

ALLIES technologies & their benefits for HSPs

George Kalpakis | Senior Research Associate

CERTH-ITI, Multimodal Data Fusion & Analytics Group

PROJECT CONSORTIUM

12 Partners | 6 Countries (BG, EL, IT, AT, ES, CY) | Coordinator: LAW and INTERNET FOUNDATION



LAW AND INTERNET
FOUNDATION
RESEARCH CENTER FOR LAW AND
TECHNOLOGY IN THE DIGITAL AGE



Information
Technologies
Institute



novelcore
RESEARCH & INNOVATION



TRANSCRIME
Joint Research Centre on Innovation and Crime



UNIVERSITÀ
CATTOLICA
del SACRO CUORE



BEST



SYNYO



mossos d'esquadra



Center for Social
Innovation



catalink
The catalyst for innovation



GRECA
Greek e-Commerce Association



webarts.



upf. Universitat
Pompeu Fabra
Barcelona



Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or European Commission. Neither the European Union nor the granting authority can be held responsible for them.

FACTS

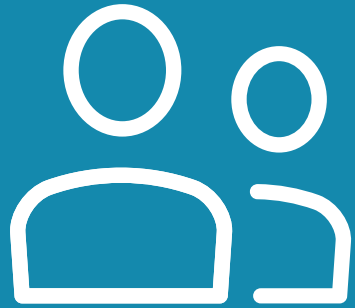
Duration 02/2023 to 02/2025

Programme ISF-2021-AG-TCO

Reference 101080090



ALLIES target group



Target Group

- **HSPs** (**micro** and **small** service providers, networks)
- **Professionals** (**LEAs**, policy makers, lawyers, prosecutors,)
- **Multiplicators** (networks, civil society organizations, lobbying groups, related projects)
- **General Public**

TCO regulation: Micro and small HSPs



- **TCO Regulation (Regulation (EU) 2021/7845)**
 - HSPs should **remove TCO within one hour**, after receiving an official removal order by a competent authority
 - Enforced to all HSPs regardless of their size and/or resources
- Limitations for **micro & small HSPs**
 - Limited **human resources**, i.e., <10/50 employees
 - Limited **operational/financial capacity**

A perspective view of a server room aisle with rows of server racks on both sides. The scene is overlaid with a complex digital network of glowing blue and orange lines, squares, and circles, creating a high-tech, data-driven atmosphere. The lighting is primarily blue, with some orange highlights from the digital elements.

ALLIES Technologies

Text Analysis

Pompeu Fabra University

- Processing of **multilingual textual content** hosted online by HSPs
- Automatic **Named Entity & Key Concept** recognition
- **Machine translation & transliteration**



ORGANISATION LOCATION DATE PERSON WEAPON

The **ISIS**^{ORG} has claimed responsibility for a suicide bomb blast in the **Tunisian**^{LOC} capital **earlier this week**^{DATE}, the **militant group**^{ORG}'s **Amaq news agency**^{ORG} said on **Thursday**^{DATE}. A **militant**^{PER} wearing an **explosives belt**^{WEAPON} blew himself up in **Tunis**^{LOC}



Speech Recognition

Pompeu Fabra University



- Multilingual **automatic speech recognition** of the **spoken content** hosted online by HSPs
- Supporting languages of interest for the ALLIES stakeholders

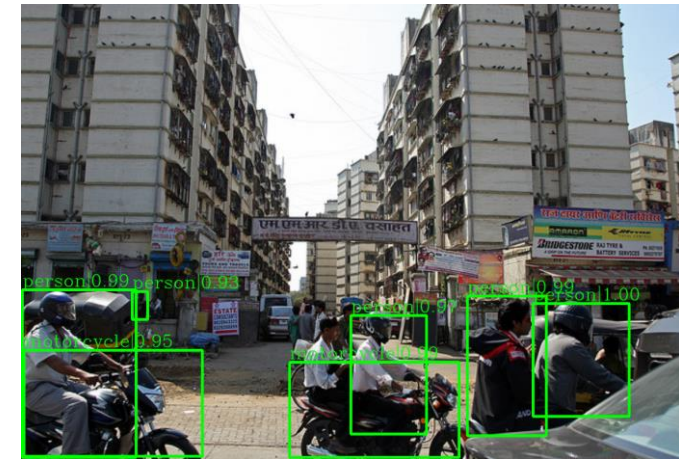


Visual Understanding

CENTRE FOR RESEARCH & TECHNOLOGY, HELLAS



- Analysing **multimedia content** hosted online by HSPs
- **Object recognition** on images/videos
 - Detection of weapons, logos, flags and/or other TCO-related classes
- **Activity/human behaviour recognition** on videos
 - Violent behaviour, bullying, etc.

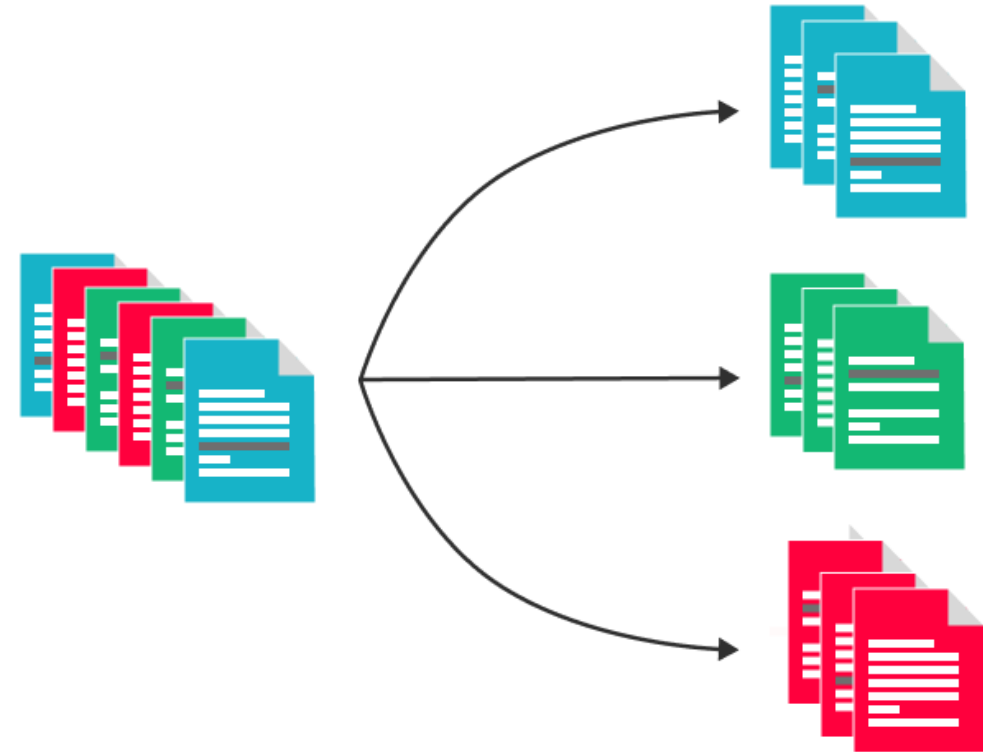


Multimodal Classification

CENTRE FOR RESEARCH & TECHNOLOGY, HELLAS



- Categorising **multimodal content** hosted online by HSPs according to its relevance to TCO
- **Modalities:** image, text
- **Training AI models** based on deep neural networks
- **Annotated datasets** provided by the ALLIES end users



Risk Assessment

Universita Cattolica del Sacro Cuore



- **Identifying** and **assessing** the **level of risk** of content hosted online by HSPs in terms of TCO
- **Explaining** the relevant outcomes

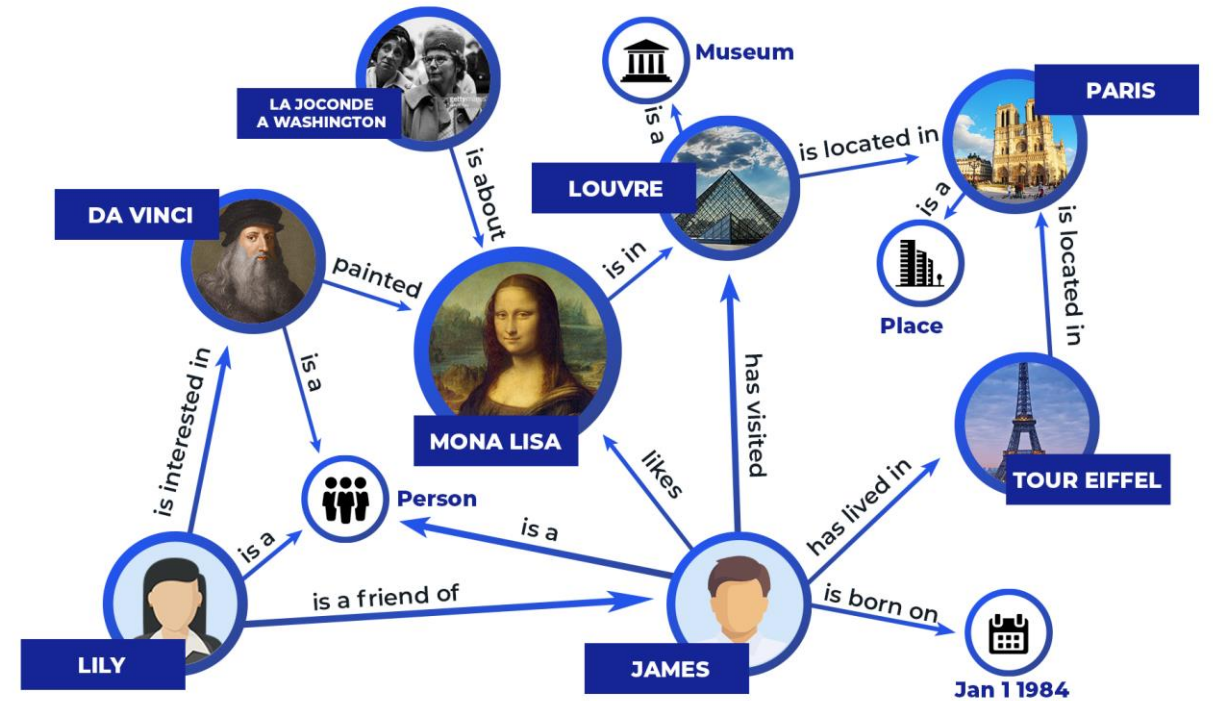


Knowledge Graph

novelcore



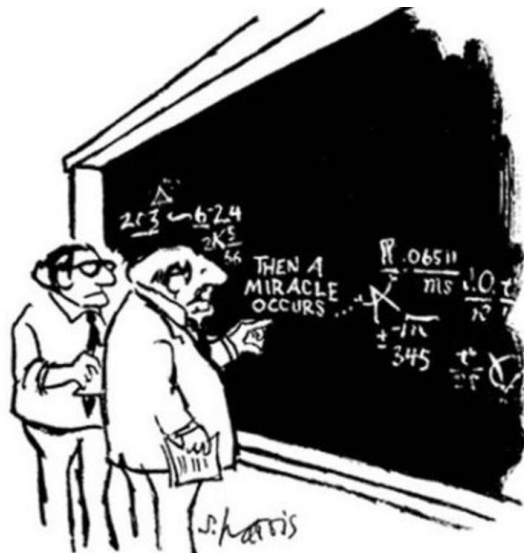
- **Decision support** based on **Knowledge Graphs**
- Creating **semantic relations** for different resources
- **Information retrieval** by matching specific criteria



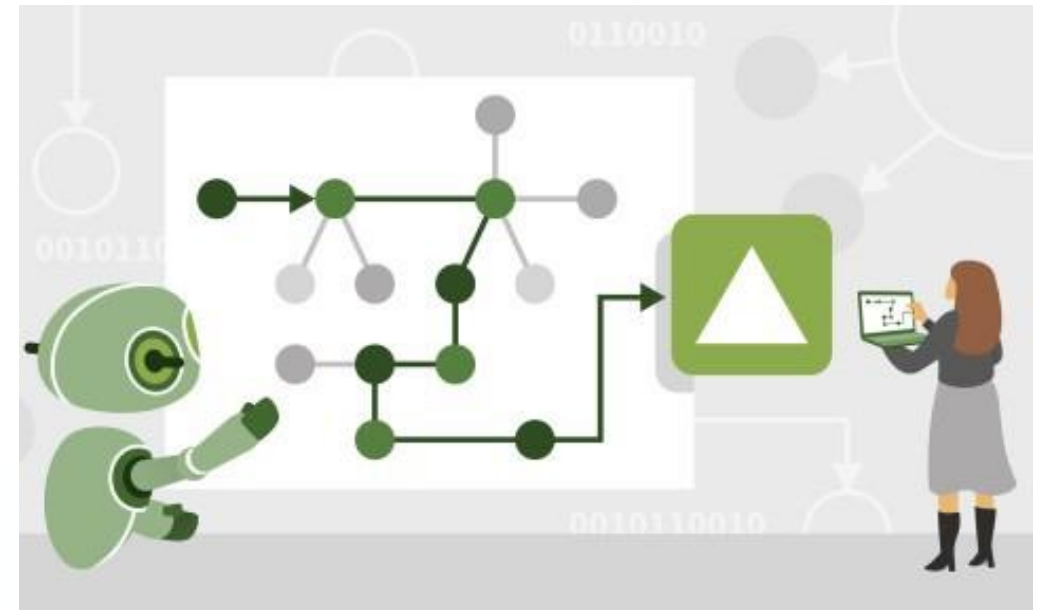
Explanation Engine

novelcore

- Diverse **explainability** and **reporting** based on AI
- Providing the **reasoning** behind the AI-based outcomes



"I think you should be more explicit here in step two."

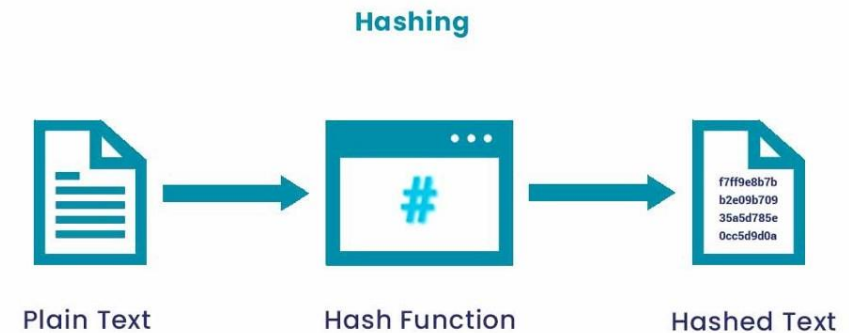


Multimodal Data Hashing

CENTRE FOR RESEARCH & TECHNOLOGY, HELLAS



- Generating **hash representations** of the multimedia content hosted online by HSPs
- **Information retrieval** of duplicate or similar content based on hash representations
- Modalities: image, video, text
 - **Unimodal** hashes (e.g., image **or** text)
 - **Multimodal** (e.g., image **and** text)



ALLIES Engagement Platform

Catalink



- **Federated UI**
 - HSP moderators access the outcomes of the analysis for data on their end
 - Outcomes accompanied by comprehensive explanations
- Platform for **removal orders**
 - LEAs submit TCO removal orders in a **standardised manner**
 - HSPs can easily respond/address removal orders
 - Efficient and effective process progress with real-time tracking of the order status
 - HSPs can proactively report TCO
- Shared **hash repository**
 - Populated with hashes of validated TCO
 - Equipped with **hash-comparison** services

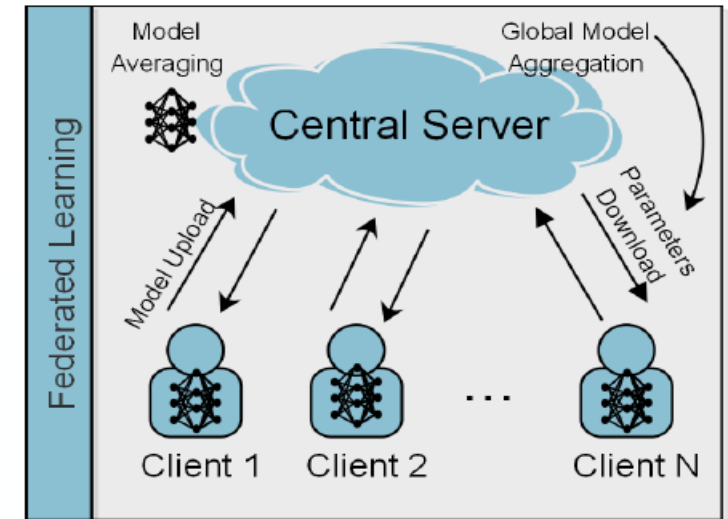


Federated Learning

Catalink



- **Decentralised** machine learning (ML) approach
 - Raw data **not exchanged** among training nodes
 - ML models **trained locally** & transferred to a central server
 - Ensures **privacy** and **security**: Data never accessed or processed by other parties, i.e., HSPs
 - **Candidate tools**: Visual understanding, Multimodal classification



A perspective view of a server room aisle with rows of server racks on both sides. The scene is overlaid with a complex digital network of glowing blue and orange lines, squares, and circles, creating a high-tech, futuristic aesthetic. The lighting is primarily blue, with some orange highlights from the digital elements.

Impact & Benefits

Benefits for micro & small HSPs (1)



- Compliance to **TCO Regulation** with limited resources
 - Response to **removal orders** within the golden hour
- **Proactive removal** of TCO
 - HSP content cross-checked against the **TCO hash-based repository**
 - **Validated TCO content** automatically removed
 - Automated content processing for the **detection** and **identification** of TCO
 - **Human-comprehensible explanations** for **content moderation**



Benefits for micro & small HSPs (2)



- Voluntary **proactive reporting** of TCO
- **Collaborative** AI model **(re-)training** without sharing raw data
- **On-premises** deployment or on the **cloud**



Questions and Discussion





Thank you!

kalpakis@iti.gr