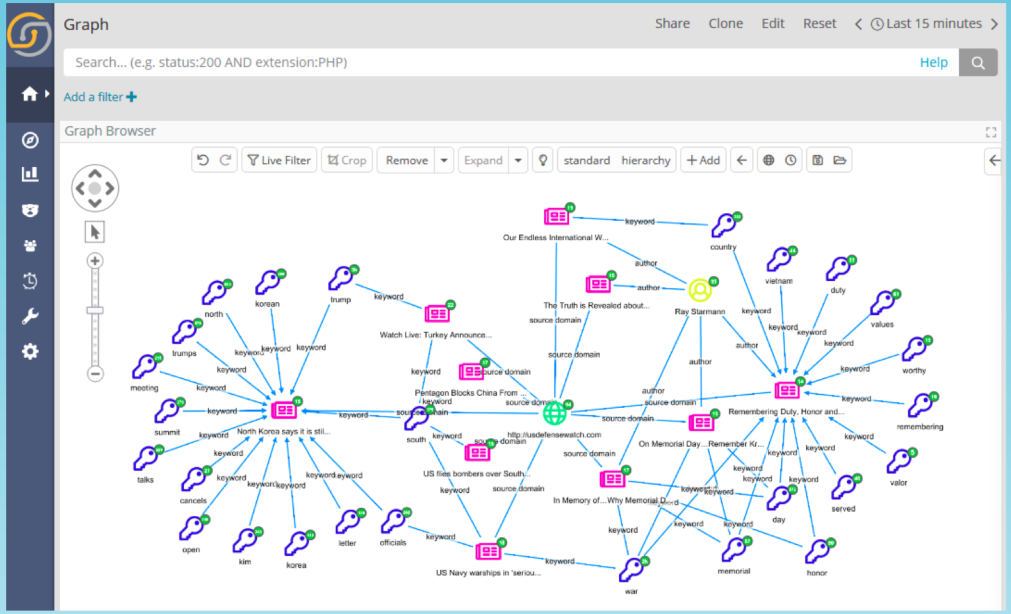
https://nationalreview.com/videos
https://counterinformation.wordpr
http://usdefensewatch.com
http://alethonews.wordpress.com
https://boilingfrogspost.com
https://express.co.uk/comment
https://us.blastingnews.com/world

Top 30 language

Top 30 version

Search Results

Time	fakeness	keywords	language	source_domain	version
May 28th 2018, 08:17:33.000	conspiracy	summit, south, trump, north, moon, korean, kim, talks, team, korea, nuclear	en	https://x22report.com/daily-alerts	1
May 28th 2018, 01:39:31.000	unreliable	america, slabinski, special, nation, war, day, memorial, memory, important, battle, honor, fal	en	http://usdefensewatch.com	1


Graph

Share Clone Edit Reset

Last 15 minutes

Search... (e.g. status:200 AND extension:PHP)

Help

Add a filter +

Graph Browser

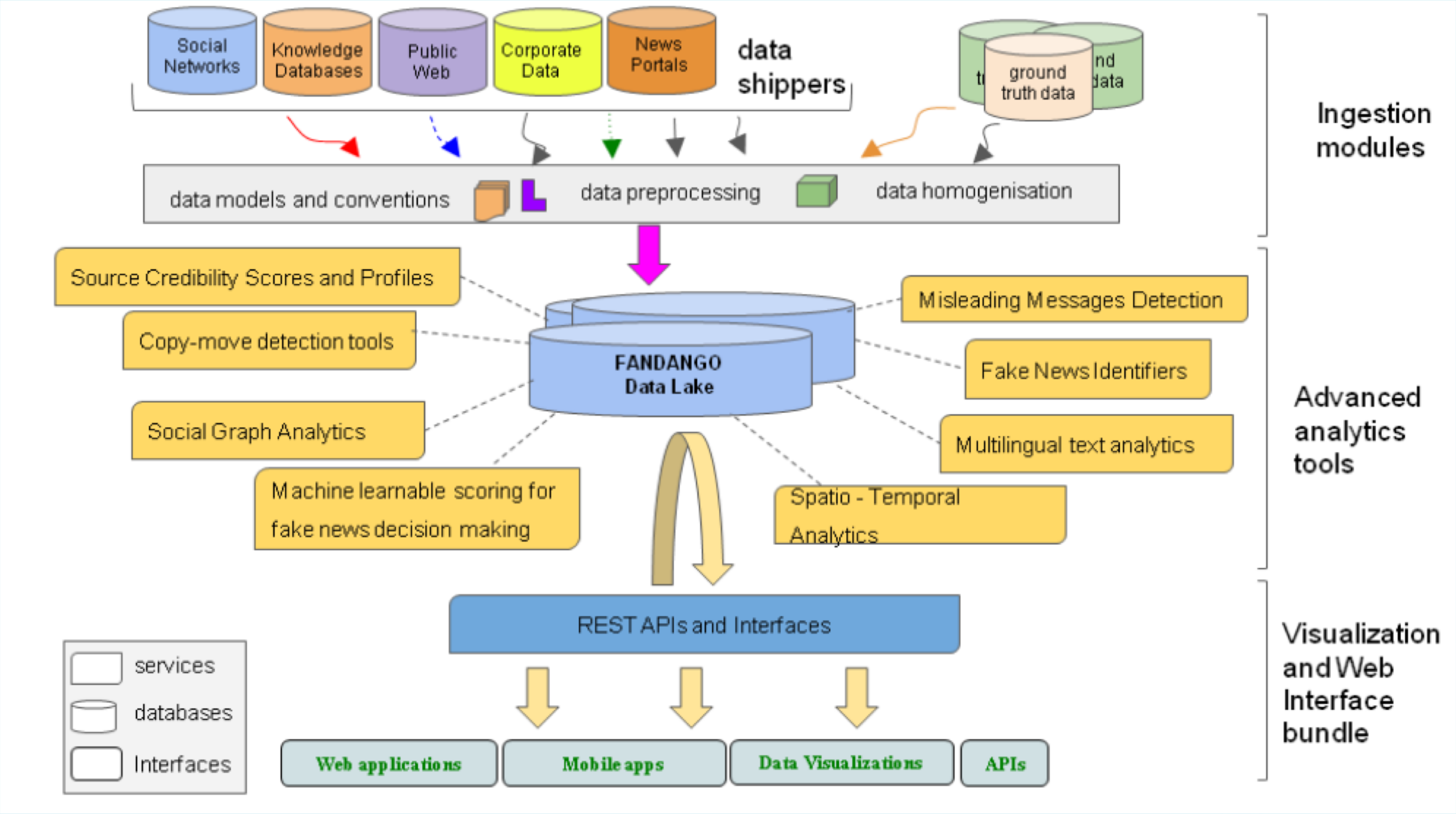
Live Filter
Crop
Remove
Expand
standard
hierarchy
+ Add
Our Endless International W...
country
The Truth is Revealed about...
Ray Starmann
vietnam
duty
values
worthy
Watch Live: Turkey Announce...
Pentagon Blocks China From ...
http://usdefensewatch.com
Remembering Duty, Honor and...
remembering
On Memorial Day....Remember Kr...
In Memory of...Why Memorial D...
US flies bombers over South...
North Korea says it is stil...
US Navy warships in 'seriou...
north
korean
trump
trumps
meeting
summit
talks
cancels
open
kim
korea
letter
officials
south
war
memorial
honor
day
served
valor
keyword
author
source domain

European Commission | This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 780355

ENGINEERING
LIVETech LIVING TECHNOLOGY
CERTH CENTRE FOR RESEARCH & TECHNOLOGY HELLAS
SIREN DATA INTELLIGENCE
vrt
CIVIO
POLITÉCNICA
ANSA

FANDANGO

Architectural approach



Social Networks

Knowledge Databases

Public Web

Corporate Data

News Portals

data shippers

ground truth data

data models and conventions

data preprocessing

data homogenisation

Ingestion modules

Source Credibility Scores and Profiles

Copy-move detection tools

Social Graph Analytics

Machine learnable scoring for fake news decision making

FANDANGO Data Lake

Misleading Messages Detection

Fake News Identifiers

Multilingual text analytics

Spatio - Temporal Analytics

Advanced analytics tools

REST APIs and Interfaces

Web applications

Mobile apps

Data Visualizations

APIs

Visualization and Web Interface bundle

services

databases

Interfaces

European Commission | This project has received funding from the European Union's Horizon 2020 research and innovation programme under **grant agreement No 780355**

ENGINEERING
LIVETech LIVING TECHNOLOGY
CERTH CENTRE FOR RESEARCH & TECHNOLOGY HELLAS
SIREN DATA INTELLIGENCE
vrt
CIVIO
POLITÉCNICA
ANSA

What we've achieved so far?

European Commission | This project has received funding from the European Union's Horizon 2020 research and innovation programme under **grant agreement No 780355**

What's next?

European Commission | This project has received funding from the European Union's Horizon 2020 research and innovation programme under **grant agreement No 780355**

Developing towards FANDANGO End Users

- **Checking Against Previous Fact Checks**: FANDANGO will match statements to previous fact-checks or consulting authoritative sources
- FANDANGO will parse statements in terms that **make sense** to a database.
- FANDANGO will help in the **verification process** itself, to find the right data, the references, assess sources/publishers, etc.
- FANDANGO is focusing on a **user friendly set of tools** to assist journalist and will leave the **final decision** about the trustworthiness of information to the journalists.

European Commission | This project has received funding from the European Union's Horizon 2020 research and innovation programme under **grant agreement No 780355**

Twitter: @fandango_eu

Facebook: @fandango.project

massimo.magaldi@eng.it

Thank you!

European Commission | This project has received funding from the European Union's Horizon 2020 research and innovation programme under **grant agreement No 780355**