reTriEval and aNalysis of heterogeneouS online content for terrOrist activity Recognition

Stefanos Vrochidis, CERTH

TENSOR in H2020

- H2020-EU.3.7. - Secure societies - Protecting freedom and security of Europe and its citizens
  - SEC-FCT-06-2015 - Law Enforcement capabilities 2: Detection and analysis of terrorist-related content on the Internet
- TENSOR
  - Innovation Action
  - 16 European partners (UK, Greece, Germany, Belgium, Italy, Spain, France)
  - Coordinated by Police Service Northern Ireland (PSNI)
  - Budget ~ 5,62 M€ / EC funding ~ 5 M€
  - Starting date: September 2016
  - Duration: 3 Years
- http://tensor-project.eu/
- http://twitter.com/TensorEU

TENSOR Consortium

Police Service Northern Ireland (PSNI)

- Project Coordinator, Security Management
- Use case creation, evaluation, dissemination, exploitation (WP2, WP8)
- Ethics requirements (WP10)

Information Technologies Institute (ITI) - Centre for Research and Technology Hellas (CERTH)

- Multimodal content analytics, domain specific search and discovery, social network analysis, semantic reasoning, multimedia forensics, visual analytics (WP4, WP5, WP6)
- Web presence (WP9)

Centre of Excellence in Terrorism, Resilience, Intelligence and Organised Crime Research. Sheffield Hallam University (SHU)

- Scientific and Technical Management
- Content acquisition, processing and indexing, concept analysis, ontology creation, mis-information indicators, secure content repository with audit trail and content source traceability (WP2, WP4, WP5, WP6)

TENSOR Consortium

Cybercrime Research Institute (CRI)
- Legal, ethical, and societal aspects, EU legislation, data management and protection (WP3)

European Organisation for Security (EOS)
- Dissemination, collaboration/clustering with other projects, Web presence (WP9)
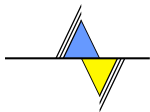
LeonardoCompany
- Innovation Management
- Exploitation coordination, market analysis, system deployment and maintenance in LEAs premises, data anonymisation (WP7, WP9)

Universitat Pompeu Fabra (UPF)
- Summarisation, text generation, concept and conceptual relation extraction from text, machine translation, dialogue-empowered bots (WP4, WP6)

European Commission

TENSOR Consortium

Fachhochschule für Öffentliche Verwaltung und Rechtspflege in Bayern Study Department Police (BayFHVR)

- Use case creation, evaluation, use training, dissemination, exploitation, content taxonomy & corpora, Web entry points & search terms (WP2, WP8)

RINICOM (RINI)

- System development and integration (WP7)
- Multimedia analysis, early warning techniques (WP4, WP5)

The Center for Security Studies (KEMEA)

- Use case creation, evaluation, dissemination, exploitation (WP2, WP8)

Thales Communications & Security (TCS)

- Content classification, clustering, and linking, credibility assessment and quality metrics, social network analysis, visual analytic tasks (WP5, WP6)

TENSOR Consortium

Police and Crime Commissioner for West Yorkshire (OPCC)
- Use case creation, evaluation, dissemination, exploitation (WP2, WP8)

Belgian Federal Police (BFP)
- Use case creation, evaluation, dissemination, exploitation (WP2, WP8)

Departament d'Interior – Generalitat de Catalunya (INT)
- Use case creation, evaluation, dissemination, exploitation (WP2, WP8)

Linguatec (LT)
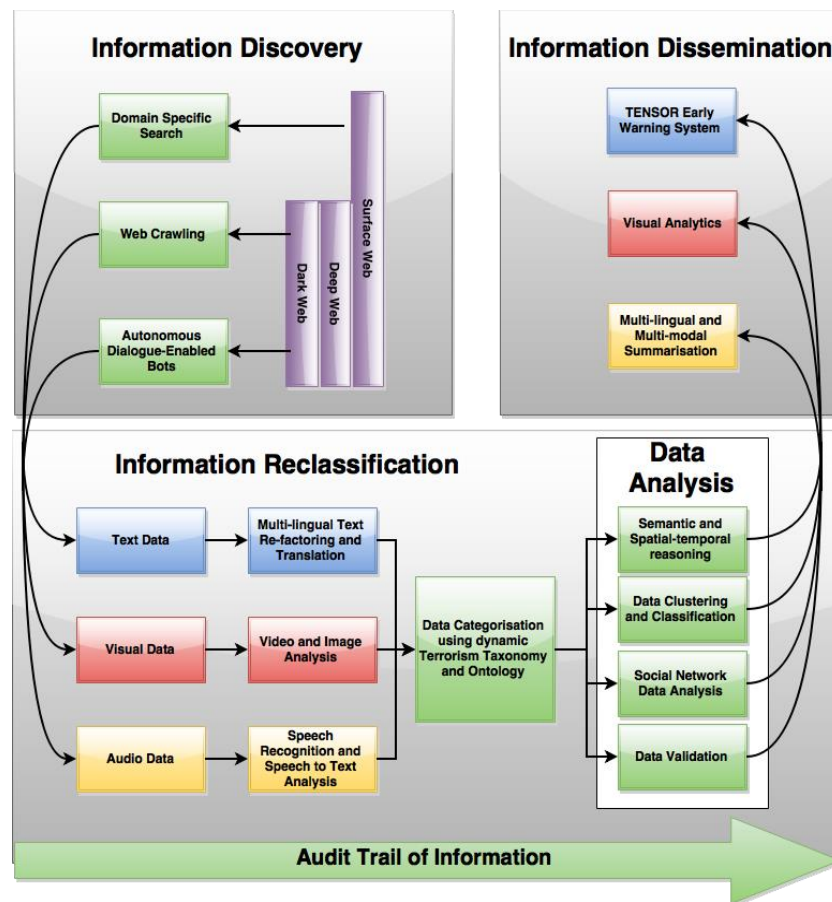- Automatic speech recognition, machine translation (WP4, WP6)

National Crime Agency
- Use case creation, evaluation, dissemination, exploitation (WP2, WP8)

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
3. perform temporal analysis of terrorism trends
4. identify mis-information
5. summarise and visualise terrorist information

Information Discovery
- Domain Specific Search
- Web Crawling
- Autonomous Dialogue-Enabled Bots
- Surface Web
- Deep Web
- Dark Web

Information Dissemination
- TENSOR Early Warning System
- Visual Analytics
- Multi-lingual and Multi-modal Summarisation

Information Reclassification
- Text Data → Multi-lingual Text Re-factoring and Translation
- Visual Data → Video and Image Analysis
- Audio Data → 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

Thank you!