TENSOR

Retrieval and Analysis of Heterogeneous Online Content for Terrorist Activity Recognition

Presenter: **George Kalpakis, MSc**

Multimedia Knowledge and Social Media Analytics Lab
Information Technologies Institute
Centre for Research and Technology-Hellas

Turkey in Horizon 2020 Phase II, International Study Visit to CERTH
Thessaloniki, Greece
June 2019

TENSOR has received funding as part of the "Secure societies – Protecting freedom and security of Europe and its citizens" challenge of the Horizon 2020 Research and Innovation programme of the European Union under grant agreement 700024

Overview (1)

TENSOR

reTriEval and aNalysis of heterogeneouS online content for terrOrist activity Recognition

CALL

H2020 - FCT 06 Law Enforcement capabilities 2: Detection and analysis of terrorist-related content on the Internet

START / END DATES

- September 2016 - August 2019

TENSOR has received funding as part of the "Secure societies – Protecting freedom and security of Europe and its citizens" challenge of the Horizon 2020 Research and Innovation programme of the European Union under grant agreement 700024

Overview (2)

- TENSOR brings to a Pan-European consortium of Subject Matter Experts form across Law Enforcement, Academia and Industry to provide a powerful **terrorism intelligence platform** offering LEAs fast and reliable planning and prevention functionalities for the early detection of terrorist organised activities, radicalisation and recruitment.

- The project will ensure that the solutions are shaped by the **privacy and data protection laws** that protect the freedom of citizens across Europe in their use of the internet.

- The Project will develop solutions that support LEAs by allowing developed data to be used in the **chain of evidence** for investigations.

TENSOR has received funding as part of the "Secure societies – Protecting freedom and security of Europe and its citizens" challenge of the Horizon 2020 Research and Innovation programme of the European Union under grant agreement 700024

TENSOR Consortium

TENSOR

Legal/ Ethics

Cybercrime Research Institute

Co-Ordinator/ End User

POLICE SERVICE NORTHERN IRELAND

End Users/ LEAs

Hochschule für den öffentlichen Dienst in Bayern
Fachbereich Polizei

Politie Police

KEMEA

mossos d'esquadra

Police & Crime Commissioner
West Yorkshire

Industry

rinicom
secure communications

LEONARDO

THALES

Academia/ Research

iti
Information Technologies Institute

CENTRIC
Centre of Excellence in Terrorism, Resilience, Intelligence and Organised Crime Research

EOS
EUROPEAN ORGANISATION FOR SECURITY

upf. Universitat Pompeu Fabra Barcelona

LEONARDO Cybercrime Research Institute EOS iti Information Technologies Institute CENTRIC Centre of Excellence in Terrorism, Resilience, Intelligence and Organised Crime Research
upf. Universitat Pompeu Fabra Barcelona Hochschule für den öffentlichen Dienst in Bayern Fachbereich Polizei rinicom KEMEA Politie Police Police & Crime Commissioner West Yorkshire THALES mossos d'esquadra

European Commission
HORIZON 2020

TENSOR has received funding as part of the "Secure societies – Protecting freedom and security of Europe and its citizens" challenge of the Horizon 2020 Research and Innovation programme of the European Union under grant agreement 700024

The Challenge

- Internet/Web technologies exploited by terrorists
  - communication, coordination, propaganda spreading, radicalisation, etc.

- Challenging for LEAs to identify & gather terrorist online content
  - heterogeneous sources: Surface/Deep/Dark Web, social media, forums, etc.

- LEAs need to interpret, extract & summarise relevant content to inform their resource deployment and investigations
  - huge amounts of heterogeneous multilingual & multimedia content

TENSOR has received funding as part of the "Secure societies – Protecting freedom and security of Europe and its citizens" challenge of the Horizon 2020 Research and Innovation programme of the European Union under grant agreement 700024

Mission

In the context of the challenges faced, the main objective of the TENSOR project is to provide a **powerful terrorism intelligence platform** offering LEAs **fast** and **reliable** planning and prevention functionalities for the **early detection of terrorist organised activities, radicalisation and recruitment**.

The objective can be achieved through bringing together industry, LEAs, legal experts and research institutions from across Europe.

TENSOR has received funding as part of the "Secure societies – Protecting freedom and security of Europe and its citizens" challenge of the Horizon 2020 Research and Innovation programme of the European Union under grant agreement 700024

Aims

Enhance LEAs capacity in the early detection of online terrorist organised activities, radicalisation and recruitment by developing a **platform** that will **integrate** a set of automated and semi-automated tools for:

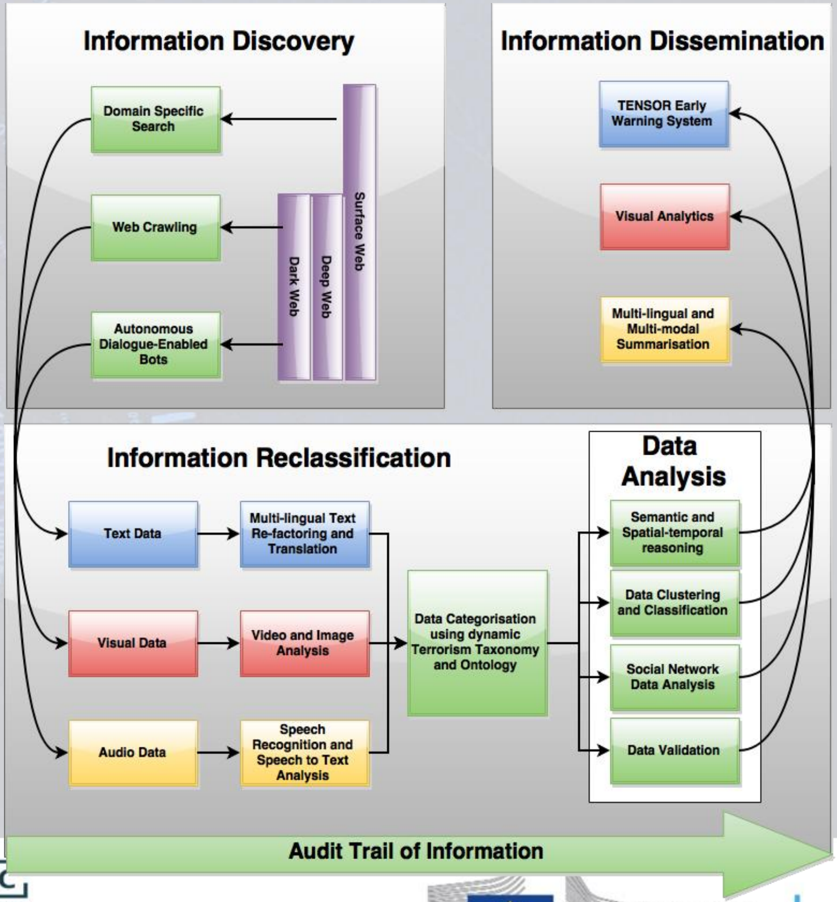
- Efficient and effective **searching**, **crawling**, **monitoring** and **gathering** online terrorist-generated content from the Surface, Deep and Dark Web;
- **Information extraction** from multimedia and multilingual content;
- Content **categorisation**, **filtering** and **analysis**;
- Real-time relevant content **summarisation** and **visualisation**;
- Creation of **automated audit trails**;
- **Privacy-by-design** and **data protection**.

TENSOR has received funding as part of the "Secure societies – Protecting freedom and security of Europe and its citizens" challenge of the Horizon 2020 Research and Innovation programme of the European Union under grant agreement 700024

TENSOR Concept

multidimensional content integration from heterogeneous online resources with a view to develop a **unified platform** to support LEAs towards:

1. efficiently and effectively categorise and analyse terrorist-generated multilingual and multimedia online content
2. detect terrorist communities and key players
3. perform temporal analysis of terrorism trends
4. identify dis-/mis-information
5. summarise and visualise terrorist information



Information Discovery

Domain Specific Search

Web Crawling

Autonomous Dialogue-Enabled Bots

Dark Web

Deep Web

Surface Web

Information Dissemination

TENSOR Early Warning System

Visual Analytics

Multi-lingual and Multi-modal Summarisation

Information Reclassification

Text Data

Visual Data

Audio Data

Multi-lingual Text Re-factoring and Translation

Video and Image Analysis

Speech Recognition and Speech to Text Analysis

Data Categorisation using dynamic Terrorism Taxonomy and Ontology

Data Analysis

Semantic and Spatial-temporal reasoning

Data Clustering and Classification

Social Network Data Analysis

Data Validation

Audit Trail of Information

European Commission

HORIZON 2020

TENSOR has received funding as part of the "Secure societies – Protecting freedom and security of Europe and its citizens" challenge of the Horizon 2020 Research and Innovation programme of the European Union under grant agreement 700024

Use Cases

TENSOR has developed Use Cases as a basis for the development of **user requirements**

The **purpose** of the use cases is to develop a series of narratives that describe the problems that LEAs face in relation to terrorist use of **Surface & Dark Web**.

Four Use Cases have been developed by the LEA partners:

1. Domestic Terrorism (e.g. Northern Ireland, the Basque Region, etc.)
2. Radicalisation (e.g. Religious fundamentalism)
3. Lone Wolf Terrorism
4. International Terrorism (e.g. ISIS or similar)

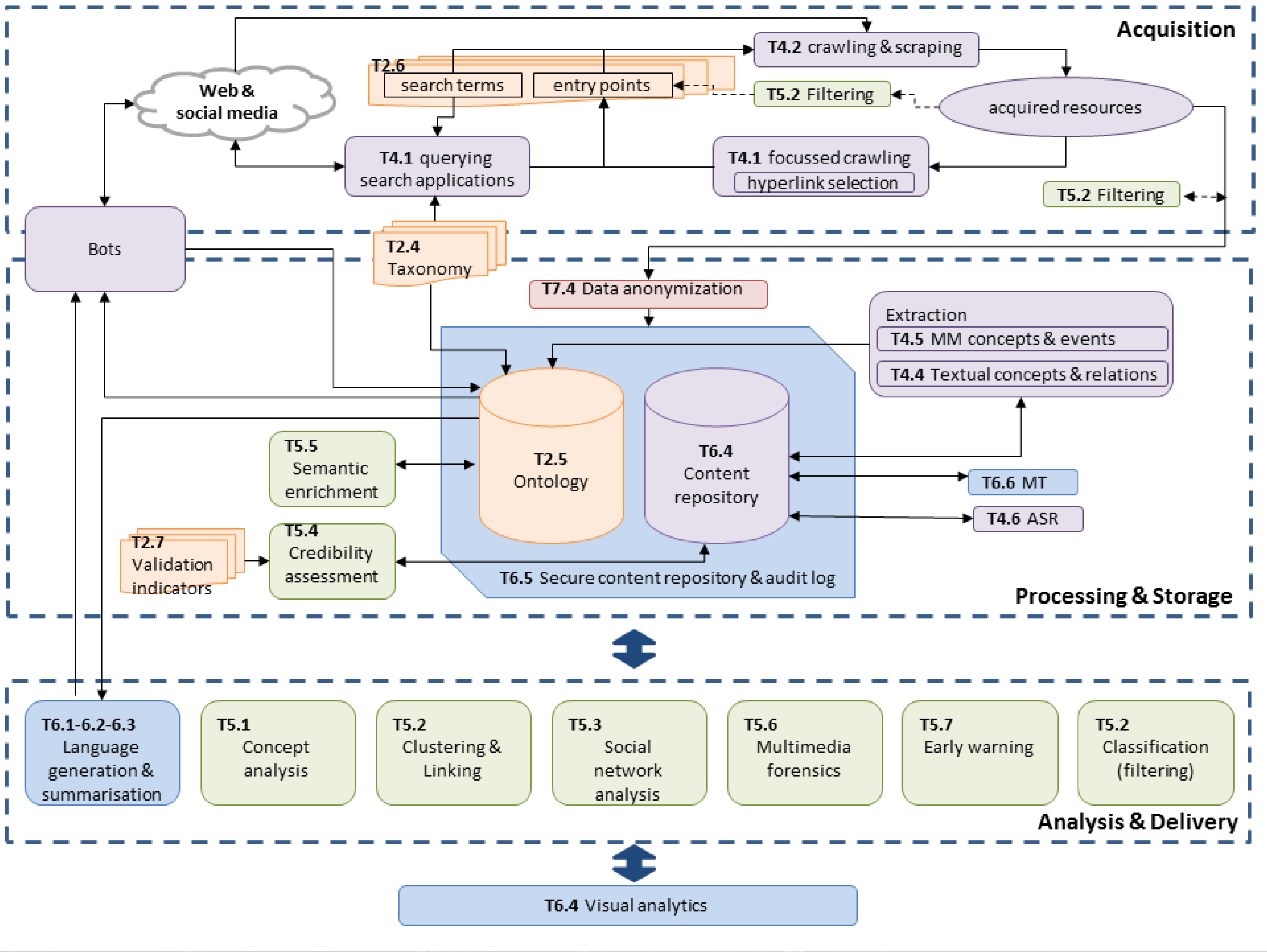
TENSOR has received funding as part of the "Secure societies – Protecting freedom and security of Europe and its citizens" challenge of the Horizon 2020 Research and Innovation programme of the European Union under grant agreement 700024

Work Packages

- Work package 1: Project Management and Coordination
- Work package 2: End-user requirements and domain modelling
- Work package 3: Legal, ethical management and data protection
- Work package 4: Terrorist-generated content acquisition, processing and indexing
- Work package 5: Multimodal content analysis
- Work package 6: Multimodal summarisation and information presentation to the user
- Work package 7: System development and integration
- Work package 8: Test cases simulation and evaluation
- Work package 9: Dissemination and exploitation

TENSOR has received funding as part of the "Secure societies – Protecting freedom and security of Europe and its citizens" challenge of the Horizon 2020 Research and Innovation programme of the European Union under grant agreement 700024

TENSOR Framework



Acquisition

Web & social media

T4.2 crawling & scraping

T2.6

search terms

entry points

T5.2 Filtering

acquired resources

T4.1 querying search applications

T4.1 focussed crawling

hyperlink selection

T5.2 Filtering

Bots

T2.4 Taxonomy

T7.4 Data anonymization

Extraction

T4.5 MM concepts & events

T4.4 Textual concepts & relations

T5.5 Semantic enrichment

T2.5 Ontology

T6.4 Content repository

T6.6 MT

T4.6 ASR

T2.7 Validation indicators

T5.4 Credibility assessment

T6.5 Secure content repository & audit log

Processing & Storage

T6.1-6.2-6.3 Language generation & summarisation

T5.1 Concept analysis

T5.2 Clustering & Linking

T5.3 Social network analysis

T5.6 Multimedia forensics

T5.7 Early warning

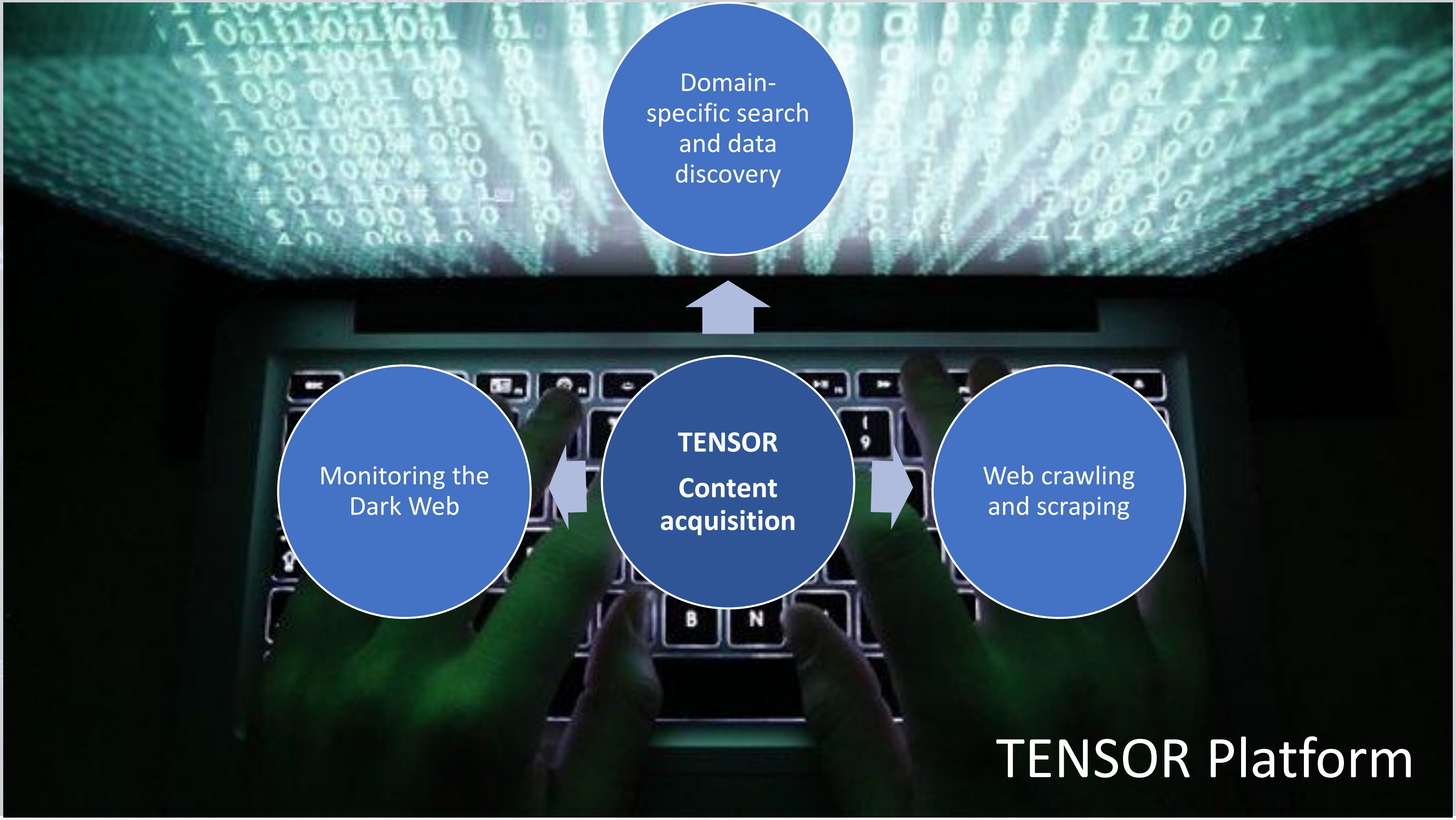
T5.2 Classification (filtering)

Analysis & Delivery

T6.4 Visual analytics

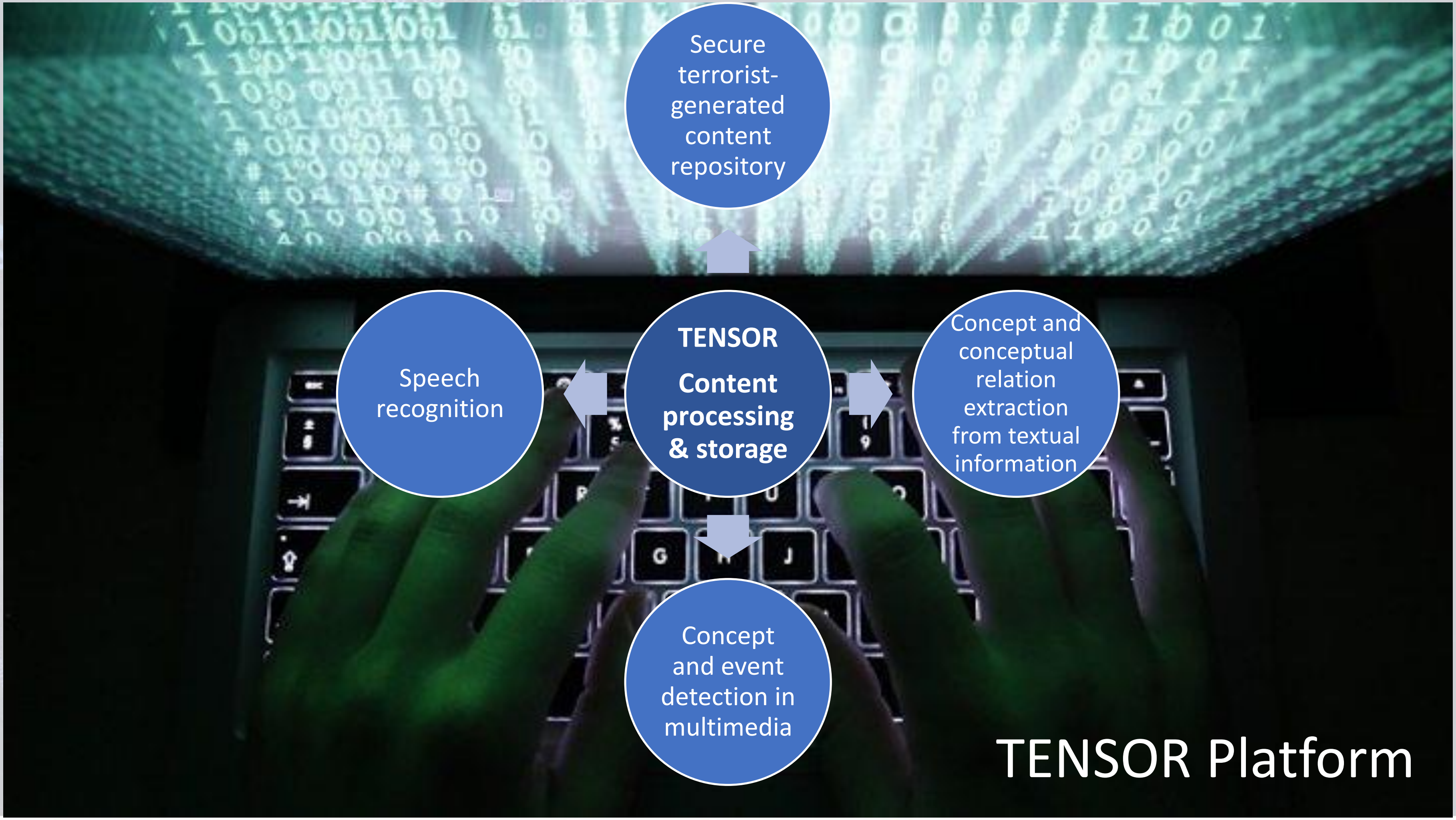
TENSOR has received funding as part of the "Secure societies – Protecting freedom and security of Europe and its citizens" challenge of the Horizon 2020 Research and Innovation programme of the European Union under grant agreement 700024

Content Acquisition


Domain-specific search and data discovery
TENSOR
Content acquisition
Monitoring the Dark Web
Web crawling and scraping
TENSOR Platform

TENSOR has received funding as part of the "Secure societies – Protecting freedom and security of Europe and its citizens" challenge of the Horizon 2020 Research and Innovation programme of the European Union under grant agreement 700024

Content Processing & Storage



TENSOR

Content processing & storage

Secure terrorist-generated content repository

Speech recognition

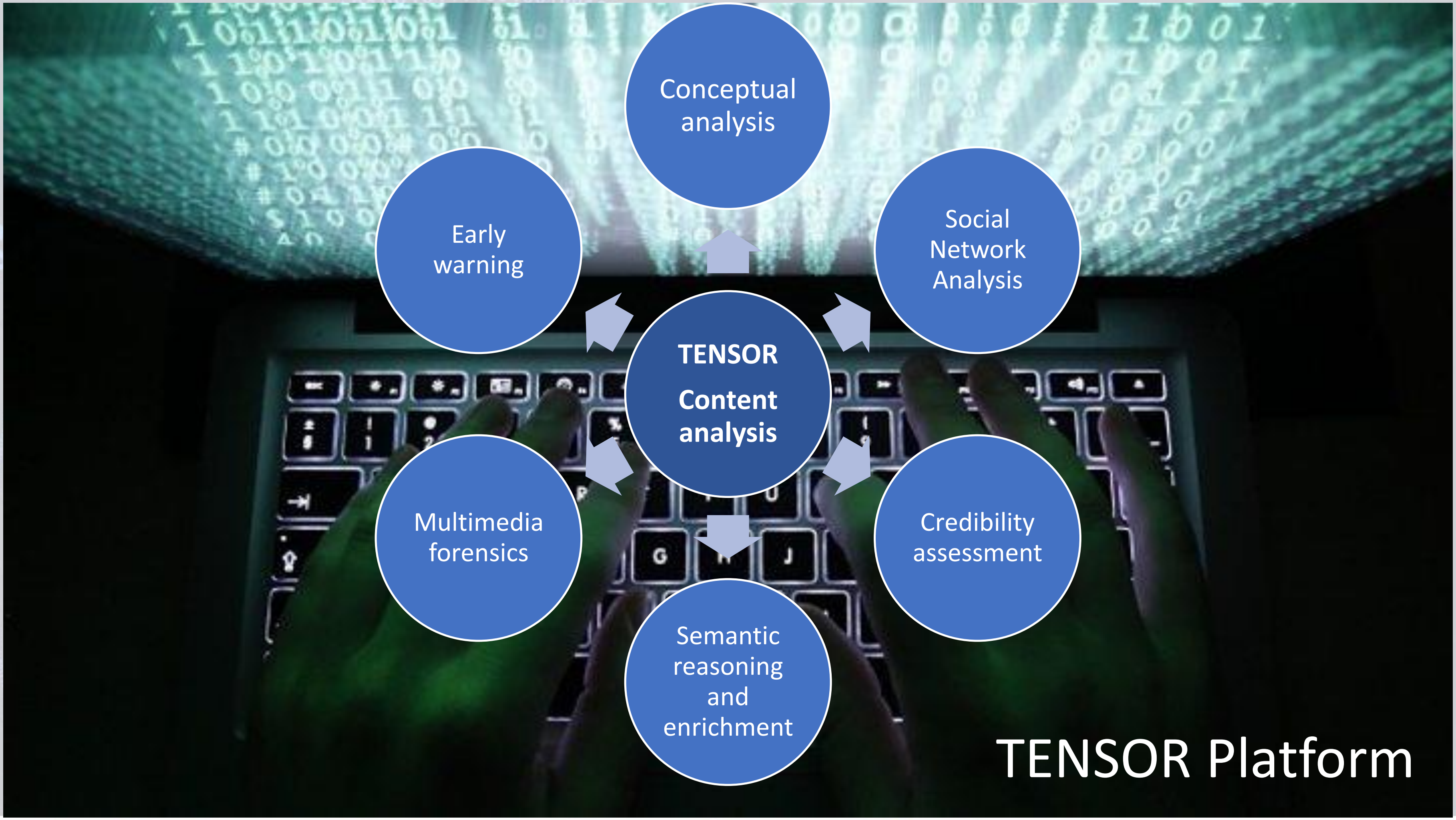
Concept and conceptual relation extraction from textual information

Concept and event detection in multimedia

TENSOR Platform

TENSOR has received funding as part of the "Secure societies – Protecting freedom and security of Europe and its citizens" challenge of the Horizon 2020 Research and Innovation programme of the European Union under grant agreement 700024

Content Analysis



TENSOR

Content analysis

Conceptual analysis

Social Network Analysis

Credibility assessment

Semantic reasoning and enrichment

Multimedia forensics

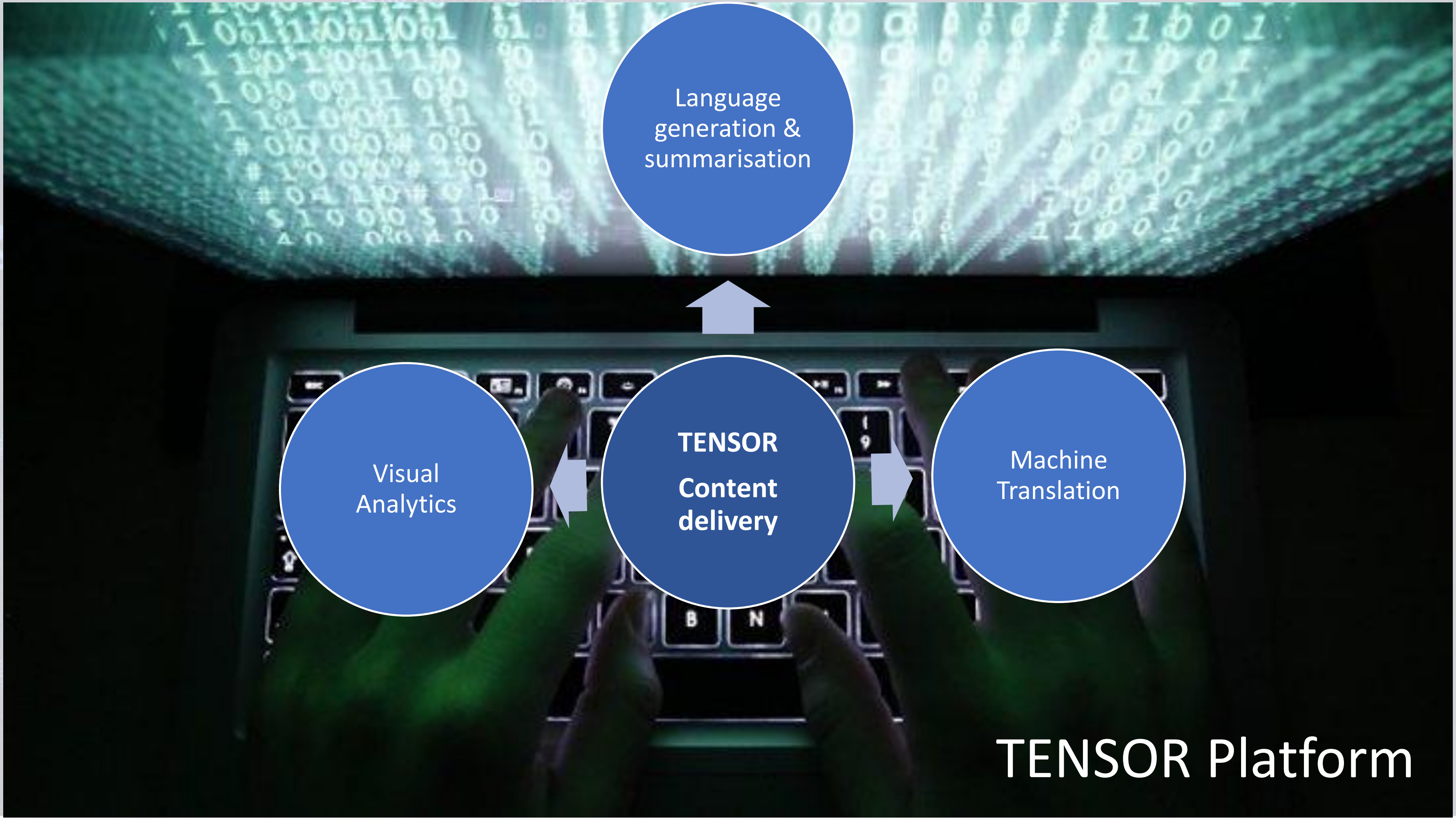
Early warning

TENSOR Platform

TENSOR has received funding as part of the "Secure societies – Protecting freedom and security of Europe and its citizens" challenge of the Horizon 2020 Research and Innovation programme of the European Union under grant agreement 700024

Content Delivery

TENSOR



LEONARDO Cybercrime Research Institute EOS Information Technologies Institute CENTRIC Centre of Excellence in Terrorism, Resilience, Intelligence and Organised Crime Research European Commission
Universitat Pompeu Fabra Barcelona Hochschule für den öffentlichen Dienst in Bayern Fachbereich Polizei rinicom KEMEA Politie Police Police & Crime Commissioner West Yorkshire THALES mossos d'esquadra HORIZON 2020

TENSOR has received funding as part of the "Secure societies – Protecting freedom and security of Europe and its citizens" challenge of the Horizon 2020 Research and Innovation programme of the European Union under grant agreement 700024

TENSOR Impact – LEAs (1)

- More effective **prevention of terrorist activities** planned/organised via the Internet through automated analysis of huge amounts of multilingual and multimedia terrorist-generated content
- Faster **detection of grassroots terrorist cells** from their online activities
- Faster and more accurate **detection and analysis of malicious content** published by terrorists
- Faster **detection and analysis of terrorism trends**
- **Reduction of the "information overload"** on Web intelligence experts due to automated summarisation of the relevant content.
- Increased **privacy** and **data protection**

TENSOR has received funding as part of the "Secure societies – Protecting freedom and security of Europe and its citizens" challenge of the Horizon 2020 Research and Innovation programme of the European Union under grant agreement 700024

TENSOR Impact – LEAs (2)

- **Scale** their effectiveness through horizontal information diffusion
- Benefit from a **greater range of operational responses** thanks to the early identification of terrorist generated content
- Employ more effective techniques for **distinguishing** non-harming religious (or other) extremist ideologies from violent radicalisation activities
- Employ more effective capabilities in **gathering data from the Dark Web**, which were previously hidden or inaccessible to them
- Establish **persistent cooperation and exchange of information** with National and European platforms, subject to national legal frameworks
- Identify patterns as well as **harmonised and uniform responses** and prevention measures, undertaken at strategic level

TENSOR has received funding as part of the "Secure societies – Protecting freedom and security of Europe and its citizens" challenge of the Horizon 2020 Research and Innovation programme of the European Union under grant agreement 700024

Contact

kalpakis@iti.gr
+30 2311 257807

TENSOR Website – www.tensor-project.eu

TENSOR has received funding as part of the "Secure societies – Protecting freedom and security of Europe and its citizens" challenge of the Horizon 2020 Research and Innovation programme of the European Union under grant agreement 700024