CHIST-ERA Call Development Conference 2021 - Topic 2 - OSNEM #8

Practice area

Practice adding notes to the MURAL board here

Click and drag the post it to move it

Drop the post it note here

Michaela

Double click to produce a post it note and type on it

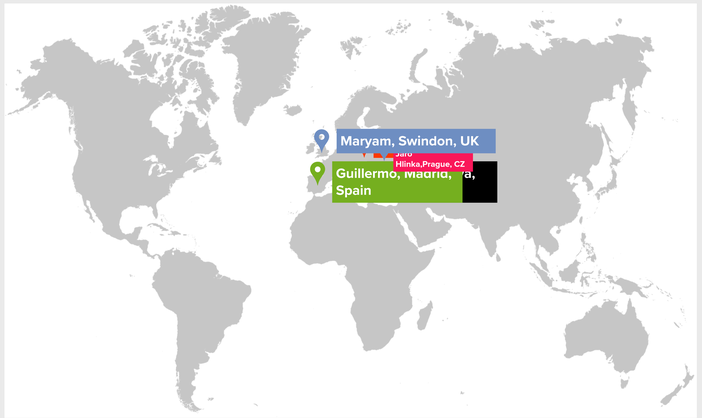
Guillermo

long live sandboxes

Double click tag to type your name and affiliation, drag to your location on the map

Tag your name and affiliation

Name, Location
Name, Location
Name, Location
Name, Location
Name, Location
Name, Location
Name, Location
Name, Location
Name, Location



Zoom in and add post it notes

Free Parking

This space is for any ideas you may have that there is no place for, or any feedback you would like to give the organisers

Acceptable alternative ways to moderate online social networks and media

Because the limitations of application of centralized content labeling and management threaten the applicability of developed algorithms

Session I

Double click where you would like to write and strta typing

Brainstorming

Well Sorted Outputs

In the mid-term, new synthetic media generation approaches may be able to create synthetic media streams that consistently represent whole alternative realities. These pose new risks and call for new detection approaches.

The simple textual format of twitter allow bots to disguise as real people & engage in misbehavior. With virtual conversational agents & deep fakes, video-based misbehavior will be possible on OSN&M beyond twitter. How do we detect & address such attacks?

As artificially created intelligent agents are becoming more easy to create it is becoming crucial to be able to recognise such activities as they can easily be used for malicious purposes

Deciding whether user-generated content is violating certain ToS or community norms is a highly subjective/complex task. For this, multimodal approaches/models are needed for detecting harmful online behavior without putting freedom of speech at risk.

It is of interest to detect the use of AI-based techniques for the automatic creation and manipulation of visual data (e.g., GANs for image/video generation, motion transfer, video translation, style transfer, automated inpainting)

Once autonomous agent and media synthesis algorithms reach new capabilities, we expect complete services and suites for creating and managing online armies for different goals (propaganda). Detecting such activities will be crucial for online safety.

It is of interest to detect re-sharing of visual media content that has not been produced/acquired natively by the user but has been already shared through the same or other web services.

The issue with fake news is not much their existence, but their (excessive) circulation. How can we leverage an interdisciplinary approach to study these dynamics, including mathematicians, computer scientists, social scientists, engineers?

Plainly false content is clearly a bad thing, but at least we can try to identify it. Instead, there is a broad grey area of low-quality content that are hard to classify. Hence the need to understand and mitigate the circulation of these "falsish" news.

Synthetic media generation techniques are getting so good that it will soon be difficult to differentiate it from authentic content. Methods and workflows are urgently needed for detecting inauthentic media content

As online content has already exploded intelligent methods are becoming essential for the automatic identification of harmful content, ensuring that this is done reliably so that freedom of speech is not harmed

Misbehaviors are made possible and sometimes favored by the way the information is shared and amplified in sought on the platforms. These dynamics can only change by chaging the platform design, which is consequence of economic incentives.

Overarching Themes

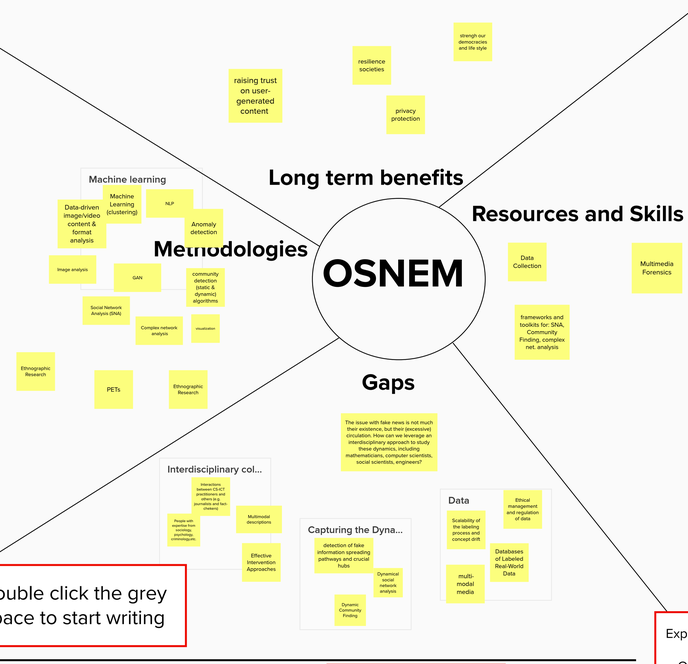
Double click where you would like to write and start typing

Overarching Theme

From the points made in Part 1, list the overarching themes that can be identified

1. Interdisciplinary synergy
2. Capturing the dynamics
3. Data
4. Multimodality
5.

Double click the grey space to start writing



Explore the Stregnths, Weaknesses, Opportunities and Threats for the Overarching Themes from the previous session

Session II

Double click the grey space to start writing

Ranking Overarching Themes

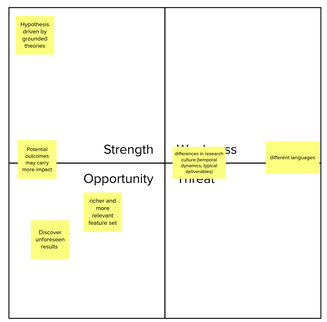
From the points made in Session I, rank and add a description to the Theme

1. Interdisciplinary synergy between computer science, political & social science,
2. Capturing the dynamics - of social network content and the network per se (quantitative and qualitative changes of information/strategies)
3. Need for Data: design of solutions requires the availability of properly generated content datasets
4. Multimodality of the analysis when it comes to different sources of information (e.g., images, video, or text) but also when it cames to the different interpretations of the sources (e.g., what is hate speech or is a meme considered hateful).

SWOT Analysis

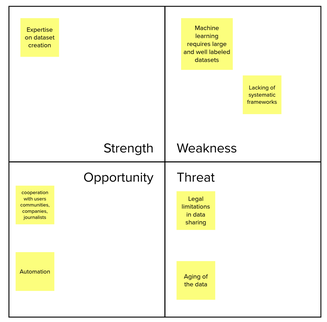
Zoom in and add post it notes

Overarching Theme #1

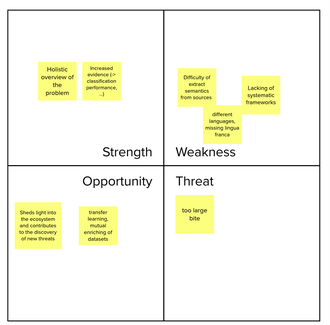


Overarching Theme

Overarching Theme #3



Overarching Theme #4



Session III

Zoom in and double click to produce a post it note

Maximising Success.....

How can Weaknesses and Threats be addressed

Refer back to the SWOT analysis and discuss how the weaknesses and threats could be mitigated

What is deliverable in a 2-3 year project

What could a CHIST-ERA Project potentially deliver?

Transnational difficulties

What challenges could be faced by a BCI CHIST-ERA project due to its transnational nature