

ANITA

Advanced tools for fighting oNline Illegal TrAfficking



- ✓ 17 partners
- ✓ from 11 countries
- ✓ delivering 1 integrated system
- ✓ tested and validated across 3 pilots
- ✓ involving 6 LEAs
- ✓ 2 SMEs and 2 NPOs

Research and Innovation Action

H2020 Call

SEC-12-FCT- 2016

Duration: 36 months

Part B – Sections 1-3

Participant No	Participant organisation name	Short Name	Type	Country
1 (Coordinator)	Engineering Ingegneria Informatica	ENG	IND	IT
2	Centre for Research and Technology Hellas CERTH - ETHNIKO KENTRO EREVNAS KAI TECHNOLOGIKIS ANAPTYXIS	CERTH	RTO	GR
3	Centro Ricerche e Studi su Sicurezza e Criminalità	RISSC	RTO	IT
4	Expert System S.p.A.	EXPSYS	SME	IT
5	AIT Austrian Institute of Technology GMBH	AIT	RTO	AT
6	Fundacio Institut de BioEnginyeria de Catalunya	IBEC	RTO	ES
7	Istituto Italiano per la Privacy	IIP	NPO	IT
8	SYSTRAN SA	SYSTRAN	SME	FR
9	Stichting Katholieke Universiteit Brabant	TIU-JADS	RTO	NL
10	Dutch Institute for Technology, Safety & Security	DITSS	NPO	NL
11	Belgian Road Safety Institute	ISBR	RTO	BE
<i>Law Enforcement Agencies (LEAs)</i>				
12	Provincial Police Headquarters in Gdansk	KWPG	USER	PL
13	Academy of Criminalistic and Police Studies – Kriminalisticko-Policijska Akademija	AoC	USER	RS
14	Home Office CAST	CAST	USER	UK
15	National Police of the Netherlands	NPN	USER	NL
16	General Directorate Combating Organized Crime, Ministry of Interior	GDCOC	USER	BG
17	Local Police Voorkempen	LPV	USER	BE

	<i>Project Coordinator</i>	<i>Scientific/Technical Coordinator</i>	<i>Innovation Manager</i>
<i>Name</i>	Ernesto La Mattina, ENG	Petros Daras, CERTH	Veronique Pevtschin, ENG
<i>Email</i>	ernesto.lamattina@eng.it	daras@iti.gr	veronique.pevtschin@eng.it

Table of Contents

1	EXCELLENCE.....	3
1.1	OBJECTIVES.....	3
1.1.1	<i>Rationale and problem statement</i>	3
1.1.2	<i>Objectives</i>	6
1.2	RELATION TO THE WORK PROGRAMME	9
1.3	CONCEPT AND METHODOLOGY.....	10
1.3.1	<i>Overall concept</i>	10
1.3.2	<i>Use case scenarios</i>	13
1.3.3	<i>ANITA approach and methodology</i>	15
1.3.4	<i>Positioning of the ANITA project</i>	16
1.3.5	<i>Building on pre-existing national or international research and innovation activities</i>	18
1.3.6	<i>Sex and gender analysis</i>	19
1.4	AMBITION.....	19
1.4.1	<i>Incorporation of implicit and explicit user feedback in deep learning representations for visual analysis</i>	19
1.4.2	<i>Extraction of new and evolving knowledge from deep neural network representations</i>	20
1.4.3	<i>Deep Web and Dark Nets sources and streams analysis</i>	21
1.4.4	<i>Adaptive user modelling for cognitive states estimation</i>	21
1.4.5	<i>Text Analysis</i>	22
2	IMPACT	23
2.1	EXPECTED IMPACTS.....	23
2.1.1	<i>Expected impacts set out in the work programme</i>	23
2.1.2	<i>Strategic impacts of the project</i>	25
2.1.3	<i>Social impact of ANITA</i>	26
2.1.4	<i>Economic impact of ANITA</i>	27
2.1.5	<i>Improving innovation capacity and the integration of new knowledge</i>	28
2.1.6	<i>Possible barriers</i>	29
2.2	MEASURES TO MAXIMISE IMPACT	30
2.2.1	<i>Dissemination and Communication Plan</i>	30
2.2.2	<i>Exploitation Plan</i>	32
2.2.3	<i>Individual Dissemination and Exploitation Plans</i>	34
2.2.4	<i>IPR, Knowledge protection, software licenses and privacy</i>	37
3	IMPLEMENTATION	38
3.1	WORK PLAN — WORK PACKAGES, DELIVERABLES	38
3.1.1	<i>Overall structure of the workplan</i>	38
3.1.2	<i>Detailed work package descriptions</i>	42
3.2	MANAGEMENT STRUCTURE, MILESTONES AND PROCEDURES.....	58
3.2.1	<i>Organisational structure and governance scheme</i>	58
3.2.2	<i>Innovation management</i>	60
3.2.3	<i>Key Roles and Committees</i>	60
3.2.4	<i>Management Issues and Procedures</i>	61
3.2.5	<i>Milestones</i>	63
3.2.6	<i>Risk Management</i>	64
3.3	CONSORTIUM AS A WHOLE	65
3.4	RESOURCES TO BE COMMITTED	67
3.4.1	<i>Table 3.4b: Other direct costs</i>	67
3.4.2	<i>Direct cost of subcontracting</i>	70
3.4.3	<i>Financial overview</i>	70
	REFERENCES	70

1 Excellence

1.1 Objectives

1.1.1 Rationale and problem statement

“The global market in drug trafficking has an estimated annual global value of between \$426 billion and \$652 billion (USD), making it the second most lucrative illicit market measured after that of counterfeit and pirated goods, which is estimated to generate as much as \$1.13 trillion annually.”

(Source¹: TALKINGDRUGS: “Report: Global Drug Trafficking Market Worth Half a Trillion Dollars”, by Avinash Tharoor, April 21, 2017)

“74 percent of guns recovered at crime scenes originated out-of-state and 57 percent of all recovered guns were out-of-state handguns. 87 percent of the guns recovered in New York City originated out-of-state.”

(Source²: Government Technology: “Tracing Data to Reduce Gun Trafficking”, by Devon Ziminski, August 3, 2017)

“The incident brings to light the complexity in tracking the finances of global terror networks as they seek channels outside of traditional banks and find a plethora of options in the booming financial technology space. In the past, ISIS has gotten away with funding itself through means such as low-level fraud including check scams and online lending, as well as social media, prepaid cards and crowdfunding sites.”

(Source³: INVESTOPEDIA: “FBI Says ISIS Used eBay to Send Terror Cash to US”, by Shoshanna Delventhal, August 14, 2017)

Over the recent years, **online illegal trafficking activities have hugely elaborated and expanded** so that to operate at global level with worldwide supply chains, production facilities and administrative offices, while their legal, economic and sustainability state is optimised. For tackling these emerging challenges, a significant part of LEAs’ efforts has been invested on training activities to equip officers and practitioners with the necessary knowledge and skills related to this emerging and continuously/rapidly evolving scenery. However, in order to efficiently understand the organisational structure, the exhibited behaviours/dynamics and their interconnections/interactions, it is of **vital importance to collect and analyse all relevant open-source information in near real-time and to combine it with closed-source information** provided by the LEAs. Another aspect that significantly adds to the complexity of the problem concerns the fact that often large political, religious and economic support networks support these online illegal trafficking activities, in order to receive funding for fulfilling their purposes or to gain political power. For this reason, the analysis of the underlying economic routes and networks can help LEAs to identify hidden relationships, recurrent strategies and emerging patterns in support of illegal activities.

Building on the aforementioned analysis, some crucial radical changes in the criminal/illegal online activities landscape need also to be taken into account:

- **Criminal actors are proliferating** (e.g. individuals, unstructured and informal organised criminal groups, structured and hierarchical organised crime, ethnic-based criminal groups, etc.), possessing a greater capacity to exploit a wide range of opportunities and resources, while maximising benefits and minimizing risks;
- **Organised criminal groups** are continuously exploiting the interdependencies among the criminal activities they are already involved in, but they are also specialising in specific typologies of emerging crimes or even in specific segments of the criminal supply chain (Europol IOCTA⁴ 2016, SOCTA⁵ 2017);
- **Online trafficking activities** are going “entrepreneurial”; in other words, they are organised in a managerial manner aiming to expand business activities by infiltrating into key legal and illegal new economic sectors/activities, on a global scale;
- **Surface Web, Deep Web and Dark Nets** can be regarded as key crime-facilitators. They offer new opportunities to the organised criminal groups, assisting them: a) to cooperate in a more effective, efficient, anonymous (e.g. anonymising software such as P2P or Tor) and secure way, although dislocated in diverse countries; b) to manage criminal business-to-business (B2B) and business-to-consumer (B2C) activities, with anonymous online payments, through crypto-currencies (e.g. Bitcoin); c) to spread properly manipulated information to camouflage/conceal criminal activities, while deceiving online consumers; and d) to enlarge the crime-as-a-service activities and the related black market (e.g. Silk Road) in the Deep Web.

Among the most emblematic, emerging, large-scale and continuously evolving aspects of online illegal trafficking are the cases of: a) **counterfeit/falsified medicines (OTFM), drugs, Novel Psychoactive Substances (NPS)**, b)

¹ <http://www.talkingdrugs.org/report-global-illegal-drug-trade-valued-at-around-half-a-trillion-dollars>

² <http://www.govtech.com/data/Tracing-Data-to-Reduce-Gun-Trafficking.html>

³ <http://www.investopedia.com/news/fbi-says-isis-used-ebay-send-terror-cash-us/>

⁴ <https://www.europol.europa.eu/activities-services/main-reports/internet-organised-crime-threat-assessment-iocta-2016>

⁵ <https://www.europol.europa.eu/socta/2017/>

weapons and firearms, and c) **terrorism funding**. The respective support networks are using different techniques to prevent them from being unveiled. **The first strategy is to present them as legal as possible.** As an example, novel psychoactive substances (NPS), the so called “legal highs”, are generated faster than the update rate of the blacklist for illegal drugs. Thus, trading and transporting of NPS is legal, even though, the psychoactive effect is often stronger, than existing illegal drugs. This “appear as legal” strategy is also followed by some specific charity organisations, which offer community services, but in fact they are organising worldwide fund raising activities for terrorism organisations. **The second strategy is to build up an environment within which identification and censorship are impossible, instead of caring about legal compliance.** The next generation of black markets will provide an infrastructure to set up anonymous companies, with invisible owners, anonymous management, employees and customer, like the Ethereum platform. The synergies between the trafficking activities concerning a) counterfeit/falsified medicine, drugs and NPS, b) weapons and firearms, and c) terrorism funding are considered as an emerging trend that, in the short term, could lead to new organized criminal arrangements as well as in an increasingly complex illegal market. Additionally, the aforementioned illegal trafficking activities will be further boosted by the next generation crypto currencies that will serve as a standard payment method for legal and illegal services, provided in the deep and dark web. They are expected to prove as a more reliable and less expensive alternative (EMCDDA⁶) to anonymous payment systems.

Regarding the identified emerging trends in online illegal trafficking, the following facts hold:

Trafficking of counterfeit/falsified medicines, drugs and Novel Psychoactive Substances (NPS)

The illegal online trade of counterfeit/falsified medicines is becoming a pandemic, because it has achieved a global scale and it potentially affects all types of therapeutic categories. At the same time, the increasing diffusion of NPS is exploiting the same weak spots of medicines, but knowledge on the criminal patterns and trends is still fragmented because of the difficulties in monitoring both the Internet and their arrival on the street markets. Specifically:

- **Online market makes available falsified medicines often at lower prices and without constraints** (50% of drugs sold over the Internet are estimated to be counterfeit⁷)
- **Illegal online pharmacies have a complex structure and rely on a solid networking** usually composed of thousands of affiliate websites, but very few anchor websites. On average, 30 fake online pharmacies can control tens of thousands of websites. The servers are located anywhere in the world, especially in countries with weak or no regulations at all (INTERPOL, Pangea Operation⁸, 2014). Affiliate networks are linked to other illegal Internet-based businesses, including those that promote pornography, pirated media and fake anti-virus software (Partnership for Safe Medicines – PSM⁹, 2013).
- **Several online marketplaces on Dark nets (e.g. Silk Road 3.0 and Evolution Marketplace) operate with virtual impunity from law enforcement control**, offering prescription medicines for sale to the public together with NPS, cannabis, “stimulants”, “relaxants” and “other” types of medication.

The weakness of laws and enforcement activities, the passivity of States to these phenomena, even the lack of awareness and knowledge by LEAs about NPS (usually perceived as less dangerous) constitute the most important criminal drivers. Moreover, the outstanding performance offered by these practices providing 20 times the profit of heroin trafficking justifies the growing interest of organised criminal groups for these crimes and for the opportunities to create interdependencies with other crimes, such as money laundering, human trafficking for sexual exploitation and weapons smuggling.

Trafficking of weapons and firearms

Illicit trafficking of weapons and firearms achieved a relevant growth in Europe during the last few years, especially as support to terrorist activities. The increasing threat that these trafficking activities constitute is currently included into the nine priority crime areas under the surveillance of Europol (EMPACT – European multidisciplinary platform against criminal threats)¹⁰. In particular:

- **Organised criminal groups exploit weapon trafficking as supplementary source of income**, together with other illicit activities, and reuse existing criminal routes for that purpose. For this reason, the monitoring of firearms trafficking activities can help to get in contact with broader criminal routes and networks.
- **Weapons and firearms are durable goods**, thus they can be circulated continuously for years. Moreover, legal firearms can be stolen or diverted during their transportation, modified or deactivated and then converted to illegal firearms; for this reason, it is more difficult to distinguish between legal and illegal firearms trafficking.

⁶ http://www.emcdda.europa.eu/system/files/publications/2155/TDXD16001ENN_FINAL.pdf

⁷ Fiona Clark, Rise in online pharmacies sees counterfeit drugs go global, The Lancet, Volume 386, Issue 10001, Pages 1327-1328,

⁸ <https://www.interpol.int/en/Crime-areas/Pharmaceutical-crime/Operations/Operation-Pangea>

⁹ <http://www.safemedicines.org/interchange-2013>

¹⁰ EU Policy Cycle – EMPACT <https://www.europol.europa.eu/crime-areas-and-trends/eu-policy-cycle-empact>

- **Firearms are typically used by all organised crime groups to facilitate illegal drugs trades and trafficking in human beings** to intimidate and coerce their victims.

The risk derived from the superficial trades of weapons has encouraged organised crime to move their commerce in deep web and dark net marketplaces, which guarantee that the exchange of information, products and payments is done anonymously. Additionally, deep and dark marketplaces allow sellers to provide buyers with a variety of weapons larger than that provided with physical trades.

Trafficking related to terrorism funding

Funding terrorist organizations from various sources is the primary mean for the sustainability of those organizations and their illegal activities. Nowadays, terrorism organizations operate in their heydays mainly because of the plethora of existing ways for raising funds, thereby supporting unceasingly their illegal activities. The illegal money trafficking to terrorist organizations can be realized through several ways, the most common ones being:

- **Terrorist-related financial networks and flows in Dark Nets are mainly used from terrorists to exchange money** either through the same organizations for their daily activities or with other terrorist organizations. It is frequent for terrorist organizations to fund each other fitfully according to their needs, expecting though some kind of assistance in another aspect of their activities. As means for money transactions are utilized cryptocurrencies and block-chains, since they provide anonymity and hence are hard to be inspected by LEAs.
- **Terrorist organizations establish fake charities as an efficient way to raising funds.** Cyber terrorists run several online charities for some world-concern humanitarian issue –e.g. provide medical assistance to war victims in Syria- that is expected to allure potential donors more easily, thus collecting a significant amount of money¹¹. For the sake of credibility, terrorists set professional websites about the charity activities, objectives and future plans, including completely falsified information. The donors provide money via legal means such as bank wire, PayPal and credit card to the well-veiled terrorist organization accounts.
- **Fund-raising campaigns over the public as well as the Deep and Dark Net** are set by terrorists to raise funds for supporting and further promoting terrorist organizations. These campaigns¹² explicitly invite other terrorists, extremists or sympathizers to donor even a small amount of money in response to the terrorists’ “holly fight” and their bigotry. These campaigns last for only very few days or even hours for avoiding LEAs inspections and in their majority they are established on the Deep and Dark Web for the same reason.

Scope

ANITA will design and develop **an innovative knowledge-based user-centered cognitive investigation system** for analysing heterogeneous (text, audio, video and image) online (Surface Web, Deep Web, Dark Nets) and offline (LEAs’ databases) resources for fighting illegal trafficking activities through an elegant combination of: a) **avant-garde data source analysis and blockchain technologies** for the analysis of crypto-currency network and transactions (supporting the cases of Surface Web, Deep Web and Dark Nets); b) **advanced Big Data analytics tools** for the automatic extraction and analysis of the vast amounts of multimodal multimedia content contained in the identified sources; c) **sophisticated methodologies for capturing, modelling, inferring, processing and storing knowledge in human understandable forms** (e.g. expressive high-level semantic ontologies), including the case of connecting the connectionist learning paradigm with the symbolic learning one (i.e. extracting relevant and new knowledge from neural networks and formally storing it in the form of ontologies); d) the development of a sophisticated, adaptive, cognitive user modelling framework that will capture, analyse, interpret, mimic and integrate key human cognitive and information processing functions for: i) incarnating the **incorporation of human perception/cognition principles in the system processing pipelines (i.e. integrating the investigators/officers/practitioners ‘in-the-loop’ of the overall analysis process)** and ii) facilitate the **transfer of domain knowledge/expertise from the expert users to the novice ones**; e) **domain-related and user-oriented intelligence applications**, able to provide users with interactive dashboards to explore, reconstruct and identify patterns for spatial, temporal and causal correlations among illegal trafficking events, entities and activities and to support decision-making processes for countermeasures to undertake. All the above functionalities will be coupled and reinforced by an **in-depth interdisciplinary analysis of the online illegal trafficking phenomenon** (including the study of reaction strategies and countermeasures) and a **thorough analysis of the online resources with respect to social, ethical, legal and privacy issues** of concern. The proposed system capabilities will be demonstrated in **multiple relevant operational environments**.

¹¹ “Charity cash ‘going to Syrian terror groups” – The Telegraph – [URL] <http://www.telegraph.co.uk/news/worldnews/middleeast/syria/10357537/Charity-cash-going-to-Syrian-terror-groups.html>

¹² “Private Donors’ Funds Add Wild Card to War in Syria” – New York Times – [URL] http://www.nytimes.com/2013/11/13/world/middleeast/private-donors-funds-add-wild-card-to-war-in-syria.html?_r=0

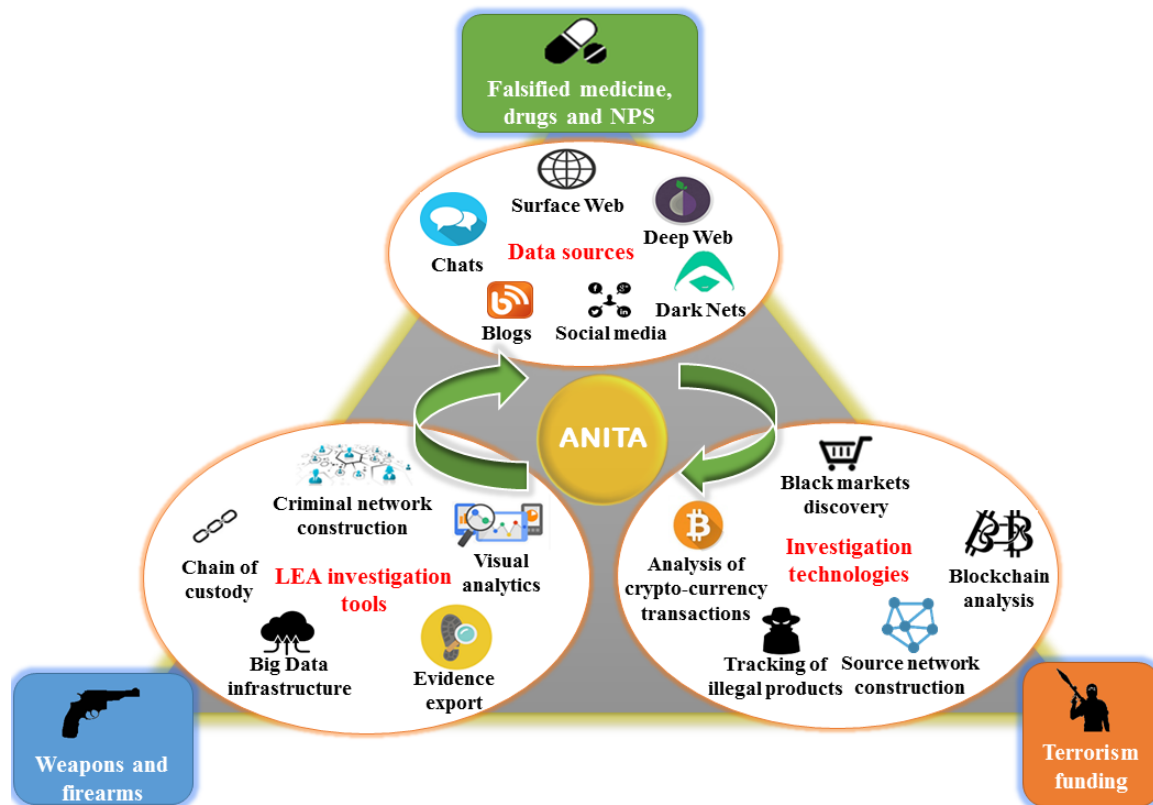


Fig. 1: Online illegal trafficking trends and LEA technological requirements

Main goals

ANITA's primary goal is twofold: a) **To boost the LEA's investigation process and to significantly increase their operational capabilities**, by introducing a set of innovative tools for efficiently addressing online illegal trafficking challenges (namely online data source analysis, blockchain analysis, Big Data analytics, knowledge modelling, incorporation of human cognitive function in the analysis pipelines, user-oriented intelligence applications), and b) **To significantly facilitate the novice officers training process and to optimize the learning curve** (by collecting, integrating and re-using knowledge from multiple expert officers and through the development of a recommendation functionality to transfer the acquired 'know-how' to the new officers).

1.1.2 Objectives

ANITA is mainly focussed on pursuing and achieving the following specific realistic **objectives (O)** within the duration of the project:

Objective 1: Enhanced investigation analytics services for Deep Web and Dark Nets

ANITA will build the next generation of investigation analytics capabilities for European LEAs. Since most of potentially useful online contents relevant for online illegal trafficking are not available in the Surface Web, the project will elaborate existing methodologies and solutions for online source identification, crawling and indexing, by making them efficient and effective for contents in the Deep Web and Dark Nets. Some of the European LEAs already have the capability to download deep and dark net content. But to address global organised crime and terrorism, this capability needs to be extended. In particular, ANITA aims to create the best possible infrastructure for near real time OSINT to find actual and new sources on the Internet, containing relevant information about organised crime support networks and illegal trafficking. The next generation of OSINT infrastructure should deliver: a) Fast and precise content identification, b) Discovery protected anonymous crawling with security by design, c) High performance download, and d) Flexible to different sources, language in depend and capable to deal with multi media.

KPIs	i) Fast and precise content identification: Content identification will be faster than today (5 relevant sources per hour, 90% relevant sources); ii) Discovery protected anonymous crawling with security by design (Crawling will not expose meaningful IP addresses, e.g. from LEAs); iii) High performance download (100.00 sites download per hour, with a test set of 10 server, extendable); iv) Flexible to different sources, language in depend and capable to deal with multimedia (Content identification will work in English, and two other languages, with text sound and video).
-------------	--

Objective 2: Accurate and efficient Big Data analytics services	
<p>In order to efficiently and robustly tackle the great challenges that arise from the extremely-large scale, fragmented, heterogeneous and distributed nature of the online contents and activities that are related to online illegal trafficking, ANITA will develop and deliver a set of sophisticated multi-modal Big Data analytics tools for analysis and revealing the information tokens of interest. In particular, the research activities will focus on delivering services for: a) Visual analysis and indexing: these will enable, on the one hand, the semantic analysis of the large quantities of visual content that is typically included in illegal trafficking sites, while, on the other hand, they will allow time-efficient search in the relevant databases based on visual queries (e.g. particular object and logo types, etc.). b) Text analysis: this will allow textual data processing in order to make the management of information - coming from Open Source Intelligence - automatic and independent on subjective criteria, c) Trend analysis: this will exploit the outputs of multilingual text analysis in order to extract information about specific trends, analytics and actionable insights on buying habits and user behaviours, d) multilingual translation: this will handle the task of automatically translating a piece of text available in the investigation data to one of the system supported languages, in order to support further investigation procedures as well as communication among LEAs from different countries, and e) ‘Speech-to-text’: this will materialize the transformation of an audio segment in one of the system supported languages to written text, in order to enable further deep natural language processing.</p>	
KPIs	<p>i) Number of visual semantic concepts detected by ANITA: >100; ii) Ability to handle truly Big Data sets: more than 1 billion data points; iii) Number of visual data points to index: more than 100 million data points; iv) Number of modalities/sources used in trend analysis: more than 5; v) Number of analytics insights to be provided: more than 4; vi) Number of entity types extracted by the semantic engine: more than 3; vii) Improving precision of text mining services for filtering, classification and events understanding: more than 5%; viii) Number of supported languages: more than 10; ix) Number of categories arranged in a 1-level taxonomy for semantic engine content categorization (for each use-case): more than 10.</p>
Objective 3: Automated domain knowledge collection, modelling, sharing and reasoning capabilities	
<p>In order to efficiently confront with the highly challenging issues that stem out from the largely unexplored, rapidly changing and continuously evolving area of online illegal trafficking, domain knowledge structures that are well-established, up-to-date, expandable, accurate and summarizing the expertise of multiple experts in the field are mandatory. Towards this goal, ANITA will rely on developing detailed, expandable and highly expressive knowledge models (in the form of ontological structures) of the domain of concern that will enable sophisticated inference procedures for solving tasks, like black-markets/illegal-shops/products tracking and criminal network construction. Particular attention will be given on exploiting and integrating knowledge that has been automatically acquired through the application of common machine learning tools (e.g. Neural Networks). More specifically, ANITA will focus on identifying prominent data relations in deep neural networks, in order to extract new and evolving knowledge from multifaceted unstructured and semi-structured information, and will subsequently utilize the latter information for refining and extending the already available knowledge structures (namely the ANITA ontologies).</p>	
KPIs	<p>i) Number of main concepts handled by the ANITA knowledge base: more than 1000; ii) Number of fully automated investigative functionalities introduced by ANITA: more than 7; iii) Automatic recommendations of concepts for black markets and Illegal shops categorization: more than 10; iv) Classes of categorization of illegal trafficking criminal groups: more than 5; v) Number of existing knowledge bases to be evolved, based on the knowledge acquired from the deep neural network: at least 3; vi) Number of application scenario for producing/formally-representing semantic knowledge from neural networks: at least 3.</p>
Objective 4: Integration of the human expert in the investigative analysis loop	
<p>In order to significantly enhance the efficiency of the proposed system and also to dramatically decrease the required time for new officers/practitioners/investigators training (i.e. enabling knowledge transfer among different users of the system), ANITA will adopt a highly novel approach towards the actual integration of the operating human experts in the investigative analysis loop. In particular, a set of user’s conscious and unconscious responses (e.g. user’s clicks, eye gaze position, eye gaze fixations, pupil dilation, facial expressions, emotions, body posture, etc.) will be continuously captured by the ANITA system. These data, together with an adaptive user modelling framework that will infer the current user’s cognitive state, will be used for iteratively updating, refining and improving the system’s Big Data analytics services; in this respect, on site domain expertise will be included in the service adaptation process. As described above, the enrichment of the system’s knowledge structures with information that has been automatically collected (O3) will also facilitate towards integrating human domain expertise coming from multiple individuals. It needs to be noted that particularly designed services will implement the transfer of the system collected knowledge on the particular application domain to the new and non-experienced investigators, based on an appropri-</p>	

ately developed recommendation scheme.	
KPIs	i) Number of types of human feedback sources to be used: >10; ii) Orders of magnitude decrease in training time of novice officers; iii) Reduction of feature detectors training time (for same recognition performance): At least 15%; iv) Relative recognition error rate reduction of feature detectors: At least 10%.
Objective 5: Understanding the phenomenon and its route causes and defining efficient counter measures and policy making acts	
In order to achieve the overall project objectives and to properly drive the technological development, a sound knowledge of criminal phenomena is fundamental. Within the ANITA project, this knowledge will be transversal to all activities, thus developing a solid teamwork between researchers, end-users and technological partners. In particular, the criminological analysis can generate a knowledge base about the phenomenon and the possible counter-measures which is of operational relevance under many respects: to better define the project scope and areas of interest; to assess the specific needs and requirements of the project target groups (especially LEAs); to give the technological partners a better understanding of the phenomena they are dealing with, so to orient the choice of the technological tools needed; to select relevant contents and sources that could be exploited by the technological partners; to contribute at thinking about new possible technological solutions to support the LEAs; to draft tailored guidelines to support the policy-making and the decision-making process; to deliver training modules using updated information/data and focusing on issues of real interest for the participants, etc. The acquired knowledge about the examined phenomenon will also be exploited to define efficient counter measures as well as targeted policy making acts.	
KPIs	i) Delivery of a theoretical and operational background to define the context in which illegal online trafficking occurs based on a criminological approach to investigate the main explanatory variables; ii) Delivery of a theoretical and operational background to define the criminal actors (new and old) mostly involved in illegal online trafficking and the peculiarities of the respective modus operandi, thus elaborating guidelines to deal with new actual and possible scenarios; iii) Delivery of guidelines about the actions needed to address illegal online trafficking, by combining knowledge and technologies.
Objective 6: Innovative TRL-6 applications for fighting illegal trafficking and criminal activities	
ANITA will end in the development of an integrated system able to provide LEAs with innovative investigation applications in support of intelligence operations oriented to the identification, analysis and monitoring of illegal trafficking activities in Surface web, Deep web and Dark nets. These applications will be designed to support investigative operations, like visual analytics on stored information, graph-based exploration of spatial, temporal and causal relationships among events, entities and activities, interactive management of investigative hypotheses, semantic-based browsing, search and retrieval of stored information. Results from O1 , O2 , O3 and O4 will be involved to realize the backend part of such applications, while innovative visual metaphors will be designed and developed as frontend counterpart starting from outcomes of O5 , in order to address LEAs needs and to support them in the deeper understanding of the illegal trafficking phenomena.	
KPIs	i) Improvement of the precision ratio of the search and retrieval mechanisms: more than 20%; ii) Improvement of the relevance of the results: more than 25%; iii) Workload reduction in relevant information retrieval: more than 15%; iv) Number of inferred patterns for linking black markets and illegal trafficking activities related to specific kinds of web sources: more than 5.
Objective 7: Evaluation in real operational environments in the context of real investigative scenarios	
The project will validate and demonstrate the project solutions, the integrated system and its functionalities through multiple pilots that will be executed in relevant environments provided by LEAs partners, in scenarios and in circumstances as close as possible to real operational conditions. Pilots will aim, on the one hand, at testing and finalizing the technological tools and, on the other, at facilitating the adoption of the ANITA system at LEAs level. Each usage scenario will have distinct characteristics and requirements to show the overall advantages provided by ANITA in illegal trafficking detection and monitoring. Pilots, hosted by the LEA partners, will be held in at least 3 different countries (namely UK, Netherlands and Belgium), with different legislation on both data integrity specifications and privacy. To this end, ANITA will work with real data from the very beginning, not only for the demonstration, but also for testing and tuning the algorithms and modules developed in the project.	
KPIs	i) Number of pilots in operational environments under the direct responsibility of the involved LEAs: more than 3; ii) Number of Law Enforcement Officer validators: more than 10.
Objective 8: Adherence to social, ethical, privacy and legal requirements	
To ensure proper management of personal data and to address all related social, ethical and legal challenges by mak-	

ing certain that the project is correctly adapted to the new legislative framework resulting from the latest EU data protection Regulation (i.e. GDPR). This objective will be achieved through the following concrete actions/processes: European data protection framework analysis, establishment of a data protection office, *analysis* of privacy, ethical and legal constraints, developing an impact assessment and defining mitigation measures. On the basis of a comparative analysis of the social, ethical, legal, privacy requirements emerged during the European data protection framework analysis, the critical aspects will be defined, including the possible impacts on the rights of individuals involved in the project activity (online users, researchers, third parties) and related solutions/mitigation measures will be defined to maintain the advanced investigation tools developed in the project within the frame of the legal and ethical standards. ANITA will also carry out a second impact assessment for researchers involved in the project. The assessment will be useful in order to balance the underlined risks with mitigation measures applying an Ethics-by-Design and Data Protection-by-Design strategy, in order to accompany the design and development of tools in a “friend of rights” manner.

KPIs	i) Ethics and legal value of produced tools/results; ii) Implementation of data protection and ethics by design; iii) Ethical and legal issues management through the use of impact assessments both for users and researchers; iv) Efficient mitigation measures to prevent and contrast any legal or ethical risk.
-------------	--

1.2 Relation to the work programme

SEC-12-FCT Challenges	ANITA contribution	Objectives
<i>Organised crime and terrorist organizations are often at the forefront of technological innovation in planning, executing, concealing their criminal activities and the revenues stemming from them.</i>	Organised crime and terrorist organisations nowadays make use of worldwide supply chains and knowledge networks to finance their activities. Surface Web, Deep Web and Dark Net offer new possibilities to organise criminal groups to proliferate. ANITA will build a novel knowledge-based system that will enable LEAs improve their picture about supply-chain and economic networks of organized crime and unveil persons, organisations, terrorist groups and activities, while focusing on the particular and rapidly evolving case of online illegal trafficking. New methodologies for online source identification, crawling and analysing content in the Deep Web and Dark Net will be introduced.	01, 02, 03, 05, 06
<i>Law Enforcement Agencies (LEAs) are often lagging behind when tackling criminal activities supported by "advanced" technologies.</i>	Although several methodologies have been recently introduced for automated processing of investigation data, they have not yet reached an acceptable level of accuracy in real operational environments. The ANITA project will provide advanced technologies, tools and services, together with an inter-disciplinary approach, to bridge the gap between metadata and features extracted by automated analysis and the natural way the investigation experts perform manual inspection. ANITA will enable LEAs to identify, collect and manage quickly, easily and automatically a massive amount of diverse typologies and sources of data/information and minimise the obstacles caused by space, time and language barriers.	01, 02, 03, 04, 06
SEC-12-FCT Scope	ANITA contribution	Objectives
<i>New knowledge and targeted technologies for fighting both old and new forms of crime and terrorist behaviours supported by advanced technologies</i>	ANITA will improve knowledge acquisition and modelling by developing a systematic knowledge model that allows to comprehend each single criminal phenomenon of interest and to analyse how criminal phenomena overlap or relate each other. Within ANITA, the domain knowledge will be incorporated into the process of automatic analysis in a way that the system's accuracy and reliability is gradually improved. Moreover, sophisticated computational schemes will enable incorporation of human perception/cognition principles in the system processing pipelines, thus, putting investigators “in-the-loop” of the overall analysis process.	01, 02, 03, 04, 06
<i>Test and demonstration of newly developed technology by LEAs involved in proposals</i>	The ANITA project will bring together major European LEAs (from 6 EU Countries) to shape, test and validate the project solutions, the integrated system and its functionalities through multiple pilots that will be executed in relevant environments provided by the LEAs.	06, 07, 08
<i>Innovative curricula, training</i>	ANITA will deliver a training programme for LEAs in order to	04, 05, 07

<i>and (joint) exercises to be used to facilitate the EU-wide take-up of these new technologies</i>	boost the integration of advanced technologies, methods and tools into standard operations and investigations. It is worth mentioning that within ANITA, a novel approach for knowledge transfer to new users/officers is envisaged, which will be based on learning from the interactions with content and the collected feedback of expert users. The latter is expected to decrease dramatically the required time for new officers/practitioners/investigators training.	
---	--	--

1.3 Concept and methodology

1.3.1 Overall concept

The struggle against illegal online trafficking activities is hindered by the presence of significant (and yet not efficiently addressed) challenges and facts that are related to the large-scale, fragmentary, rapidly-changing and mystical nature of the problem, such as: a) Complex and continuously updated production, distribution and consumption patterns of goods, b) Extensive and fragmented supply chains, which also make extensive use of Web resources (especially in the Deep Web) and crypto-currencies; c) The nature of the related crimes that renders them difficult to detect and justify; d) Multi-level, scattered structure of criminal groups and organizations, typically composed of both physical and digital Web identities; and e) Diverse and often not converging national approaches that leave space for displacement of harmful products/substances between Member States and cause a disruption of the legitimate trade. For efficiently encountering the problem intense, multi-national and well-organized activities are required towards implementing appropriate policy-making acts, health/legal approaches, law enforcement activities, crime prevention initiatives, information exchange, cooperation among national authorities and EU institutions/agencies. Towards this goal, ANITA introduces a **multi-disciplinary approach that targets the delivery of an on-line investigation analysis system for illegal online trafficking (including the popular trends related to: a) counterfeit/falsified medicine, drugs and NPS, b) weapons and firearms, and c) terrorism funding), which covers all aspects ranging from Surface-Web/Deep-Web/Dark-Nets source discovery/analysis to sophisticated Big Data analytics and high-level semantic reasoning services**. The aforementioned approach will also incorporate means for realizing an interdisciplinary analysis of the phenomenon and study of strategies/countermeasures, while particular attention will also be given for addressing social, ethical, legal and privacy issues that are related to the analysis of the online sources.

Based on the abovementioned fundamental scope, the project services and applications will be specifically designed and adapted for the analysts and investigation-related stakeholders, with the aim of supporting them (through Big Data collection, analysis and management) in the most demanding investigation processes. Discovering cues, facts, links and events of interest is one of the key challenges of the proposed solution, which will be used for preventing potential threats and discovering existing illegal activities and their relationships with organised crime. The overall project goal is to detect such relevant online contents in the Surface Web, but also in the Deep Web, including the Dark Nets. Based on advanced blockchain analytics, a core element of ANITA will be to collect all relevant open source information from Internet, to identify typical patterns of organized crime and to combine this information for unravelling the illegal trafficking networks. The ultimate goal of the system is to combine block-chain information from crypto currencies with additional information from deep web and dark nets, in order to make it possible to a certain degree to reveal the true identities of the stakeholders, customers of large scale black markets (like silk-road) and terrorists. Additionally, web-mining of organizations' communication in deep web and dark nets will provide the necessary information to enable the generation of a multidimensional situational awareness picture of the organized crime.

For addressing the above challenges, ANITA will design and deliver **an automatic, large-scale, user-driven, cognitive system for the efficient detection, tracking, monitoring and eventual prevention of the illegal online trafficking activities, while particular attention will be given on collecting knowledge for the application domain and re-using it for training new/novice officers**. A high-level graphical representation of the proposed system can be seen in Fig. 2.

The first necessary steps in the analysis chain comprise the detection, assessment and analysis of potentially interesting sources that can be found on Surface/Deep/Dark Web. In particular, a new generation of data collection tools will be developed, specific to LEAs needs. Specifically, dedicated services will be responsible for: a) **Anonymous identification of new relevant content with a balance between speed and precision**; b) **High performance download and storing in a secure repository**; and c) **Assessing the importance (i.e. level of relevance and dangerousness) of the examined Web sources**, the discovery of black markets, **block chain analysis for revealing cues about illegal transactions** tracking and the construction of a source network (that will include multi-level information for every source as well as the interconnections/interrelations among the identified sources).

Having identified and collected vast amounts of multimedia material related to illegal trafficking, ANITA will apply a set of sophisticated **Big Data analytics** services for efficiently manipulating the acquired information and robustly detecting meaningful events. In particular, text analysis services will deliver, through the usage of a semantic based engine, capabilities of **automatic categorization and entity extraction** of the contents coming from all the different sources (Social, Deep/Dark Web, etc.) used as input to the ANITA platform.

Content categorization will be designed by defining taxonomies -arranged in a tree of categories- fully customized on ANITA domain of interest and end user requirements. Entity extraction will instead allow to extract from text not only standard entities (i.e. people, organizations, places, etc.) but also specific and domain related entities (i.e. different types of drugs and NPS), including the capability to detect the usage of specific slangs, jargon acronyms and abbreviations. Besides, the extraction process will allow the retrieval of more sophisticated information, like **relationships among the extracted entities** and capabilities to extract time reference to perform **advanced information correlation**. In addition, the usage of an avant-garde semantic-based technology for text analytics will allow the implementation of advanced behavioral capabilities, like **stylometric analysis** (a.k.a. writeprint) for supporting clustering and linking of criminal and terrorist groups.

In parallel, **visual content analysis** will also be performed. A set of visual analysis tools (namely **object, concept** and **event** extraction modules) will be applied in order to identify potentially interesting pieces of information or evidence in the formed databases. For addressing the particular challenge of accurate and time-efficient search in the large-scale, **advanced visual indexing techniques** (namely deep hashing approaches) will be developed to recognise semantic entities at multiple scales (i.e. global and local concepts) and different levels of granularity. Additionally, for supporting the processing of documents written in different languages and facilitating the communication among the LEAs of different countries, a **multilingual translation service** will materialize the automatic translation of a speech segment to another language (supported by the system). For exploiting the information contained in the audio stream, **speech-to-text** methodologies will perform the transformation of a speech segment to the form of a written document. Moreover, **illegal trafficking trend analysis** will exploit the outputs of multilingual text analysis in order to extract information about specific trends, analytics and actionable insights on buying habits and user behaviours. Processing will involve statistics for highlighting actions (e.g. if the request for a specific substance/drug is increasing in a particular location at a specific time frame and probably correlate it with other similar actions), the most preferred products at a specific location and time, etc. These will assist investigators in prevention of emerging criminal behaviours, get actionable insights and respond rapidly to critical situations. The module will use state of the art **analytics** frameworks and **machine learning algorithms** to mine data from noisy input. It needs to be highlighted that all the above multi-modal Big Data analytics services will in principle adopt the **deep learning paradigm** and so the use of deep Neural Networks (NNs).

Apart from the inevitable large-scale data-driven analysis, the ANITA system will be grounded on appropriate semantic knowledge structures (in the form of expressive and exhaustive ontologies) that will summarize explicit domain expert knowledge regarding the application field and which will enable the realization of **high-level semantic inference** tasks (e.g. inconsistency checking, reasoning, outlier detection, etc.).

The collected knowledge (i.e. application domain expertise) will render feasible the realization of complex and highly demanding tasks, like **criminal network construction**, **illegal shops tracking** and **knowledge-based search and retrieval** that are of vital for analysing different aspects of the illegal trafficking incidents. The aforementioned tasks will be achieved in ANITA, with the support of multi-level, multi-source and multi-modal analysis services and tools, by recognizing specific types of online trafficking activities. The capability to monitor and model these activities will allow finding similar patterns during the crawling of unknown/unidentified sources, offering the possibility to improve **foresight, investigation and mitigation activities** on illegal trafficking and related terrorist and organised criminal groups.

Towards these goals, a highly novel methodology will also be considered that will target the bridging of the gap between the **connectionist learning approaches** (i.e. Neural Networks) and the **symbolic learning ones** (i.e. ontologies). More specifically, ANITA will pursue the evolution of ontologies through the extraction of patterns (in terms of network components' sequences) formulated within the available Neural Networks; thus, allowing for the acquisition of new, generic and contextual knowledge. The latter will be represented in uniform, machine understandable formalizations (e.g. propositional knowledge bases, ontologies), paying particular attention to expressing uncertainty and probability.

Current security systems for online source management and analysis exhibit two main limitations: a) They typically rely on a set of feature detectors (visual, audio, text processing) that are difficult and time-consuming to adapt to the application domain of concern; hence, the frequent failure of such detectors in real operational environments typically leads to failure of the overall system. b) When explicit high-level semantic knowledge is required, the cor-

responding knowledge structures are generally difficult to be well developed, maintained, updated and expanded; hence, often leading to a semantic gap between the output of the detectors and the high-level inference procedures.

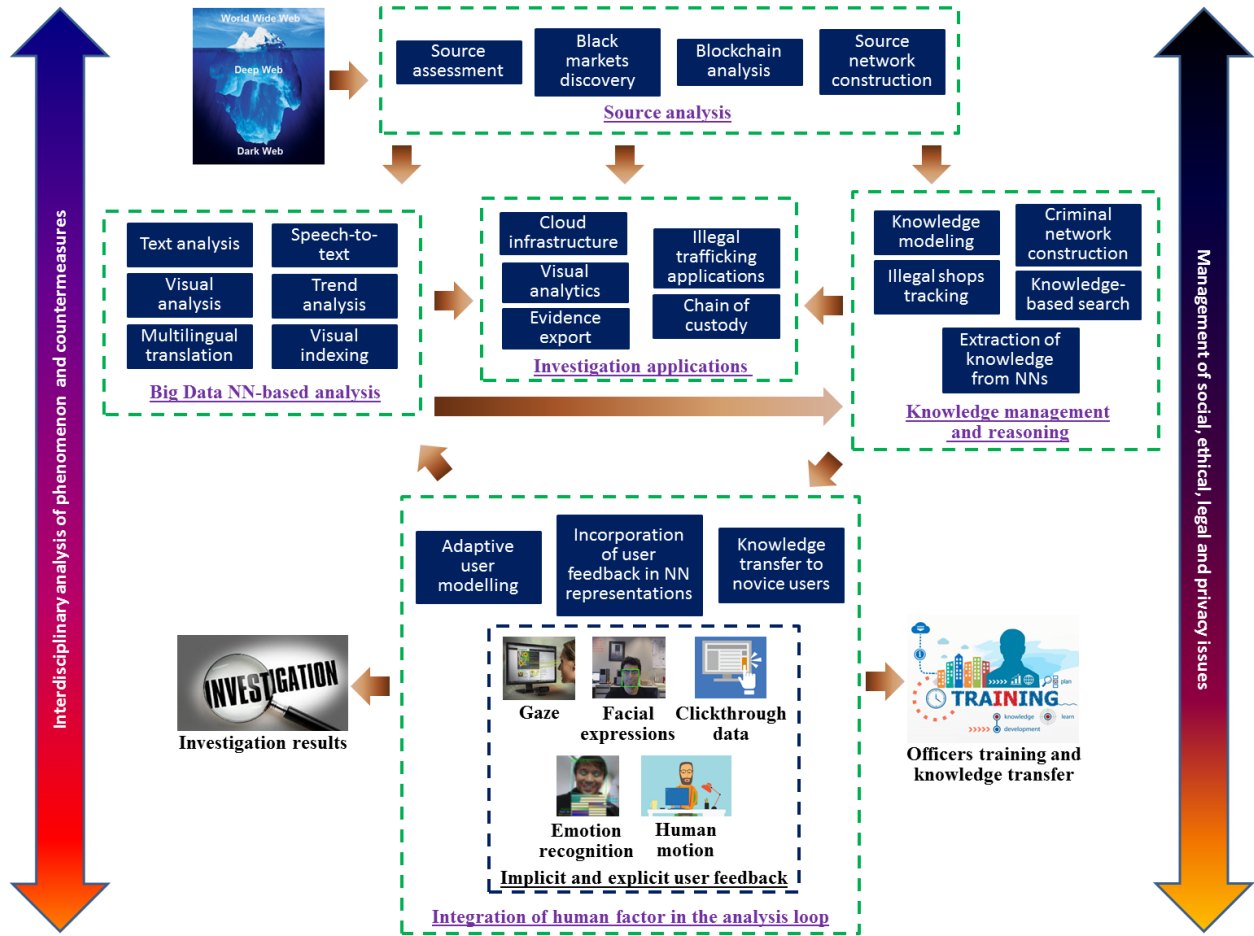


Fig. 2: Graphical representation of the ANITA system.

For overcoming the above limitations, ANITA will implement a revolutionary approach by integrating the implicit and explicit responses of the human user (i.e. officer, practitioner, investigator, etc.) in the analysis process; thus, forming an iterative closed-loop circuit, comprising machine (connectionist) data-driven learning, symbolic representation and reasoning, and incorporation of human conscious and unconscious feedback. Central to this approach will be the development of an adaptive user modelling framework that will infer the current user's cognitive state and performance from a combination of features obtained while interacting with the system. In particular, low-level primitives composed of task-related user actions (e.g. user's clicks, eye gaze position and eye gaze fixations, user presence, self-reports, etc.), outputs of the system (e.g. information returned by the system, etc.) and unobtrusive psychophysiological data (e.g. pupil dilation, facial expressions, body posture, etc.) will be translated into more abstract, high-level descriptions derived from theoretical frameworks of learning and development such as the three stages of learning (resistance, confusion, abduction) postulated by the Distributed Adaptive Control theory [Verschure11] and the concept of Zone of Proximal Development [Vygotsky78]. The captured user feedback, along with the abovementioned user modelling interpretation and corresponding computational models, will be subsequently utilized for: a) Improving the robustness and performance of the employed feature detectors. The latter will be realized by exploiting the large quantities of user responses (that will be automatically and non-invasively generated) about the particular application domain. Especially for the case of NN-based system components (i.e. the most typically met connectionist learning method nowadays), different approaches for incorporating human feedback in the analysis process will be investigated (e.g. human response driven training functions, regularization operators, representation layers, etc.). b) Update of the system knowledge structures and, hence, continuous integration of new domain specific knowledge. Specifically, through the developed methodologies for integrating the connectionist learning paradigm in the symbolic representations described above, the ANITA knowledge structures will be continuously updated, while forming a centralized knowledge repository integrating cues coming from many different human users. A highly innovative and outstanding characteristic of the ANITA system will be the transfer of knowledge that is collected from multiple users to new/novice ones. Specifically, the model of the

user, which is continuously learnt over multiple interactions with different expert users (i.e. by keeping track of the successful courses of actions and the associated implicit responses), will serve as a recommender system to replicate the strategy from expert users when novice users are interacting with the ANITA system to maximize the outcome and to significantly reduce the learning curve.

All the above-mentioned system functionalities will drive the design and will support the operation of a set of **novel investigative applications** to be delivered to the project stakeholders. In particular, the ANITA system will be based on the design and implementation of a scalable and **Big Data oriented infrastructure**, able to analyse large volumes of data in near real time and to summarize analysis results to provide LEAs with relevant insights on illegal trafficking related phenomena. The system will be based on a modular architecture, in order to include the possibility of integrating new tools and services that could be developed in the future to address the evolving domain of illegal online trafficking. The system will provide a set of **advanced functionalities for boosting LEAs investigations and supporting decision-making processes**. Such applications will include the possibility to manage different **investigative hypotheses** on events, activities, people and groups, the **graph-based knowledge exploration** at different levels of granularity (from a higher level in which relationships between key elements of the criminal networks are highlighted, to a lower level in which details of each entity or event can be examined and compared), the semantic search and retrieval of stored information, **notification mechanisms** that will alert users in case of relevant content detection, and the **geo-temporal visualization of correlations** among trafficking events and people to offer predictive support to the investigative operations. In addition, a visualization-as-a-service approach will provide users with an innovative framework, which will include a customizable and interactive dashboard to benefit from innovative **visual analytics** over stored information and to identify patterns and trends on trafficking events. Moreover, **chain of custody** and **evidence data export** will facilitate towards presenting investigation outputs in suitable way for the court of law.

Overall, the fundamental consideration of the ANITA system to integrate the human user in the analysis pipeline serves the following two fundamental project goals (and simultaneously main outputs of the system): a) **To significantly boost the efficiency of the investigation process**, by continuously improving the robustness of the feature detectors through the incorporation of the explicit and implicit user feedback, while also updating and expanding the knowledge infrastructure for the selected application domain. b) **To remarkably speed up the training process of new/novice investigators, practitioners and officers** for the application domain at hand, by re-using and transferring knowledge that has been collected and combined from multiple expert users.

ANITA will also pay particular attention to criminological-related issues. In particular, ANITA will focus on the following two main aspects: a) There will be a mutual cooperation and a sound understanding between all the most relevant components (research, enforcement, technological development); thus, ensuring that the project will develop actual, efficient and effective results, with operational relevance; b) The knowledge base generated by the criminological analysis will be fundamental to outline the possible actual and potential scenario, to single out the main related threats and factors, and to finally elaborate detailed guidelines to support the policy-making process and the enforcement activities, but also to propose new issues to be further investigated by researchers.

Beside the technological solution described above, ANITA will also develop an accurate set of tools responding to an **Ethics-by-Design** and **Data Protection-by-Design** strategy. In fact, through the evaluation of data protection issues and impacts on the project with regard to the processing of personal data pursuant Regulation UE 2016/679 (General Data Protection Regulation) and Directive 2016/680, ANITA will promote cybersecurity and information security as priority, avoiding potential misuse of the tools by balancing LEAs requests with data protection/ethics-by-design and by-default approach. This will contribute to establish a brand new relationship between the Dark Web users (represented by the member of the Consortium) and LEAs (participating in the project). For what concerns the ethical and legal aspects, ANITA will consider and protect also the researchers, by carrying out an impact assessment concerning the possible ethical and legal risks for researchers involved in the project, in light of the likely sensitiveness and possible side effects of the envisaged research activity in Dark Web. At the same time, the users' perspective will be evaluated in terms of impacts of the project on human rights. Overall, ANITA will not only develop technologies, but also **identify and tackle all potential legal and ethical obstacles** of its tools and approaches, drafting a comprehensive legal and ethical foundations report that takes into consideration all the above described perspectives.

1.3.2 Use case scenarios

The ANITA pilots reflect the classification of the means by which the Internet is often utilized to promote and support online illegal trafficking acts. A well-organised testing methodology has been designed to validate the proposed solutions, consisting of three different pilots, corresponding to three appropriately selected Use Cases. These use cases have already been identified from the LEAs partners of the project according to their experience and taking into account the current state of practices and will be further refined and detailed, during the first months of the

project, upon consultations with the end-user and advisory board members. A detailed description of the three main scenarios that correspond to nowadays popular trends related to illegal online trafficking and which ANITA will address is given below.

1.3.2.1 *Use Case 1: Counterfeit/falsified medicine, drugs and NPS*

Supported by: GDCOC, NPN, LPV

Scenario Description: The market of medicine and drugs holds a remarkable sharing among the world trade, either purporting for pharmaceutical purposes or as narcotics. Besides, the increasingly worldwide interest for the New Psychoactive Substances (NPS) that nowadays thrive and have become a global phenomenon -over 100 countries have reported one or more NPS substances- tends to even proliferate the total amount of transactions around medicine-related substances. On the other hand, considering the growth of Internet services that enables a profusion of opportunities for anonymous and quick online illegal transactions, medicine trafficking costs millions to the European Nations, having also adverse impact on the public health of European citizens. The ANITA system key purpose is to automatically detect and further monitor medicine trafficking activities from online transactions that are realized through the Public Web, as well as Deep and Dark Nets. Medicine trafficking is mainly realized using e-payments, pre-paid mobile phones, social media profiles and cryptocurrency markets. To this end, the system will automatically detect, monitor and analyze heterogeneous multimedia large-scale data from online retrieved sources for identifying the real people groups behind, thus disrupting their illegal activities. The ANITA system is anticipated to address the high complexity of detecting financial networks –e.g. blockchain analysis- through the Surface and Deep Web that is currently performed manually from the LEAs’ investigators, by incorporating high automation levels in its services. Finally, the system is expected to increase its accuracy in medicine trafficking detection by learning from the investigator experience, as well as provide backwards feedback to the system user for enhancing his investigation expertise in a double-way knowledge transfer manner.

1.3.2.2 *Use Case 2: Weapons and firearms*

Supported by: KWPG, GDCOC, AoC

Scenario Description: Illegal trade of weapons and firearms is quite spread among several illegal activities, including terrorism and criminal actions. Living into the zenith of terrorism across Europe and reportedly having records for criminal activities, the weapon trade market is increasingly growing. Criminal activities are significantly sustained from the weapon trafficking and therefore the contemporary Internet structure that promotes the easy and anonymous trading of guns and firearms. Therefore, towards the disruption of criminality, obviating the weapon trafficking activities will be particularly helpful for confining the means for access in illegal equipment. Once weapon and firearm trafficking can be realized across several individuals or criminal groups, the possibility for a murder or attack gets significantly decreased and hence it is imperative to control in some way those activities. The ANITA system aims at providing automated services for detecting weapon trafficking activities through different levels of online sources. Besides detection, the system will convey functionalities for automatically monitoring such activities and eventually identifying the individuals that trade illegal weapons and firearms. ANITA will collectively integrate knowledge from online retrieved Big Data sources for assisting LEAs’ investigations that are currently performed by hand and thereby are by far more complex and time-consuming.

1.3.2.3 *Use Case 3: Terrorism funding*

Supported by: CAST, KWPG, NPN, LPV, GDCOC

Scenario Description: Fund-trafficking activities are the key element for every terrorist organization, in order to collect funds for sustaining its operational needs. Therefore, the control and disruption of those activities is casted as the most consequential part of any fight against terrorism. The rapid growth of the Internet which comprises several services that enable the anonymity of online transactions has been quite popular for terrorists. Online fund trafficking methods provide great means for terrorists to raise funds with numerous ways (e.g. illegal trafficking of goods, money transfer in Deep Web, money laundering, illegal transactions, etc.), posing significant burdens to LEAs’ officers for potentially detecting and identifying the involved people. The ANITA system will enhance the officer’s investigations, as it will provide automated services for the detection of financial networks and flows through the Deep and Dark Nets, since most of the fund trafficking activities are performed in the private zone of the Web. ANITA will provide means for retrieving relevant online information through the Web crawling services and those tools that are needed for the analysis of heterogeneous collected multimedia information. Therefore, the system will detect and monitor financial activities based on multi-source Big Data information and thus will draw inferences in a holistic way, using all the evidence that is pertinent to the subject of the investigation.

1.3.3 ANITA approach and methodology

The H2020 work programme has an increased focus on innovation capacity of projects in addition to requiring scientific excellence from research. ANITA as a Research and Innovation Action carefully positions its activities to maximise the potential impact and accelerate the transition of research results conducted in the lab-based experimental facilities to industry wanting to exploit them. Research is the process of turning money into knowledge, whereas Innovation is the process of turning knowledge into money. The overall approach combines a well-defined scientific and technical strategy (while considering business constraints and opportunities) to ensure project objectives are achieved. The ANITA consortium will put first the effort to successfully carry out the needed research in the fields of Dark Net sources and streams analysis, Big Data analytics, deep learning, knowledge management and user modelling and will deliver a new line of technologies beyond the state of the art. This will foster the competitiveness of European industry in the European (initially) and global (subsequently) levels, and could introduce new market segments. The ANITA consortium is confident that successful completion of the project will foster the effectiveness of the Law Enforcement Agencies to address illegal online trafficking in Europe (initially) and globally (subsequently), substantially reducing the loss in terms of time and human effort.

1.3.3.1 Business Strategy

While pursuing a concrete advancement in the scientific and technical state of the art, ANITA aims to achieve a strong market position for the new and highly innovative tools resulting from the project. All project activities and resources are allocated so as to fully achieve this goal. The breakthrough technologies to be delivered by ANITA require investigation and in-depth analysis of complex trade-offs including technical (e.g. hardware constraints, algorithmic constraints, etc.), operational (e.g. costs, quality of service, etc.) and socio-economic (e.g. business models for sustainability, cross-jurisdictional, etc.) aspects. A decision-making framework will be provided to drive technical and operational decisions within the project allowing for consideration of such complex trade-offs and associated innovation. Key elements of this business-driven strategy include:

- Use of partners' available proof-of-concept or market-ready technologies and solutions to be evolved and innovated with breakthrough technologies;
- Tight collaboration with end-users to promote the results and technologies across Europe and globally;
- Allow each partner an opportunity to exploit their contribution separately, where possible, as well as for partner groups to exploit sub-components of the overall framework;
- Create a unique market proposition, a pronounced innovation strategy and sustainable business plans;
- Build a large community of stakeholders and interested parties for efficiently communicating them the outcomes of the project and finding possible technology consumers. The presence of LEAs partners in the consortium will allow validating and optimising the solution and boost its adoption in their operational context.

1.3.3.2 Technical Strategy

The technical strategy is derived from the above-mentioned business-driven strategy. It will be used to guide the technical ambition of the project. For achieving the scientific goals and delivering high-quality results, ANITA will adopt an iterative development, integration, and prototyping strategy, where the various framework components will be integrated and combined in several phases to implement the overall approach. The project will adopt a service oriented integration approach towards the delivery of the ANITA set of tools. The integration process of the individual ANITA components and services will be incremental, while, in parallel, thorough evaluation and unit tests of the individual services and implementation of the defined use-cases will be performed. Development will progress in parallel for the various tasks (see the work plan), allowing researchers and developers to optimize their work while keeping the project aspects in sight during the implementation.

ANITA will follow the iterative and incremental development proposed by the **SCRUM** scientific and technological methodology that is based on four phases. In more detail:

- **Phase 1** spans over the first 12 months of the project. It will include definition of the use cases and scenarios, elicit the user requirements, the ethics/regulation/social acceptance and system requirements, model the dynamic assessment and define the system architecture and the system integration plan.
- **Phase 2** starts at month 13 and finishes at month 18 and is focused on first system validation in lab. During this phase, the core developments take place, regarding the data sources and streams analysis, Big Data analysis and analytics. The 1st integrated ANITA prototype will be released, following integration guidelines. Evaluation will start via lab prototyping and testing, identifying potential limitations and drawbacks (M15-M18).
- **Phase 3** starts at month 19 and finishes at month 27 and is based on the feedback from Phase 2. New prototype development, including complete functionality for all ANITA services and tools, will be evaluated at both usability and performance levels. Lab testing will be exponentially increased, including initial user evaluation (M19

- M24). During this phase the first version of the pilots will take place so that usability and impact creation will be analysed and assessed (M25 – M27) in order to achieve a complete validation in relevant environment.

- **Phase 4** (M28 – M36) is based on the evaluation results of Phase 3, leading to the refinement of the final set of tools after a first validation stage in relevant environments and taking into account feedback and improvements required by end users and security experts. During this phase the second stage of pilots will take place (M34-M36) in order to achieve a complete demonstration in relevant environments.

Thus, the ANITA methodology will allow for:

- **Overlap, parallel execution and optimal coordination between activities** like requirements extraction, specifications, prototyping and real-time operation testing and validation. This ensures both faster project execution and avoids the syndrome of “throwing things over the wall” to the next department and process phase, without any chance of influencing outputs of preceding phases.
- **Involvement of users and relevant stakeholders from different areas of expertise and/or organizational units** from earlier project activities in the different phases, by adopting a **User Centred Design** approach towards user-centred validation, where “user” within ANITA is any Collaborating Member, ranging from LEAs to end users, as well as a **Privacy-by-Design** approach ensuring the compliance with ethics and regulations at national and European level. The key benefit of the ANITA user centred design is that stakeholders and policy makers along or across value chains can contribute with their insight at a sufficiently early stage, so that their input can be taken into consideration; for example, changing the design of a part leads to a much more efficient production process, without drastically changing its properties.

1.3.4 Positioning of the ANITA project

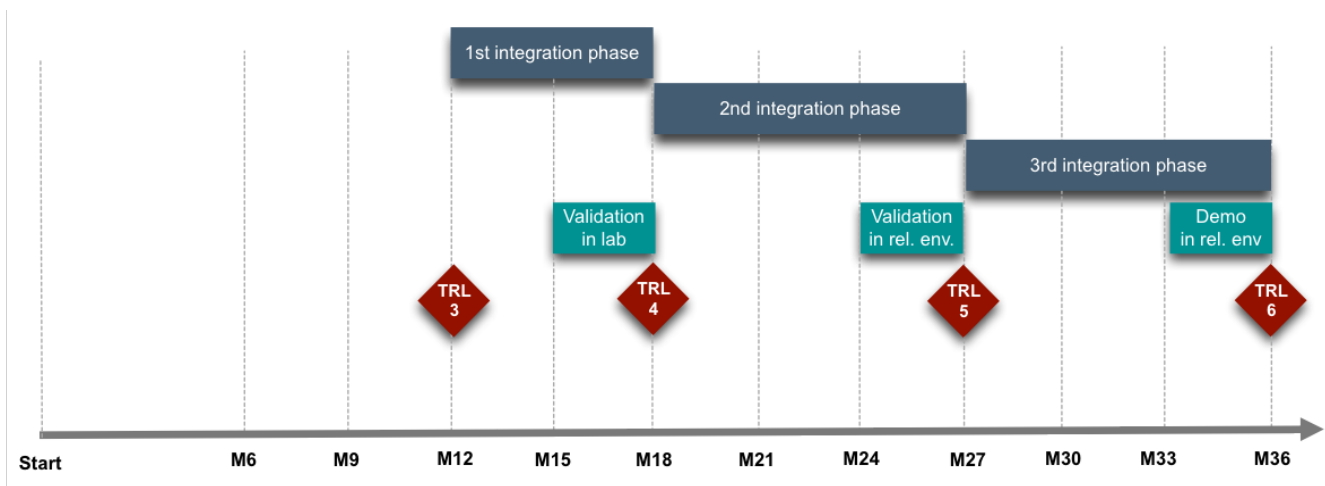


Fig. 3: ANITA system integration/validation phases and TRLs.

Using the Technology Readiness Levels (TRLs) (annex G of the H2020 work programme), ANITA covers the full spectrum from TRL3 (experimental proof-of-concepts) to TRL6 (technology demonstrated in relevant environment). The ANITA services and the final set of tools and applications will be demonstrated in the relevant environments provided by the LEAs involved in the project (KWPG, AoC, CAST, NPN, GDCOC, LPV) using data provided or validated by them. Three different integration and validation stages will be carried out in order to achieve the ANITA objectives. **During the first integration phase (M13-M18) of ANITA**, a preliminary integration will take place, using the initial versions of the tools and services. At this phase, the developed technology will be validated in lab achieving at least TRL-4. **During the second phase (M19-M27)**, refined, almost final services will become ready. Validation in relevant environments will take place, reaching TRL-5. **Finally, during the third integration phase (M28-M36)**, after demonstration of final versions of tools in relevant environments, feedback from LEAs will be taken into account and final improvements will reach at least TRL-6 in all components. Specific workshops will be also organised to validate the TRL at each stage of the project. The first one for the validations in lab will involve only the ANITA consortium; the second and the third will also involve external LEAs and advisory board members. A list of proposed ANITA key technologies along with the expected TRLs is given in the table below.

ANITA key technology	Description	Starting TRL	Expected TRL
Data source risk assessment	The service will provide a fast and easy preliminary classification algorithm to identify homepages at crawling time, which are relevant for online illegal	3	6

	trafficking. This is necessary for focused crawling.		
<i>Black markets discovery</i>	This will identify new homepages with illegal services, like black markets for drugs and weapons, at crawling time to support automatic reasoning for focused high performance crawlers.	3	6
<i>Blockchain analysis</i>	The blockchain analytics service will provide network analytics on existing (e.g. Bitcoin) and upcoming next generation cryptocurrencies to identify financial networks.	5	7
<i>Source network construction</i>	The network construction service adds high performance network forming attributes to the crawling results to visualize the identified sources as a temporal geographic network.	4	6
<i>Filtering, classification, grouping</i>	The service will provide capabilities for rapid data analysis oriented to filtering online trafficking activities related contents, while it will provide semantic characterization and categorization.	5	7
<i>Semantic natural language processing</i>	This will support avant-garde natural language processing heuristics able to manage knowledge about slangs, acronyms and abbreviations; hence, activating the analysis of deep and dark web sites and transforming their content into usable information.	5	7
<i>Stylometric analysis</i>	This will comprise an advanced write print component that will expose to the LEA analysts an advanced socio-cultural analysis tool related to content biometrics.	4	6
<i>Authorship verification</i>	This will use results of stylometric analysis and a supervised approach for authorship verification of new and anonymous text.	4	6
<i>Visual analytics</i>	This will automatically detect a wide set of meaningful semantic entities (i.e. objects, concepts, events) in the available visual content.	4	7
<i>Multilingual automated translation</i>	The aim of this module is to enable the automatic translation of a speech segment to another language (supported by the system).	5	7
<i>Multilingual speech to text</i>	This module will perform the transformation of a speech segment to the form of a written document.	5	7
<i>Visual indexing</i>	This module will enable large-scale indexing of images and videos, using global as well as local information, so as to realize fast and accurate search and retrieval tasks in vast visual content collections.	4	6
<i>Black markets and illegal shops-tracking</i>	Metadata and textual content coming from heterogeneous resources, present in black markets and illegal shops, will be deeply examined to discover possible connections and will be merged together to reconstruct complex illegal events.	4	6
<i>Criminal network reconstruction</i>	This module will deliver the reconstruction of criminal networks, by the application of advanced reasoning mechanisms for the discovery of hidden relationships on monitored sources.	4	6
<i>Extraction of knowledge from deep neural networks</i>	This module will identify prominent data relations in deep neural networks, in order to enable the extraction of new and evolving knowledge, where both multifaceted unstructured and semi-structured information will be considered.	3	6
<i>Adaptive user modelling</i>	This module will comprise an adaptive user modelling framework to infer the user's cognitive state and performance from a combination of implicit and explicit features.	2	6
<i>Incorporation of user feedback in deep learning representations</i>	This module will incorporate the captured explicit and implicit human responses (e.g. gaze signal, facial expressions, click-through data, user emotions, user motion, etc.) in the actual deep-learning-based analysis process.	3	6
<i>Knowledge transfer to new officers</i>	This module will utilize the expert's user model and will transfer the knowledge to novice users. This component will act as a recommender system that will learn from the interaction with content and the collected feedback of the expert users.	2	6
<i>Novel applications for illegal trafficking</i>	Innovative applications to support LEAs in discovering and monitoring illegal trafficking activities will make LEAs able to explore, monitor, reconstruct and identify spatial, temporal and causal correlations among illegal trafficking events, entities and activities.	4	7
<i>Big Data visual analytics</i>	The analytical framework will include an interactive dashboard and several visualization patterns oriented to the illegal trafficking domain, in order to provide LEAs with the capability to explore information at multiple levels of details.	4	7
<i>Chain of custody and evidence export</i>	This module will allow maintaining integrity of stored information and it will also be able to export stored information in the most common formats.	3	6

Table 1.3.a: Starting and expected TRLs of ANITA key technologies.

1.3.5 Building on pre-existing national or international research and innovation activities

The ANITA Consortium recognizes the opportunity of exploiting partners experience and know-how acquired in past projects and to collaborate with other national and international projects. Hereafter, the relevant EU projects that will enhance the ANITA activities and with which ANITA can establish a liaison relationship are listed below.

Project	Description	ANITA Partners	Relation to ANITA
<i>National projects</i>			
SINTESYS Security INTElli-gence SYStem (Italian National Project)	Advanced framework for analysing, investigating and correlating vast amounts of data from a variety of heterogeneous, multichannel and multimodal (text, image, video, audio) sources. (http://sintesys.eng.it)	ENG EXPSYS	The source integration framework, the integration middleware and the knowledge browsing and search applications developed in SINTESYS will be adopted in ANITA as starting points for services.
BITCRIME (Binational project Germany-Austria)	Advanced framework for identification of organized crime networks, based on block chain analytics and dark net crawling. (https://www.bitcrime.de/)	AIT	Relevant research results from BITCRIME comprise a unique method to identify financial networks. This will be used as starting point for research and development in ANITA.
<i>International projects</i>			
FAKECARE (2012 – 2015)	Developing expertise against the online trade of fake medicines, by producing and disseminating knowledge, counterstrategies and tools across the EU. (www.fakecare.com)	RiSSC	The achieved research results for the analysis of the “hot products” will be further improved for achieving a better understanding of the reasons why some medicines are more commonly falsified and their online supply chain.
EPS/NPS Enhancing Police Skills concerning Novel Psychoactive Substances. (2015-2017)	Co-financed by DG Justice-DG Migrations and Home Affairs/Targeted call on cross border law enforcement cooperation in the field of drug trafficking. (www.npsproject.eu)	RiSSC	ANITA will build on the project’s interdisciplinary research/analysis on Internet and Deep Web + virtual library, cooperation platform addressed to LEAs and relevant public/private stakeholders.
SICH Semantic Illegal Content Hunter	The project aims to create a smart tool to support investigation's activities in order to discover illegal contents on the Net (related to three separate areas: Xenophobia/Racism, Online Illegal Gambling and NPS, capture "weak signals" and identify trends.	EXPSYS RiSSC	Resources customized and know-how developed in the SICH project will be used as a starting point for developing new “semantic resources” (ontologies, taxonomies, semantic rules) more precise and more effective in the ANITA project.
IANCIS Indexing of Anonymous Networks for Crime Information Search	The project aims to building a tool, able to crawl and semantically index and cluster Onion websites. The criminal scope targeted is related to different crimes existing on the onion network.	EXPSYS	Resources customized and know-how developed in IANCIS project will be used as a starting point for developing new “semantic resources” (ontologies, taxonomies, semantic rules) more precise and more effective in the ANITA project.
DANTE Detecting and analysing terrorist-related online contents and financing activities	DANTE aims at performing large-scale online multi-modal (visual, audio, text, social media) content analysis for tracking terrorist related activities (fund-raising, propaganda, training) in the surface and the deep/dark Web. (http://www.h2020-dante.eu/)	CERTH ENG EXPSYS RiSSC AIT	The expertise obtained within DANTE concerning large-scale visual content analysis, text processing, knowledge modelling, on-line content manipulation, financial network tracking and criminological analysis will be transferred and elaborated within the ANITA project.
ASGARD Analysis System for Gathered Raw Data	ASGARD focuses on creating, maintaining and evolving the best possible set of tools for the extraction, fusion, exchange and analysis of Big Data including cyber-offenses generated data for forensic investigation. (http://www.asgard-project.eu/)	ENG, CERTH, AIT, CAST	The technologies developed within ASGARD, related to data acquisition from Dark Net, multimedia analysis, indexing and knowledge modelling, will be exploited as background knowledge in ANITA.
TRILLION Trusted, Citizen-LEA collaboration over social networks	TRILLION provides a novel community policing platform that gathers and processes information (text, image, video, data, etc.) in real time through various sources (social, web, mobiles, wearables, etc.). (http://trillion-project.eng.it/)	ENG CERTH DITSS	The ANITA project will be built upon the outcomes of TRILLION in text and visual analytics and data mining to ensure that relevant information is exploited in the analytics cycle.
CEEDS The Collective Experience of Empathic Data Systems	The CEEDS project has developed novel, integrated technologies to support human experience, analysis and understanding of very large datasets. CEEDS has developed tools to exploit theories showing that discovery is the identification of patterns in complex data sets by the implicit information processing capabilities of the human brain.	IBEC, CERTH	In ANITA we will capitalize on the results of the CEEDS project which constituted a successful attempt toward the integration of human factors and the analysis of implicit user’s states in the discovery process of neuroscientific datasets.
EASEL Expressive Agents	EASEL has explored and developed a theoretical understanding of human-robot symbi-	IBEC	In ANITA we will further explore the perspective of developing adaptive tutoring system

Project	Description	ANITA Partners	Relation to ANITA
for Symbiotic Education and Learning	otic interaction (HRSI), incorporating key features of human tutors and other proven approaches capable to instruct a human user and learn from their interactions during large time scales.		incorporating concepts that are grounded on theoretical models of human experience (DAC and Zone of Proximal Development).
Privacy Flag Enabling Crowd-sourcing based privacy protection for smartphone applications, websites and Internet of Things deployments	The Privacy Flag project researches and combines the potential of crowdsourcing, ICT technologies and legal expertise to protect citizens' privacy when visiting websites, using smartphone applications, or living in a smart city.	IIP	ANITA will benefit of the legal analysis on GDPR and of the Data Protection Office method already tested and validated in Privacy Flag
KNOWMAK Knowledge in the making in the European society	The aim of this project is to create a methodology and web-based platform that presents data on institutional, regional and national knowledge co-creation hot-spots, enabling web-based search for co-creation of hot spots addressing societal challenges or Key Enabling Technologies.	AIT	Methods of KNOWMAK will be used in ANITA to identify innovation networks of organised crime and terrorism.

Table 1.3.b: Relevant projects.

1.3.6 Sex and gender analysis

The consortium acknowledges the “gender issue” as stated in the EU regulation 1291/2013 as of 11. December 2013 establishing Horizon 2020 and supports the objectives of the framework programme on: a) Gender balance in research teams, b) Gender balance in decision-making and c) Integrating gender/sex analysis in R&I content. The companies and organisations involved in the project are committed to encouraging equal opportunities of career among women and men in their staff according to national and European laws and corporate ethical code. The project basically addresses women’s needs, as much as men’s needs. Women partners are fairly involved in the project. The partners will promote gender equality within the frame of the project, addressing the European Commission goal of reaching a 50% participation of women at all levels, in implementing and managing research programmes. In particular, the Consortium, which is in charge of monitoring the progress of the overall project, will also have the task of: a) Adopting the appropriate measures encouraging women participation in the management of the project, in order to achieve a balanced representation; b) Solving any gender-related issue; c) Supporting the implementation of the recommendations produced by the European Technology Assessment Network (ETAN) as well as by the “Helsinki Group” on the development and production of statistics and indicators, about the situation of women in scientific research. ANITA is, in general, gender-neutral, because the same requirements for audience measurement will apply for both genders without any reason to discriminate them. Nevertheless, a particular attention in all the phases of the project will be paid to gender-related issues that will be considered while investigating scenarios and user needs.

1.4 Ambition

1.4.1 Incorporation of implicit and explicit user feedback in deep learning representations for visual analysis

Data-driven techniques and in particular the so called Deep Learning (DL) paradigm have recently shown outstanding performance in a broad range of visual analysis application tasks, including object detection, concept detection and image classification. However, the fundamental prerequisite for the application of any DL technique constitutes the availability of vast amounts of annotated training data, which are typically associated with binary annotation tags (i.e. presence/absence of a semantic concept in the content). On the contrary, visual perception and cognition in humans involve complex information processing procedures, which cannot be summarized in the form of binary tags. Prior to the DL era, several studies have been proposed for incorporating captured user feedback in the analysis process, with human gaze being by far the most widely used feature. In particular, Shcherbatyi et al. [Shcherbatyi15] introduce an early integration approach within a deformable part model, which constitutes a joint formulation over gaze and visual data. Visual fixations are used in [Papadopoulos14], as a form of weak supervision, for forming object detectors. However, in the context of DL techniques, little work has been conducted so far that exhibits the following main limitations: a) The literature focuses on saliency detection methodologies, i.e. avoiding the development of feature detectors or modelling recognition mechanisms, and b) No end-to-end systems have been proposed so far, i.e. the captured human response has not been used for guiding the actual NN training process, but only as a separate processing step. A Convolutional Neural Network (CNN) approach is introduced in [Roman17] that selects object proposals on the basis of visual attention maps computed over gaze fixations, using the computed saliency as a pre-processing filtering step. Additionally, Sugano et al. [Sugano16] utilize a split atten-

tion model for image captioning that integrates human gaze information into an attention-based Long Short-Term Memory (LSTM) architecture. Moreover, a gaze pooling layer is introduced in [Sattar17a] that integrates gaze information into CNN-based architectures as an attention mechanism for predicting categories and attributes of visual search targets. Furthermore, Kloss et al. [Kloss16] present a top-down saliency model, which is trained on human fixation data and predicts relevant image locations when searching for specific objects.

Ambition: Within ANITA, novel DL schemes will be developed targeting the incorporation of implicit (subconscious) and explicit (conscious) feedback of human users (e.g. officers, practitioners, investigators, etc.) in the visual analysis process. In particular, ANITA will exploit non-invasive user responses that can be efficiently captured in real operational environments (e.g. gaze signal, facial expressions, clickthrough data, user emotions, user motion, etc.). These sensorial signals will be used for guiding the training process of the NN-based visual analysis modules; hence, integrating detailed expert knowledge in the ANITA feature detectors and generally the analysis process. For realizing the latter, different neuronal implementations will be investigated for each type of sensorial signal, such as adaptive human response guided NN loss functions, regularization operators, intermediate representation layers and deep auto-encoders (for multi-level sensorial signal modelling and interpretation).

Innovation Potential: Despite the numerous works that have been presented so far that are based on the deep learning paradigm, these are purely data-driven and require large amounts of data for training purposes. ANITA will aim at incorporating human behavioural information, which will render the Neural Network training process more intuitive and efficient; hence, introducing a whole new generation of deep learning methods for multimedia information processing.

1.4.2 Extraction of new and evolving knowledge from deep neural network representations

For the efficient representation, interoperability, management, search and retrieval of data/meta-data (especially in the Big Data era), it is of paramount importance to construct and utilize a comprehensive, holistic, dense and well-structured semantic vocabulary. In this respect, formal knowledge representations and formalisms (with ontologies being among the most common ones) provide the required expressivity and conceptual basis, and in extent enable advanced decision-making through semantic inference procedures. However, despite the fact that ontological representations enable the generation of accurate semantic descriptions, as well as search and inference over multipart and intricate knowledge, they are hampered by the knowledge acquisition bottleneck [Hovy13]. To this end, intense research effort has recently been devoted towards developing methods for automatic knowledge acquisition. In most cases where automatic knowledge retrieval from unstructured or semi-structured resources was pursued, it was limited to shallow taxonomy induction, sometimes coupled with the induction of pre-defined, high-level, non-taxonomical relations between concepts, or obscure connective relations between entities, lacking understanding and representation of meaning [Cambria14]. The latter render automatically retrieved information insufficient in modelling meaningful, holistic, complex knowledge structures. Recently, the exploration of the deep learning paradigm with implications to pattern extraction and meaningful knowledge acquisition comprises an emerging and highly challenging topic [Wang15], where the deep neural network internal representations are used in order to emulate and transcend human cognitive abilities towards enabling machines to capture critical aspects of data. The latter task actually refers to the highly ambitious and yet largely unexplored research challenge of bridging the two dominant learning paradigms in Artificial Intelligence (AI): a) Connectionist learning, with Neural Networks being the most commonly met model, and b) Symbolic learning, where ontologies have a decisive role. In particular, it was early on recognized that deep learning can drive knowledge acquisition beyond taxonomy induction and lexical analysis, while it can “learn distributed representations of structure, such as logical terms”. In this respect, Tran et al. [Tran13] introduce a rule extraction method for discovering knowledge from deep networks, aiming at a better understanding and explanation of the relations between network layers. Additionally, Petrucci et al. [Petrucci16] exploit advances in the deep learning field for tailoring the ontology learning problem as a transductive reasoning task that learns to convert knowledge from natural language to a logic-based specification.

Ambition: ANITA will focus on identifying prominent data relations in deep neural networks, in order to extract new and evolving knowledge from multifaceted unstructured and semi-structured information, taking into account the neural network internal representations. Additionally, representing this knowledge in uniform, machine understandable formalizations (e.g. propositional knowledge bases, ontologies, etc.) will be pursued, with respect to expressing uncertainty and probability over extracted knowledge. Analysis and knowledge retrieval on these networks will be based on graph theory, set theory and machine (deep) learning. The project will also integrate reasoning techniques for ontology verification, i.e. to automatically assess and rectify the produced knowledge.

Innovation Potential: The scalable and re-usable knowledge that will be produced in the ANITA project will enable more efficient representation of and access to complex and intrinsic investigation information. Such data-driven knowledge acquisition will allow for interweaving disperse information across voluminous and heterogeneous data,

and for capturing and representing underlying meaning, complex notions and experiences that govern behavioural and situational models of investigation analysis. The acquisition mechanisms will support for both the augmentation of human-generated knowledge with collective and tacit knowledge, as well as the creation of standalone ontologies, formalised in standardised, expressive knowledge bases that will support solving complex problems through logical inferencing, thus enabling advanced analytical investigation services.

1.4.3 Deep Web and Dark Nets sources and streams analysis

In the last years, Open Source Intelligence (OSINT) has become a very important part in investigations of global organised crime networks. Crawling the web for investigation purposes has become a high technology capability of state of the art police work [Sameer17]. In 2015, the Justice and Home Affairs (JHA) Council of the EU agreed to set up, Europol's Check-the-Web (CTW) service, provided by the EU Internet Referral Unit [EU-IRU15]. With about 21 employees, the IRU managed to identify 9 platforms with terrorist activities in 2015 and 70 platforms in 2016. An open question is, whether manual scanning or machine scanning is better. Actual publications about web crawling point to the direction that machine crawling will win the innovation race [Akshay17]. But practical experience shows that this is not that clear. Automated crawling is fast and can handle huge amounts of information [Parigha16]. Methods like CrowdOSInfo, invented by AIT and CDRC, shows that using human crowd intelligence outpaces every deep learning relevance reasoning mechanism. In ANITA, the deep web and dark net analysis research is based on the experience from a national horizon scanning center (CDRC, Cyber Documentation and Research Center), developed by AIT for the Austrian Ministry of Defence. Operating since 2013, the main experience in effective web scanning is that neither manual scanning nor automatic scanning is the best solution. The best solution is a combination of manual scanners, supported by software with artificial intelligence. This experience will contribute to develop very effective scanning capabilities for European LEAs. Another experience from the CDRC is that valuable data comes in very different data formats. It is important to identify different relevant data sources. Even if the data sources have a very different data format, then the usual web sources, it is important to be not restricted by these formats. Google only gets about 5% of all possible data sources, but for investigations it is very important to get all relevant data as fast as possible. Existing solutions for web crawling focus on surface web and social media. To unmask global organised crime networks it is important to use surface web, deep web and dark net data, as well as unusual data sources, like blockchain data from existing and next generation crypto currencies [Jacyncz16]. Therefore, the ambition from ANITA is to include the other sources, like deep web, dark net sources, as well as data from crypto currencies into a system of high performance data gathering and streaming.

Ambition: Having in mind that existing Web investigation software focus on usual web sources from surface web and social media, the ambition of ANITA is to include the usual and the unusual sources from deep web and dark nets, as well as unusual data sources like blockchain information from existing and next generation crypto currencies. To support real crime investigations, it is essential to get this data fast (in near real time), secure and complete, but only if it is relevant. Thus, the main ambition in ANITA is to provide high performance crawling in combination with a secure infrastructure (up to EU secret) and anonymous data collection.

Innovation Potential: The main innovation is to provide a usable data collection service, optimised to the needs for European LEAs. Having in mind, ANITA will be the only one tool used investigation environments, so that the data is collected in near real time. For this, it is necessary to create a new crawler design, based on private and secure cloud solutions. To collect only relevant data, an innovative solution will be created, which will combine human relevance judgement and artificial intelligence to generate the best possible relevance classification.

1.4.4 Adaptive user modelling for cognitive states estimation

There have been several attempts in modelling human behavior in the human-computer interaction literature, such as the Model Human Processor, the GOMS (Goals, Operators, Methods, and Selection rules) or the Human Action Cycle, which are mainly limited to describe how an expert user interacts with the system. Recently, a trend has emerged towards adaptive and customizable tools, which use modelling and reasoning about the domain, the task and the user to offer individualized feedback and support to each potential user. Several studies provide empirical support to the effectiveness of such adaptive interfaces in increasing user's performance and engagement [Trumbly94]. A grand challenge of such adaptive interfaces is therefore to exploit various models to more effectively process input, generate output, and manage the dialog and interaction between human and machine to maximize the efficiency and the effectiveness of the interaction. Although there have been many advances, the level of adaptability in current systems is rather limited and there are many challenges left to be investigated. Latest applications like Google Now (<http://www.google.com/landing/now/>) adapt information or system settings according to particular situations. However, these applications include very limited adaptation features that are implemented in proprietary ways for specific use cases. One central feature of such adaptive systems is the way they make use of the learned knowledge. These systems may suggest information or generate actions that the user can accept or reject. In this

way, the system can receive feedback to update its knowledge base either implicitly or explicitly [Jaimes07]. The system should carry out online learning, in which the knowledge base is continuously updated while an interaction occurs. Because of the online nature of the learning process we'll need to initially advance learning methods and algorithms that can achieve high accuracy and optimal information provided from small training sets. Learning styles models have been suggested as a useful framework for adaptive system development among other sources, such as previous knowledge and user background. Other personalization sources, including background knowledge, intelligent capability, cognitive traits, goals, language and motivation level, can also be used for personalization purposes. Many different internal states of users can be inferred using physiological signals. For example, mental workload has been successfully extracted from scalp electroencephalography (EEG) and pupil dilation. This information has been used to adapt the content or the interface that is being presented to the user and maintain an optimal level of mental workload, resulting in improved performance [Chen04]. The affective states of the user, like positivity of emotions and arousal level, can also be inferred from other physiological measurements, like electrodermal activity (EDA) and electrocardiogram (ECG) [Agrafioti12], which can also improve the interaction.

Ambition: In ANITA we will move beyond state of the art user modelling approaches, by incorporating concepts grounded on theoretical models of human experience. In particular, ANITA will elaborate a unique and advanced model of the user grounded in a number of well-established computational models of perception, cognition and action [Verschure16]. In particular, current user's states (e.g. cognitive effort, attentional allocation, emotional salience) and performance (e.g. time to complete a task, number of items found, etc.) will be inferred from a combination of implicit and explicit primitive features extracted from a range of sensors (e.g. eye tracking, EDR, facial expressions, pupil dilation, user clicks, etc.). These low-level signals will be pre-processed using mono and multivariate analysis methods to obtain metrics (e.g. eye fixation time/frequency, tonic/phasic electrodermal variations, etc.) that are directly associated with the subjective states. The DAC theoretical framework will be responsible for linking the extracted metrics to the higher concepts by bootstrapping meaningful associations between stimuli and responses and, at the highest level, concepts such as the ZPD will be used to control the interaction by providing information that is specifically tailored to the current needs and goals of each particular user.

Innovation Potential: The introduction of an adaptive cognitive user modelling framework in ANITA for observing, monitoring and controlling the information that is presented to the user (in order to maximize his/her ability of data examination and consumption) will inevitably exhibit tremendous innovation potentials for similar systems or extensions in multiple application domains in the Big Data era.

1.4.5 Text Analysis

The three types of knowledge that can be extracted from unstructured text are objective (who, what, where, when), subjective (who has which opinion about what) and meta-knowledge (extra content, e.g. about its author) [Daelemans13]. Relevant tasks on objective knowledge extraction for online trafficking and crime-related content detection and analysis include automated classification [Toti12] and understanding of domain-specific language (including slang, spellings, etc.). On the other hand, stylometric feature belongs to the latter category; in fact meta-knowledge refers to psychological, sociological properties and ultimately to the identity of the author. In the Machine Learning approach, stylometric research is still linked to the idea of automatic text categorization. It is mainly based on maximizing performance, rather than increasing of understanding of meanings and it's necessary a lot of work to configure systems (complete texts versus short ones, weighting and methods, algorithm optimization, etc.) and it needs a large quantity of manually tagged content to be tuned and configured properly [Verhoeven14]. So, many open issues from a research point of view are still open.

Ambition: Within the context of ANITA, the creation of an innovative semantic text analysis component on top of current Expert Systems Cogito® technology can lead to performing advanced automated text analytics tasks to: a) classify the contents under a strongly customized taxonomy; b) extract relevant information like standard entities (people, organization and places) but also domain specific entities (e.g. drug and firearm types) and their mutual relationships aiming to connect the dots; and c) better understand natural language, in particular slang, acronyms, spelling mistakes and abbreviations, enabling a high precision analysis of blogs and social media sites. Additionally, the integration in this textual analysis process of objective and subjective knowledge extraction will enable the extraction of meta-knowledge. Another goal of research will focus on understanding the meanings to have added value of understanding stylometry and types of writers.

Innovation Potential: Within ANITA, innovative semantic algorithms on top of Cogito® technology will be designed to extract and classify data regarding online trafficking, crimes and related entities. Specifically, ANITA will strongly tune NLP semantic algorithms to extract specific slang, acronyms, spelling from sources, to create new taxonomies to categorize and correlate the text with high precision. Deep semantic approach and disambiguation process can improve performance in terms of precision and recall measures, developing an avant-garde approach that will allow a better context aware understanding and a better linking between the events and entities in-

volved. Such deep semantic text analysis will be readily applicable to a very wide range of textual analysis potential cases.

2 Impact

2.1 Expected impacts

2.1.1 Expected impacts set out in the work programme

ANITA addresses all topics set out in the work programme under SEC-12-FCT-2016-2017 – “Technologies for prevention, investigation, and mitigation in the context of fight against crime and terrorism” Sub-topic: Others¹³:

Expected impact	Short term: Improved investigation capabilities
ANITA contribution	In the short term, ANITA aims at improving investigation capabilities of LEAs by delivering a set of tools and techniques to efficiently address online illegal trafficking of counterfeit/falsified medicines, NPS, drugs, and weapons. This is achieved through appropriate knowledge modelling and reasoning services; discovery and monitoring of new and existing online marketplaces; resolving criminal identities in social networks and web and identification of authors in fora and web contents; unmasking of fake information, disinformation and camouflage of the real nature of information; insights on criminal groups relevant and related to trafficking of illegal products; discovery and understanding of trends and behavioural patterns; revealing, tracking, and monitoring of payments and transactions in crypto-currency networks; interoperability with available relevant investigation systems already in place and operation at and for LEAs. This will support the LEAs in <i>more effective investigation activities by using online contents and information obtained under a lawful warrant</i> . Specifically, we aim at supporting LEAs activities with innovative tools that operate in an efficient and easy-to-use way . In addition, ANITA will produce a <i>training and dissemination programme for LEAs</i> with the aim to improve investigators’ capabilities (T11.4) based upon the project results and those currently in use by LEAs. Specifically ad-hoc training sessions will be organized and delivered both via web and face-to-face. Face-to-face trainings will be mainly based on the train-the-trainers approach with the intent to select, involve and train key-figures among LEAs with the aim of providing them with the knowledge, skills, materials and attitudes needed to facilitate the transfer of information in their own organisations. Finally, ANITA will develop a formative path addressing LEAs and representatives from public and private sectors in order to increase awareness and knowledge, but also exchange of good/best practices and experiences. Specifically, we aim at improving LEAs knowledge and operations, in order to enforce the law in an efficient and precise way by simplifying the learning and acting process .
KPIs	<ul style="list-style-type: none"> • New functionalities to tackle online trafficking delivered to the LEA partners ≥ 2 per LEA • Average reduced human effort for legal investigation $> 30\%$
Objectives WPs	O1, O2, O3, O4, O5 WP5, WP6, WP7, WP8, WP11
Expected impact	Medium/Long term: Crimes solved more rapidly reducing societal distress, investigative costs and the impact on victims and their relatives
ANITA contribution	ANITA is structured to automatize numerous investigation tasks that otherwise would be accomplished by a lot of human teams of investigators resulting in exorbitant investigation costs and shortage of investigators to be dedicated to other criminal-fighting activities: <ol style="list-style-type: none"> 1. Detection and monitoring of relevant sources in the surface Web and in the Deep Web, including the Dark Web and the proper management of the monitored sources. 2. Superficial analysis of all identified and extracted contents (e.g. textual filtering, classification, object/person detection/identification in images and videos, language identification in audio and speech-to-text sources). 3. Deep analysis and understanding of relevant contents (e.g. Extraction of evolving knowledge from deep neural network representation; reasoning mechanisms for criminal network reconstruction; implicit and explicit user feedback capturing; adaptive user modelling for cognitive states estimation; Incorporation of conscious and subconscious user feedback in deep learning representation).

¹³ <http://ec.europa.eu/research/participants/portal/desktop/en/opportunities/h2020/topics/sec-12-fct-2016-2017.html>

	As a result, ANITA aims at reducing the costs related to investigations on illegal crimes of online trafficking . Relying on mature and well established technologies provided by the ANITA partners, our approach will make faster the investigation activity, and therefore a faster solving of crimes. This is expected to contribute significantly to reduce the impact on victims and their relatives, as well as reducing the societal distress as a whole (e.g. by reducing effect on victims' emotional or psychological wellbeing, reducing risks and effect on citizens' physical wellbeing and health, increasing the confidence on LEAs thanks to faster crime investigation, resolution and therefore judicial prosecution).
<i>KPIs</i>	<ul style="list-style-type: none"> Increased percentage of closing cases due to unchallengeable evidence > 20% Average reduction of human work/effort in investigations > 25%
<i>Objectives</i> <i>WPs</i>	O1, O2, O3, O4 WP5, WP6, WP7, WP8
Expected impact	Medium/Long term: Prevention of more terrorist and criminal endeavours
<i>ANITA contribution</i>	The ANITA solution will contribute to unmask terrorist-related activities of online illegal trafficking with the aim of supporting preventive analysis and understanding of more terrorist endeavours and behaviours. This is specifically done by the knowledge mining and reasoning services for revealing and tracking terrorist long-term activities reconstruction, linking people and digital identities, terrorist group discovery (based on social relationships analysis as well as activity analysis), detection and tracking of financial transactions in crypto currency systems, identification of new and emerging forms of online trafficking of illegal products. In particular, the automated analysis tools delivered by the project for text analysis, image/video analysis, and audio analysis will be exploited for the detection of financial networks and flows through the Deep and Dark Nets, since most of the fund trafficking activities (the key element for every terrorist organization in order to raise funds for sustaining its operational needs) are performed in the private zone of the Web.
<i>KPIs</i>	<ul style="list-style-type: none"> Increase in success rate of evidence searching for terrorist and criminal endeavours by 30% Supporting automated analysis and monitoring of 10 times more online sources related to terrorist endeavours than currently performed by LEA partners.
<i>Objectives</i> <i>WPs</i>	O1, O2, O3, O6 WP5, WP6, WP7, WP9
Expected impact	Medium/Long term: LEA officers provided with better tools to help them on their (specialized) daily work
<i>ANITA contribution</i>	ANITA will offer to LEA officers a web-based investigation system to process online contents related to online illegal trafficking in order to better support analysis operations and improve <i>investigation capabilities</i> , thus, facilitate their (specialized) daily work. The peculiarity of the ANITA system is represented not only by the innovative technological approach/solutions. In fact, it will be based on a solid inter-disciplinary analysis and knowledge of the criminal phenomena, as well as of the specific needs and requirements of LEAs and judicial authorities. This will allow developing tailored tools, capable of effectively supporting investigators and facilitating their activities (in many cases, technologies used in the prevention/fight against crime still use a generalist approach towards criminal activities and require men to be at their disposal and not vice versa). The automated analysis tools for text analysis, image/video analysis, and audio analysis will be specialized to handle with specific contents, languages, terminology, vocabulary, symbols, pictures, etc. related to the considered crimes. This will increase the effectiveness of the automated and assisted investigations of relevant online trafficking crimes to current practices and systems available in LEAs.
<i>KPIs</i>	<ul style="list-style-type: none"> LEA perceived innovation rate > 60 % Reaching a precision of 70%¹⁴ in the automated unmasking of relevant online contents
<i>Objectives</i> <i>WPs</i>	O2, O3, O6 WP6, WP7, WP9
Expected impact	Medium/Long term: Better identification and understanding of criminal activities
<i>ANITA con-</i>	Through ANITA, we specifically aim at making better (faster and accurate) (wrt. current

¹⁴ Of course it is not easy to measure a recall ratio for such task, as it is not possible on real data to measure the false negative.

<i>tribution</i>	practices of the LEAs partners) the process of investigation of online illegal trafficking activities. Specifically we aim at developing an accurate set of tools responding to an Ethics-by-Design and Data Protection-by-Design strategy . The increased awareness and knowledge of such criminal phenomena, their trends and mostly their dramatic impact on public health will certainly trigger more attention and commitment from LEAs, as well as from judicial authorities. The information extracted from identified relevant contents will be integrated, analysed, aggregated, and correlated to build <i>crime and criminal dossiers</i> (origin, style, threats, activities, previous events, etc.); the effective integration, management and visualization of the knowledge extracted from online contents will improve LEAs' understanding of these criminal activities. The tight collaboration with many LEAs (partners, advisory board, and user community) will generate a comprehensive and completed set of real world and relevant requirements and needs that will support not only the development and delivery of ANITA system and services, but also a better understanding of the requested functionalities and capability on Internet investigations for these crimes. Finally, the report D2.3 on <i>Criminal profiling and possible scenarios. Guidelines for action</i> (which is expected to be public) will also include the list of final operational LEA requirements, while the report D4.2 on ANITA system specifications (which is expected to be confidential, restricted under conditions set out in Model Grant Agreement) will include a more detailed elaboration of such requirements.
<i>KPIs</i>	<ul style="list-style-type: none"> • Evidence gathering, consolidation and profiling for at least 3 criminal activities • Behaviour modelling for at least 3 criminal activities
<i>Objectives</i> <i>WPs</i>	O3, O5, O7, O8 WP2, WP3, WP4, WP7, WP10

2.1.2 Strategic impacts of the project

There is no debate over the importance of investigating and countering online illegal trafficking of counterfeit/falsified medicines, NPS, drugs and firearms. However, the lack of awareness and limited knowledge among LEAs on the online illegal trafficking, the limited availability of technological tools to support online analysis and monitoring activities and the difficulties in establishing effective cooperation between LEAs as well as between LEAs and judicial authorities represent significant obstacles. The solution proposed by ANITA will improve awareness and knowledge among LEAs in order to enhance the investigative and prosecution capabilities. Moreover, it will improve the Internet investigation capabilities and tools to harmonise/standardise the approach and procedures among LEAs as well as increase the exchange of best practices and information sharing among LEAs, thus facilitating also the investigative processes of trans-national (criminal) cases.

ANITA will impact the EU Internal Security Strategy's goal of protecting its citizens and their rights and freedoms of EU citizens and those residing, staying in or visiting the EU (see WP3 and Section 5). Consideration of different legislations, languages and ways of working facilitates a European shared vision, which is also assured by the direct involvement and experience of the partner LEAs in ANITA. Contributions to the development of a comprehensive model for information exchange are also expected, as well as to the professional training of analysts in Europe, facilitating transnational cooperation with increasingly efficient technologies. The active participation of end-users, research and development industry in ANITA will impact reducing the gap between research and market (end-users), addressing one of the main problems faced by the EU Security Industry, according to the Action Plan for an Innovative and Competitive Security Industry¹⁵ (4.2.1 aligning funding programmes, exploiting IPR routes). An optimal use of public funding and defragmented consideration of the EU security scenario will also be targeted by taking into account ongoing and concluded R&D projects, Open Source Initiatives and what is available.

ANITA will contribute to the Treaties of the Union and Charter of Fundamental Rights¹⁶. This structure, complemented by a "privacy-by-design" and ethics awareness approach, clearly contributes to Action 8 in the Action Plan for an Innovative and Competitive Security Industry under the Security Industry Policy, and the Internal Security Strategy. In this way, societal and fundamental rights will be considered throughout the project and enshrined in the conditions of use beyond the project as a covenant of use reflected in the license terms under which the results will be made available. These terms shall reflect the recommendations of legal and ethical experts and those Model Grant Agreement Articles, as applied, related to compliance with national and EU Law, Restriction on Dual Use and Export, and Data Privacy.

¹⁵ <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2012:0417:FIN:EN:PDF>

¹⁶ http://ec.europa.eu/justice/fundamental-rights/charter/index_en.htm

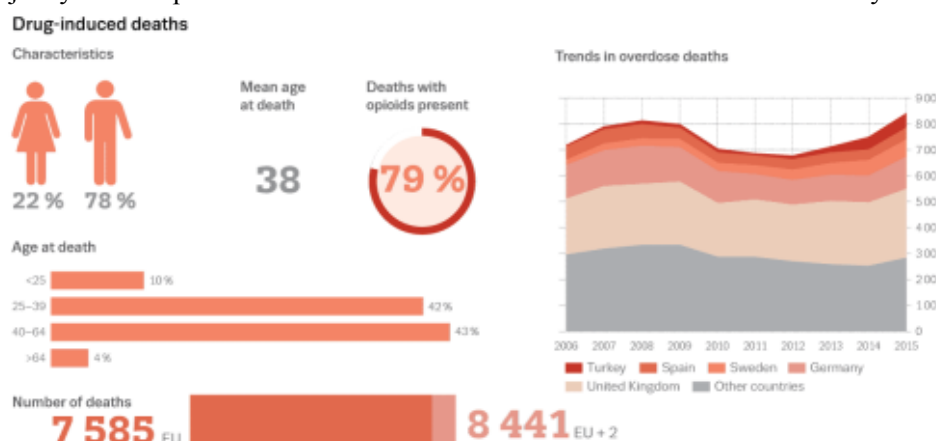
2.1.3 Social impact of ANITA

A large number of sectors of society will benefit from tools and techniques that would tackle online-based activities related to pharmaceutical crimes and illegal trade of NPS, drugs, and firearms. First of all, society at large needs reassurance that such systems and procedures that assist criminal investigations are ‘fit-for-purpose’ and that LEAs across Europe have the necessary skills, competence levels as well as collaborative structures to efficiently and effectively target criminality. Inability to properly address the above problems could cause detrimental effects.

Illegal trafficking of falsified medicines is becoming a serious risk for global public health. Based on a report (“The Counterfeiting Superhighway”) by European Alliance for Access to Safe Medicines (EAASM), 62% of medicines bought online were either substandard or counterfeit, while the World Health Organization (WHO) suggested that over 50% of medicines purchased from online sites concealing their location may be counterfeit. Increased use of counterfeit medicines can have incalculable consequences. As an example, the International Policy Network estimates that globally, 700,000 deaths a year are caused by fake malaria and tuberculosis drugs¹⁷. Seizures involving counterfeit pharmaceuticals rose by 34% in 2015 compared with the previous year; 33% of the shipments captured were of ‘commercial size’, containing more than 1,000 doses. In 2015, INTERPOL’s Operation **Pangea VIII** seized 20.7 million illicit and counterfeit medicines, which is more than twice the amount confiscated during a similar operation in 2013. The same INTERPOL **Pangea VIII** seizure amounted for USD 81 million worth of potentially dangerous medicines, the suspension of 550 online adverts for illicit pharmaceuticals and 2,414 websites taken offline¹⁸. According to the World Trade Organization (WTO), up to 50,000 Internet pharmacies are in operation, around 95% of which don’t comply with the laws and industry standards created to protect patients. Additionally, 50% of websites worldwide that hide their physical address are selling illicit pharmaceuticals, including those labelled with counterfeit trademarks, while 90% of drugs purchased online come from a different country than the website claims. Despite the warnings about the risks, consumers are still willing to buy medicines online. It is worth mentioning that the vast majority of Europeans still know little about fake medicines online and very few – 20% – associated the word counterfeiting with medicines (Sanofi 2015).

The situation is similar in illegal trafficking of drugs and NPSs. Drug abuse continues to be a significant challenge for European societies. The number of drug overdose deaths in Europe has risen for the third consecutive year. A total of 8441 overdose deaths, mainly related to heroin and other opioids, are estimated to have occurred in Europe in 2015, a 6% increase on the estimated 7950 deaths in the 30 countries in 2014 (EMCDDA 2017).

Regarding NPS, EMCDDA monitoring showed that in 2012 the number of online shops offering new psychoactive substances (as well as ‘legal highs’) for sale to consumers in the EU more than doubled to 693 shops compared to 314 in 2011 (source EUROPOL). In 2013, EMCDDA online monitoring identified 651 websites selling ‘legal highs’ to Europeans (source EMCDDA), while by the end of 2016, the EMCDDA was monitoring more than 620 NPS (compared with around 350 in 2013). The rise in trafficking of NPSs can be also justified by the fact that the use of encrypted browsers on the deep web and the use of cryptocurrencies minimize the risk of being caught by LEAs. Firearms are the lifeblood of organised crime in Europe as elsewhere in the world. They facilitate the trade in illegal drugs and trafficking in human beings. The supply of stolen, smuggled and converted replica guns fuels urban gang conflict. The damage caused by criminal use of firearms is direct and indirect. Direct impact includes the number of killings and injuries – over 10,000 in the EU alone over the last decade, in addition to over 4,000 suicides by firearm each year. Indirect impact cannot be quantified, but it is vast: firearms are typically used by all organised crime groups involved in illegal drugs trade and trafficking in human beings to intimidate and coerce their victims (UNODC 2012). The EC Taxation and Customs Union affirms that in 2015, there was an increase in weapons (76%) and ammunition (230%) seizures, amounting to 7300 weapons and 3.2 million ammunition seizures. The health risks for citizens, the increase in the number of drug-related crime, the increased violent



¹⁷<http://www.un.org/africarenewal/magazine/may-2013/counterfeit-drugs-raise-africa%E2%80%99s-temperature>

¹⁸<http://www.interpol.int/News-and-media/News/2015/N2015-082>

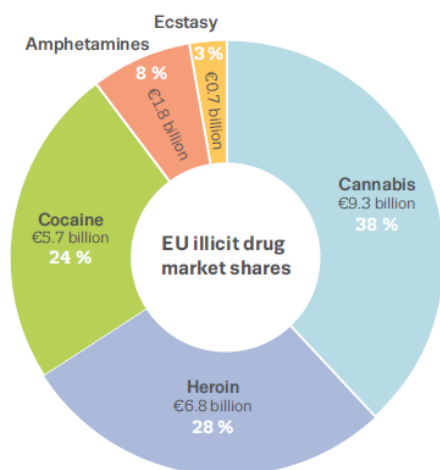
crimes and risk for terrorism attacks due to more availability of illegal weapons result in a higher perception of insecurity among citizens and increased demand for security services from LEAs. By increasing the number of solved crimes and the delay in which they are solved, **ANITA will indirectly contribute to furthering trust in LEAs**. By bringing key stakeholders together to share best practices, improve competency via certified/accredited training programmes as well as fostering collaboration and data exchange across national borders, **ANITA will aid investigations and unearth evidence which will serve to increase the probability that victims of organised crime will see offenders prosecuted** for their crimes. We expect also that the increased effectiveness of ANITA will have a positive impact not only on the actual level of security but also on the psychological feeling of security, resulting in a positive effect on the day to day quality of life of society. In addition, privacy-preserving techniques will be developed within the project to ensure that the exchange of data between EU states will not have a negative impact on the personal privacy. In addition, the inclusion of certification and trainings tasks will make it possible to instruct end-users in what are permitted uses of the developed techniques and those that are not. WP3 will work closely with the external ethical expert to oversee these aspects and incorporate data control reports in deliverables.

2.1.4 Economic impact of ANITA

The first and direct economic impact will come from the **fight against criminal activities**. It is not trivial to measure the economic effects of illegal trafficking of drugs, NPS and firearms. The scope and duration of the effects are often arbitrary and do a poor job at quantifying the non-tangible costs to the person or society. In this subsection, the most notable economic effects in several sectors are reported.

Illegal trafficking of counterfeit medicines affects strongly the public health system. More specifically, the EC estimated projected **costs incurred** to the EU from **falsified medicines** from 2008¹⁹ until 2020 are **€1.8bn to €22bn for hospitalization** as a consequence of treatment using counterfeit medicines, **€93m to €1.1bn avoidable medical**

Estimated minimum retail value of the illicit market for the main drugs in the EU



Note: Percentages do not add up to 100 % due to rounding.
Source: EMCDDA.

treatment at primary healthcare doctors and **€7.65bn to €93bn of non-action based on QALY** (Quality-Adjusted Life Years - combined life expectancy and quality of life). The **Global counterfeit pharmaceuticals** market is rapidly increasing; it is estimated between \$75bn and \$200bn, i.e. 8.5% and 22% of the global pharmaceutical industry market. According to estimates by the European Union Intellectual Property Office (EUIPO), the **EU's pharmaceutical industry** losses **€10.2bn** because of counterfeit drugs. The costs and effort for **licit medicine development** is huge: according to the National Academy of Sciences, the estimated cost of developing a new drug is more than \$1.5bn and the average time for drug discovery and development from target identification to approval is 10-14 years. At the same time, **counterfeit medicines** generate from \$200,000 to \$450,000 per \$1,000 invested – 10/20 times more than heroin and 5/10 times more than counterfeit cigarettes (IRACM). The unfair competition by criminals and lost revenues for licit companies results in direct loss of jobs in the pharmaceutical sector and related sectors (EU experts estimate that **37,700 jobs** have been lost directly as a result of the pharmaceutical counterfeiting racket and a further **53,200** have been lost in other sectors indirectly²⁰).

Regarding trafficking of illicit drugs, the most used ones in Europe are marijuana, heroin and cocaine. Cocaine is the most consumed stimulant drug in Europe, with a retail market value estimated by the European Monitoring Centre for Drugs and Drug Addiction (EMCDDA) to be about €5.7bn a year. The most seized drug in Europe is cannabis with more than 500000 seizures and 600 tons, followed by cocaine with 78000 seizures and 61.6 tons in 2014. The number of seizures for amphetamine in Europe is almost 40000 with approximately 9 tons and for methamphetamine 12000 seizures and 0.9 tons (EMCDDA). The important contribution of the drug trade to illicit financial flows was also highlighted in a recent report, which estimated **that illicit drug markets in the EU accounted for one-quarter of the proceeds from all illicit retail markets**. The illicit drug market absorbs resources from the legal economy: money spent on illicit drugs is denied to governments in taxes and to producers of other consumer goods. Moreover, the market for products or services can be distorted when business and investment decisions are

¹⁹ http://ec.europa.eu/health/files/pharmacos/pharmpack_12_2008/counterfeit-ia_en.pdf

²⁰ <https://www.euractiv.com/section/health-consumers/news/fake-drugs-cost-eu-billions-of-euros-its-health-and-its-jobs/>

made not on a commercial basis but to reduce the risk of detection or facilitate drug-related activities, thus undercutting legitimate businesses operating in the same field and effectively pricing them out of the market²¹.

Through the provision of an advanced set of tools, **ANITA aims to significantly facilitate the investigation process of the analysts, increase the number of resolved cases, and, therefore, lower the cost of investigation. Moreover, by increasing the number of resolved cases, the detrimental effects of illicit trafficking on the economy will be mitigated.**

ANITA will develop original state of the art know-how (and later on skills) in avant-garde data source analysis and blockchain technologies, advanced Big Data analytics, modelling knowledge in human understandable forms and incorporation of human perception/cognition principles in the system processing pipelines. These areas have significant economic potential themselves. The ANITA results will be achieved in the broader area of security and data analysis processing, which is a powerful driver of economic growth (and employment) in today knowledge-based economy.

2.1.5 Improving innovation capacity and the integration of new knowledge

In the domain of the online illegal trafficking in counterfeit/falsified medicines, NPS, drugs, and weapons, ANITA results and knowledge will be able to boost the innovation capacity of LEAs and other relevant stakeholders in related knowledge and legal issues, as well as in LEAs capabilities/tools and approaches for countering online trafficking. ANITA will be able to improve the **current situation** as reported below in detail.

Areas	Current situation	ANITA Innovation Capacity and new knowledge
Related knowledge and legal issues	Fragmentary knowledge base on most vulnerable, trafficked and emerging products; Limited information on OC groups involved and their activities; Lack of reliable statistics and figures and qualitative/ quantitative analysis.	Enhanced inter-disciplinary knowledge on the online illegal trafficking; Systematisation of the sources of information/materials , facilitating fruition by stakeholders; Exploitation of the knowledge base achieved by ANITA to develop a stable collaborative knowledge management process; Implement “friend of rights” tools considering the different users that make use of dark web, also for non-criminal and legitimate reasons and activities (e.g. hacktivism or vigilantism);
	Difficulties in implementing a harmonised approach to regulate the online legal market as well as to fight back against the online illegal market; Fragmentary and diversified legal framework , limited capacity for identification, risk assessment and control; Limited efficacy of the Early Warning Systems and the alert circulation system.	Support to the identification and analysis of the critical issues , elaboration of possible guidelines of reference to support policy makers at EU and national level thus impacting on the harmonisation of legislative initiatives; Adoption of measures that might be taken to prevent abuse such as interference with the private sphere of innocent individuals; risks of discrimination or stigmatisation of individuals/groups suspected as criminals; Increased commitment/opportunities to circulate the Project results among key-stakeholders.
LEAs capabilities/tools and countering approaches	Widespread lack of awareness and limited knowledge among LEAs on the online illegal trafficking in medicines and especially NPS; limited number of LEAs’ specialised cyber-crime units.	Enhancement of both awareness and knowledge of LEAs and other stakeholders through dissemination and training activities (WP11); support from Turin Local Police (N.I.S.T.), Security Science Center of Óbuda University, Polícia Judiciária, Métropole Nice Côte d’Azur, Romanian Directorate General Logistics and Baden-Württemberg Police, Policia Municipal de Madrid.
	Limited availability of technological tools to support online analysis and monitoring activities.	Specialized investigation tools and services developed in WP5-8 ; support LEAs to extract, compare, correlate, filter, reason over and interpret suspect information, data, communications stored or transferred (in Deep Web/ Dark Nets); Improved capacities for sampling, analysing, evaluating, interpreting and recording evidence for the court.

²¹ http://www.emcdda.europa.eu/publications/eu-drug-markets/2016/online/ramifications/legal-business-and-economy_en

<p>Limited adoption of the multi-agencies approach at a national level; poor close communication between law enforcement officials and public health employees;</p> <p>Difficulties in establishing effective cooperation between LEAs as well as between LEAs and judicial authorities.</p>	<p>Community building (Task 11.2); Creation of a positive, collaborative and operative environment to promote cooperation between LEAs, judicial authorities, public and private sectors and the organisations giving external support to the project.</p>
<p>Limited cooperation between LEAs and the private sector; proliferation of initiatives at public and private level, which are often not coordinated.</p>	<p>Increased opportunities for mutual understanding, networking and cooperation thanks to the initiatives included in the Community building activity (Task 11.2) and in particular to the Cooperation platform.</p>
<p>Limited/sometimes misleading mass-media representation of criminal phenomena, which is likely to generate contradictory convictions among citizens.</p>	<p>Substantial increased visibility, awareness and education thanks to the visibility and dissemination actions (WP11); involvement of a wide range of organisations already active globally in spreading the word on the phenomena of online illegal trafficking and their related risks.</p>

Innovation within ANITA will be pursued not only through the consortium's efforts and success in researching and developing new products and services, but also through demonstration, piloting, proof of-concept and validation activities operating close to all end-users and the market. Special attention will also be given to the widest possible use of knowledge generated by the supported activities up to the commercial exploitation of that knowledge and accelerate the deployment and diffusion of innovative products and services into the market. Through the geographical dispersion of ANITA consortium, a transnational research and innovation network is already established and will be further expanded throughout the project, and will integrate the knowledge triangle - research, researcher training and innovation in areas of key societal challenges and promote international cooperation with non-EU countries.

2.1.6 Possible barriers

A number of external factors that can affect the achievement of the proposed impacts have been identified below:

Barrier	Mitigation Plan
Heterogeneous Legal & regulatory EU frameworks	Before the design of the tools and of the overall ANITA platform is finished, a complete assessment of the compliance with current (and legislation due to come into force during the project) EU and national laws will be assessed under WP2 and WP3 in order to comply with all heterogeneous legal and regulatory frameworks across Europe.
Interoperability with legacy systems	Interoperability spans both the interoperability of data types as well as data model, but also how the solution will interface allowing for simple integration with existing systems. ANITA addresses and analyses existing legacy systems to provide solutions based on multiple data formats without changing existing data models and interoperability (through APIs) with existing systems of LEAs.
Current LEA culture and approach to collaboration	The LEAs that are taking part in ANITA as full partners have successfully collaborated among them and with the technical partners of the consortium in similar past and/or ongoing research and development projects. These previous experiences have already smoothed the path but, in order to boost multidisciplinary collaboration of all partners even more, the project includes frequent meetings and workshops with users throughout the project lifecycle.
Lack of human resources on behalf of LEAs	Due to the heterogeneity of the size and resource of LEAs (even departments belonging to the same Agency), not all LEAs have the computing resources to mount full-scale intelligence operations; or alternatively the existing (distributed) computing re-sources are saturated with the support of daily tasks. In this view, ANITA will provide a cloud/web-based solution, available to all involved agencies through strict security mechanisms. This hardware and software infrastructure will be able to support sophisticated intelligence and investigation operations supported by the ANITA tools.
Refusal on behalf of the end user to leverage the full potential of the ANITA toolset	Many of the end-users (LEA Officers) have either limited IT background or are reluctant to adopt new tools and practices. This is partially justified due to the typically heavy schedule and the heterogeneous tasks they perform. In order to confront this challenge, we plan to pay special attention through an entire work package (WP9 Integration of human factor in the analysis loop) in order to design and adapt services and User Interface to the human factor of the end user, and not the user that need to adapt to the proposed interface.
Societal acceptance	Data extraction and analysis along with video surveillance and analysis often meet scepticism or even reaction from the society due to privacy protection aspects. ANITA fully respects such con-

	cerns and has dedicated ethical and privacy (WP3) – related activities in order to verify societal acceptance for the resulting data and video analysis solution, and the make sure ANITA solution abide by the actual laws in terms of privacy and data protection.
--	--

2.2 Measures to maximise impact

The dissemination of the project results and the exploitation of the achievements of ANITA are the key elements for the success of ANITA and the main activities to maximize the expected impacts. Dissemination and exploitation of results is grounded in linkages to the LEAs, security agencies, international supply chain movements organizations, universities, standardisation bodies, policy-makers, and end-users in the field.

Special attention will be also given to the widest possible use of knowledge generated throughout the project up to concrete exploitation of that knowledge to accelerate the deployment and diffusion of innovative products and services into the market. Through the geographical dispersion of ANITA consortium, a transnational research and innovation network is already established and will be further expanded during the project lifetime, and will integrate the knowledge triangle - research, researcher training and innovation - in the key area of Fight against Crime and Terrorism and promote international cooperation with non-EU countries.

2.2.1 Dissemination and Communication Plan

ANITA, in order to fulfil successfully its dissemination and exploitation goals, sets up a comprehensive dissemination and communication plan, which will be proactively implemented from the early stages of the project (D11.2, M6). The plan will provide exhaustive details on the following key dissemination aspects: (a) *target groups of dissemination*, (b) *dissemination content/materials*, (c) *dissemination measures/channels*, (d) *project events and participation in events*, and (e) *partner specific dissemination plans*. Here we provide the initial draft of this plan that will be updated during the first months of the project.

2.2.1.1 Dissemination Strategy

The ANITA consortium is targeting all the groups of stakeholders that are likely to affect or to be affected by the project in relevant sectors, such as law enforcement and forensics, the pharmaceutical industries and security as well as ICT and web-related industry. ANITA will follow a phased approach to define, plan, organize and exploit communication channels and dissemination events:

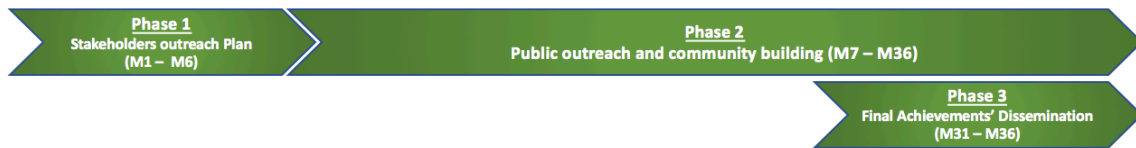


Fig. 4 : Dissemination and Communication Logic Plan

Phase 1 – Stakeholders outreach plan:

- **Project WEB Site Set-up:** the ANITA WEB site will be set up as soon as the project is started. It will describe the project, the challenges and the goals, project members and will provide information on news/conferences/workshops, etc. The website will also include a reserved area for internal documentation exchange/versioning and User space.
- **Community established:** initiate community building and inform the market regarding the scope and the objectives of ANITA. The target groups for the dissemination activities, beyond the members of the Advisory Board, are the End Users, mainly LEAs – especially counter terrorism and departments fighting organised crime – and other potential customers (policy and decision makers).
- **Attraction of Target Audience:** press releases will be prepared to inform the public and relevant audience of the launch of the ANITA project, specific objectives and challenges, involved partners and project plans.
- **Creation of an ANITA “brand identity”:** ANITA logo, templates for project documents and presentations.
- **Project Information Material:** leaflet, brochures (paper and electronic versions) and introductory presentation.

Phase 2 – Public outreach and community building:

- **Project Information material:** Newsletter, press releases, presentations, leaflets, brochures and videos will be updated with achieved results to keep the audience informed and increase the interest. All project information material will be maintained in the ANITA WEB Site, including public deliverables.
- **Scientific papers and articles:** Scientific papers will be written and submitted to conferences and workshops on specific ANITA topics to further raise awareness among the scientific and industrial community. Scientific papers and articles will also promote possible follow-up research activities.
- **Conferences and Workshops:** Conferences and Workshops will be organized during the course of the project to

disseminate the ANITA initiative goal and results to the widest audience possible. Conferences and workshops will increase the interest, promote the discussion on ANITA topics and provide feedbacks.

- **Social Media & Blog:** Social media (Twitter/LinkedIn) and the relevant Blog communities will be frequently posted with ANITA news and information material to spread information and increase interest and audience.

Phase 3 – Final Achievements’ Dissemination:

- **Promotion of Project results.** Final presentations, videos, newsletter and press releases will be prepared promoting ANITA achievements, results and measured KPIs.
- **Exploitation Activities:** the adoption of ANITA approaches and technical solutions will be encouraged, recommendations and guidelines will be provided for follow-up activities. A final ANITA workshop will be held at completion of the project focusing on exploitation of results.

The aim of ANITA communication activities is not only to provide individuals interested or concerned with the topics addressed by the project (pharmaceutical crime, intelligence, law enforcement fighting organised crime, etc.) with information about the results of ANITA research, but also to **spread information** among the general public, addressing those segments of population which may have been completely unaware of the topics and the related results. In order to reinforce the capability and the opportunities to ensure proper visibility and dissemination, ANITA can also rely on the external support (as stated in the official supporting letters) of key organisations which are active at a global level. In particular, a major role will be played by: Provincial Police Headquarters in GDANSK (Poland), Belgian Road Safety Institute, Academy of Criminalistic and Police Studies - KRIMINALISTICKO-POLICIJSKA AKADEMIJA (Serbia), Home Office (UK), General Directorate Combating Organized Crime, Ministry of Interior (Bulgaria), and Local Police Voorkempen (Belgium). The main communication target groups for ANITA include the **EC community**, with the aim to raise awareness regarding the goals and results of ANITA project and to foster collaboration with other projects sharing similar or complementary objectives, **practitioners and scientific community**, with the purpose to spread project results and to receive feedback from field experts, researchers, academic and other relevant stakeholders in related research areas, **general Public and end users**, with the goal to raise overall awareness of project objectives and potential benefits. The latter target group includes the potential end users of products/services based on ANITA technologies that are its main target audience. Online mainstream media such as online newspapers, blogs, video platforms and social networks will also be targeted for reaching high volumes of General public. A dedicated section in the **project website** will target dissemination of project activities and highlight societal impacts.

2.2.1.2 Dissemination and communication actions

Dissemination Actions	KPIs
Peer-reviewed publications: Among journals targeted by ANITA it is worth mentioning: Interdisciplinary scientific journals on addictions (e.g. Addiction Journal etc.); Medico-legal journals (e.g. Journal of Forensic and Legal Medicine; Medico-Legal Journal), criminological journals (e.g. International Journal of Cyber Criminology), Cyber/Internet journals (e.g. Cyberpsychology, Behavior, and Social Networking; International Journal on Cyber-Security and Digital Forensics; Journal on Cyber Policy; ICT law Journal and magazines; International Journal on Cyber Behavior, Psychology and Learning). Neuroscience Journals: Journal Neuron, Journal Cognition, Cognitive Processing (International Quarterly of Cognitive Science); Journals of Experimental Psychology: Human perception and performance, Journal of Human Computer Studies Cognitive Science; Engineering, artificial intelligence and Media related Journals: IEEE Trans. On Pattern analysis and Machine Intelligence, International Journal of Computer Vision, IEEE Transactions on Information Forensics and Security. Among conferences targeted by ANITA: IEEE International Conference on Computer Vision, IEEE Conference on Computer Vision and Pattern Recognition, International Conf. on Forensic Science – Criminalistics Research, SPIE International Conf. on Multimedia Storage and Archiving Systems, International Conference on Imaging for Crime Detection and Prevention. Other local conferences include Archibald Reiss Days and Traditional International conference at Tara Mountain (Serbia).	- <i>At least 24 peer-reviewed publications:</i> - <i>12 peer-reviewed journal papers</i> - <i>12 peer-reviewed conference papers</i> - <i>Organization of 6 round tables for discussing on the regional or national level results of the project</i>
Targeted articles: Publication of news articles about ANITA as well as features, interviews, articles about ANITA on sector magazines and popular online magazines such Security Magazine, Wired, IEEE Security & Privacy Magazine. IBEC will seek publication in international popular science magazines: New Scientist, Scientific American, Spektrum der Wissenschaft.	- <i>At least 6 articles.</i> - <i>At least 1 newsletter every 3 months</i> - <i>At least 6 Pamphlets’ from the roundtables</i>

Dissemination Actions	KPIs
A set of promotional material will be produced (using a common project logo and standard graphic presentation) which includes leaflets, brochures and DVDs. Moreover, specific press releases will be realised to target General Press, following the latest requirements of the European Commission.	- <i>Project logo, leaflet and brochure</i> - <i>At least 8 press releases</i>
Setting-up of a dedicated web site , to be considered as a public dissemination and communication channel, that will host news content, details of events as well as public deliverables (that can be disseminated without violating the IP and © of ANITA). All partners are committed to increase the visibility of the project web site and will do so using social media and their individual networks.	- <i>Web site at M3; at least 2500 visitors.</i> - <i>At least 200 organizations receive the e-newsletter</i>
Presentation of ANITA project results at conferences, seminars and workshops targeting both the industry, users and academic communities; some of the international conferences and actors targeted by these ANITA presentations could be: International Conference on Terrorism and Organised Crime, ASIS European Security Conference & Exhibition, EMCDDA organised and promoted events, ITU Forum, IEEE World Forum, International conferences on knowledge processing and extraction, International conferences on data mining and big data analytics, Initiatives and events organized by the organization members of the Advisory Board supporting the project, such as ASOP EU, ASOP Global, IFPMA, IRACM, ALIMIS. Furthermore, ANITA will be presented also during the initiatives organised by project end users. IBEC Organizes the Barcelona Cognition, Brain and Technology school (BCBT) every year (http://bcbt.upf.edu); JADS actively contributes to the organization of ICSOC (International Conference on Service-Oriented Computing). Next to this, the SummerSOC summer-school could provide in dissemination of the project to prospective data students.	- <i>Presentation of results at least in 8 international events</i> - <i>Distribution of at least 200 leaflets in such international events</i>
Presentation of ANITA project results at business sector fairs and events as well as performing events . Piggy-backing on events where partners already participate will be leveraged. Indicatively ANITA could be presented at the following business events: MILIPOL, LT-Innovate Summit.	- <i>Presentation of results at least in 6 events</i> - <i>Distribution of at least 200 leaflets</i>
ANITA events : 1 final conference, 1 international workshop (WS) for LEAs, 3 national training WSs (addressed to LEAs + policy makers) and 3 national WSs (informative national workshops for policy makers, public authorities, media, academics/researchers, practitioners, health/drug social services...) will be realised by ANITA partners during the project. The WSs will be realised in Italy, Poland, Spain, Austria, Serbia, France, UK, Netherlands, Bulgaria, Greece, and Belgium involving main stakeholders and End-users.	- <i>1 international WS (~75 attendees)</i> - <i>3 national training WSs (~30 attendees)</i> - <i>3 national WSs (~ 50 attendees)</i>
Dissemination of ANITA through Social Networks channels to communicate concepts and results in social networks such as Twitter and LinkedIn. Moreover, videos will be uploaded in YouTube. Also blogs related to security will be targeted in dissemination actions. Social media channels of SPECS group at IBEC: twitter (https://twitter.com/SPECS_lab), Youtube channel (https://www.youtube.com/user/SpecsUPF), website (http://ibecbarcelona.eu/specs). JADS will use its twitter (@JAdatascience) and YouTube channel (Jheronimus) to communicate results, events, articles and publications regarding ANITA.	- <i>Creation of Facebook and LinkedIn Group, Twitter groups. At least 1000 followers/likes.</i> - <i>Active Participation in at least 2 relevant Blogs</i>

2.2.2 Exploitation Plan

2.2.2.1 ANITA Exploitation Strategy

Exploitation is seen as a key enabler for the success of the ANITA project. Hence all partners are aware of, and are committed to, the exploitation of the project's results. It is a paramount principal in exploitation activities to make use of the results of the project to create value within the participating organizations and thus to improve their competitive advantages. The ANITA exploitation strategy will be composed of three building blocks:

Building Block 1: Initial Market Analysis and Business Model(s)

- **Market analysis:** The potential market will be reviewed and updated in the light of the project's objectives and intended outputs, as well as in the light of latest available market data. The market analysis will define, characterize and segment the potential opportunities for different solutions developed in the project, as well as describe the value chains serving each technology.
- **Business model(s):** preliminary business models will be reviewed and consolidated taking into account the outcomes of the market analysis, the identified value chains and the intended customers for each developed prod-

uct. The business model(s) will map how ANITA will create and deliver value by identifying: (a) the relevant customers segments; (b) the value proposition; (c) the channels to deliver the value proposition to customers; (d) the resources required, (e) the activities to be performed and (f) the required partnership. The elements above will be linked to revenue streams, which will be further quantified in the business plan.

Building Block 2: IPR Management

- **IP Management Plan:** According to the specifications of Horizon 2020 a specific IP Management Plan will be elaborated for ANITA project in order to protect the property of the project's background and results. The IPR management plan will also manage IPR and ownership distribution among project partners.
- **Final IP Protection:** Activities and methodology defined in the IP Management Plan will be fine-tuned and agreed among partners towards the end of the project when final products have been generated.

Building block 3: Strategy for market take-up and long-term sustainability

- **Definition of a Marketing Plan:** The marketing strategy will be elaborated in the marketing plan e.g.; in terms of price, promotion, distribution approaches etc. Taking into account final analysis and synthesis of project' exploitable assets, the correct positioning for the products will be defined in this phase, based on their capability to satisfy user needs vis-à-vis the offer of competitors. Based on this assessment, coherent positioning choices (branding, product description) will be performed too.
- **Standardization Roadmap:** A standardization roadmap will be defined according to the standardization needs identified in the previous phase. This roadmap will define the timescales for the standardization actions, and will identify the relevant standardization bodies that should be approached.
- **Business Plan:** A business plan shall be defined including identified costs, the market analysis and the refined business models. Special attention shall be paid to value proposition, business models, target markets, timescales, competitors and risks assessment and estimated costs to assess quantitatively the financial feasibility of ANITA quantifying initial CAPEX and OPEX vs. generated revenues through the time, based on realistic estimations of price and penetration. Different scenarios will be developed for the business plans (optimistic, realistic and pessimistic).

2.2.2.2 The ANITA Joint Business Plan

ANITA partners have carried out an initial **business plan for the joint exploitation of the system:**

Step 1 – mission: to provide **an innovative knowledge-based user-centred cognitive Investigation system** for analysing heterogeneous (text, audio, video and image) online (Surface Web, Deep Web, Dark Nets) and offline (LEAs' databases) resources for fighting illegal trafficking.

Step 2 – consortium background: The ANITA consortium inherits significant technological and marketing/business background from the companies that establish it. ENG, the coordinator of ANITA, will lead the transfer to the market, as it is an industrial partner with strong expertise in the field.

Step 3 – product description – competitive advantages: at the proposal stage, the ANITA consortium has identified at least 15 exploitable assets (derived from key technologies – Section 1.3.4), which will be further evaluated during the project under exploitation task (T11.5).

Step 4 – market analysis: ANITA will evaluate the relevant market (IT solutions for big data analytics, deep web/dark net crawling, knowledge representation, use modelling) and analyse the competitors during the project lifetime in the framework of WP11. However, a first **SWOT (Strengths Weaknesses Opportunities Threats) analysis** has already been attempted. This analysis will be refined based on thorough investigations in T11.5. **Step 5 –**

Strategy for market take-up: Some of the most critical tasks to be carried out in the project are related to the joint-exploitation and market take-up and the deployment plan for ANITA, once the project has finished. WP11 (task 11.5) will define the roadmap for the future exploitation of the ANITA Framework and its services: which service models will be delivered first, which target end users will be addressed during the initial stage of the project, which technical parts will be prepared first, etc. Different scenarios will be drafted depending on the amount of investment potentially obtained. The time to market could be divided into two separated phases:

PHASE 1. This stage consists in a continuous testing and validation of the ANITA services, by replicating among other stakeholders across Europe (e.g. through pilots) and the new user groups identified, with the objective of having as many reference implementations in different domains as possible.

- Continue with reference implementation phase and wide spreading the knowledge among European stakeholders (months 0-6 after the end of the project).
- Look for additional investment for a global deployment of ANITA. Time to market of the procedures to be implemented – 6 months after the end of the project.

PHASE 2. System ready to be widely implemented in the market.

- Implementation of the market and legal procedures validated in phase 1 by the intermediary-users participating in the project and other stakeholders (months 6-12 after the end of the project).
- Creation and packaging of ANITA specific new products and services following the feedback and recommendations from end users. Time to market of the system – 12 months after the end of the project.

Step 6 – Financial planning: Given that ANITA project will lead to a working prototype, a significant part of the design and development will be accomplished through EU funding. Nevertheless, additional funding for the commercialisation of ANITA system will be necessary. ANITA partners will seek funding through EU-initiated activities, venture capitals and networks of business angels or the European Investment bank. They have reached this decision during proposal preparation time, based on the fact that, when it comes to accessing private finance, Europe lags well behind the US regarding venture capital, according to the Innovation Union Competitiveness 2011 report. Early stage venture capital funds in the EU are at less than half the level in the US.

Strengths <ul style="list-style-type: none"> • Main actors of the value-chain are represented in the consortium: LEAs, private security companies, technology providers, RTOs; • World-class multidisciplinary partnership covering all necessary technical, business and sectorial aspects for success; • Clear vision for architecture and mature technology assets already available as platform building blocks; • Innovative approach for integrating big data analytics, blockchain technologies, knowledge modelling and user modelling. 	Weaknesses <ul style="list-style-type: none"> • Non-homogeneous regulation on privacy content protection at European level; • Complexity of a pan-European consortium; • Growing market with a significant percentage of the target market not aware of the ANITA existence; • Lack of thoroughly tested proof that a full and multi-faceted network of the ANITA building blocks can address the problems of on-line illegal trafficking.
Opportunities <ul style="list-style-type: none"> • Inexistence of such an approach able to combine the connectionist learning paradigm with the symbolic learning one and to integrate the practitioners “in-the-loop” of the analysis process; • Market demand on applications to tackle on-line illegal trafficking; • Big Data analytics technologies easy to transfer to other sectors; • Create best practices 	Threats <ul style="list-style-type: none"> • Movements of major market players in particular from the US, Israel, and Japan; • Being the pioneer increases risks due to market uncertainty level; • Complex bureaucratic procedures in the adoption of ANITA solutions; • Refusal of personnel to leverage the full potential of the project.

2.2.3 Individual Dissemination and Exploitation Plans

ENG	
Dissemination	Exploitation
ENG will leverage on its expertise and consolidated network to disseminate project results through the following channels: i) the ENG official web site (http://www.eng.it); ii) the official ENG web magazine (http://www.ingenium-magazine.it/en/); iii) the official Internal Blog “IN-SIDE”; and iv) ENG social networking activities (ENG LinkedIn Account and ENG Twitter Account).	The ENG R&D Lab of Intelligent Systems and Social Software has the mission of developing new solutions and promoting research outcomes to business units. The results of this project will be integrated in current ENG’s offerings or in new solutions to convey them to several existing (and potential) customers.
CERTH	
Dissemination	Exploitation
CERTH is focused on research and dissemination by publishing results in well-known and widely read international venues, performing presentations in international scientific workshops and exhibitions, participating in seminars and talks organised for specialised audiences and publishing on Web and social media channels.	The research group of CERTH will carefully study the ANITA project results and investigate the technologies that will help to strengthen D-cube’s (CERTH’s spin-off company targeting the development of immersive solutions) products and to give added value to its products.
RISSC	
Dissemination	Exploitation
RISSC will disseminate the project results to the scientific community as well as the LEAs community, since it has an	RISSC will exploit the ANITA project results mostly to elaborate new issues to be further investigated by

extensive network of contacts at EU level. Additionally, RISSC will disseminate the project results during workshops and seminars, trainings, conferences and journals. Also Web and social channels will be used.	researchers. RISSC will also use the results to publish reports and articles, as well as to develop initiatives in the field of trainings and awareness-raising activities addressed to key stakeholders.
EXP	
Dissemination	Exploitation
EXP dissemination plan will involve: i) Invitation to project events of government institutions interested in security issues; ii) Dissemination of project related information via social media channels (EXP Twitter and LinkedIn accounts); iii) Dissemination of project results in related conferences and events.	By collaborating with other ANITA partners, EXP will be able to develop new algorithms to identify and extract metadata connected to the project domain in its product COGITO®, i.e. an advanced semantic text analysis engine.
AIT	
Dissemination	Exploitation
AIT will disseminate the project results in prestigious scientific platforms, like the European Foresight Platform with more than 4000 members, European OSINT and Public Safety Communication Europe (PSCE Europe).	AIT will use the project results to support public organisations directly with knowledge and new methods or to initiate spin-offs for a more intense support. As operator of the European Foresight Platform (EFP), with more than 4000 members, AIT is prepared to distribute the research results from ANITA with policy briefs into this high level community.
IBEC	
Dissemination	Exploitation
The SPECS group at IBEC will focus on: i) Scientific Publications; ii) Scientific events, specifically the Barcelona Cognition, Brain and Technology school (BCBT) organized every year (http://bcbt.upf.edu); iii) Social media channels: twitter (https://twitter.com/SPECS_lab) and YouTube (www.youtube.com/user/SpecsUPF); iv) Dissemination on website (http://ibecbarcelona.eu/specs); and iv) Dissemination in general medial (magazines, radio, tv).	IBEC plans to exploit the results through the creation of a spin off or the transfer through licences to a company. Initially IBEC will work with industrial partners in the consortium in the transfers of the new methods and technology for the terrorism case. IBEC will also explore how to apply the technology in other sectors.
IIP	
Dissemination	Exploitation
IIP will support the following activities: i) organization of workshops, events, roundtables and conferences for the dissemination of the project's results; ii) spreading of results to the wide public (through website, articles, etc.); exploitation of Web media (Web channel, LinkedIn group and Twitter).	The main purpose of IIP exploitation plan is to create contents addressed to its different stakeholders with whom it has already a relationship of trust and confidence, activating a worldwide network that can spread the word about the project.
SYSTRAN	
Dissemination	Exploitation
SYSTRAN's dissemination plan will be based on: i) disseminate the project's achievements through scientific publications; ii) communicate the results at business oriented events and conferences; iii) the project will be promoted at specific communication events organized by SYSTRAN, such as the SYSTRAN Community days.	SYSTRAN will exploit the ANITA outcomes into its current translation products, including SYSTRAN Enterprise Server (SES). The research and development carried out in ANITA will help SYSTRAN to improve its translation products and models, especially in the Defence and Security domains.
TIU-JADS	
Dissemination	Exploitation
Through the 'JADS Crime Room', JADS will perform, organize and assist in the following activities: i) Publishing articles in academic journals and conferences; ii) Academic and non-academic sectorial seminars; iii) Bringing LEA's and relevant governments and organisations together in the JADS Crime Room; iv) updated about the project through the JADS intranet; v) The JADS Twitter and	Through ANITA, JADS hopes to use the EU cooperation to professionalize and intensify the analytic and predictive capabilities of the Crime Room cooperation in Noord-Brabant. Alternatively, JADS has developed a large network of data-startups and spin-offs, which can be included in the results of the project and create a long-term of the ANITA-project in the region of

YouTube will update the public.		Noord-Brabant.	
DITSS			
Dissemination		Exploitation	
DITSS will maximise the ANITA results via an active dissemination strategy: i) Support the creation of dissemination material; ii) Be actively involved in ANITA Community Building activities; iii) Organize a Dutch ANITA related event; iv) Supporting social media coverage as well as press publications; v) Publish articles; vi) Organize knowledge exchange in the Netherlands.		DITSS organizes and facilitates however on a regular bases Fieldlab projects and LivingLab environments, where our network partners develop and test new innovative products & services. Together with JADS we will facilitate and operate a so called 'Crime Room' and use it for dissemination, validation and exploitation goals within ANITA.	
ISBR			
Dissemination		Exploitation	
ISBR will implement the following communication aspects: i) Determining the content of the communication in relation to the target groups; ii) Preparation of materials to be used for distribution (e.g. seminar, journal article, blog, E-newsletter); iii) Planning of project events; iv) Participation in external events; iv) Involvement of external partners (e.g. Belgian Circle of Police Leaders, Center for Policing & Security (formerly known as the Center for Police Studies).		ISBR will aim at overcoming current obstacles in the field (e.g. that society is reluctant to accept new approaches, products and technologies) by organizing POC for targeted police to demonstrate the added value for police purposes. ISBR can therefore rely on its privileged relationships with government, federal and local police forces. ISBR will bring together public and private partners to set the agenda for Future Innovation and Security Strategy.	
KWPG			
Dissemination		Exploitation	
KWPG will realize: i) Promoting knowledge obtained during conferences and targeted workshops; ii) Fostering internal dissemination; iii) Publication in Police magazines; iv) Dissemination through Website (www.pomorska.policja.gov.pl) and social networks (such as Facebook); v) Establishing a link with similar broader initiatives; and vi) Demonstration of results to visitors from other EU countries.		KWPG is interested in adopting ANITA project results to improve its operations, in particular its fight against online trafficking of counterfeit/falsified medicines, NPS, drugs, and weapons. In addition, the Provincial Police Headquarters in Gdańsk will collaborate with project partners to test the platform and propose it in its national and international networks.	
AoC			
Dissemination		Exploitation	
AoC will make active use of the following channels: i) web site (http://www.kpa.edu.rs); ii) the official AoC web magazine and MOIs official magazine Bezbednost and; iii) the official social networks profiles of AoC (Facebook, LinkedIn, Twitter); iv) organisation of events focused on stakeholders' enlargement; v) the design and creation of dissemination material.		AoC and partner subjects could organize different trainings online via web (through webinar/web meetings and e-learning platform) or on the ground (face to face) with LEAs. Strategic analysis of identified and extracted contents would be taught through advanced trainings and the focus groups should be senior officers.	
CAST			
Dissemination		Exploitation	
CAST will ensure that the results from ANITA will be disseminated to law enforcement partners, through bulletins that notify stakeholders of progress and outputs, as well as direct communications with law enforcement partners. CAST will host a workshop that targets specific stakeholders interested in ANITA results. In addition, we will use the ENLETS network to communicate project results to LEAs in all 28 member states.		CAST will work with commercial partners to ensure that the solutions are fit for purpose and operationally viable. Our remit will be to provide the consortium with a link to operational capability and procedures which will ensure maximum exploitation opportunities. CAST will participate in developing the curriculum to ensure that it meets with LEA needs thereby maximising exploitation opportunities.	
NPN			
Dissemination		Exploitation	
NPN will disseminate and promote project activities and results to targeted LEAs (such as INTERPOL and EUROPOL) and in organisations where it is member (such as ENLETS and DITSS). NPN will help promote ANITA		NPN is interested in testing and possibly adopting ANITA project results which can improve its operations, in particular its fight against online trafficking of counterfeit/falsified medicines, NPS, drugs, and	

within its organisation and other related partners to encourage participation to ANITA workshop & training events.	weapons. In addition, NPN will collaborate with DITSS, JADS and partners to test the platform and propose it in the national and international networks.
GDCOC	
Dissemination	Exploitation
GDCOC will disseminate the project results through its website, professional journal and participation to workshops for law enforcement and security actors. Additionally, due to its connections to other judicial institutions – prosecutor office, other LEA bodies, etc., the knowledge on project results will be transferred to other relevant actors.	GDCOC as a LEA is interested in testing and possibly adopting ANITA project results which can improve its operations, in particular its fight against online trafficking of counterfeit/falsified medicines, NPS, drugs, and weapons.
LPV	
Dissemination	Exploitation
LPV activities will include: i) Present ANITA to Belgian LEAs; ii) Present ANITA to the Ministry of Interior, the Federal Police; iii) Present the project during the annual ASTRID-user days and the INFOPOL conference; iv) Demonstrate NEWTON-HiPeC results to public and private companies in the Security domain; v) Support organisation of ANITA Workshops; vi) Disseminate ANITA via LPV web-site and LPV social media channels.	LPV will feed the findings of the project to its extensive end user network to provide direct contact with the LEA end users with direct interest in the field of ANITA.

2.2.4 IPR, Knowledge protection, software licenses and privacy

2.2.4.1 Knowledge management and IPR

ANITA is expected to produce significant research outcomes and technological innovations. Therefore, special attention will be given to knowledge management and protection issues from the beginning, and during the whole lifetime of the project. All details regarding management and protection of knowledge created within ANITA will be specified in the Consortium Agreement (CA), following well-known models, such as the DESCA Horizon 2020 document. The Consortium Agreement will address (a) confidentiality of the information disclosed by partners during the project, ownership of results resulting from the execution of the project, (b) legal protection of results resulting from the execution of the project through patent rights, (c) commercial utilisation of results, also taking into account joint ownership of the results, (d) patents, know-how and information related to the use of knowledge owned by one of the partners, resulting from work carried out prior to the agreement and (e) sublicenses to third parties within clearly defined limits. The cited Consortium Agreement will implement the Grant Agreement defining the distribution of the EU funds and the Intellectual Property. The principle basis of the information and know-how exchange will be based on free access right. The general outline of the IP rules is as follows:

- Pre-existing partners' knowledge (background) will be specified in the Consortium Agreement.
- Knowledge that is generated within the framework of ANITA (foreground) shall remain the property of the partner that generated it. If more than one partner generates that knowledge and it is not possible to separate their contributions, the knowledge will be jointly owned.
- Access rights to knowledge that is needed by a partner for the execution of its part in ANITA shall be granted to the partner on a royalty-free non-transferable basis, unless otherwise agreed before the signature of the CA.
- A partner will not publish any knowledge provided by another partner and identified as confidential, without the other Partner's prior written approval. However, if open source software licenses apply, the open source software license rules will apply for publishing knowledge.
- To meet the need of both industrial partners with commercial and IP interests and research partners in the project, which have a major role in ensuring results are widely, disseminated, dissemination assets will be submitted to the Project Coordinator and the Dissemination Leader and distributed to the other relevant partners who may object within a small-time period (no more than a month) which is agreed upon in the CA. Otherwise the dissemination may proceed.

To ensure smooth development of software components, an IPR Repository with all software licenses used will be created to aid partners in identifying possible licencing issues, selecting the best licensing models for the software developed during the project and ensuring on the one hand that no IPR issues hinder the exploitation potential of the software assets produced during ANITA lifetime, on the other hand IPR is fully honoured. Exploitation Managers, with the support of Project Coordinator and contributions from all partners, will provide periodically (at least once a year) a review of the relevant IPR development published and advice on this issue to all participants.

2.2.4.2 *Open Source approach and Standards*

One of the aims of the ANITA project is to provide benefits to European citizens through increased security and therefore to support customs and border guard authorities through innovative IT. To this end project partners in the consortium believe in the **value of adopting Open Source** and standards during the project: in fact various project partners may be either using Open Source code in their deliverables or contributing their deliverables to the Open Source communities. In fact, many project partners have a strong commitment to Open Source and has launched various successful Open Source commercial products. For this reason, the **ANITA platform will be released as Open Source**. Details concerning open source code use and standard contributions will be addressed in the Consortium Agreement.

Regarding parts (in particular partners' background) being brought to ANITA which are proprietary, this is a typical scenario in collaborative projects of this kind: a well-tested solution to address this situation is on the one hand to design the **platform architecture** in a very **modular** manner so that components are clearly identifiable by functionality and possibly interchangeable, on the other hand to adopt an **As-a-Service approach** where components even if closed source expose open APIs (e.g. RESTful), enabling other further services to be connected to the platform and in turn increase its potential impact. A key principle of the open source strategy for the project will be offering of the project's results in a **business friendly manner**, i.e. in a way that makes it possible for third-party European organisations (e.g. LEAs, government agencies, NGOs, companies including SMEs) to benefit from the project's results. To this end, partners commit to the selection of business friendly licenses, i.e. Academic licenses (such as Apache, Eclipse and LGPL) for the open source results of the project. To better facilitate this process the IPR monitoring activities will include setting up an IPR Repository and supporting partners in license management for software development. In addition, existing standards that will most likely impact ANITA will be clearly identified, evaluated, then applied/followed when possible in order to ensure widen interoperability. A potential list of standards is those related to extraction of evidence to ensure compatibility with other systems used in courts of law. These include Europol's UMF XML Schema, ETSI TS 1.

2.2.4.3 *Open Access to project results*

To be in line with the Open Science priority under H2020 and local legislations, and in particular with the Citizen Science term, **mandatory publication in Open Access journals (gold way) or in open access repositories (green way) will be adopted to increase the access of researchers and general public to the results of the project**. But before dissemination and communication actions, research results will be evaluated by the Security Advisory Board (see section 3.2.3.6) to determine its possible protectability, working always in accordance to Horizon2020/EC contract rules regarding background and foreground and taking also into account national regulations and the guidelines of the IPR-Helpdesk funded by the European Union (www.ipr-helpdesk.org).

2.2.4.4 *Data Management*

As part of task T.3.3, D3.3 Data Management Plan (DMP) will be elaborated by M6. The DMP will describe how the handling of research data during and after the project will be done, what data will be collected, processed or generated, what methodology and standards will be applied, whether data will be shared or made open access and how, and how data will be curated and preserved. The DMP will be updated as needed after its initial submission, but at minimum it will be updated at the end of the project.

3 Implementation

3.1 Work plan — Work packages, deliverables

3.1.1 Overall structure of the workplan

In order to achieve the objectives and expected results discussed in the sections above, the activities of the project are organized into eleven work packages (WP), each providing one or more specific main outputs, contributing to either the macro-areas of the project scope or other supporting activities (such as quality management and control, project management, dissemination and exploitation, and so forth).

Specifically, **WP1 – Project management** deals with the planning, management and control of project activities throughout its duration, including the management of the consortium, the report of project outcomes and status, and the implementation of quality control system to ensure the quality of project deliverables and the activities for controlling data usage during the project. **WP2 - Interdisciplinary analysis of online illegal trafficking and study of strategies and countermeasures** includes the analysis of activities, processes, trends and human factors that affect online illegal trafficking, in order to increase the awareness of the stakeholders (i.e. LEAs) on the potential risks and provide deep knowledge, insights and recommendations on effective countermeasures to adopt to combat such phenomena. **WP3 – Social, Ethical, Legal and Privacy issues of online sources analysis** aims at studying ethical, legal and privacy constraints in order to design and develop a system completely respectful of the human rights and do not generate ethically unwanted personal or social effects. **WP4 – Use cases, Requirements and System Architecture** will deal with uses cases and scenarios and with the elicitation of the user requirements for defining,

modelling and documenting the system specifications and the ANITA system architecture. **WP5 – Data sources and stream analysis** provides capabilities for integrating, monitoring and linking online sources (Surface and Deep Web and Dark Nets), focussing particularly on the black markets and on blockchain analysis. **WP6 – Big Data analysis and analytics** deals with the implementation of advanced and multi-modal (text, image, video and audio) Big Data analysis and analytics services for extracting relevant low-level information and emerging trends in illegal trafficking context from the selected data sources. **WP7 – Knowledge management and reasoning** aims at implementing fusion, reasoning and knowledge mining services and tools for correlating high-level information, events and facts and support investigators in understanding dynamics of illegal trafficking activities and reconstructing criminal and terrorist groups involved. **WP8 – Integration of human factor in the loop** aims at developing and deploying an adaptive user modelling framework that exploits both explicit and implicit human factors to optimize discovery and understanding of illegal trafficking data and re-using the acquired knowledge for efficient training of novice officers/practitioners/investigators. **WP9 – Novel applications for LEAs and System Integration** provides intelligence applications and analytics for illegal trafficking by exploiting results from WP5, 6 and 7 and integrates applications, analytics and analysis services into the final ANITA system. **WP10 – Demonstration in relevant environment** deals with defining, planning, preparing and executing the demonstration and validation of ANITA system in relevant operational environments provided by LEAs. **WP11 – Dissemination, Training and Exploitation** aims at managing activities for disseminating and creating awareness around the project results, building a stakeholder community, arranging training activities and workshops, and preparing business plans and exploitation activities for results coming out from the project.

The Table 3.1.a provides a summary of the above-mentioned WPs, highlighting the corresponding leader partner, the effort in person/month and the start/end month. The following table shows the correlations between project objectives (as described in Section 1.1) and above detailed work packages (bigger dots represent main contributions for the referenced objectives, whereas smaller dots refer to minor contributions). It is worth noting that all objectives are fully addressed by at least one WP.

	WP2	WP3	WP4	WP5	WP6	WP7	WP8	WP9	WP10	WP11
Obj. 1			•	●						
Obj. 2			•		●					
Obj. 3			•			●				
Obj. 4			•		•	•	●			
Obj. 5	●		•							●
Obj. 6			•	•	•	•	•	●		
Obj. 7								•	●	
Obj. 8		●	•							

The expected duration of the project is **36 months**. Milestones summarized in Table 3.2.a represent the key points of the planned approach and strategy for the project implementation, defined according to the technical strategy in Section 1.3.3. The planning of ANITA project is detailed at task level in Fig. 6, in which internal deliverables (represented with “I”) have to be intended as draft versions of official deliverables (represented with “D”). The full list of official deliverables is reported in Table 3.1.b.

High-level dependences among WPs are summarised in the following picture.

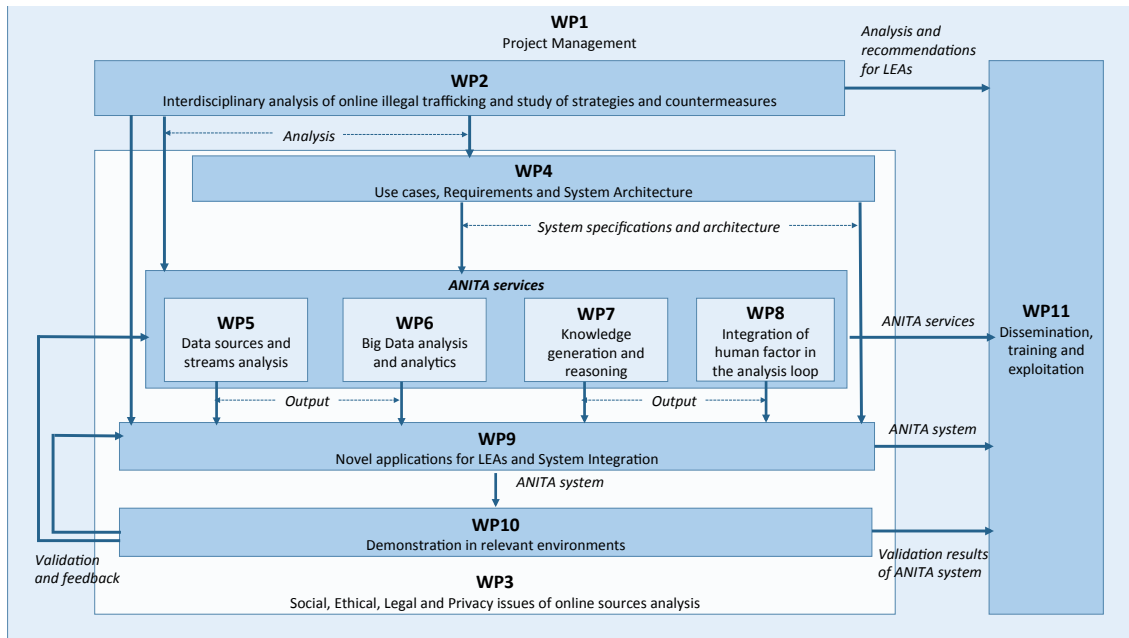


Fig. 5: Interdependencies among WPs.

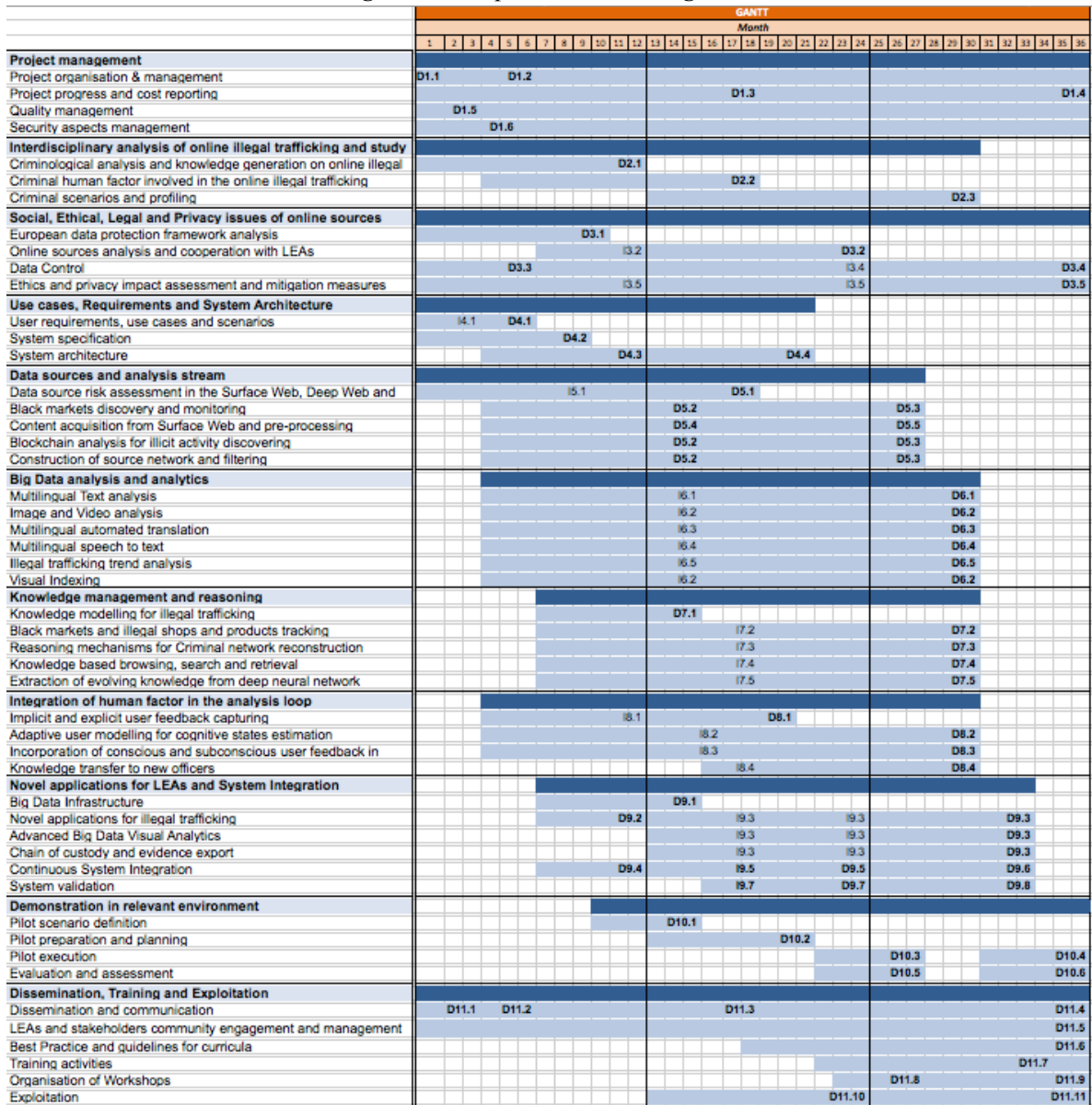


Fig. 6: ANITA Gantt chart

WP Number	Work Package Title	Lead part. Number	Short Name	Person-months	Start month	End month
WP1	Project management	1	ENG	47	1	36
WP2	Interdisciplinary analysis of online illegal trafficking and study of strategies and countermeasures	3	RISSC	53	1	30
WP3	Social, Ethical, Legal and Privacy issues of online sources analysis	7	IIP	28	1	36
WP4	Use cases, Requirements and System Architecture	2	CERTH	36	1	21
WP5	Data sources and streams analysis	5	AIT	47	1	27
WP6	Big Data analysis and analytics	2	CERTH	68	4	30
WP7	Knowledge management and reasoning	4	EXPSYS	65	7	30
WP8	Integration of human factor in the analysis loop	6	IBEC	52	4	30
WP9	Novel applications for LEAs and System Integration	1	ENG	58	7	33
WP10	Demonstration in relevant environment	10	DITSS	87	10	36
WP11	Dissemination, Training and Exploitation	13	AoC	79	1	36
TOTAL Person-months				620		

Table 3.1.a : List of work packages.

Del. Num.	Deliverable name	WP	Short name of lead part.	Type	Diss. level	Delivery date
D1.1	Project Reference Manual and Tools	WP1	ENG	R	CO	M1
D1.2	IPR Plan	WP1	ENG	R	CO	M6
D1.3	First Periodic Management Report	WP1	ENG	R	CO	M18
D1.4	Second Periodic Management Report	WP1	ENG	R	CO	M36
D1.5	Project Quality Plan	WP1	CERTH	R	CO	M3
D1.6	Best practice and guidelines for handling ANITA resources	WP1	ENG	R	CO	M5
D2.1	Trafficking Internet. The role of Surface and Deep Web – Case studies from ANITA Project	WP2	RISSC	R	PU	M12
D2.2	Exploring online traffickers. The criminal human factors and the role of new and old criminal groups	WP2	RISSC	R	PU	M18
D2.3	Criminal profiling and possible scenarios. Guidelines for action	WP2	RISSC	R	PU	M30
D3.1	European data protection framework and Ethical requirements analysis	WP3	IIP	R	CO	M10
D3.2	LEAs cooperation policy	WP3	IIP	R	CO	M24
D3.3	Data Management Plan	WP3	CERTH	R	CO	M6
D3.4	Data protection office	WP3	IIP	R	CO	M36
D3.5	Ethics and privacy impact assessment and mitigation measures	WP3	IIP	R	CO	M36
D4.1	Requirements, use cases and scenarios	WP4	CAST	R	CO	M6
D4.2	System specification	WP4	CERTH	R	CO	M9
D4.3	System architecture	WP4	ENG	R	CO	M12
D4.4	System architecture - Update	WP4	ENG	R	CO	M21
D5.1	Data source risk assessment in the Surface Web, Deep Web and Dark Nets Report	WP5	TIU-JADS	R	CO	M18
D5.2	Black markets and Blockchain analysis	WP5	AIT	R+DEM	CO	M15
D5.3	Black markets and Blockchain analysis - Update	WP5	AIT	R+DEM	CO	M27
D5.4	Content acquisition from Surface Web and pre-processing	WP5	EXPSYS	R+DEM	CO	M15
D5.5	Content acquisition from Surface Web and pre-processing - Update	WP5	EXPSYS	R+DEM	CO	M27
D6.1	Text Analysis services	WP6	EXPSYS	R+DEM	CO	M30
D6.2	Image and Video Analysis services and Indexing	WP6	CERTH	R+DEM	CO	M30

D6.3	Multilingual automated translation services	WP6	SYSTRAN	R+DEM	CO	M30
D6.4	Multilingual speech to text services	WP6	SYSTRAN	R+DEM	CO	M30
D6.5	Illegal trafficking trend analysis services	WP6	TIU-JADS	R+DEM	CO	M30
D7.1	ANITA Knowledge modeling and representation	WP7	EXPSYS	R+DEM	CO	M15
D7.2	ANITA services for linking and tracking illegal trafficking activities	WP7	EXPSYS	R+DEM	CO	M30
D7.3	ANITA Knowledge mining services	WP7	AIT	R+DEM	CO	M30
D7.4	Search and retrieval services	WP7	ENG	R+DEM	CO	M30
D7.5	Knowledge extraction from neural networks	WP7	CERTH	R+DEM	CO	M30
D8.1	Implicit and explicit user feedback acquisition and first prototype of the user model	WP8	IBEC	R+DEM	CO	M20
D8.2	Adaptive modeling of user cognitive states	WP8	IBEC	R+DEM	CO	M30
D8.3	Incorporation of user feedback in deep learning representations	WP8	CERTH	R+DEM	CO	M30
D8.4	Transferring domain knowledge to novice officers	WP9	IBEC	R+DEM	CO	M30
D9.1	ANITA Big Data Infrastructure	WP9	ENG	R+DEM	CO	M15
D9.2	ANITA mock-ups	WP9	ENG	R	CO	M12
D9.3	Novel applications for illegal trafficking and chain of custody	WP9	ENG	R+DEM	CO	M33
D9.4	ANITA Integration Plan	WP9	ENG	R	CO	M12
D9.5	ANITA System 1st release	WP9	ENG	R+DEM	CO	M24
D9.6	ANITA System 2nd release	WP9	ENG	R+DEM	CO	M33
D9.7	ANITA Validation report 1st release	WP9	EXPSYS	R	CO	M24
D9.8	ANITA Validation report 2nd release	WP9	EXPSYS	R	CO	M33
D10.1	Pilot scenarios	WP10	CAST	R	CO	M15
D10.2	Pilots plan	WP10	DITSS	R	CO	M21
D10.3	Pilots execution report 1st release	WP10	DITSS	R	CO	M27
D10.4	Pilots execution report 2nd release	WP10	DITSS	R	CO	M36
D10.5	Pilots evaluation and assessment 1st release	WP10	CAST	R	CO	M27
D10.6	Pilots evaluation and assessment 2nd release	WP10	CAST	R	CO	M36
D11.1	ANITA Website	WP11	EXPSYS	R	PU	M3
D11.2	Dissemination and communication plan	WP11	EXPSYS	R	PU	M6
D11.3	Report on Dissemination Activities 1 st release	WP11	EXPSYS	R	PU	M18
D11.4	Report on Dissemination Activities 2 nd release	WP11	EXPSYS	R	PU	M36
D11.5	LEAs and stakeholder community building and management	WP11	DITSS	R	PU	M36
D11.6	Guidelines for curricula definition	WP11	CAST	R	PU	M36
D11.7	Training materials	WP11	AoC	R	PU	M34
D11.8	1st Workshop Report	WP11	AoC	R	PU	M27
D11.9	2nd Workshop Report	WP11	AoC	R	PU	M36
D11.10	Market Analysis and Initial Exploitation Plan	WP11	ENG	R	CO	M24
D11.11	Market Analysis and Final Exploitation Plan	WP11	ENG	R	CO	M36

Table 3.1.b: List of deliverables.

3.1.2 Detailed work package descriptions

3.1.2.1 WPI - Project Management

WP Number	1	Leader:			ENG	Start month:		1	End month:		36
WP Title	Project Management										
Participant number	1	2	3	4	5	6	7	8	9	Total	
Particip. short name	ENG	CERTH	RISSC	EXPSYS	AIT	IBEC	IIP	SYSTRAN	TIU-JADS		
Person/Months	23	5	1	1	1	1	1	1	1		
Participant number	10	11	12	13	14	15	16	17			
Particip. short name	DITSS	ISBR	KWPG	AoC	CAST	NPN	GDCOC	LPV			

Person/Months	1	2	1	1	4	1	1	1		47
Objectives										
The objective of this WP aims is to coordinate the project and monitoring its evolution. The expected objectives of this work package are the achievement of the project objectives within budget and scheduled deadlines, periodic reporting, quality management, achievement of milestones, handling of IPR issues, arrangement of periodic meetings and the overall monitoring of project development status and cost reporting. Data control is also included to ensure reliable data processing throughout the duration of the project.										
Task 1.1 Project organisation and management (Leader: ENG; M1-M36)										
This task deals with the organization, planning, monitoring and execution of project management activities. It includes the preparation of project kick-off and periodic plenary and boards restricted meetings, the timely control and submission of deliverables to the EU Commission, monitoring of project status according to the selected milestones and KPI, the risk assessment and mitigation management, the assurance of the proper function of the project boards. In addition, this task will setup the initial IPR Plan in order to tackle confidentiality issues and potential contractual arrangements as well as to identify intellectual property usage guidelines and steps for its protection. This task contributes to the deliverable D1.1 and D1.2.										
Task 1.2 Project progress and cost reporting (Leader: ENG; Participants: All partners; M1-M36)										
This task coordinates all activities related to the assessment of project development. This includes scheduling, organizing and reporting periodic review reports, coordinating the information flow, organizing periodic project board meetings for project progress review, establishing and maintaining financial records, coordinating costs submission, conducting preliminary checks of individual costs against known criteria (contractual commitments, progress reports, delivery of results) and consolidating cost, following-up of EC payments, and distributing partner shares. Periodic reports will describe project progress and identify significant risks with proposed contingency plans. This task contributes to deliverables D1.3 and D1.4.										
Task 1.3 Quality management (Leader: CERTH; Participants: ENG; M1-M36)										
The work in this task involves the actions to be performed in order to organise quality control, to establish/benchmark project milestones and implement mitigation plans in case of identified risks. They are mainly referred to: i) Quality control of the planned actions and time schedules, definition of project requirements/specifications and quality objectives and definition of responsibilities and authorities; ii) Establishment and benchmark of project milestones and periodic monitoring towards their fulfilment; iii) Check of timely production of deliverables and establishment of quality/acceptance criteria to be implemented; iv) Control of quality and consistency against technical and contractual aspects. This task contributes to deliverable D1.5.										
Task 1.4 Security aspects management (Leader: ENG; Participants: ISBR, CAST; M1-M36)										
This task refers to the activities that will be performed in order to ensure the correct management of the project resources and results. A Security Advisory Board (SAB) will be established using proven experience people in the consortium in order to assess the proper execution of security and protection measures and procedures for accessing and handling, results of the ANITA project. SAB will monitor regularly the project activities and assess the sensitivity of information handled by participants providing continuously support to the consortium. This task contributes to deliverable D1.6.										
Work package deliverables										
<i>D1.1 – Project Reference Manual and Tools [M1]</i>										
<i>D1.2 – IPR Plan [M6]</i>										
<i>D1.3 – First Periodic Management Report [M18]</i>										
<i>D1.4 – Second Periodic Management Report [M36]</i>										
<i>D1.5 – Project Quality Plan [M3]</i>										
<i>D1.6 – Best practice and guidelines for handling ANITA resources [M5]</i>										

3.1.2.2 WP2 – Interdisciplinary analysis of online illegal trafficking and study of strategies and countermeasures

WP Number	2	Leader:	RISSC	Start month:	1	End month:	30			
WP Title	Interdisciplinary analysis of online illegal trafficking and study of strategies and countermeasures									
Participant number	1	3	10	12	13	14	15	16	17	Total
Particip. short name	ENG	RISSC	DITSS	KWPG	AoC	CAST	NPN	GDCOC	LPV	

Person/Months	2	30	3	3	3	3	3	3	3	53
Objectives										
<p>This WP aims at generating the knowledge to achieve the overall proposed results and to contribute at improving its efficiency/effectiveness. The following objectives are targeted: to generate in-depth and inter-disciplinary knowledge about online illegal trafficking, with regards to the domains and the use-cases of interest for the project, to the interdependencies between the online and the offline dimensions, and the human factors involved; to develop a knowledge-generation process to continuously update the ANITA knowledge base and to make it available for the partners and the stakeholders; to provide feedback and propose operational inputs to contribute to overcoming the actual difficulties in understanding, investigating and contrasting online illegal trafficking; to contribute at increasing awareness on the dramatic risks/impact of online illegal trafficking and train key stakeholders at public and private level; to assess critical issues and guidelines for LEAs and stakeholders.</p>										
<p>Task 2.1 Criminological analysis and knowledge generation on online illegal trafficking (Leader: RISSC; Participants: DITSS, KWPG, AoC, CAST, NPN, GDCOC, LPV; M1-M12)</p>										
<p>This task focuses on the generation of new knowledge about domains and use cases of interest for the project, with specific regards to key-factors characterising the phenomena from the criminological standpoint. As regards the online dimension, both Surface and Deep Web will be investigated in order to assess the actual criminal dynamics, the correlation between the online and the offline dimensions, the virtual and physical routes of trafficking, the actual and potential trends and related risks. Desk research and literature analysis, interviews with experts, direct contributions from/exchange of information and brainstorming with LEAs, data crawling, innovative methodologies for analysis will be a key part of the approach/work. This task contributes to deliverable D2.1.</p>										
<p>Task 2.2 Criminal human factor involved in the online illegal trafficking (Leader: RISSC; Participants: DITSS, KWPG, AoC, CAST, NPN, GDCOC, LPV; M4-M18)</p>										
<p>This task deals with the human factors involved in the online illegal trafficking with the intent to further integrate and complement the analysis developed in Task 2.1, thus generating a comprehensive description of the actual criminal scenario of illegal trafficking over the Internet, considering both Surface and Deep Web. The research activities will be mostly oriented to understanding the nature of the criminal human factors and the possible connections with organised crime. As for Task 2.1, desk research and literature analysis, interviews with experts, direct contributions from/exchange of information and brainstorming with LEAs, data crawling, innovative methodologies for analysis will be a key part of the approach/work. This task contributes to deliverable D2.2.</p>										
<p>Task 2.3 Criminal scenarios and profiling (Leader: RISSC; Participants: ENG, DITSS, KWPG, AoC, CAST, NPN, GDCOC, LPV; M13-M30)</p>										
<p>This task further exploits the knowledge base generated by Task 2.1 and 2.2 with the intent to enhancing the capacities to investigate and classify available information, to develop a continuous knowledge-management process and to develop inputs/guidelines for improving intelligence activities and processes. The guidelines will be elaborated to support the policy-making process and the definition of those countermeasures relevant to face the threat of online illegal trafficking, with specific regards to the domain of interest for the ANITA Project. The guidelines will consider both the human and the technological factors and will be segmented according to the diverse stakeholders, which could be involved in the prevention and reaction against these types of crime (e.g. public authorities, LEAs, private organisations/companies, NGOs, citizens...). Furthermore, it will provide inputs for innovative curricula, training and (joint) exercises in favour of LEAs and other key stakeholders. The work will be mostly analytical in nature and will select the most innovative methodologies for analysis available in the scientific domain. This task contributes to deliverable D2.3.</p>										
Work package deliverables										
D2.1 – Trafficking Internet. The role of Surface and Deep Web – Case studies from ANITA Project [M12]										
D2.2 – Exploring online traffickers. The criminal human factors and the role of new and old criminal groups [M18]										
D2.3 - Criminal profiling and possible scenarios. Guidelines for action [M30]										

3.1.2.3 WP3 – Social, Ethical, Legal and Privacy issues of online sources analysis

3.1.2.5 WP3 Social, Ethical, Legal and Privacy issues of online sources analysis										
WP Number	3	Leader:		IIP	Start month:		1	End month:		36
WP Title	Social, Ethical, Legal and Privacy issues of online sources analysis									
Participant number	1	2	7							Total
Particip. short name	ENG	CERTH	IIP							
Person/Months	3	3	22							28

Objectives
<p>This WP aims at identifying and tackling all potential legal and ethical issues of the project and its tools. So as to deal with the highlighted challenges, the following specific objectives are targeted: to evaluate data protection issues and impacts on the project with regard to the processing of personal data pursuant Regulation UE 2016/679 (General Data Protection Regulation) and Directive UE 2016/680; to evaluate the Cybersecurity Strategy for the European Union (e.g. NIS Directive) and European Union Agency for Network and Information Security activities; to define the cooperation between the Consortium and LEAs; to carry out an impact assessment concerning the possible ethical and legal risks for researchers involved in the project, in light of the likely sensitiveness and possible side effects of the envisaged research activity in dark web; to evaluate the impacts of the project on human rights with regard on how users make use of dark web, also for non-criminal and legitimate reasons and activities and evaluation of measures that might be taken to prevent abuses; to evaluate possible ethical issues that may arise in connection to potential misuse of the tools employed by the project made by undemocratic governments.</p>
Task 3.1 European data protection framework analysis (Leader: IIP; M1-M10)
<p>The first step of the analysis is to perform a general assessment of the ethical, privacy and legal concerns. To achieve this goal, the relevant European legislation will be studied in detail. The elements of analysis will be: Directive 2016/680 on the protection of natural persons with regard to the processing of personal data by competent authorities for the purposes of the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, and on the free movement of such data; the General Data Protection Regulation (GDPR); the ethical principles constituted in the Charter of Fundamental Rights of the European Union and the European Convention on Human Rights and all other applicable international, EU and national legislation in order to ensure end-user acceptance and ethics compliance. This task will identify some functional and non-functional requirements to be addressed by the ANITA tools. It shall synthesize those requirements to support the architecture design and the subsequent work of the other WPs. The output of the previous mentioned analysis will be a set of requirements which balance the legal basis and ethical one in order to define both the legal and ethical requirements of the whole project, with a focus on the tools development so as to integrate an ethics/data protection by design and by default approach. This task contributes to deliverable D3.1.</p>
Task 3.2 Online sources analysis and cooperation with LEAs (Leader: IIP; Participants: ENG; M7-M24)
<p>This task is concerned with the monitoring of the Cybersecurity Strategy for the European Union (e.g. NIS Directive) and European Union Agency for Network and Information Security activities. It will help in defining the effective cooperation with LEAs involved in the project (e.g. cooperation agreements or “informal” trusted network). Moreover, this task will evaluate ethical issues that may arise in connection to potential misuse of the tools employed by the project made by undemocratic governments in order to better understand how to cooperate against illegal trafficking crimes with law enforcement agencies without compromising the democratic European asset. The aim of the task is defining <i>ex ante</i> and producing the policies for boosting cooperation with LEAs and how to deal with their requests, in order to avoid misuse of project tools. This task contributes to deliverable D3.2.</p>
Task 3.3 Data Control (Leader: IIP; Participants: CERTH; M1-M36)
<p>This task will involve all necessary actions to be taken by the ANITA project in order to guarantee a lawful and accountable data processing throughout the duration of the project, considering that the identifiability (and the identity) of users is not a goal of the project and, thanks to the anonymity of the dark web, ANITA will monitor anonymous groups only. The establishment of a Data Protection and Ethics Office will avoid the potential misuse of research results and tools, both from partners of the Consortium and from external malicious actors by monitoring the researcher actions, defining a procedure for incidental personal data collection, define and implement ethics and privacy policy etc. Through the application of the accountability principle, by producing all the relative documents (e.g. records of processing activities, processors designation, data breach procedures etc.) the Controller (i.e. the consortium) will be responsible for and be able to demonstrate compliance with the GDPR and Directive 680/2016. The output of the task will be the drafting of all the compliance documents and internal procedures in order to be fully aligned with EU law and Ethics principles and coordinate Consortium’s compliance actions as defined during T3.1 with the set of requirements. Moreover, the researchers will be trained with a set of materials (e.g. presentations during the meetings) defined to maintain the advanced tools developed in the project within the frame of the legal and ethical standards. This task contributes to deliverables D3.3 and D3.4.</p>
Task 3.4 Ethics and privacy impact assessment and mitigation measures (Leader: IIP; Participants: ENG, CERTH; M1-M36)
<p>In this task, on the basis of a comparative analysis of the social, ethical, legal, privacy requirements emerged in Task 3.1, the critical aspects will be defined during the tools development. In order to do so, all the possible impacts on the rights of individuals involved in the project activity (on line users, researchers, third parties) will be</p>

assessed, so as to identify related solutions/mitigation measures. The impact assessment will be carried out also for researchers involved in the project. In fact, there are possible ethical and legal risks for them, in light of the likely sensitiveness and side effects of the envisaged research activity in dark web. The task aims to do a legal and ethical assessment of the various components of the project (i.e. tools, systems) following the outputs of Tasks T3.1 and 3.2 and a) searching the impact of the technology; b) searching the extent in which fundamental social, ethical and legal goods (including privacy) may be infringed; c) searching the impact on the limitations of rights and freedom generated by the tools developed in the project. As output of the task, at the end of the assessment, an awareness report will be produced, in order to balance the underlined risks with mitigation measures applying an Ethics-by-Design and Data Protection-by-Design strategy, in order to accompany the design and development of tools in a “friend of rights” manner. This task contributes to deliverable D3.5.

Work package deliverables

D3.1 – European data protection framework and Ethical requirements analysis [M10]

D3.2 – LEAs cooperation policy [M24] (Internal release [M12])

D3.3 – Data management plan [M6]

D3.4 – Data protection office [M36] (Internal release [M24])

D3.5 – Ethics and privacy impact assessment and mitigation measures [M36] (Internal release [M12], [M24])

3.1.2.4 WP4 – Use cases, Requirements and System Architecture

WP Number	4	Leader:	CERTH	Start month:	1	End month:	21			
WP Title	Use cases, Requirements and System Architecture									
Participant number	1	2	3	4	5	6	7	8	9	Total
Particip. short name	ENG	CERTH	RISSC	EXPSYS	AIT	IBEC	IIP	SYSTRAN	TIU-JADS	
Person/Months	6	8	1	2	2	2	1	2	1	
Participant number	10	11	12	13	14	15	16	17	18	
Particip. short name	DITSS	ISBR	KWPG	AoC	CAST	NPN	GDCOC	LPV		
Person/Months	2	1	1	1	3	1	1	1		36

Objectives

This WP will manage all activities related to the elicitation of user requirements and the definition of use cases and scenarios, analysing and turning them into the system specifications, with the contribution of all technical partners, in order to design the reference architecture of the ANITA system. The following specific objectives are targeted: to describe use cases and scenarios that will guide ANITA system design and implementation; to elicit user requirements; to study and specify system level requirements; to provide ANITA system specification, to design the overall system architecture taking into account user and system requirements.

Task 4.1 User requirements, use cases and scenarios (Leader: CAST; Participants: CERTH, RISSC, DITSS, ISBR, KWPG, AoC, CAST, NPN, GDCOC, LPV; M1-M6)

This task aims at defining the use cases, scenarios and user requirements that will guide the elicitation of system requirements and the development of ANITA system. With the proactive support of LEAs, use cases and scenarios will be defined by adopting a real-life perspective, in order to realise a final system that is compliant with the real needs of end users and that is able to adapt itself to the continuously changing environment, necessities and circumstances linked to the evolving domain of illegal trafficking and organised crime. A specific workshop will be organised at the end of M3 and will involve all LEAs of the Consortium in order to foster user requirements elicitation. User requirements will concern both operational needs (i.e. how to boost the officers everyday work), but also requirements and best practices that will speed up the training procedure of novice officers; hence, fulfilling both fundamental goals of the ANITA system. The outcomes of this task will also provide relevant inputs for the definition of pilots’ scenarios (T10.1). This task contributes to deliverable D4.1.

Task 4.2 System specification (Leader: CERTH; Participants: ENG, EXPSYS, AIT, IBEC, IIP, SYSTRAN, TIU-JADS; M1-M9)

Starting from the results of T4.1, the objective of this task is to produce a clear system specification that will guide the design of the system architecture. Several technical issues will be investigated starting from the high-level functionalities that the system will have to satisfy, such as interoperability, robustness, maintainability, security and scalability, search and retrieval mechanisms, as well as non-technical issues, like usability and ethical/legal compliance. The outcome of the task will be the definition of the system level requirements for the design of the

overall system architecture and the specification of the indicators for the validation of user scenarios. This task contributes to deliverable D4.2.
Task 4.3 System architecture (Leader: ENG; Participants: CERTH, EXPSYS, AIT, IBEC, SYSTRAN; M4-M21)
Starting from system specification defined in T4.2, this task will deal with the design of the overall ANITA system architecture. A privacy-by-design approach will be adopted to ensure ethical and legal compliance of the final system. In addition to identified system requirements, the architecture will be designed in order to be scalable for managing increasing amounts of information, flexible for allowing additional sources, modules and services to be added easily to the system, and interoperable for enabling information exchange with external information systems already used by LEAs. To this end, a Service Oriented approach will be adopted in order to facilitate the integration and orchestration of existing services and of future services that could be required outside ANITA project. Starting from the first released version (M12), the system architecture will be enhanced and improved according to feedback provided by system validation (T9.6). This task contributes to deliverables D4.3 and D4.4.
Work package deliverables
<i>D4.1 – Requirements, use cases and scenarios [M6] (Internal release [M3])</i>
<i>D4.2 – System specification [M9]</i>
<i>D4.3 – System architecture [M12]</i>
<i>D4.4 – System architecture – Update [M21]</i>

3.1.2.5 WP5 – Data sources and streams analysis

3.1.2.5 WP5 Data sources and streams analysis

WP Number	5	Leader:		AIT	Start month:		1	End month:		27
WP Title	Data sources and stream analysis									
Participant number	1	3	4	5	9					Total
Particip. short name	ENG	RISSC	EXPSYS	AIT	TIU-JADS					
Person/Months	8	2	9	20	8					47

Objectives
This WP will develop a suitable infrastructure for collecting relevant resources from Surface Web, Deep Web and Dark Nets in the context of illicit trafficking, to find actual and new sources containing relevant network information about organised crime support networks and illicit goods trafficking. The following specific objectives are targeted: fastest and most precise content identification; protecting the infrastructure from discovery through anonymity-oriented development; defining strategies to assess the risk level of sources in Surface Web, Deep Web and Dark Nets; developing methodologies and tools that collect only relevant information; discovery of black markets in Deep web and Dark nets; content acquisition from Surface Web; blockchain analysis; reconstruction of links between sources and their filtering.

Task 5.1 Data source risk assessment in the Surface Web, Deep Web and Dark Nets (Leader: TIU-JADS; Participants: AIT, RISSC; M1-M18)

This task will focus on the definition of methods and indicators for the risk assessment of new data sources. Risk level is referred to a different aspects strictly connected: on one hand, the risk level is due to the involvement of source in illegal trafficking activities; on the other hand, the possibility that criminal groups that own or manage sources monitored by LEAs could recognize anomalous activities due to the crawling process and block them, compromising the investigations. Taking into account that risk level of a source can increase with its depth in the Web, such methods will then consider the position of the source (in Surface Web, Deep Web or Dark Nets), metadata that can be extracted in order to define which policies should be enabled to perform the crawling and with which degree of anonymity. Finally, websites appearance, their update frequency and refresh in data will be taken into account during the assessment stage. The outcome of this task consists in a set of guidelines, methods and principles to establish the risk level of sources in Surface Web, Deep Web and Dark Nets, and the related policies to be applied for ensuring safe and secure crawling. This task contributes to deliverable D5.1.

Task 5.2 Black markets discovery and monitoring (Leader: AIT; Participants: ENG; M4-M27)

This task deals with the design and implementation of crawlers for Deep Web and Dark Nets focused on the discovery and monitoring of black markets. By applying data mining techniques within selected sources, it is possible to identify products, volumes, behaviours of buyers and sellers that could be related to illegal trafficking activities. Monitoring tasks on suspected source are then started in order to alert LEAs in case of detection of relevant events. Furthermore, similarity algorithms based on types of product and involved virtual identities will be used to recognize the presence of the same markets under different names and thus to track their evolution, providing inputs for

the source network construction (T5.5). This task contributes to deliverables D5.2 and D5.3.									
Task 5.3 Content acquisition from Surface Web and pre-processing (Leader: EXPSYS; Participants: ENG; M4-M27)									
This task will be dedicated to the development of a crawler for structured and unstructured content extraction based on open technologies, like Apache Nutch and Apache Camel. This crawler will be able to acquire content both from open sources on Surface Web, like RSS, Social Networks (Twitter, Facebook), Web, Blogs and from local sources like files in the most common formats (i.e.: pdf, word, excel, power point, mail etc.) stored in folders or repositories. Heterogeneous resources (image, video, text) will be extracted from contents acquired by the targeted sources. The adopted approach will also include a preliminary study for selecting relevant data sources to be monitored for intercepting illegal trafficking activities. Specific taxonomies will be defined for supporting contents pre-processing. This task contributes to deliverables D5.4 and D5.5.									
Task 5.4 Blockchain analysis for illicit activity discovering (Leader: AIT; M4-M27)									
The purpose of this task is to develop methods for the detection, tracking, and monitoring of monetary transactions (e.g. between black markets or online traders), with the purpose of recognizing illegal trafficking activities. Given the new opportunities of crypto-currency networks, the task will focus on crypto currency transactions, but will be not limited to this. Through the combination of blockchain data and information extracted from Deep Web and Dark Nets, it is in some cases possible to de-anonymise users of black markets. In this task, blockchain analytics will be combined with mining for products and technologies to get insights into the pharmaceutical supply chain and the financial network in behind. As a result this task produces a database about the ownership of pharmaceutical organisations, their products and financial transactions between these organisations, based on open source information. This task contributes to deliverables D5.2 and D5.3.									
Task 5.5 Construction of source network and filtering (Leader: AIT; Participants: ENG, EXPSYS; M4-M27)									
Starting from outcomes of T5.1, T5.2, T5.3 and T5.4, this task will be responsible for the reconstruction of links and relationships among sources discovered in Surface Web, Deep Web and Dark Nets, including different accounts that can be linked to the same criminal group. The IP address discovery of the monitored sources will also allow investigators to track the information flow and to identify when the illegal trades move from Surface to the Deep Web and Dark Nets. The reconstruction of source network constitutes the first step towards the discovery of underlying criminal networks and will enable the automatic discovery of new relevant sources, linked to those already known. In order to limit the amount of sources to consider during the network construction, specific filtering mechanisms will be implemented and applied to each source, in order to provide a preliminary overview of its relevance in illegal trafficking domain, taking into account outputs from analysis tools (developed in WP6) and its position in the source network. This task contributes to deliverables D5.2 and D5.3.									
Work package deliverables									
<i>D5.1 – Data source risk assessment in the Surface Web, Deep Web and Dark Nets report [M18] (Internal [M9])</i>									
<i>D5.2 – Black markets and Blockchain analysis [M15]</i>									
<i>D5.3 – Black markets and Blockchain analysis – Update [M27]</i>									
<i>D5.4 – Content acquisition from Surface Web and pre-processing [M15]</i>									
<i>D5.5 – Content acquisition from Surface Web and pre-processing - Update [M27]</i>									

3.1.2.6 WP6 – Big Data analysis and analytics

WP Number	6	Leader:	CERTH	Start month:	4	End month:	30			
WP Title	Big Data analysis and analytics									
Participant number	1	2	4	8	9					Total
Particip. short name	ENG	CERTH	EXPSYS	SYSTRAN	TIU-JADS					
Person/Months	6	19	18	17	8					

Objectives

The aim of this WP is to manipulate, analyse and semantically organize the vast amount of multimedia content that have been acquired from Surface/Deep Web, Dark Nets, Social media and other online sources. For achieving this, the following specific objectives are targeted: to realize multilingual text analysis for processing textual resources and revealing meaningful information that they contain; to semantically analyse the visual content (image, video) for detecting and extracting valuable information cues that will boost the investigation process; to develop multilingual translation and speech-to-text services for effectively processing the audio stream; to develop a visual indexing infrastructure for enabling time efficient search and retrieval in the large-scale using queries in the form of visual content. This off-line analysis will be based on the use of machine learning techniques and in principle on

the exploitation of the deep learning paradigm. The modules developed in this WP will be further reinforced, taking into account the methodologies of WP8 that aim at including the system user in the analysis loop.

Task 6.1 Multilingual Text analysis (Leader: EXPSYS; Participants: ENG; M4-M30)

This task will be focused on designing and evolving an avant-garde semantic based text analytics engine for automatic content categorization and entities extraction. The categorization process will be implemented according to taxonomies will be specifically defined around the online illegal trafficking domain. The engine will be enriched with capabilities to extract advanced information, like temporal references and relationships among the extracted entities. Besides, an adaptation to short texts and misspellings will be considered, to make the semantic engine suitable to deal with not conventional textual resources, like those that can be retrieved from deep web and dark nets and that are strongly characterized by this kind of language. Dedicated rules for data normalization will be defined and implemented as well, enabling the training of the semantic engine with standard terminology typically used in the communication between buyers and sellers in the field of illegal trafficking. With this normalization rules inside, the semantic engine will allow processing input sources using a LEA common language. Stylometric analysis (a.k.a. writeprint) will be also included in order to support illegal trafficking clustering. This task will deliver an advanced and Big Data oriented textual resource analysis framework that will integrate services for categorization and semantic information extraction from huge amounts of textual resources, while it will be based on common machine learning techniques (including deep learning ones). This task contributes to deliverable D6.1.

Task 6.2 Image and Video analysis (Leader: CERTH; M4-M30)

The goal of this task is to deliver the algorithms that are going to extract relevant information from the visual content (images, videos) that is available in online (Surface Web, Deep Web, Dark Nets, social media, etc.) and offline (LEAs' databases) resources. In particular, novel Deep Learning (DL) architectures will be developed for automatically detecting potentially interesting pieces of information or evidence (i.e. objects, concepts and events) in the formed databases. For achieving this goal, ANITA will concentrate on the followings: a) Object recognition: The goal is to locate and identify the real-world objects that may be presented in the visual medium. ANITA aims at introducing effective and time-efficient generic object localization and recognition methods, exploiting multiple-level contextual information in a complementary way to global and local visual information. b) Concept: detection: The objective is to robustly detect a wide range of high-level semantic entities that are included in the visual medium. These entities may correspond to many different levels of semantic granularity or abstraction, e.g. ranging from specific object types to individual scene categories. In particular, ANITA will develop a large-scale visual-based semantic concept detection framework, taking significantly into account the respective scalability issues that are expected to be present in the domain of illegal trafficking-related activities. c) Event detection: This aims to deliver a service to identify both high-level as well as primitive semantic events related to the activities of the depicted individuals. ANITA will concentrate on the implementation and training of classifiers for static images (i.e. one-shot event detection). This task contributes to deliverable D6.2.

Task 6.3 Multilingual automated translation (Leader: SYSTRAN; M4-M30)

This task aims at developing multilingual translation services based on Neural Machine Translation. Two classes of models will be developed: one consists in developing multilingual translation models into English, the second in developing translation models from English to many languages. The first step is to create a model trained on multiple related source languages translated into English. A model trained on these languages will allow the easy integration of additional languages by simply continuing training on bilingual data belonging to the new language. Moreover, the model will enable the translation of new languages without additional training if the new language is sufficiently close to a language included in the previous training. This approach achieves two goals at the same time. It reduces the amount of required training data and simultaneously makes it faster to create a model for a new language. Having a set of pretrained models for all major language families (Indo-European, Sino-Tibetan, Turkic, etc.) will allow reacting fast to new language needs. One challenge in this approach is to handle slightly different spellings in different languages. The second step is to develop multilingual models translating from English. This task is more challenging since the decoder now needs to know which language to produce. To this end, a model will be pretrained with all related languages, then either the low resource language will be overweighed or the training will be finished using only the real target language. This task contributes to deliverable D6.3.

Task 6.4 Multilingual speech to text (Leader: SYSTRAN; M4-M30)

The objective of this task is to develop speech-to-text services for many languages to enable the analysis of audio data. This task relies on three specific sub tasks: multilingual and multi-dialect features and acoustic modeling; use of a unified semantic space for Automatic Speech Recognition and Machine Translation; acoustic data augmentation. In multilingual/multi-dialect features and acoustic modelling, when too little speech data is available to build reliable acoustic models, sharing commonalities with other languages can be the most effective way to construct

models for the targeted task. In unified semantic space, neural networks with word embeddings, used to reduce the impact of sparseness of the language modeling training data since 2002, will be expanded by taking advantage of a semantic multilingual representation used for translation. Data augmentation methods will be used to reduce the impact of data sparseness and increase the acoustic model robustness regarding various sources of variabilities and noises. The original training data will be modified using additive and convolutional noises, as well as vocal tract length and speech rate perturbations and other voice morphing methods. This task contributes to deliverable D6.4.
Task 6.5 Illegal trafficking trend analysis (Leader: TIU-JADS; Participants: CERTH; M4-M30)
The task is responsible for providing a set of tools for detecting trends and analysing the incoming information with respect to illegal trafficking. The tools will exploit the outputs of multilingual text analysis in order to extract information about specific trends, analytics and produce actionable insights on buying/transaction habits and user behaviours. Processing will involve extraction of statistics in order to guide and support investigators for early detection of criminal behaviours and respond rapidly on critical situations. Big data analytics will be used and new algorithms will be implemented to handle the noisy data from different sources. Graph analysis will be used to detect key nodes in the exchange network, while multimodal deep learning architectures will be investigated to classify behaviours and flag illegal activities. Finally, an iterative map-reduce based algorithm will be provided to generate daily, weekly and monthly reports on the extracted analytics. This task contributes to deliverable D6.5.
Task 6.6 Visual Indexing (Leader: CERTH; M4-M30)
This task will provide an end-to-end visual indexing framework based on deep learning and the learning to hash approach. The visual indexing framework will enable indexing of images and videos using global as well as local information within each visual frame to incorporate contextual knowledge and concept analysis in the indexing procedure. The task will also investigate the usage of other contextual cues such as captions or content metadata that will guide the creation of effective hash functions while keeping the dimensionality of the visual descriptors low. Finally, object-based visual indexing approaches will also be investigated to provide sub-frame indexing and object re-identification functionalities. This task contributes to deliverable D6.2.
Work package deliverables
D6.1 – Text Analysis services [M30] (Internal release [M15])
D6.2 – Image and Video Analysis services and Indexing [M30] (Internal release [M15])
D6.3 – Multilingual automated translation services [M30] (Internal release [M15])
D6.4 – Multilingual speech to text services [M30] (Internal release [M15])
D6.5 – Illegal trafficking trend analysis services [M30] (Internal release [M15])

3.1.2.7 WP7 – Knowledge generation and reasoning

WP Number	7	Leader:	EXPSYS	Start month:	7	End month:	30		
WP Title	Knowledge generation and reasoning								
Participant number	1	2	4	5					Total
Particip. short name	ENG	CERTH	EXPSYS	AIT					
Person/Months	24	12	24	5					65

Objectives
This WP aims at modelling knowledge for illegal trafficking domain and at implementing reasoning mechanisms for criminal network reconstruction, in order to support LEAs work in discovering hidden networks, events, people, groups and their relationships. The following objectives are targeted: to model illegal trafficking ontology to describe crimes and related events; to design and develop a Machine Learning based framework able to integrate services for recognizing black markets and illegal shop and products; to define Reasoning mechanisms and services for Criminal network reconstruction; to design and implement the Knowledge Management Framework to deal with knowledge by ANITA tools and services and from deep neural network representations.
Task 7.1 Knowledge modeling for illegal trafficking (Leader: EXPSYS; Participants: ENG; M7-M15)
Starting from the background analysis performed in WP2, and taking into account use cases and scenarios analysed in WP4, this task will focus on knowledge modelling for illegal trafficking. A proper set of ontologies will be defined in order to model all crime aspects including events, suspicious and illegal activities, threats, people, organisations, places, black-markets and illegal shops, products and their relationships. The defined ontologies will constitute the starting point for the definition of inference and reasoning rules and mechanisms that will be developed throughout WP7. This task contributes to deliverable D7.1.
Task 7.2 Black markets and illegal shops and products tracking (Leader: EXPSYS; Participants: ENG, AIT;

M7-M30)
<p>This task aims at delivering inference and reasoning services to analyse relevant information extracted by services developed in WP6, in order to discover, correlate and track the evolution of black markets, illegal shops and products involved in illegal trafficking activities. Metadata related to several fragments of information coming out from heterogeneous resources (text, audio, image and video) will be deeply examined to discover possible connections and merged together in order to reconstruct complex illegal events. Activities related to different stages of monitored events will be linked together to reconstruct the crime storyline. Event calculus formalism will be applied to represent illegal actions in narrative way, thus enabling events reconstruction in time and space. Moreover, advanced techniques for information clustering will be involved to discover hidden relationships and similarities among monitored markets and products. The outcome of this task will be a semantic framework for new knowledge discovery consisting in a hybrid approach that merges available knowledge and information, linguistic and semantic characterizations of resources extracted from monitored sources and Machine Learning algorithms. This task contributes to deliverable D7.2.</p>
<p>Task 7.3 Reasoning mechanisms for Criminal network reconstruction (Leader: ENG; Participants: EXPSYS, AIT; M7-M30)</p> <p>Illegal markets that appear in deep web and dark nets under different names are in most cases linked together, because criminal organisations behind them are the same. For this reason, this task deals with the reconstruction of such criminal networks by the application of advanced reasoning mechanisms for extracting hidden relationships on monitored sources, available knowledge about illegal trafficking activities and suspected people, and information extracted from relevant resources. These reasoning mechanisms will also involve the concepts of “smart crawling” and “smart analysis” to boost networks reconstruction: in case of lack of information during the execution of the reasoning mechanisms, new targeted crawling tasks and analyses will be automatically set up and started to search and/or extract the missing information within stored sources and resources. Such information will then be used to fully understand how criminal organisations and people are correlated and linked among them, from the point of view of both hierarchical (vertical) and matrix (horizontal) structures. Moreover, the task will focus on the detection of typical behavioural patterns and activities within and among criminal organisations and illegal trafficking activities, in order to identify similar behaviours in new sources and enable the early detection of new criminal networks (or new branches of those already known). This task contributes to deliverable D7.3.</p>
<p>Task 7.4 Knowledge based browsing, search and retrieval (Leader: ENG; M7-M30)</p> <p>This task will develop advanced semantically based service for search and retrieval, assisting analysts in any stage of the retrieval process. A multilingual virtual assistant will guide users during the query formulation phase, providing them semantics-aware suggestions. The service will find out context-aware information throughout the query expansion performed by using lexical-semantic relations according to the user request. Performances will be enhanced by involving query expansion mechanisms and knowledge-based recommendations. Furthermore, natural language understanding will be also integrated to answer queries expressed in natural language with targeted and punctual results (opposed to keyword-based search) to assist users in the reconstruction of illegal trafficking phenomena. To this end, the semantic representation and indexing of data will be specifically studied to address illegal trafficking domain. Such services will be included in the knowledge management framework that will rely on ANITA Knowledge Base, which will be able to manage different kinds of information through the integration of multiple storage systems (RDFs, graph-based and document-oriented), and to support multimodal annotations over stored data. This task contributes to deliverable D7.4.</p>
<p>Task 7.5 Extraction of evolving knowledge from deep neural network representations (Leader: CERTH; M7-M30)</p> <p>This task will focus on identifying prominent data relations in deep neural networks, in order to enable the extraction of new and evolving knowledge. Both multifaceted unstructured and semi-structured information will be considered, while taking into account the neural network internal representations. The research goal will be to represent the extracted knowledge in uniform machine understandable formalizations (e.g. propositional knowledge bases, ontologies, etc.), while particular attention will be paid for expressing uncertainty and probability. The overall analysis will rely on the mathematical tools stemming out from graph theory, set theory and machine (deep) learning. Additionally, the task will integrate reasoning techniques for ontology verification, i.e. to automatically assess and rectify the produced knowledge, as well as for validating it against existing knowledge bases and towards its compliance to the world it models, through semantic inference in real-world tasks. Key characteristic of the developed framework will be its capacity to learn novel hypotheses from the observed data and to infer new beliefs based on these hypotheses. This task contributes to deliverable D7.5.</p>
<p>Work package deliverables</p>

D7.1 – ANITA Knowledge modeling and representation [M15]

D7.2 – ANITA services for linking and tracking illegal trafficking activities [M30] (Internal release [M18])

D7.3 – ANITA Knowledge mining services [M30] (Internal release [M18])

D7.4 – Search and retrieval services [M30] (Internal release [M18])

D7.5 – Knowledge extraction from neural networks [M30] (Internal release [M18])

3.1.2.8 WP8 – Integration of human factor in the analysis loop

WP Number	8	Leader:	IBEC	Start month:	4	End month:	30		
WP Title	Integration of human factor in the analysis loop								
Participant number	2	6	4	8					Total
Particip. short name	CERTH	IBEC	EXPSYS	SYSTRAN					
Person/Months	18	32	1	1					52

Objectives

This WP aims to integrate the implicit and explicit responses of the human user in the analysis process by developing an adaptive user modelling framework able to interpret the states of the user and use it to iteratively refine the system's Big Data analytics services and enable rapid knowledge transfer among different users of the system. The following specific objectives are targeted: to develop a model of the current user's cognitive states and performance from a combination of explicit/implicit features obtained while operating with the system and based on a theoretical framework of learning and development; to incorporate the outcome of the user model into deep learning representations; to enable human interaction with the system in machine understandable form (e.g., in the form of free text); to transfer the collected knowledge on the particular application domain to the new and non-experienced users. The captured user feedback and the user model interpretation will be used to improve the robustness and performance of the feature detectors and update the knowledge structures of the system.

Task 8.1 Implicit and explicit user feedback capturing (Leader: IBEC; Participants: CERTH; M4-M20)

This task will define the framework to capture, analyse and use data from the implicit and explicit interaction of the user with the ANITA system. First it will define the hardware needed for capturing reactions: i) an optical eye tracker attached to the screen to capture gaze and pupil size, ii) a camera to track head movements, body actions and face expressions, iii) the mouse of the computer to capture hand movements and actions in the interface. Secondly it will create a middleware software that integrates all the signals and adds high level interpretation such as speed of eye movements, cognitive load (from pupil size), fixations (from eye movements and from the mouse), and emotional state (from facial expressions and human motion). The MIDAS²² (Modular Integrated Distributed Analysis System) system will be used as an open source framework for integration of signals. Finally the signals will be stored to feedback the user model, contextualize with the information and to input the neuronal network. This task contributes to deliverable D8.1.

Task 8.2 Adaptive user modelling for cognitive states estimation (Leader: IBEC; M4 –M30)

The overall objective of the task is to develop and deploy an adaptive model of the human user able to infer the states and the performance of the user while interacting with the system. Low-level primitives (metrics) composed of task-related user actions (e.g., user's clicks, eye gaze position and eye gaze fixations, user presence, self-reports), outputs of the system (e.g., information returned by the system, etc.) and unobtrusive psychophysiological data (e.g., pupil dilation, facial expressions, body posture, etc.) will be translated into more abstract, high level descriptions derived from theoretical frameworks of learning and development (i.e., Distributed Adaptive Control theory and Zone of Proximal Development) which will constitute the basis of the adaptive user model/profiling. This task will concentrate on the definition of the state space to be explored and its main operationalization. The high level concepts allow for an evaluation of the challenges that the user is facing, the interest/attention and the actual performance in the task. These variables could enter a self-regulatory system (based on the DAC allostatic control regulation) that balances those concepts according to the individual to operate in an optimal range defined by the pedagogical framework. This task contributes to deliverable D8.2.

Task 8.3 Incorporation of conscious and subconscious user feedback in deep learning representations (Leader: CERTH; Participants: EXPSYS, IBEC, SYSTRAN; M4-M30)

The aim of this task is to incorporate the explicit and implicit human responses (T8.1) that are captured in real operational environments (e.g. gaze signal, facial expressions, click-through data, user emotions, user motion, etc.) in the actual analysis process (i.e. for reinforcing the purely data-driven methods developed in WP6), taking into ac-

²² <https://github.com/bwrc/midas/wiki>

count the developed adaptive user model (T8.2). For achieving this, the aforementioned conscious and subconscious user feedback information is going to be exploited, in addition to the multimedia data (visual, audio, text). In particular, the captured sensorial signals, their high-level interpretations and their analysis based on the user model will be incorporated in the developed deep learning representations (WP6 modules). More specifically, different neuro-physiologically and neuro-biologically grounded neural network architectures will be designed and implemented, where sensorial signals will be used for reinforcing the learning process. To this end, standard Convolutional Neural Network (CNN) architectures will be tailored to the sensorial signals, making use of adaptive human response guided loss functions. In addition, custom intermediate representation layers will be evaluated that either convolve user feedback information with the data activation maps (correlation layer) or boost certain features on corresponding human response areas. Furthermore, different types of deep auto-encoders (e.g. sparse, adversarial, etc.) will be explored in order to model and interpret lower and higher-level correlations among sensorial signals and multimedia information. This task contributes to deliverable D8.3.

Task 8.4 Knowledge transfer to new officers (Leader: IBEC; Participants: CERTH; M16-M30)

This task extends the framework that collects implicit and explicit user feedback with a new component that will use the expert's user model and will transfer the knowledge to novice users. The new component will act as a recommender system that will learn from the interaction with content and the collected feedback of expert user. The system will keep the history of expert users using a graph database that relates different actions and content with positive outcomes (i.e. discovery of new links in the data) with the possibility of creating different user profiles and strategies. The recommender system will then replicate those strategies when novice users interact with the ANITA system to maximize the outcome and reduce the learning curve. Operationalizing on the ZPD concept, the amount of assistance will be targeted to the specific needs of each user to reach the desired level of expertise. This task contributes to deliverable D8.4.

Work package deliverables

D8.1 – Implicit and explicit user feedback acquisition and first prototype of the user model [M20] (Internal [M12])

D8.2 – Adaptive modelling of user cognitive states [M30] (Internal release [M16])

D8.3 – Incorporation of user feedback in deep learning representations [M30] (Internal release [M16])

D8.4 – Transferring domain knowledge to novice officers [M30] (Internal release [M18])

3.1.2.9 WP9 – Novel applications for LEAs and System Integration

WP Number	9	Leader:	ENG	Start month:	7	End month:	33		
WP Title	Novel applications for LEAs and System Integration								
Participant number	1	2	4	5	6	8	9		
Particip. short name	ENG	CERTH	EXPSYS	AIT	IBEC	SYSTRAN	TIU-JADS		
Person/Months	40	2	4	3	5	2	2		

Objectives

This WP is responsible for planning, monitoring and managing the integration, deployment and validation activities towards the overall ANITA system. The main objectives of this WP are: to prepare and configure a Big Data infrastructure that constitutes the architectural skeleton of the final ANITA system; to plan the integration and the validation activities; to design and implement novel applications for supporting LEAs activities for preventing, investigating and mitigating illegal trafficking activities; to provide LEAs with Big Data visual analytics capabilities; to implement a complete chain-of-custody to maintain the integrity of stored data and deliver services for evidence export; to develop an iterative integration strategy, by including all services developed in other technical WPs; to validate and refine the integrated system through the inclusion of internal (in-lab) and external (end users) feedback and suggestions. A continuous integration and refinement approach will be followed to include all services developed from WP5 to WP9 and all feedback and suggestions received from the multiple validation stages (from in-lab validation to validation in relevant environments).

Task 9.1 Big Data infrastructure (Leader: ENG; M7-M15)

Starting from non-functional requirements and technological constraints identified in WP4, this task will deal with Big Data infrastructure preparation and configuration. The infrastructure will be compliant with the Big Data management principles, in order to be scalable and flexible, and will be based on a Service Oriented Architecture, including interfaces to facilitate the integration of services developed in WP5 to WP9 and the parallel implementation and testing of the individual services. The infrastructure will also support services orchestration and the execution of complex workflows needed for crawling, analysis, fusion and reasoning. A Privacy-by-Design approach

will be adopted in order to ensure compliance with legal and ethical constraints and regulations at both National and European level. Finally, if needed, the infrastructure will be subject to iterative refinements according to the outcomes of task 9.5 on validation phases, in order to include feedback from both technical and user-centred perspectives. This task contributes to deliverable D9.1.
Task 9.2 Novel applications for illegal trafficking (Leader: ENG; M7-M33)
This task aims at designing and developing novel applications to support and speed up LEAs prevention, investigation and mitigation processes for fighting illegal trafficking activities. Such applications will exploit knowledge stored into the ANITA Knowledge Base and relevant information extracted by the analysis tools and services implemented in WP5, WP6 and WP7. The envisaged applications will be equipped with user-friendly and interactive widgets able to fulfil LEAs' requirements and boosting investigative operations allowing officers to verify and validate Big Data analysis results, report events, browse and analyse illegal trafficking activities and potential links among them. These applications will make available useful suggestions to users coming from reasoning algorithms in order to provide insights for the reconstruction of temporal and causal sequence of events and activities in order to understand the specific criminal dynamics. Collaborative services to enable and improve active collaboration within and among LEAs will be integrated into the applications in order to coordinate investigative teams against the selected crimes. This task will provision early mockups of the applications to be validated by LEAs involved in the project by M12. This task contributes to deliverables D9.2 and D9.3.
Task 9.3 Advanced Big Data Visual Analytics (Leader: ENG; M13-M33)
The main objective of this task is to design and implement advanced visual analytics services to represent patterns and trends of online illegal trafficking crimes. An innovative analytical framework will be defined starting from the analysis of the most advanced visual analytics models and paradigms. Specific visualisation metaphors will be studied and implemented to improve LEAs capabilities in understanding the dynamics of the phenomena, including geo-spatial distribution of the illegal trafficking activities, main causes, modus operandi, countermeasures adopted to elude LEAs, channels used for advertise and sell counterfeit products, NPS, drugs and firearms. Visual Design Patterns will be applied in order to deliver user-friendly dashboards to effectively present of different aggregation views and perspectives. Interactive widgets for big data visualization will support investigation activities and will help LEAs during the decision-making processes. Furthermore such widgets will provide users with analytics on-demand services. According to the Visualization-as-a-Service (VaaS) approach, services for customizable Big Graph perspective views will be implemented grounded on search and retrieval mechanisms and on analytics results. Such widget will allow users to compare results of multiple analytics highlighting specific correlations. This task contributes to deliverable D9.3.
Task 9.4 Chain of custody and evidence export (Leader: ENG; M13-M33)
This task will provide the ANITA system with capabilities for enabling chain of custody and export of evidence. All phases of the information workflow will be analysed in order to define the chain of custody architecture, which will ensure the integrity of information stored into the system, according to the guidelines and best practices provided by LEAs. Mechanisms to ensure data protection, according to the access control policies and taking into account legal, ethical and privacy regulations and constraints will be defined and implemented. Maintenance of integrity will be ensured following specific documented policies. If data integrity is correctly respected, users will be able to export stored information as evidence to use in court, with all its chain of custody history, or to be shared with other LEAs. For this purpose, this task will also study standards and formats for information exchanging among LEAs. This task contributes to deliverable D9.3.
Task 9.5 Continuous System Integration (Leader: ENG; Participants: CERTH, EXPSYS, AIT, IBEC, SYSTRAN, TIU-JADS; M7-M33)
This task is related to the adoption of a Continuous Integration/Continuous Delivery (CI/CD) approach, in order to speed up and automatize testing and integration of services and deployment of stable versions of ANITA system. A Source Code Repository (SRC) will be set up to allow developers to share their code at regular intervals in a unique repository. An automated building system, which downloads the source code from SRC, performs all tests and finally executes the deployment of the new stable version of the code, will be also included. In this way, possible errors can be detected and fixed immediately, and the new software version can be released at any time because ensured to be stable (since all tests have been executed successfully). Integration will be divided in three stages: the first stage (M18) will be related to in-lab validation, whereas the second (M24) and the third (M33) ones will be related to the validation results of pilots' executions. This task contributes to deliverables D9.4, D9.5 and D9.6.
Task 9.6 System validation (Leader: EXPSYS; Participants: ENG, CERTH, AIT, IBEC, SYSTRAN, TIU-JADS; M16-M33)

This task regards the validation of the ANITA system at different integration phases, in order to guide system development from TRL-3 (at M12) to TRL-4 (first integration stage at M18) for in-lab validation, to TRL-5 (second integration stage at M24) for first pilots' execution, and finally to TRL-6 (third integration stage at M33) for final pilots execution. A specific validation plan will be properly defined and executed and some validation indicators will be identified in order to provide concrete KPI for evaluating the improvements reached at each integration stage and those required for the next integration stage. This task contributes to deliverables D9.7 and D9.8.

Work package deliverables

D9.1 – ANITA Big Data Infrastructure [M15]

D9.2 – ANITA mock-ups [M12]

D9.3 – Novel applications for illegal trafficking and chain of custody [M33] (Internal release [M18],[M24])

D9.4 – ANITA Integration Plan [12]

D9.5 – ANITA System 1st release [M24] (Internal release [M18])

D9.6 – ANITA System 2nd release [M33]

D9.7 – ANITA Validation report 1st release [M24] (Internal release [M18])

D9.8 – ANITA Validation report 2nd release [M33]

3.1.2.10 WP10 – Demonstration in relevant environment

WP Number	10	Leader:		DITSS	Start month:	10	End month:		36	
WP Title	Demonstration in relevant environment									
Participant number	1	2	3	4	5	6	7	8	9	Total
Particip. short name	ENG	CERTH	RISSC	EXPSYS	AIT	IBEC	IIP	SYSTRAN	TIU-JADS	
Person/Months	4	4	4	4	4	4	2	2	2	
Participant number	10	11	12	13	14	15	16	17		
Particip. short name	DITSS	ISBR	KWPG	AoC	CAST	NPN	GDCOC	LPV		
Person/Months	11	7	8	10	6	5	6	4		87

Objectives

This WP aims at defining and executing pilot scenarios for the validation and demonstration of the overall ANITA system in relevant environment. The following specific objectives are targeted: to define the pilot scenarios; to prepare and plan the execution stages; to perform the validation and the demonstration of ANITA system in relevant environment; to evaluate and assess the results from pilot executions and provide feedback to technical partners for refinements and/or improvements of the system.

Task 10.1 Pilot scenario definition (Leader: CAST; Participants: ENG, CERTH, RISSC, EXPSYS, AIT, IBEC, IIP, DITSS, ISBR, KWPG, AoC, NPN, GDCOC, LPV; M10-M15)

This task deals with the definition of scenarios for ANITA system validation. Starting from use cases defined in WP4, each scenario will be specified with the following information: storytelling, initial datasets, involved actors, system functionalities that will be involved and benefits that are expected from the use of the ANITA system. Such benefits will constitute the basis for system evaluation and assessment in task 10.4. This task contributes to deliverable D10.1.

Task 10.2 Pilot preparation and planning (Leader: DITSS; Participants: ENG, CERTH, RISSC, EXPSYS, AIT, IBEC, SYSTRAN, TIU-JADS, ISBR, KWPG, AoC, CAST, NPN, GDCOC, LPV; M13-M21)

In this task, detailed plans for the validation and demonstration of ANITA system will be defined. Plans will describe actions required for the preparation and execution of pilots, datasets will be used, tuning of ANITA services for the purposes of the pilots, role of each actor, infrastructure to set up, potential involvement of external people and organisations. Pilot plans might be updated according to the needs emerging from the previous validation phases. This task contributes to deliverable D10.2.

Task 10.3 Pilot execution (Leader: DITSS; Participants: ENG, CERTH, RISSC, EXPSYS, AIT, IBEC, SYSTRAN, TIU-JADS, ISBR, KWPG, AoC, CAST, NPN, GDCOC, LPV; M22-M27; M31-M36)

Starting from the plans defined in task 10.2, this task will focus on managing and monitoring the stages of pilots' execution. LEAs will be directly involved in pilot execution and will provide opinions and suggestions on its benefit, performance and usability. During the execution, feedback will be collected following the methodology defined in task 10.4. This task contributes to deliverable D10.3 and D10.4.

Task 10.4 Evaluation and assessment (Leader: AoC; Participants: ENG, CERTH, RISSC, EXPSYS, AIT,

IBEC, IIP, DITSS, ISBR, KWPG, CAST, NPN, GDCOC, LPV; M22-M27; M31-M36)
This task deals with the definition of a methodology and the implementation of the related protocol for the collection of feedback provided by end users during pilots' execution and for the evaluation of the quality, relevance and benefits of ANITA system. Starting from the use cases and pilot scenarios, a set of quantitative and qualitative indicators will be identified in strict collaboration with LEAs and will be adopted as evaluation criteria. After the pilots' execution, collected feedback will be analysed in terms of usability, performances, accuracy, user acceptance and strategic and socio-economic impact of ANITA solution on LEAs relevant environments. The results of this analysis will constitute the inputs for technical partners for the refinement and improvement of ANITA system, services and applications. This task contributes to deliverables D10.5 and D10.6.
Work package deliverables
<i>D10.1 – Pilot scenarios [M15]</i>
<i>D10.2 – Pilot plans [M21]</i>
<i>D10.3 – Pilots execution report 1st release [M27]</i>
<i>D10.4 – Pilots execution report 2nd release [M36]</i>
<i>D10.5 – Pilots evaluation and assessment 1st release [M27]</i>
<i>D10.6 – Pilots evaluation and assessment 2nd release [M36]</i>

3.1.2.11 WP11 – Dissemination, Training and Exploitation

5.1.2.11 WP11 – Dissemination, Training and Exploitation										
WP Number	11	Leader:		AoC	Start month:		1	End month:		36
WP Title	Dissemination, Training and Exploitation									
Participant number	1	2	3	4	5	6	7	8	9	Total
Particip. short name	ENG	CERTH	RISSC	EXPSYS	AIT	IBEC	IIP	SYSTRAN	TIU-JADS	
Person/Months	7	4	8	6	3	4	1	3	3	
Participant number	10	11	12	13	14	15	16	17		
Particip. short name	DITSS	ISBR	KWPG	AoC	CAST	NPN	GDCOC	LPV		
Person/Months	6	3	5	11	5	3	4	3		79

Objectives

This WP deals with dissemination, exploitation, training and community building activities. Best practices and training curricula will be defined and used in workshops and demonstrations, in order to allow trainees to acquire the specific knowledge and competencies to get benefits from ANITA system. Main objectives are: To develop a dissemination; to establish a stakeholder community in order to benefit from its experience throughout the ANITA project; to identify best practices, develop tailored training curricula and set up training activities to develop a strategy and prepare a plan for the exploitation of results of ANITA project.

Task 11.1 Dissemination and communication (Leader: EXPSYS; Participants: All partners; M1-M36)

The task will develop and promote means to raise awareness among interested parties and communities that shall be impacted by ANITA outcomes. Dissemination and communication activities will focus on: defining techniques and media as appropriate for fostering project results; targeting specific audiences that can benefit from the ANITA results; showing the project results to the relevant stakeholders; participating to relevant initiatives in order to guarantee a wide visibility of the project objectives and outcomes. This task will also create and keep updated the dissemination material and project branding: ANITA logo, tagline, ANITA public and internal website, slide and paper templates, merchandise for events, social network presence. A dissemination and communication plan will be established at the beginning of the project, in order to provide all partners with guidelines for maximizing project visibility into their communication channels and networks.

This task contributes to deliverables D11.1, D11.2, D11.3 and D11.4.

Task 11.2 LEAs and stakeholders community engagement and management (Leader: DITSS; Participants: ENG, RISSC, TIU-JADS, ISBR, CAST; M1-M36)

This task aims to support the ANITA user community, both the pre-existing users as well as reaching out to more users through the establishment of a network of LEAs, public entities and private organisations, in order to benefit from the experiences of a community of contacts in this area. To achieve these objectives, the following actions will take place: (i) Desk-based research: targeted review of sector reports, organisation websites and social media platforms to identify additional relevant stakeholders and individuals in this area; (ii) Contribute to the communication strategy to be implemented in this work package based on invitations sent to stakeholders to join the

ANITA network and on implementation of communications campaign throughout the project including, but not limited to, surveys and other research tools. This network will facilitate activities both in this work package, to generate information on best practices (T11.3) and training activities (T11.4), as well as other work packages, particularly for criminological analysis (WP2), use cases and requirement elicitation (WP4). This task will be responsible of the management of Advisory Board area. A common platform will be also available to promote community building among LEAs and relevant public/private stakeholders. This task contributes to deliverable D11.5.
Task 11.3 Best practices and guidelines for curricula (Leader: CAST; Participants: CERTH, RISSC, IBEC, TIU-JADS, KWPG, AoC, NPN, GDCOC, LPV; M18-M36)
This task aims at promoting the development of educational paths oriented primarily to LEAs and representatives from public and private sectors in order to increase awareness and knowledge, but also exchange of good/best practices and experiences. Existing gaps between current practices and LEA needs will be analysed and exploited to define ad-hoc training curricula for the setup of qualified officers with expertise in illegal trafficking area. Furthermore, this task will focus on the optimisation of the use of resources in administrations/enforcement units dealing with the considered crimes, by introducing in the LEA operational activities the trained key figures when setting up a team/taskforce for counteracting or investigating this kind of crime. The curricula definition process will consider the existing experiences in Europe, and will be part of the discussion in the interactive training activities (T11.4), prior to execution of the demonstrations, with the goal of gathering feedbacks useful for the improvement of the proposals. Starting from the crucial role of research in the knowledge generation process, this task will try also to outline possible suggestions for Universities in order to set up specific professional courses on the issues of interest for the project. This task contributes to deliverable D11.6.
Task 11.4 Training activities (Leader: AoC; Participants: ENG, CERTH, EXPSYS, AIT, IBEC, SYSTRAN, KWPG, CAST, NPN, GDCOC, LPV; M22-M36)
This task aims at organising training activities for officers of different LEAs and other relevant stakeholders, in order to equip them with comprehensive knowledge and effective skills to recognize and address illegal trafficking activities and to facilitate cooperation among LEAs. Such capacity-building activities will be supported by the introduction of ANITA system as tool to boost knowledge acquisition processes and expertise reinforcement. Training activities will be realised in the form of workshops, webinars, professional courses and face-to-face meetings, at individual-level (individual expertise and competency achievement), institutional-level (support in defining organisational structures and operational methods to handle effectively illegal trafficking investigations), and societal-level (guidelines for enabling cooperation among institutions and for adopting tailored mitigation measures). Training activities will be based on the training curricula defined in task 11.3 and might be subject to changes whenever some refinements are applied to such curricula. This task contributes to deliverable D11.7.
Task 11.5 Organisation of Workshops (Leader: AoC; Participants: ENG, CERTH, RISSC, EXPSYS, DITSS; M23-M36)
This task deals with the organisation of two ANITA workshops oriented to LEAs and stakeholders involved in monitoring of illegal trafficking and organised crime activities. Such workshops are aimed to increase LEAs background and awareness on illegal trafficking crimes and at the same time to raise and engage interest on the potential of ANITA system among LEAs as tool to improve their work. To this end, some showcase sessions of ANITA system will be arranged, supported by dissemination material prepared for training activities and distributed among participants. This task contributes to deliverables D11.8 and D11.9.
Task 11.6 Exploitation (Leader: ENG; Participants: CERTH, RISSC, EXPSYS, AIT, IBEC, SYSTRAN; M13-M36)
This task is designed to address all the issues related to the exploitation strategy of the ANITA results. In particular, it conducts the market analysis composed of identification of the national and international target markets, market segmentation, analysis of user needs and preferences, as well as competition to ensure that real needs are better addressed and best possible solution developed targeting relevant stakeholders and potential business partners in the most appropriate ways. Moreover, in order to identify the most appropriate approach for the exploitation of ANITA tangible and non-tangible results, this task will explore relevant business models addressing the complete value chain of stakeholders involved. Business models, based on the CANVAS Business Model, will be defined taking into account the possible different modality of service offering towards end users and related adoption and pricing models. Finally, this task will identify ANITA Key Exploitable Results (KER), formulate the exploitation strategy both individual to each partner and jointly for the consortium as a whole, and then elaborate ANITA Exploitation Plan. Analysis of ownership and governance aspects related to the exploitation of project results will be made, and the IPR management strategies and plan for future exploitation will be established based on the initial IPR plan (D1.2) delivered in WP1. This task will contribute to deliverables D11.10 and D11.11.

Work package deliverables

D11.1 – ANITA Website [M3]
 D11.2 – Dissemination and communication plan [M6]
 D11.3 – Report on Dissemination Activities 1st release [M18]
 D11.4 – Report on Dissemination Activities 2nd release [M36]
 D11.5 – LEAs and stakeholder community building and management [M36]
 D11.6 – Guidelines for curricula definition [M36]
 D11.7 – Training materials [M34]
 D11.8 – 1st Workshop Report [M27]
 D11.9 – 2nd Workshop Report [M36]
 D11.10 – Market Analysis and Initial Exploitation Plan [M24]
 D11.11 – Market Analysis and Final Exploitation Plan [M36]

3.2 Management structure, milestones and procedures

This section describes the management structure and procedures defined by ANITA consortium to ensure successful project execution and delivery of results, as well as collaborative communication between all partners.

3.2.1 Organisational structure and governance scheme

The ANITA integration management structure has been designed to ensure that the project attains the established project objectives and, to the best possible level, fulfils partners' individual expectations, maintains full process quality assurance, takes advantage of using proper monitoring and self-assessment procedures, incorporates the monitoring of alignment to legal, ethical, security and privacy guidelines as defined by WP3. The project is managed at four different levels, with four different committees:

- Strategic and vision management: **General Assembly**;
- Administrative, financial, overall coordination: **Project Management Committee**;
- Scientific and Technical and content coordination: **Technical Management Committee**;
- User community coordination: **User Community Committee**;
- Data protection management: **Data Protection Office**.

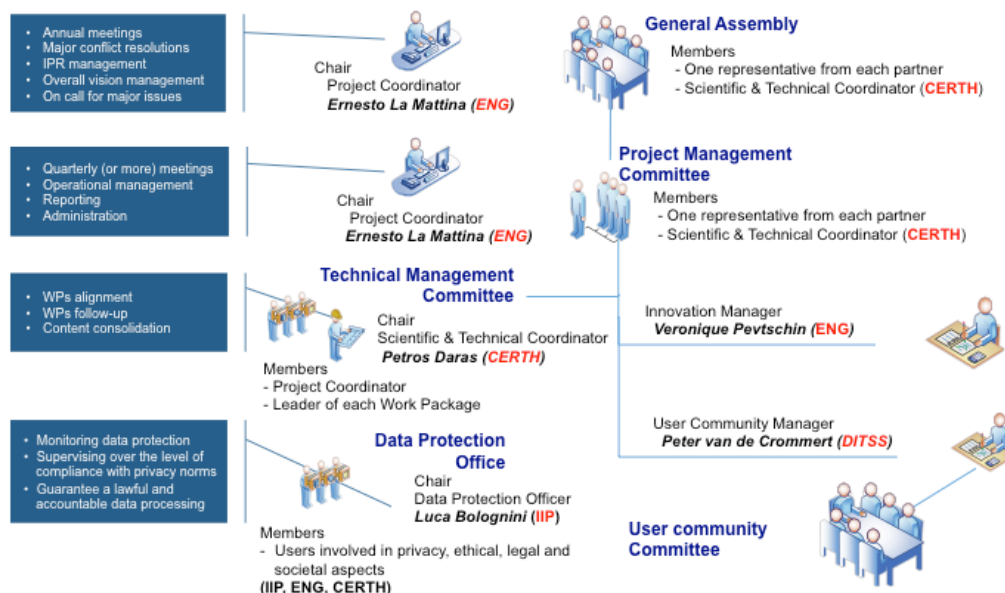


Fig. 7: ANITA Management structure

These committees meet at timely intervals (face-to-face, telephone or video conference). This organisational structure has been agreed to optimise planning, monitoring and coordination of ANITA project activities & tasks performance, reporting and accountability. ANITA will be steered and guided in a top-down fashion, while issues are to be raised and problems solved in a bottom-up manner. **Error! Reference source not found.** summarizes the anagement structure and the associated leaders. This is complemented by the innovation management support, with

the link being implemented through the TMC. The ANITA Project Management structure has been agreed among the partners and is based on a similar model already implemented with success in past EU projects.

3.2.2 Innovation management

ANITA will deploy best practice innovation management methods to help output new and improved products, processes, methodologies and standards closely oriented to market needs. The project will connect with customers throughout the value chain and with emerging technical developments to create insights. The workplan for the ANITA project have explicit provision for creating ideas, prioritising them and evaluating scientific and technical feasibility mainly in WP11, *Task 11.6: Exploitation*. Innovation Management will be coordinated by an Innovation Manager and carried out day to day by exploitation managers identified in each beneficiary organisation, with oversight by the Technical Management Committee (TMC). Exploitation managers are responsible for ensuring capture of any relevant ideas created during their project work, prioritising and developing these to a point where they can present them to the TMC for a decision about inclusion within existing work-packages and any associated changes to the delivery plan.

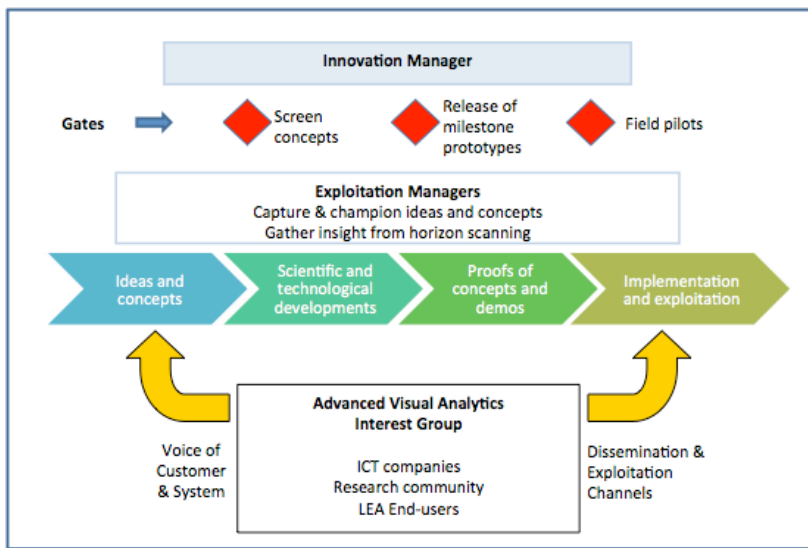


Fig. 8: Overall innovation process operating across ANITA

The innovation management activities will be performed by partners in WP11.

To ensure that relevant external or internal opportunities are identified and responded to, all consortium members will contribute to exploitation meetings (coinciding with TMC meetings) with a horizon scanning report, which will be important in informing development decisions throughout the project and will draw on inputs from the Market interest group. The diagram illustrates the overall innovation process operating across ANITA. Champion ideas and concepts, scientific and technological developments, and proofs of concepts and demos are challenged by the TMC at checkpoints called Gates, which are activated to ensure appropriate levels of rigour in technical and commercial feasibility assessment and risk management. Exploitation managers will act as champions at these gates, proposing targeted development activities.

To support connection to market stakeholders, in the context of Task 11.6, an inclusive Counter Illegal Trafficking Interest Group will be established with members representing all functions involved in LEAs, research, and ICT solution providers. This group will serve to provide insight for early stage creativity and on-going development, validation or design principles and finally provide channels to support exploitation and dissemination of outputs. This group will act in part as a focus group but will also provide connectivity to the full range of stakeholder groups to enable deployment of best practice “voice of customer” methods such as lead user analysis. By including representatives of the SME supply and research communities this group will also provide “voice of system” inputs that help identify technical opportunities for improvement.

3.2.3 Key Roles and Committees

The **Project Coordinator (PC)** will ensure the overall co-ordination and daily supervision and monitoring of operations and official communication with the European Commission and with other parties (e.g. relevant EU and national projects, relevant bodies and other related activities). He has the overall responsibility for the financial and contractual obligations with the Commission. The PC (with the support of the PMC) is responsible for preparing the reports, cost statements and project documents required by the EC, as well as for obtaining Certificates on Financial Statements and bank guarantees (if needed) from the consortium members and for distribution of all payments received from the EC. The PC will also oversee the promotion of gender equality in the project, organise EC review meetings, coordinate knowledge management and represent the consortium in events. The PC (with the support of IPR experts) will be responsible for making sure that IPR terms and conditions specified by the Consortium Agreement are properly interpreted and adhered to. The Project Coordinator is Ernesto La Mattina from ENG. ENG will provide a Project Office (PO) comprising staff familiar with administrative, legal, financial, communication and IPR issues to support the operational groups in all their responsibilities. The PC will be assisted by the PO.

The **Scientific & Technical Coordinator (STC)** has the overall responsibility of ensuring content synchronisation between the different ANITA outcomes, as well as alignment to the overall ANITA goals. The STC is Petros Daras from CERTH.

The **User Community Manager (UCM)** is the central contact point for users, and is responsible for the clarity and quality of information and interaction with users. The UCM is Peter van de Crommert from DITSS.

The CVs of the coordinators can be found in section 4.

The **General Assembly (GA)** is the highest decision-making body of the project. It is responsible for major administrative decisions that require a vote; it is the deciding body on potential disputes and also manages IPR issues. Each partner is represented with one person in the GA, which will be chaired by the PC. It will meet at least every 10 months and can be called to extraordinary session at the request of the Project Management Committee (PMC).

The GA will be involved in the overall plans, the alleviation of bottlenecks, the resolution of conflicts, general monitoring and control of progress of the project, the management of resources and the acceptance on the integration or withdrawal of partners and personnel. Most issues of the GA will be decided by votes (one vote per partner). The whole GA's role, responsibilities, rules, and decision-making procedures will be extensively detailed in the Consortium Agreement (CA).

The **Project Management Committee (PMC)**, led by the Project Coordinator, is the executive body of the project and will meet face-to-face or by teleconference at least every three months, for the following tasks: approve the financial plan and work plan; evaluate the progress; discuss any management issues; resolve sensitive technical, administrative or contractual issues; monitor coherence and integration in the project; identify and analyse potential risks; discuss/approve changes in the consortium; take decisions on proposed changes to CA; conflict resolution.

PMC will comprise one representative from each partner, the PC and the STC. The PMC reports to GA.

The **Technical Management Committee (TMC)** has the role to ensuring the timely progress of the project and the high quality of the results. The TMC reports to the PMC. It will meet (face to face / teleconference) monthly at the end of the project to review the success of the project. It will permanently stay in contact, mainly using e-mail and regular teleconferences. Its responsibilities are: reviewing the work plan and the overall project progress; approving the reference architecture of the ANITA system; defining responsibilities for components; determining conventions to be followed; following up the timely achievement of milestones and deliverables; managing the day-to-day activity of the WPs; coordinating the integration according to the work plan; reporting to the PMC about the progress; coordinating the preparation/distribution of deliverables; maintaining a quality plan; analysing the technical environment to assess risks to project impacts; forming ad-hoc Technical Task Forces, if necessary. The TMC will be chaired by the STC. The leaders of all work packages will be members, along with the PC. The committee activities are supported through WP1.

The **User Community (UC)** has the assignment of ensuring clarity and consistency in communicating with the users (LEAs) involved in the pilots and validating feedback information flowing between pilots and between pilots and design phases. The UC reports to the PMC. It will meet every three months. Its responsibilities are: reviewing the pilots approach and deployment plan; reviewing the feedback from each pilot phase and presenting it to the GA and to the TMC as needed; monitoring pilots results and ensuring consistency across reporting approaches from trials; highlighting dissemination and training opportunities towards users. The UC will be chaired by the UCM.

The committee activities are supported through WP10 and WP11.

Work package leaders: Each work package will have a **work package leader (WPL)**, who will have the responsibility and be the contact person for that work package (WP). Responsibilities include organising and managing the work within the WP, detailed planning of tasks and activities, and the control of results from the activities. In addition, they will provide deliverables and other reporting required. The WP team will communicate regularly (face-to-face or teleconference). WP leaders take responsibility for presenting to the TMC the progress and successful completion of the work within the ANITA project.

The **Security Advisory Board (SAB)**, chaired by the Project Security Officer, will consist of consortium members with sufficient knowledge of security issues related to project resources as well as datasets, and resources extracted from Surface Web, Deep Web and Dark Nets and have the main responsibility to assess the proper execution of security and protection measures and procedures. SAB will monitor regularly the project and assess the sensitivity of all information handled by participants and all deliverables. SAB will report to the PC. SAB responsibilities include validate, refine and assess the security and protection measures, review project deliverables and resources and support Consortium in the correct use of the resources.

3.2.4 Management Issues and Procedures

Consortium Agreement. Prior to the start of the project, the PC will ensure that all partners have signed the project "Consortium Agreement" (CA). This document will set out all the internal rules of the project and will be signed

by all partners. The CA will include the Conflict Resolution and Relationship breakdown, IPR Management, etc. The GA will control the respect of IPR rules and information dissemination procedures defined in the CA. The obligations and rights of the participants will be detailed in the Consortium Agreement. This Consortium Agreement will make explicit reference to decision procedures, risk management strategies, the right of each partner in the exploitation of results, equal opportunities and gender equality policies.

Continuous planning. Over the first three months of the project start, a consolidated and detailed plan will be produced and maintained by the PC with support of the STC and the PMC. A detailed work plan will be produced and continuously updated (at least every 2 months) within each WP, to ensure all tasks are covered.

Quality Management. Within ANITA, a Quality Assurance System will be established in the earliest stages of the project, with a Project Quality Plan to be delivered at Month 3, to be used internally by the Consortium. It will describe the guidelines adopted by the project on preparation and validation of deliverables, internal peer reviewing, periodic reporting, preparation of financial statements, as well as risk management.

SELP Compliance Management. Within ANITA, privacy, security, ethical, usability and legal monitoring has been introduced as part of the management activity (WP3). This monitoring is a compliance management process through which selected deliverables are reviewed to ensure their alignment to the different of WP3.

Deliverables. Each work package leader is responsible for the organization of the deliverables for his work package, including internal review of all reports. Once a report has passed internal work package review, it must pass project-level review. The PC and the STC may review any deliverables before submission and may return them for additional refinement if they feel this necessary.

Reporting and communication. In addition to the deliverables, there will be an internal reporting mechanism and annual cost statement financial report to be submitted to the European Commission. *Every two months*, each WP leader will write and send an internal management report to the PC, by collecting contribution from partners in the WP. The report will describe the technical and management work done. *Every 6 months*, each partner submits a report detailing progress and effort expenditure. These reports are collected by the PC. Each WP leader will in addition answer a questionnaire, structured in three major sections: (1) assessment of the work done; (2) key issues for the development; (3) on-going results of evaluation indicators. Based on the questionnaires, ENG will analyse all the self-assessment questionnaires and summarise the overall project progress, to be presented as reports. *At M12, M24 and M36*, each partner prepares a financial statement, which is collected and presented to the EC by ENG. The self-assessment questionnaires mentioned in previous point are used to generate the annual project management report to the EC and to be presented also in the General Assembly of ANITA.

Tools and Services. Several tools and services will be made available to the ANITA partners to stimulate their co-operation. A *web based internal communication and collaboration tool* will be introduced and managed by ENG. This will include: a web-based tool with private access for consortium members, containing all related documents, a news board and information exchange; a collaborative Wiki; dedicated mailing lists. This web-based tool will simplify the collection of information across different teams and work packages in an easy manner. This tool is included in the report Project Reference Manual.

Conflict Resolution. When a conflict occurs in a WP, consensus should be sought to solve the problem. If the problem cannot be solved, the WP Leader prepares a description of the problem and its possible solutions, and transmits it to the TMC. If consensus cannot be reached within the TMC, the Project Leader escalates it to the PMC and a vote occurs, requiring a simple majority. Ultimately, any unresolved dispute is referred to the GA who, in the event that an amicable agreement cannot be reached within that body, will have the right to appoint external independent arbitrators.

IPR Management. In ANITA we expect that all project partners agree on explicit rules concerning IP ownership, access rights to any Background and Results for the execution of the project and the protection of intellectual property rights (IPRs) and confidential information before the project starts. Therefore, in order to properly address such issues, the CA (based on DESCA model) between all project partners established a legal framework aimed at providing clear regulations for issues (within the consortium) on the work, IP-Ownership, Access Rights to Background and Results for the duration of the project and any other matters of the consortium's interest.

Meetings. The project will be launched by a consortium plenary kick-off meeting. GA is expected to meet annually. Intermediate meetings and conference calls are held by the PMC. It is anticipated that working meetings will be held on demand as face-to-face meetings or as virtual meetings. The consortium may decide to support a series of joint meetings to speed up the development and integration process across the ANITA development teams of the different partners, as well as to foster joint effort and to support the involvement of users in ANITA. The detailed travel costs information (below in Section 3.4) describes, on a per partner basis, the foreseen meetings.

3.2.6 Milestones

Milestone num.	Milestone name	Related WPs	Est. date	Means of verification
MS1	Project successfully started and preliminary use cases analysis performed	WP1, WP11	M3	Project Reference Manual (D1.1), and Project Quality Plan (D1.5) completed. Project Public Website (D11.1) available. Preliminary analysis of use cases and scenarios performed (I4.1)
MS2	Use case analysis and user requirements elicitation completed	WP4	M6	D4.1 completed.
MS3	System specification completed	WP4	M9	D4.2 completed.
MS4	Analysis of the European data protection framework	WP3	M10	Requirements from European data protection framework completed (D3.1).
MS5	Main system concepts analysed	WP2, WP3	M12	D2.1, I3.2, I3.5 completed
MS6	First version of ANITA system architecture defined	WP4	M12	D4.3 completed
MS7	ANITA mock-ups validated by LEAs	WP2, WP4, WP9	M12	D2.1, I3.2, I3.5, D4.3 and D9.2 completed
MS8	First version of data source and data stream analysis services ready to be integrated for validation in Lab	WP5	M15	D5.2 and D5.4 completed.
MS9	First version of the ANITA Big Data analysis and analytics services ready to be integrated for validation in Lab	WP6	M15	I6.1, I6.2, I6.3, I6.4, I6.5 completed
MS10	First integration completed for validation in lab	WP7, WP9	M18	First version of knowledge mining services (I7.2, I7.3, I7.4, I7.5) and applications (I9.3) integrated (I9.5).
MS11	ANITA prototypes validated in lab	WP9	M18	Internal reporting on system validation in lab (I9.7)
MS12	Final version of system architecture defined	WP4	M21	D4.4 completed
MS13	ANITA services and applications ready for validation in relevant environment	WP9	M24	Updated versions of applications (I9.3) ready for the second integration phase (D9.5).
MS14	ANITA prototypes validated in relevant environment	WP9, WP10, WP11	M27	D10.3 and D10.5 completed
MS15	Final version of data source and data stream analysis services	WP5	M27	D5.3 and D5.5 completed
MS16	Final version of the ANITA Big Data analysis and analytics services delivered	WP6	M30	D6.1, D6.2, D6.3, D6.4, D6.5, D8.2, D8.3, D8.4 completed
MS17	Final version of the knowledge management services	WP7	M30	D7.2, D7.3, D7.4, D7.5 completed
MS18	Final version of ANITA services and applications ready for demonstration in relevant environment	WP9	M33	D9.3 and D9.6 completed
MS19	ANITA system demonstrated in relevant environment	WP9, WP10, WP11	M36	Pilot execution, evaluation and assessment (D10.4, D10.6). Training to LEAs (D11.7) completed.

Table 3.2.a : List of milestones.

3.2.7 Risk Management

The ANITA project management team will perform continuous evaluation throughout the project, identifying any possible problems/risks at an early stage so that solutions can be elaborated in time. Specifically, effective control and management of the various types of risks will be performed in order to successfully accomplish the project objectives on time and according to budget. A first list of risk factors together with their contingency plans is reported in Table 3.2.b. A simple but state-of-the-art Risk Management Plan will be developed and implemented as an internal deliverable within the first three months. It will build upon the risks identified, meaning that it will identify the risk of any nature that might occur in the project; assess the likely severity of each risk and its potential impact on the project; assess the potential probability of the risk factor; identify the measures that may be necessary, if relevant, to offset or prevent the occurrence or limit the impact of that risk; and identify the measures that may be necessary, if relevant, to minimise the impact of the risk should it nevertheless occur. The accuracy of identified risks will be reviewed quarterly and the plan will be improved and completed accordingly.

Description of risk		WP Involved	Counter-measures
Managerial	A Partner withdraws or is unable to provide a foreseen contribution	WPs in which that partner is involved	A strong and thorough Consortium Agreement before the Grant Agreement ensures keeping original commitments on track and managing such unlikely situation professionally. The team composition of the consortium ensures a balance of skills. The very experienced Project Management team is committed to tightly monitor the progress of each partner and promptly react where required. In the unlikely case of all measures failing, the consortium will consider bringing a new partner into the consortium, upon consultation with the EU Project Officer. The consortium members have a deep knowledge of the stakeholders in the field, and can easily find a replacing partner.
	Lack of coherence in project development and lack of cooperation among partners	All WPs	Most of consortium's members are very familiar and experienced in working with one another thanks to collaborations in previous projects. The careful selection of partners was based on this consideration too. In addition, the tight monitoring of each partner's progress by the strong project management team will ensure coherence in the overall project. The existence of four leads (one managerial, one technical, one for innovation, and one for end users) will help in this.
	Project overspending in travel or equipment	All WPs	Managerial and progress review meetings will coincide with the identified milestones, so as to be performed in the same location. Most of the technical collaboration will be based on email, the project website and online cooperation tools, as well as Skype videoconference. The project coordinator will be informed in advance of WP leaders' intentions for meeting and travel, as well as for equipment spending, so as to make full use of synergies.
	Critical deliverables are delivered too late and milestones are missed.	All WPs with those deliverables	Management processes include specific roles for the monitoring and management of general, technical and human/legal/privacy/end-user issues and tasks, so as to efficiently monitor and respond to any issues on the timely delivery of results. The risk is a general risk (not specific only to ANITA, so previous management experience will be helpful for minimizing delays, and the selected people at management level have long proven expertise in EU projects). Partners involved in delayed tasks will allocate additional resources to meet the planned deadlines. Agile project management and frequent status calls will mitigate the risk of falling behind schedule.
	Insufficient involvement of stakeholders	WP1, WP2, WP4, WP8, WP9	One of ANITA's significant strengths is the active participation and strong commitment of end users as full members of the consortium, hence equally responsible in terms of their contracted contributions. The letters of support from a good range of additional end users ensure that there will be sufficient coverage of the different types of community policing stakeholders. A strong dissemination strategy will ensure a wide coverage of related stakeholders in Europe.
	Problems related to IPR	All WPs	The consortium agreement will constitute the primary source to resolve IPR issues. The General Assembly may be involved if conflicts remain.

Technical	The required services cannot be implemented within the time and cost constraints.	WP4, WP5, WP6, WP7, WP8	The ANITA services will rely on existing assets and solutions, as explained in Section 1.3. The partners will prioritise applications and services development to ensure maximum deployment and evaluation within project constraints. The selection of a staged approach including mockups/early prototypes (see task 9.2 and deliverable D9.2) and two pilot phases will ensure that if such a risk is encountered, it will be adequately and early assessed so as to be efficiently confronted.
	The project does not address issues that are relevant for illegal trafficking domain.	All WPs	Continuous monitoring of the changing state of the art and practice in Intelligence and illegal trafficking will support agile adoption of changes. Having the project firmly grounded in use cases (defined by experts and LEAs) that have high value for LEAs and society will make sure that ANITA remains strategic both during the course of the project and for years after.
	The required components are not available.	WP5, WP6, WP7, WP8	Selection of partners and allocation of technical tasks has been made on the basis of previous research and development expertise in each area so as to ensure experience and early detection of technical problems. The partners are already familiar with the relevant technology and are confident that all major items of the required infrastructure can either be sourced or developed within the project.
	Identified user requirements cannot be met within the scope of the project	WP4	Design of WP structure and the work plan have been made in a way that ensures that user requirements collection and end user involvement is adequately linked to the technical specifications delivery and architecture design. There is special provision in the management structure for efficient monitoring of human related and technical related tasks and the compatibility of their results.
	Deep Web and Dark Nets technologies are not mature and stable	WP5	The crawling of the Deep Web and Dark Nets is one of the key innovations introduced by this project. To this aim, the partner involved in these activities (AIT) has a strong experience in this specific topic and technologies. It will bring existing and working solutions for monitoring and crawling the Deep Web and Dark Nets as starting points to be evolved during ANITA.
	Final product is not satisfactory to Stakeholders	WP4, WP9	ANITA's strong emphasis on active participation of users (and external stakeholders) from the early stages of requirements specification throughout the project and until the final phase of pilots, ensures that the results will be based on users' and stakeholders' feedback and tested with their active participation.

Table 3.2.b : Critical risks for implementation.

3.3 Consortium as a whole

The ANITA consortium was built on the basis of the following criteria:

- The consortium provides a full coverage of all expertise and links needed to ensure that ANITA will not miss a key competency for its execution;
- Each partner brings in the project its specialized and relevant prior expertise in order to improve the overall excellence of the consortium, to speed up the execution, and to increase the scope of ANITA, building on prior activities directly relevant to the project at technical, legislative, ethical and privacy levels;
- Each partner is involved only in the activities and topics in which the partner has a recognized expertise, experience, and (for technical partners) solutions (in some cases already working and/or in the market) to be put as a starting point for the delivery of ANITA services;
- A significant (though manageable) number of relevant LEAs having the counter terrorism in their core activities, bringing in the project relevant experience and current practices to innovate and improve, and offering the opportunity to validate and demonstrate the project, its approach and results across very different contexts in Europe;
- The consortium as a whole provides a wide coverage across Europe.

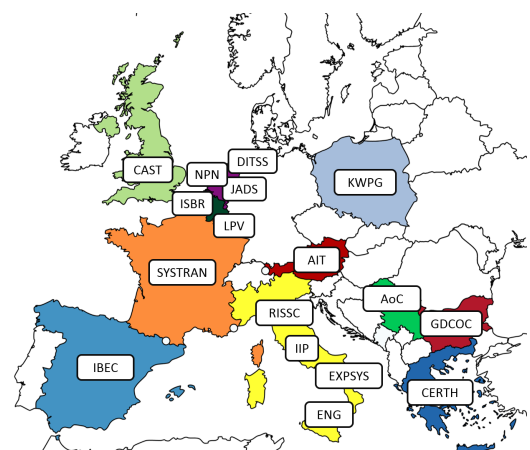


Fig. 9: ANITA consortium

Fulfilling the above criteria, ANITA is proposed by a really unique consortium with the following partners from 11 EU countries:

- The biggest Italian system integrator (**ENG**) with a huge experience in coordinating European research projects and in developing systems supporting LEAs in security, intelligence, and forensics;
- Important European-level technology research centres specialized in image processing (**CERTH**), Deep Web and Dark Nets and financial networks (**AIT**) and bioengineering (**IBEC**);
- A research centre specialized in criminology and illegal trafficking (**RISSC**);
- SMEs leaders in the market of semantic text analysis and natural language processing (**EXPSYS**) and automated multilingual translation (**SYSTRAN**);
- An NPO specialised in social, ethical, legal and privacy aspects (**IIP**);
- National foundations made by companies, government and research institutions from The Netherlands (**DITSS**) and Belgium (**ISBR**), specialised in technologies and innovations for safety and security of communities;
- A research academy specialised in data science and its applicability in real-world solutions (**TIU-JADS**);
- 6 Law Enforcement Agencies (LEAs) very active in illegal trafficking activities such as **KWPG**, **AoC**, **CAST**, **NPN**, **GDCOC** and **LPV**.

These partners bring in the consortium all the competencies and experience needed to fulfil the ambition project goals. The following table reports the main role of each partner in the ANITA project.

Partner	Main role in ANITA
ENG	Project Coordinator; architecture definition; leader of “Novel applications for LEAs and System Integration” WP, especially for Big Data Infrastructure, Novel applications for illegal trafficking, Advanced Big Data Visual Analytics, Chain of custody and evidence export, Continuous System Integration. Leader of the exploitation activities. Supporting technological partner for the field demonstration.
CERTH	Scientific and Technical coordinator; Data controller; leader of “Social, Ethical, Legal and Privacy issues of online sources analysis” WP especially for system specification; Leader of “Big Data analysis and analytics” WP especially for Image/Video processing and Video Indexing; Leader of Incorporation of conscious and subconscious user feedback in deep learning representations
RISSC	Leader of the interdisciplinary analysis of online illegal trafficking and study of strategies and countermeasures (WP2); Leader of training activities and end user workshops.
EXPSYS	Leader of semantic text analysis; Leader of “Knowledge management and reasoning” WP, especially for knowledge modelling and reasoning on black markets and illegal shops; Leader of System validation and Dissemination and communication
AIT	Leader of “Data source analysis and analytics” WP especially for crawling from Deep Web and Dark Nets, black markets discovery and monitoring, blockchain analysis and source network reconstruction.
IBEC	Leader of “Integration of human factor in the analysis loop” WP, especially for Implicit and explicit user feedback capturing, Black markets and illegal shops and products tracking and Knowledge transfer to new officers.
IIP	Leader of social, ethical, legal and privacy aspects (WP3)
SYSTRAN	Leader of multilingual translation and speech-to-text services.
TIU-JADS	Leader of data source risk assessment in Surface web, Deep web and Dark nets; Leader of illegal trafficking analysis.
DITSS	User coordinator, Leader of demonstration of ANITA system in relevant environment and Leader of LEAs and stakeholders community building.
ISBR	Expert in user management, LEAs and stakeholders community building, demonstration of ANITA system in relevant environments and training activities.
KWPG	Users of the project. They will participate from the early stages and will play a central role in the project: they will help in identifying and prioritizing the required functionalities, thus participating in and supporting the design phase, will participate in evaluation and refinement of resulting solutions, through a demonstration of the delivered system in their relevant environment and will provide contributions in defining training activities and organising workshops for the assessment of ANITA results.
AoC	
CAST	
NPN	
GDCOC	
LPV	

At the level of **geographical outreach**, ANITA extends from North to South and from West to East, another unique feature of the consortium. Countries involved in ANITA include: Austria, France, Belgium, Bulgaria, Greece, Italy, Netherlands, Poland, Serbia, Spain and UK.

ANITA is supported by further LEAs and experts, which will co-operate in the full implementation of the project. Some of them are part of the ANITA Advisory Board (AB) that comprehends people with strong expertise that will advise during the user requirements analysis, test phase and tuning of the tools, disseminate project results, and also provide relevant data and information. The current list of LEAs and experts is reported below:

- Gianfranco Todesco (Turin Local Police – Nucleo Investigazioni Scientifiche e Tecnologiche), Italy;
- Prof. Tibor Babos (Security Science Center of Óbuda University), Hungary;
- José Maria de Almeida Rodrigues (Polícia Judiciária), Portugal;
- Christian Estrosi (Métropole Nice Côte d'Azur) France;
- Ion Peligrad (Directorate General Logistics), Romania
- Frank Winterhalter - Cybercrime and Digital Traces Department (Baden-Württemberg Police), Germany
- Joachim Fischer – Organised Crime and Drug Crime Department (Baden-Württemberg Police), Germany
- Matthias Jehle - Organised Crime and Drug Crime Department (Baden-Württemberg Police), Germany
- Antonio Gomez Montejano, Deputy Commissioner, Policia Municipal de Madrid

3.4 Resources to be committed

The Table 3.4.a reports the distribution of the effort per partner and per WP.

Participant	WP1	WP2	WP3	WP4	WP5	WP6	WP7	WP8	WP9	WP10	WP11	TOTAL
1 ENG	23	2	3	6	8	6	24		40	4	7	123
2 CETH	5		3	8		19	12	18	2	4	4	75
3 RISSC	1	30		1	2					4	8	46
4 EXPSYS	1			2	9	18	24	1	4	4	6	69
5 AIT	1			2	20		5		3	4	3	38
6 IBEC	1			2				32	5	4	4	48
7 IIP	1		22	1						2	1	27
8 SYSTRAN	1			2		17		1	2	2	3	28
9 TIU-JADS	1			1	8	8			2	2	3	25
10 DITSS	1	3		2						11	6	23
11 ISBR	2			1						7	3	13
12 KWPG	1	3		1						8	5	18
13 AoC	1	3		1						10	11	26
14 CAST	4	3		3						6	5	21
15 NPN	1	3		1						5	3	13
16 GDCOC	1	3		1						6	4	15
17 LPV	1	3		1						4	3	12
TOTAL	47	53	28	36	47	68	65	52	58	87	79	619

Table 3.4.a: Number of person/months required

3.4.1 Table 3.4b: Other direct costs

3.4.1.1 Other costs - Cost category (B): Equipment

Participant	Costs (€)	Justification
2 CETH	15.000	TobiiX2-30 eye-tracker. Hardware for capturing the human gaze signal (T8.1) will be subsequently integrated in the deep learning paradigm (T8.3)
2 CETH	8.000	2 PC Tower high performing PCs for implementing Deep Learning approaches (WP6)
2 CETH	4.000	2 PC Workstation for software development
5 AIT	52.000	Costs of the hardware used for building the secure environment for ANITA pilots. The system is composed by one protection gateway and 5 clients to access this secure environment. Specifically: 1 SB50.76 - SINA L3 Box S 200M-2, VPN gateway enabled to transmit data via cryptographic VPN tunnels (7.000€) 5 SB70.267 - SINA WS S T460s 512 SWTS US, multi-session crypto client (for mobile use)

			enabled to process public data and RESTRICTED data in parallel sessions (25.000€) 1 SB80.13 - SINA Management Software, 9 SC, 4 SC-Reader, 3y Supp (20.000€)
6	IBEC	20.000	Remote Eye Tracking - Tobii Pro X2-60 eye tracker or SMI Red. Hardware for capturing eye movements; to develop research studies it's required an eye tracker of minimum 60Hz
6	IBEC	500	Depth camera. (Intel Real Sense or Microsoft Kinect) o perform body - face motion analysis and detection of emotions
6	IBEC	2.000	Computer to <u>perform experiments of Human Activity Recognition</u> ; to present content to users (Anita use cases), processing of reactions in real-time, and perform detailed behavioural experiments
6	IBEC	5.000	2 High-end GPU powered workstations for running big data analytics and machine learning to support Human Activity Recognition experiments.
6	IBEC	3.000	2 <u>Workstations to perform offline analysis of experimental results</u> , post-processing of data, creation of reports / plots and for presenting results in events

3.4.1.2 Other direct costs - Cost category (B): major costs for organization of workshop costs and training/dissemination materials

During the ANITA project, the following kinds of workshops and events will be organized: user workshops for collecting and eliciting requirements, field demonstration and training, public dissemination and presentation of project results (e.g. workshops organized in co-location with major relevant events and conferences), and final training exercises. Two working days are planned for each event: the first day will be devoted to technical preparation; the second day is for the operational session. ANITA workshops will be organised by Home Office CAST and by the 3 research centres on security issues: DITSS, ISBR and RISSC. The attendees for the workshops will be: max 6 representatives of the host, 2 representatives from each of the other LEAs partners and the other partners of the consortium, 10 external invited representatives from other LEAs and LEA academies (external to the ANITA consortium). The attendees for the preparation meetings will be project partners and some members of the Advisory Board. The cost for the workshop logistics and organization is on the host, including the catering: estimated 4,000 euro for each event. The cost for the advisory members attending the user workshops will be covered by the cost for “Advisory Board Management” (in the DITSS budget). Details about the planned workshops are briefly reported below:

Participant	Costs (€)	Justification
3 RISSC	4.000	Organisation and hosting of 1 workshop in Italy for Public Demonstration
10 DITSS	4.000	Organisation and hosting of 1 workshop in Netherlands for requirements elicitation
10 DITSS	4.000	Organisation and hosting of 1 workshop in Netherlands for Public Demonstration
11 ISBR	8.000	Organisation and hosting of 2 workshops for field demonstration in Belgium
14 CAST	8.000	Organisation and hosting of 2 workshops for field demonstration in UK
13 AoC	8.000	Organisation and hosting of 2 workshops for training in Serbia

ANITA will hold at least 1 **workshop for the elicitation and collection of the user requirements**: this workshop will be organized during the first project meeting (with all partners attending), at the end of month 3. The expected attendees are: project partners involved in the activity and some members of the Advisory Board. The user workshops for requirements elicitation and collection are in the scope of *task 4.1*. No additional travel costs (with respect to the planned first project meeting) will be needed for partners attending these user workshops.

The project will organise 2 **workshops for training sessions**, which will be led by the Academy of Criminalistics and Police Studies (AoC, participant 13) and will also include online webinars and **training materials**. These activities and events are in the scope of *task 11.4*.

ANITA will also organise 2 **workshops for field demonstrations**. The planned locations for the field demonstrations are Netherlands, Belgium, and UK.

Finally, 2 workshops for **public demonstration of the ANITA results** will be organised in the scope of *task 11.1* (individual presentation of project results, dedicate workshops requested/organized within main events and conferences), *task 11.2* (it could be during meeting or events organized by LEA organizations such as ENFSI, ENLETS, etc.) and *task 11.5*. The planned locations for the field demonstrations are Netherlands, and Italy.

3.4.1.3 Other direct costs: other goods and services (cost category B)

Participant	Costs (€)	Justification
1 ENG	5.000	Audit certificates
2 CERTH	1.000	Audit certificates

2	CERTH	9.000	Publications in international journals. Tentative publication venues: Conference on Computer Vision and Pattern Recognition (CVPR), International Conference on Computer Vision (ICCV), European Conference on Computer Vision (ECCV), IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), IEEE Transactions on Circuits and Systems for Video Technology (TCSVT)
3	RISSC	1.500	Conference fees
4	EXPSYS	4.000	Audit certificates
4	EXPSYS	5.000	Dissemination materials
5	AIT	4.000	Audit certificates
5	AIT	11.500	Costs for training on the use (installation, configuration and administration) of the one way gateway used to build the secure environment
6	IBEC	5.000	Publications in international journals.
8	SYSTRAN	4.000	Audit certificates
9	TIU-JADS	1.500	Conference fees
9	TIU-JADS	5.000	Publications in international journals
10	DITSS	4.000	Audit certificates
10	DITSS	32000	Cost for end user advisory board (EUAB) management: flights and accommodation for the member of the EUAB to attend the meetings and participate to project activities.
13	AoC	1.000	Training materials
13	AoC	12.000	Organization and logistics of training sessions

3.4.1.4 Other direct costs: travel costs (cost category B)

The travel costs have been computed based on a task based participation level and country of origin. The consolidated travel costs are provided in the first section lines of the table, the detailed computation is provided in the second section.

Travel costs per partner:

Partner	1	2	3	4	5	6	7	8	9
	ENG	CERTH	RiSSC	EXPSYS	AIT	IBEC	IIP	SYSTRAN	TIU-JADS
Travel cost	28.100	25.200	17.600	21.400	21.400	21.400	17.600	17.600	17.600
Partner	10	11	12	13	14	15	16	17	
	DITSS	ISBR	KWPG	AoC	CAST	NPN	GDCOC	LPV	
Travel cost	22.500	19.750	22.250	22.250	19.750	19.750	22.250	19.750	

Detailed computation:

Partner	1	2	3	4	5	6	7	8	9
	ENG	CERTH	RiSSC	EXPSYS	AIT	IBEC	IIP	SYSTRAN	TIU-JADS
Average plane costs per meeting	350€	400€	350€	350€	350€	350€	350€	350€	350€
Food & lodging (per day)	200€	200€	200€	200€	200€	200€	200€	200€	200€
Partner	10	11	12	13	14	15	16	17	
	DITSS	ISBR	KWPG	AoC	CAS T	NPN	GDCOC	LPV	
Average plane costs per meeting	350€	350€	450€	450€	350€	350€	450€	350€	
Food & lodging (per day)	200€	200€	200€	200€	200€	200€	200€	200€	

Partner	1	2	3	4	5	6	7	8	9
	ENG	CERTH	RiSSC	EXPSYS	AIT	IBEC	IIP	SYSTRAN	TIU-JADS
Conferences	2	4	2	2	2	2	2	2	2
Project meetings (including kick-off)	12	10	9	9	9	9	9	9	9
WP meetings	6	6	4	4	4	4	4	4	4
Reviews (including preparation)	6	3	3	3	3	3	3	3	3
User workshops and training meeting	4	5	2	2	2	2	2	2	2
Field trials and training	8	5	4	8	8	8	4	4	4
Partner	10	11	12	13	14	15	16	17	
	DITSS	ISBR	KWPG	AoC	CAST	NPN	GDCOC	LPV	DITSS

Conferences	2	2	2	2	2	2	2	2	2
Project meetings (including kick-off)	9	9	9	9	9	9	9	9	9
WP meetings	4	1	1	1	1	1	1	1	4
Reviews (including preparation)	3	3	3	3	3	3	3	3	3
User workshops and training meeting	4	2	2	2	2	2	2	2	4
Field trials and training	8	8	8	8	8	8	8	8	8

Partner	1	2	3	4	5	6	7	8	9
	ENG	CERTH	RiSSC	EXPSYS	AIT	IBEC	IIP	SYSTRAN	TIU-JADS
Conferences	1.500	3.200	1.500	1.500	1.500	1.500	1.500	1.500	1.500
Project meetings (including kick-off)	9.000	8.000	6.750	6.750	6.750	6.750	6.750	6.750	6.750
WP meetings	3.300	3.600	2.200	2.200	2.200	2.200	2.200	2.200	2.200
Reviews (including preparation)	4.500	2.400	2.250	2.250	2.250	2.250	2.250	2.250	2.250
User workshops and training meeting	2.200	3.000	1.100	1.100	1.100	1.100	1.100	1.100	1.100
Field trials and training	7.600	5.000	3.800	7.600	7.600	7.600	3.800	3.800	3.800
Partner	10	11	12	13	14	15	16	17	
	DITSS	ISBR	KWPG	AoC	CAST	NPN	GDCOC	LPV	
Conferences	1.500	1.500	1.700	1.700	1.500	1.500	1.700	1.500	1.500
Project meetings (including kick-off)	6.750	6.750	7.650	7.650	6.750	6.750	7.650	6.750	6.750
WP meetings	2.200	550	650	650	550	550	650	550	2.200
Reviews (including preparation)	2.250	2.250	2.550	2.550	2.250	2.250	2.550	2.250	2.250
User workshops and training meeting	2.200	1.100	1.300	1.300	1.100	1.100	1.300	1.100	2.200
Field trials and training	7.600	7.600	8.400	8.400	7.600	7.600	8.400	7.600	7.600

3.4.2 Direct cost of subcontracting

Not envisaged at the time of the proposal.

3.4.3 Financial overview

The overall financial overview corresponds to 620 man months over 36 months duration. The ANITA's budget distribution is well balanced at the partner profile level. The budget rate for industries (including large solution and technology providers and SMEs) is 33.78% (around 1/3 of the overall budget), whereas the budget rate is around 39.01% for research organizations and 12.96% for non-profit organisations. The budget for SMEs is 17.58% of the total budget. Moreover ANITA reserves a significant budget (14.25%) for user partners (LEAs), who are strongly committed and interested to adopt and deploy project results in their operational environments. The ANITA's funding highlights a balance both at European level, across the 11 Member States represented by the 17 partners. Indeed requested grant is distributed as follow: Italy 35.51%, Greece 9.99%, Spain 7.06%, Austria 11.34%, France 6.61%, Netherlands 15.24%, Belgium 5.34%, UK 3.84%, Poland 1.46%, Serbia 1.94%, Bulgaria 1.68%.

References

- [Agrafioti12] Agrafioti, F., Hatzinakos, D., & Anderson, A. K. (2012). ECG pattern analysis for emotion detection. *IEEE Transactions on Affective Computing*, 3(1), 102–115. <https://doi.org/10.1109/T-AFFC.2011.28>
- [Akshay17] Akshay Bharambe, Rishab Dey, Vishal Bahadur, Bhavana Tanpure, Rashmi Ramteke, A Survey on Web Crawler Approaches, *International Journal of Innovative Research in Computer and Communication Engineering*, Vol. 5, Issue 2, February 2017.
- [Cambria14] Cambria, Erik, Daniel Olsher, and Dheeraj Rajagopal. "SenticNet 3: a common and common-sense knowledge base for cognition-driven sentiment analysis." *Proceedings of the twenty-eighth AAAI conference on artificial intelligence*. AAAI Press, 2014.
- [Chen04] Chen, D., & Vertegaal, R. (2004). Using mental load for managing interruptions in physiologically attentive user interfaces. *Extended Abstracts of the 2004 Conference on Human Factors and Computing Systems - CHI '04*, 1513. <https://doi.org/10.1145/985921.986103>
- [Daelemans13] Walter Daelemans *Explanation in Computational Stylometry CLiPS*, University of Antwerp, Belgium.
- [EU-IRU15] EU Internet Referral Unit, YEAR ONE REPORT, HIGHLIGHTS, 2015, Europol.
- [Hovy13] Hovy, Roberto Navigli, and Simone Paolo Ponzetto. 2013. Collaboratively built semi-structured content and Artificial Intelligence: The story so far. *Artif. Intell.* 194 (January 2013), 2-27.
- [Jacynycz16] Jacynycz, Viktor; Calvo, Adrian; Hassan, Samer; et al., A Distributed Bounty-Based Crowdfunding Platform over Ethereum, *Conference: 13th International Conference on Distributed Computing and Artificial Intelligence (DCAI)* Location: Sevilla, SPAIN Date: JUN 01-03, 2016, Sponsor(s): IBM; Indra; Fidelia; IEEE SMC Spain.

- [Jaimes07] Jaimes, Alejandro, and Nicu Sebe. "Multimodal human–computer interaction: A survey." *Computer vision and image understanding* 108.1 (2007): 116-134.
- [Kloss16] Kloss, Alina, et al. "Learning where to search using visual attention." *Intelligent Robots and Systems (IROS)*, 2016 IEEE/RSJ International Conference on. IEEE, 2016.
- [Papadopoulos14] G. Papadopoulos, K. Apostolakis, and P. Daras. Gaze-based relevance feedback for realizing region-based image retrieval. *ACM-MM*, 16(2), Feb 2014.
- [Parigha16] Parigha V. Suryawanshil, D. V. Patil, Enhancing Crawler Performance for Deep Web Information Extraction, *IJETT* ISSN: 2455-0124 (Online) | 2350-0808 (Print) | GIF: 0.456 | September 2016 | Volume 3 | Issue 3 | 6013.
- [Petrucchi16] Petrucci, Giulio, Chiara Ghidini, and Marco Rospoche. "Ontology Learning in the Deep." *Knowledge Engineering and Knowledge Management: 20th International Conference, EKAW 2016, Bologna, Italy, November 19-23, 2016, Proceedings 20*. Springer International Publishing, 2016.
- [Roman17] de San Roman, Philippe Pérez, et al. "Saliency Driven Object recognition in egocentric videos with deep CNN: toward application in assistance to Neuroprostheses." *Computer Vision and Image Understanding* (2017).
- [Sameer17] Sameer Jadhav, Crawling Web Information for Forensic, *International Journal of Innovative Research in Computer and Communication Engineering*, Vol. 5, Issue 1, January 2017.
- [Sattar17a] Sattar, Hosnieh, Andreas Bulling, and Mario Fritz. "Predicting the Category and Attributes of Visual Search Targets Using Deep Gaze Pooling." *arXiv preprint arXiv:1611.10162* (2017).
- [Shcherbatyi15] Shcherbatyi, Iaroslav, Andreas Bulling, and Mario Fritz. "GazeDPM: Early integration of gaze information in deformable part models." *arXiv preprint arXiv:1505.05753* (2015).
- [Sugano16] Sugano, Yusuke, and Andreas Bulling. "Seeing with humans: Gaze-assisted neural image captioning." *arXiv preprint arXiv:1608.05203* (2016).
- [Toti12] IEEE A knowledge discovery methodology for semantic categorization of unstructured textual sources D Toti, P Atzeni, F Polticelli, (2012 Eighth International Conference on Signal Image Technology and Internet Based Systems.
- [Tran13] Tran, Son N., and Artur d'Avila Garcez. "Knowledge extraction from deep belief networks for images." *IJCAI-2013 Workshop on Neural-Symbolic Learning and Reasoning*. 2013.
- [Trumbly94] J. Trumbly, K. Arnett, P. Johnson Productivity gains via an adaptive user interface. *Journal of Human–computer Studies*, 40 (1994), pp. 63-81
- [Verhoeven14] "CLiPS Stylometry Investigation (CSI) Corpus: A Dutch corpus for the detection of age, gender, personality, sentiment and deception in text", Verhoeven, B., & Daelemans W., 2014
- [Verschure11] Verschure, P. F. (2011, September). The complexity of reality and human computer confluence: stemming the data deluge by empowering human creativity. In *Proceedings of the 9th ACM SIGCHI Italian Chapter International Conference on ComputerHuman Interaction: Facing Complexity* (pp. 36).ACM.
- [Verschure16] Verschure, P. F. M. J. (2016). Synthetic consciousness: the distributed adaptive control perspective. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 371(1701), 263–275. <http://doi.org/10.1098/rstb.2015.0448>
- [Vygotsky78] Vygotsky, L.: Interaction between learning and development. *Read. Dev. Children* 23(3), 34–41 (1978)
- [Wang15] Wang, Hao. "Semantic Deep Learning." (2015).