Intelligent Information System Supporting
Observation, Searching and Detection for
Security of Citizens in Urban Environment

European Seventh Framework Programme
FP7-218086-Collaborative Project

Evaluation of Components, D9.4 WP9

The INDECT Consortium

AGH – University of Science and Technology, AGH, Poland

Gdansk University of Technology, GUT, Poland

InnoTec DATA GmbH & Co. KG, INNOTEC, Germany

Grenoble INP, France

MSWiA[1] - General Headquarters of Police (Polish Police), GHP, Poland

Moviquity, MOVIQUITY, Spain

Products and Systems of Information Technology, PSI, Germany

Police Service of Northern Ireland, PSNI, United Kingdom

Poznan University of Technology, PUT, Poland

Universidad Carlos III de Madrid, UC3M, Spain

Technical University of Sofia, TU-SOFIA, Bulgaria

University of Wuppertal, BUW, Germany

University of York, UoY, Great Britain

[1]MSWiA (Ministerstwo Spraw Wewnętrznych i Administracji) – Ministry of Interior Affairs and Administration. Polish Police is dependent on the Ministry

Technical University of Ostrava, VSB, Czech Republic

Technical University of Kosice, TUKE, Slovakia

X-Art Pro Division G.m.b.H., X-art, Austria

Fachhochschule Technikum Wien, FHTW, Austria

Document Information

Contract Number	*218086*
Deliverable name	*Evaluation of components*
Deliverable number	*D9.4 (WP9)*
Editor(s)	*Jan Derkacz*
	Andrzej Dziech
	Jacek Dańda
	Magdalena Maderska
Author(s)	*Nils Johanning*
	Mikołaj Sobczak
	Andrzej Figaj
	María José Martínez Gil
	Jan Derkacz
Reviewer(s)	*Mariusz Ziółko*

Dissemination level	*Public*
Contractual date of delivery	*31.12.2009*
Delivery date	*31.12.2009*
Status	*Final version*
Keywords	*INDECT, components, modules, description*

This project is funded under 7th Framework Program

Table of Contents

(This page is left blank intentionally)

1 Executive Summary

This deliverable presents an initial overview of components within the system under development in INDECT project. In the research carried out by INDECT partners, new concepts and technological results are being delivered as parts of integrated INDECT security system. This system will have impact on security of citizens in the European Union.

Most of components that are developed within the Project are in initial phases of planning or research. Therefore the current possible evaluation is limited to description of objectives only. For some Workpackages, additional information on work in progress is available now and it is presented in this deliverable.

2 Introduction

Present-day security of citizens is threatened mostly by terrorism, organized crime, Internet crime, financial crime, and common crimes. Popularization of information technology and availability of high capacity data networks, mentioned threats appear in a different way in the Internet. Present-day criminals use complex operational methods based on IT techniques. Methods applied by the police are a step back as compared to those used by criminals. It is caused by two most important factors:

- Level of finance, which is fast and precisely allocated.
- Police have to respect the rule of law and human rights.

The Internet is an important source of information about committed and planned crimes. The most important issues are that there are no technological solutions that allow extraction of this information and using it in security maintenance. Such solutions, that are developed within INDECT, as separate modules of the INDECT system, can enhance, and integrate the arsenal of the police.

The document is divided into 9 chapters that describe objectives of each Workpackage and modules developed within this Workpackage (if any), related to this objectives.

3 Work-Package 1

The main goal is to provide methods for the monitoring of public areas, and to develop a pilot system that enables detection of threats. Another functionality of the pilot system being in development is operational data exchange and user registration. Operational data collected in the system include audio, images, video, raw and processed alphanumeric data. The proposed solution enables intelligent data processing based on the latest achievements in multimedia processing techniques and delivers functionality of announcing public safety services of potentially dangerous conditions. The objective of WP1 is to design and create the Node Stations (NS) hardware and software, both serving as a platform for performing the intelligent monitoring and detection of threats. Another objective of WP1 is to create and integrate audio and video processing algorithms for event detection

3.1 Objective 1.1: Conceptions, Methods and Measures for the Monitoring of Buildings and Areas

Automatic event detection algorithms developed within WP1 are meant to aid a person operating the monitoring system, allowing concurrent analysis of a couple of orders greater number of audio and video streams, than what is currently possible. Content analysis algorithms can automatically protect content recognized as a private, including faces, car plate numbers, house windows, etc. Computer analysis can provide further tools – prediction of dangerous events, and detection of incorrectly ignored events. The first step is identification of the most important events for automatic detection and selection of proper hardware for acquisition and processing of media streams.

The objective of WP1 is to design and create the Node Stations (NS) hardware serving as a platform for performing the intelligent monitoring and detection of threats. NS performs unattended acquisition of multiple media streams, processes them, and in case of dangerous event being detected sends the result to the Central Station (CS), accompanied with textual description of event type, location, time, date, and a media stream from the camera/microphone.

3.2 Objective 1.2: A pilot system for detecting dangers

The objective of WP1 is to create and integrate audio and video processing algorithms for event detection, particularly:

- Acquisition and pre-processing of audio and video
- Recognizing predefined audio events
- Tracking of moving objects in the image

- Recognizing types of moving objects

Results of practical evaluation of elaborated strategies will be presented in the D1.3, a document reporting on acquired results of pilot trial.

3.3 Evaluation of Components related to Objective 1.1

For the purpose of End-User requirements analysis an End-User Questionnaire was established, created with cooperation of all INDECT Project Partners. This document, Deliverable D1.1., describes End-User Questionnaire structure, its purpose from the point of view of WP1, outcomes of analysis of answers related to WP1 work, and preliminary specification of functionality and hardware of the system fulfilling the requirements for intelligent monitoring and automatic detection of threats.

Quality of Experience methodology is employed for providing three important aspects of system functionality, i.e. defining measurable requirements factors on an automatic recognition of events during testing of created algorithms for assurance of decision quality, optimizing system/algorithms parameters on-the-fly for assuring quality of audio and video media, and guaranteeing sufficient conditions for media processing. All algorithms have their limitations related to lowest possible quality for which they can deliver proper results, e.g. resolution and frame rate of video, and sound sampling frequency, bit resolution and number of audio channel.

3.4 Evaluation of Components related to Objective 1.2

Within the WP1, an initial system specification of system hardware was defined. The INDECT concept of the multimedia platform assumes the elaboration of a distributed system whose principal element is an autonomous Node Station (NS). This automatic data acquisition station will be used to acquire data, signals, and images from the surveyed area, then to pre-process the data intelligently and transmit the gathered information to the remote servers. It will cooperate with cameras, sensors, and microphones located within the range of its operation through wired or wireless connections, and it will pass the collected and partially processed information through the gateway of a computer network. The distributed data processing system, provided with huge computational power and a vast repository of knowledge connected also to spatial information system, will be programmed in a way that will allow the automatic detection of events that could pose a potential threat to security and safety. The NS can be equipped with megapixel, wide angle, fixed cameras or moving PTZ (Pan Tilt Zoom) cameras as well as microphones and speakers. It is assumed that WP1 solutions are aimed at analysis of high definition video, and the algorithms are provided with direct video stream from camera with high frame rate, high resolution, low noise video and high quality compression (e.g. without colour artefacts, noise or blocking). It is technically

feasible as a result of locating processing unit directly near the cameras and microphones. Acquired streams are either transmitted by wire, or short distance wirelessly, therefore high capacity connection is available, allowing high data rate of media.

The video and audio data are analyzed by NS and alerts accompanied with Metadata (i.e. text description, geo-location, time and date, etc.) are sent to the Central and mobile terminals by any available network. The live audio and video streams can be coded in NS for adaptation to different transmission medias, and terminals. Analysis algorithms in the NS are designed to communicate with other Stations, for detection and tracking of particular objects (cars, persons) in large areas covered by number of NSs. All communication is performed through Central Server for backup, storage, and control. Databases are distributed among NSs but the significant data are also backed up in Central database along with streams (audio and video).

The NS is developed in WP1, as a multifunctional platform for acquisition and processing of audio and video streams. WP1 is dedicated to creation of algorithms for NS for audio and video processing and automatic intelligent detection of threats. Next in WP7 that NS is to be incorporated within INDECT Platform and the communication protocols and streaming procedures for NS/Platform are to be created.

3.5 State-of-the-Art Solutions Analysis

Current monitoring systems depend heavily on operator's attentiveness. His/her task is to watch numerous video monitors, in search of dangerous or untypical events, and then perform some reactive actions. Most of monitoring systems are focused on registration of the video material serving as post-factum evidence.

Commercially available monitoring hardware is often a closed solution, hard or impossible to extend with additional cameras or other sensors. Moreover only image is being processed, while the sound is being omitted. Single monitoring "station", defined as a processing unit for analysis of video stream(s) is located inside the monitoring centre, while the cameras are installed in remote locations. That requires high quality and high speed connection for streaming video from the camera to the monitoring centre.

State-of-the art monitoring systems utilize image analysis for automatic detection of simple events. Commercial systems for the intelligent monitoring perform motion detection for recognition of: appearance or disappearance in/from the area, people counting, passing a border, left or removed objects. That is performed by simple motion analysis, finding regions in the image containing changes in time, but interactions between objects are not taken into account. Therefore it is very hard or impossible to define other events.

Commercially available monitoring systems do not perform automatic sound analysis for event detection. Even if a system is equipped with the microphone it is dedicated for voice

communication between bystanders on the street and the system operator, therefore sound quality is no sufficient for digital signal processing and sound classification and these algorithms are not implemented.

Output of these system is presented to the operator on the monitor screen, often numerous video streams, textual information, visual indicators of detection are played simultaneously on single screen, therefore constant effective comprehension of that image by the operator is considered as hard.

3.6 Progress beyond State-of-the-Art Solutions

3.6.1 Progress in Terms of Scientific and Research Advance

The proposed solution enables intelligent data processing based on the latest achievements in multimedia processing techniques and delivers functionality of announcing public safety services of potentially dangerous conditions.

Node Station (NS) is designed as a research platform for remote audio and video processing. It is intended to implement procedures for distant maintenance, remote software updates and algorithms testing, gathering of processing results, acquisition and processing parameters changing, protected by high security transmission protocols. NS is designed as an open platform for processing of arbitrary number of digital streams (limited by computational power only, not by purchased license).NS can be equipped with high-end cameras, microphones, and other sensors, for high quality digital stream acquisition, and processing, allowing utilization of sophisticated and demanding algorithms. NS encourages new technologies such as stereovision achieved by utilization of stereo-camera setup, and introduces a domain of 3D vision to monitoring.

The new solution of event detection algorithms (for the video image) running on dedicated Node Stations (NS) introduces parallel: object detection, motion estimation, tracking of moving objects, classification of that objects, analysis of interactions between objects of given types. That allows defining and recognizing complex events, not limited to: appearance or disappearance, people counting, passing a border, left or removed objects. The range of detected events is significantly extended.

Other media analyzed by the NS are sounds from attached microphones. Utilization of numerous microphones allows: sound direction estimation, sound classification, recognition of dangerous events (shouts, gunshots, sound of broken glass), and therefore introducing new paradigm in the monitoring – sound monitoring.

Privacy is better protected by the concept of image processing on-site, and transferring of only important video clips, categorized as containing dangerous event. If the live view mode is required by the operator the image acquisition procedures can automatically obscure

people faces and car license plates, as long as no dangerous events are detected and normal activity is performed.

3.6.2 Progress in Terms of Engineering and Integration Advance

Automatic event detection algorithms that will be developed in WP1 are meant to aid a person operating the monitoring system, allowing concurrent analysis of practically any number of audio and video streams (limited by computational power, which is easily extendable). The operator's work will be verification of alarms instead of inspection of multiple numbers of streams in the same time, resulting in increase of effectiveness of threat detection.

Node Stations (NS) hardware is integrated platform for acquisition and processing of audio and video. These two functions are not passed to separate components, therefore, by combined processing; a more effective analysis is possible. Processing algorithms have direct access to high quality media streams (audio and video), and high speed network connection is not required. Therefore the high resolution cameras and high sampling rates for microphones can be utilized.

Multimedia streams are not transmitted to the Central Station (CS), unless are requested by the operator, or dangerous event is detected, and operator must be informed of that occurrence. Therefore single operating centre, with CS, can cooperate with numerous Node Stations (the number is yet to be estimated, possibly dozens up to one hundred NS for single CS), and remain effective, and not overwhelmed by transferred data. The operator is not expected to watch and control all media streams; his/her task is only to verify incoming alerts.

Almost all dangerous events contain some sound indicators, and the omni-directional permanent sound "observation" can be performed. Localization of detected sound event can be utilized for automatic pointing the cameras of the system onto the event.

Output of the system can be presented on the monitor screen, or on the mobile terminal, therefore the operator can also be mobile. For example, a security officer in the field can request a direct video feed from selected camera to the mobile terminal in his possession.

The major difference to the commercially available systems is that there is no need to present all streams simultaneously and for the operator to watch them. If the system detects dangerous event the alert is presented on a single dedicated monitor, and operator's task is only to verify the alarm, not to look for it inside numerous streams.

3.7 Hardware Evaluation

Within WP1, the preliminary specifications were made. These specifications are related to the list of events to be recognized and the hardware features for audio and video acquisition,

processing, and storage aimed specifically at effective automatic and intelligent recognition. It should be treated as a road map for further work, and it is assumed that all of the requirements are meant to be reconsidered in a time span of INDECT Project. Final specification of these features will be provided in next deliverables -D1.2 Report on NS and CS hardware construction.

Designed Node Station is based on desktop PC enclosed in a weatherproof casing. It is dedicated to be located in close proximity to cameras and microphones of the monitoring system, with fast network connection for acquiring high quality digital streams from these devices. Designed Central Station is based on desktop PC. Designed Mobile Terminal is based on PDA preferably operating Windows Mobile OS.

The system for automatic detection of threats comprises of mandatory: Node Stations, IP cameras, Microphones, and Central Station. Optionally one can introduce Mobile Terminals, and other sensors, such as smoke detectors, infrared cameras, etc.

3.8 Description of Non-Labour Costs

Non labour costs in case of WP1 include:

- Purchase of hardware platform for the system
- Purchase and installation of software, including OS for the platform
- Travel expenses for installation and trainings of users
- Trainings of users and system administrators

The above costs may be estimated in the order of tens of thousands of Euros.

The system can be reconfigured depending on end-user requirements. Main parts costs are as follows:

- Single NS hardware - 6000€
- Single CS hardware - 15000€
- Single camera - 2400€
- Single microphone - 300€

Estimated cost of the system with 10 NSs, 1 CS, 15 cameras, and 10 microphones is 114.000€.

Depending on technology advances finally these costs can change up to 150%.

Audio, video and audiovisual stream processing software licensing costs are based on number of NSs, CSs, and mobile terminals in the system.

- analysis software for single NS - 500€
- server software for single CS - 2000€

- communication software for single mobile terminal - 150€

3.8.1 OPEX (only non-labour)

Hardware Maintenance, including powering and inspections:

- Single NS maintenance costs (assuming 3 cameras and 1 microphone connected) 150€ per month
- Single CS maintenance costs - 200€ per month
- Single mobile terminal maintenance costs - 10€ per month

Software Maintenance:

- Single NS analysis software maintenance costs - ca. 50€ per operation (reinstalling, reconfiguring)
- Single CS server software maintenance costs - ca. 50€ per operation (reinstalling, reconfiguring)
- Single Mobile Terminal communication software maintenance costs - ca. 50€ per operation (reinstalling, reconfiguring)

3.9 Summary of the preliminary results

For WP1 the first step is to gather End-User requirements helping to define functionality of the system, specifically for task related to automatic detection of events. For that purpose the End-User Questionnaire was established, created with cooperation of all INDECT Project Partners. The objective of the End-User Questionnaire and outcomes of the analysis were presented, resulting in preliminary specification of the functionality for event detection and hardware specification. Automatic event detection algorithms that will be developed in WP1 are meant to aid a person operating the monitoring system, allowing concurrent analysis of practically any number of audio and video streams (limited by computational power, which is easily extendable). The operator work will be verification of alarms instead of inspection of multiple numbers of streams in the same time, resulting in increase of effectiveness of threat detection.

Other added values were analyzed within WP1, such as reduction of storage space, automatic protection of content recognized as a private, prediction of dangerous events, and detection of previously overlooked events.

The described solution is expected to remain unchanged in most important aspects during the project. The factors likely to change are related to technology advancements: performance of computer hardware and speed of data transmission networks.

4 Work-Package 2

The main goal of WP2 is to develop the prototype of the integrated, network-centric system supporting the operational work of police officers, providing techniques and tools for the observation of various mobile objects. Equipped with advanced devices, in permanent wireless connection with the central module of the system, police officers will be able to receive detailed information on tracked individuals, vehicles and objects location. This part of system will employ mobile wireless devices, integrated on the basis of a common commanding component. As one of the core elements of the system, so called Unmanned Aerial Vehicles (UAVs) of different classes will be developed and constructed. Their purpose will be to make reconnaissance flights for police officers working in the field.

4.1 Objective 2.1: Mobile system for identification and positioning

The main goal of this objective is to achieve the idea of a multi-mode system for positioning and tracking moveable objects on the basis of wireless communication networks. The devised algorithms will enable location of tracked objects and advanced navigation. Delivered methods and tools will enable officers in the field to follow traced objects, to observe their location and history of movement on digital vector maps, to predict the location of tracked users and to support the control of observational tasks.

4.2 Objective 2.2: Unmanned Aerial Vehicles

The Integrated Air Surveillance System for police departments will work on the basis of autonomous, intelligent, unmanned vehicles. The main goal of the activity is to develop algorithms for planning and commanding patrols, advanced navigation and aircraft piloting. Examinations of the usage of unmanned aerial vehicles in police operations will be carried out, with the support of officers working in the field and their integration with multi-task sensor network systems.

4.3 Objective 2.3: System for Positioning and Guiding

This objective includes construction of prototypes for mobile objects tracking. A mobile sensor network will consist of tiny computers containing different sensors and integrated network interfaces. These elements will create self-configurable, ad-hoc architecture network. Algorithms for routing and state will be designed and implemented. Travel route optimization will consider the priority of signalized junctions for a single vehicle, different possible transport modes and their limitations, and the positioning of specified objects with location attributes.

5 Work-Package 3

The major objective is to develop techniques supporting surveillance of Internet, analysis of acquired information, and detection of criminal activities and threats, as well as appropriate software tools. This part of system is based on an agent paradigm, flexible enough to model large, distributed, complex unpredictable systems. It is effective for systems of a hybrid technical/human nature. Different agent's roles allow three virtual subsystems to be distinguished which are responsible for monitoring, information processing, and decision support for police and security services.

5.1 Objective 3.1: Monitoring of Internet Services and Acquisition of Information on Identifying Criminal Behaviour

Functionality oriented towards collection of information about users of computer networks and their tasks is indispensable for the detection of criminal activities in the global network. It aims to construct agents assigned to continuous, automatic monitoring of public resources such as: web sites, discussion forums, UseNet groups, file servers, p2p networks as well as individual computer systems. Both information content and traffic data will be monitored.

5.2 Objective 3.2: Content Analysis of Information from Distributed Heterogeneous Sources

Agents and their subcomponents that help detect signs of criminal activities and threats will be dedicated to particular analysis tasks. To increase effectiveness of the detection algorithms, content analysis techniques will be applied to the processed information.

5.3 Objective 3.3: Criminal Analysis and Decision Support for Operational and Investigation Activities

Criminal analysis is a complex process involving information gathered from different sources, mainly of a quantitative character (billings, bank account transactions, etc.), but also of a qualitative character such as eyewitnesses testimonies. Because of the massive quantity of such information, operational or investigation activities will be vastly improved when supported by intelligent techniques, based on behavioural patterns, and dedicated mechanisms provided by the agents to be developed.

5.4 Objective 3.4: Knowledge-Based Support for Organization of Security Tasks

Security related work is usually connected with collecting, searching, and making available the huge number of documents. To help a police analyst in such situations, it is necessary that these documents be annotated in a possibly rich semantic way. Furthermore, an appropriate subsystem would support access to the documents, and all related data.

6 Work-Package 4

The aim of WP4 is to develop key technologies to facilitate the building of Internet based intelligence gathering system by combining and extending existing techniques in natural language processing and text mining. There is very little work done in proposed topics, and extending existing methods is challenging. Being able to learn inference patterns, draw temporal/spatial connections and discover un-usual information is vital.

6.1 Objective 4.1: Relationships through Websites and Social Networks

This objective involves putting together bits of information into patterns that can be analyzed and classified. A relationship mining system will generate a labelled graph where the nodes correspond to individuals and edges correspond to weighted relationship types and edge attributes correspond to supporting evidence.

6.2 Objective 4.2: Learning Behavioural Profiles of Known Criminals

Detecting behaviours that are considered criminal/dangerous would provide the law enforcement agencies with a powerful tool against crime and a method for gathering evidence that would be useful in a legal case. Once the activities of a known criminal are available to the system, it can be used to analyze behavioural patterns to learn key determining factors that can be considered dangerous.

6.3 Objective 4.3: Development of Tools to Find Specific Information on the Internet

To help intelligence officer to efficiently find specific information we plan to build an enhanced search tool that can conduct search using syntactic and semantic information patterns and latent semantic analysis LSA (Latent Semantic Analysis). Such a tool is intended to allow meaning-based querying by permitting both conventional text queries and more importantly structured queries. The work will build upon our existing work in pattern based information extraction and LSA based text information similarity recognition.

6.4 Objective 4.4: Mining For and Detecting Suspicious Websites

Intelligence agencies have lists of known suspicious websites, but such websites may change their identities and new sites come into the picture. Hence, any such list be-comes

outdated quickly. It is necessary to build systems that will crawl the net to automatically find websites that have possible criminal uses.

6.5 Objective 4.5: Combining Confidence-Rated User Supplied Knowledge from Diverse Sources

Information is collected from diverse sources such as intelligence officers, police databases, public records and other databases of various kinds. An information analysis system is only useful if it permits easy integration and assimilation of such knowledge. The ability to integrate diverse forms of information will be a key requirement for the proposed system. The proposed system will allow confidence values to be attached to natural language text and to database records.

6.6 Objective 4.6: Tools to Process the Information Provided by Citizens

This objective concerns tools to process the information provided by citizens via emergency web-site: tools for filtering, validation and classifying textual information. Citizen provided information can vary from a hole on the road reports to reports on supposed criminal activity, etc. This task will combine a CMS (Content Management System) with a KMS (Knowledge Management System) incorporating intelligent information processing tools based on knowledge engineering.

6.7 Evaluation of Components related to Objective 4.1

6.7.1 State-of-the-Art Solutions Analysis

State-of-the-Art was analyzed in Deliverable D4.2. It provides a critical survey of the field of relation extraction. Relation extraction methods are divided into three categories, i.e. supervised, weakly-supervised and unsupervised approaches. This categorization is based on the ability of approaches to overcome the knowledge acquisition bottleneck. This common problem appears in most areas of NLP (Natural Language Processing) and it is caused by the lack of adequate training data for building classifiers using learning methods. These classifiers are then applied to unseen data in order to perform their task e.g. relation extraction, word sense disambiguation, question answering and others.

6.7.2 Progress beyond State-of-the-Art Solutions – In Terms of Engineering and Integration Advance

The importance of relation extraction results from its potential applications. In particular, it can be used to improve the accuracy of search engines by allowing them to handle complex queries, i.e. queries whose correct answers depend on their semantic interpretation.

6.8 Evaluation of Components related to Objective 4.2 – State-of-the-Art Solutions Analysis ¶

D4.2 has provided a thorough overview of the current-state-of-the-art on the annotation schemes employed for the identification of entities and the attributes that characterize them. The survey part focused on the annotation schemes used publicly and under license available data sets.

7 Work-Package 5

The main purpose of WP5 is to offer a complete solution for searching, identifying and storing documents and multimedia contents to make the process of identifying criminals, hot objects, or just finding document, easier and faster. In addition, the high security for information flow needed by police and prosecutor offices will be offered.

7.1 Objective 5.1: Designing New Algorithms

One of WP5 objectives is designing new algorithms both for metadata storage in multimedia content, which utilizes a watermarking-based approach with a background in current standards i.e. MPEG-7 and MPEG-21) and fast algorithms designed for searching such compressed and encrypted content with automatic authenticity checking. It aims at high encryption and redundancy whilst preserving perceived quality.

7.2 Objective 5.2: High Capacity Watermarking Technology

Another WP5 objective is development of the high capacity watermarking technology using fast algorithms of selected transforms for the purpose of fast searching.

7.3 Objective 5.3: Software Assistance Support for Searching

The next goal is designing and implementing the method of software assistance support for secure searching of required persons and documents, including any security activities of police and prosecutors.

7.4 Objective 5.4: Assorted Repositories

Facilitating the use of existing assorted repositories designed for police and prosecutor offices by adding new interfaces and integrating currently separate systems is another goal of the WP5.

7.5 Objective 5.5: Coherent Distributed System for Identification of Criminals

One of the most important objectives is establishing a coherent distributed system, this assists police and prosecutors in identifying criminals and hot objects in real time.

7.6 Objective 5.6: Semantic Search Engine

The last goal is developing a semantic search engine for local and wide urban areas including Query by Example (QbE) techniques.

8 Work-Package 6

The INDECT Portal developed in WP6 is an intelligent information tool intended to contribute to the detection and prevention of threats in order to preserve the security of citizens.

Work package 6 will be aimed at designing and implementation of an intelligent portal acquiring heterogeneous data, supporting decision making and functioning as ‘a single point of access’ supporting various locations of users equipped with various terminals, featuring role-based data verification, distinguishing the basis of facts correlation, helping users to identify potentially criminal situations, threats and attacks, visualizing (in a comprehensive way) results of the monitoring of indicated important parts of the urban infrastructure (crucial from the security of citizens’ point of view). Portal will be tested against unauthorized access to any privacy-related content, mostly focusing on the ethics-related content. All operations in the Portal will be recorded automatically. Only authorized, selected police officers will have access to selected, privacy-sensitive data.

From the system integration perspective, the main aim of INDECT Portal will be to integrate the services developed in the project and to provide them to the end-users.

In other words, INDECT Portal will be a joint information tool performing user-centric cross-WPs integration, where each WP co-operates as a high-level logical ‘super-component’ providing its share in functionality, according to the widely approved approach of Service Oriented Architecture (SOA). Thus, the INDECT Portal’s behaviour responds to users’ expectations regarding interactive support to daily tasks of police forces and the protection of citizens against threats. All these objectives for the crisis management portal - as one of the strongest points of the INDECT Project – will be approached through various tasks whose outcomes are presented below.

MOVIQUITY as leader of WP6 has contributed in the description of context for WP6 and the evaluation of the components considered for the WP6 Portal presenting the following approach: Presentation of progress achieved in the decomposed tasks for WP6 with regard to the specific objectives that were fixed for this WP; and its progress towards the overall goals of the INDECT Project.

Presentation of the work done within the corresponding deliverables.

8.1 Objective 6.1: User Requirements in Crisis Management Portal

The crisis management portal, developed in Work Package 6, has to respond to user needs, tailored in terms of features related to presentation and interaction with various INDECT data sources. The portal provides specific functionalities that are accompanied by security safeguards compliant with Privacy regulation.

As a first approach to accomplish requirements INDECT has to address the needs of two different classes of users:

- General citizens that have to be protected against threats and to whom refers the information handled (so that their Privacy concerns have to be respected). They have an open access to the website where aide search mechanisms have been developed.
- Police officers with their job profiles that give them personalised access rights to the portal authorising them secure interactions with the information that has been stored. The intelligent portal is conceived as a tool to help them to perform their work.

For the presentation and the interaction with the information an adaptive and interactive user interface has to be implemented.

As a first step to provide services and relevant information tailored to all the target groups and personalised for police officers as end users of the system all functional requirements have to be collected. This is done through the elaboration and distribution of End Users' Questionnaires considering several categories of questions that are specific to respondents. Furthermore, specific usage scenarios have to be modelled in order to provide guidance in the design of the architecture in a way that fulfils all service behaviour expectations.

Analysis of user requirements through End-user questionnaires and preparation of scenarios as input for functional and non functional requirements to be considered in the design of the architecture. Accomplishment of the goals after delivery of the document and link with future work.

8.2 Objective 6.2: System architecture for functional integration in the Crisis Management Portal

WP6 is focused on data fusion and presentation, which corresponds to the INDECT Portal platform‘s ability to support storage and exchange of different types of data, while always considering the requirements of privacy (secure access and transmission) related to sensitive information with regulation protection.

Moreover, regarding the data fusion and exchange in relation to an electronic data registry, various characteristics should be considered at the same time:

- The possibility of combining multiple formats for storage, data exchange and remote access, supporting share/reuse of the already acquired information (captured or obtained through prior activities).
- Standardised interfaces for internal and external data exchange and manipulation need to be defined using XML format for messages in Web services provided by each WP. The semantic integration of heterogeneous information sources where data has to be correlated and interpreted will have to be achieved by means of decision

support applications.

- The existence of requirements for privacy and security (when there is sensitive information involved) with restrictions that obey a very strict regulatory protection.
- The implementation of an adaptive and interactive user interface for the web based interface of the INDECT Portal.

Consideration of user requirements defined for the design of the architecture (business, logical, module, behavioural and deployment aspects). The system will provide its global functionality by means of three layers that have been implemented as part of logical SOA-based architecture. WP6 Portal will integrate heterogeneous data sources; where different streaming data can be acquired and processed.

8.3 Evaluation of Components Related to Objective 6.1

In the D6.1 it was detailed how needs of the system users are being analysed in-depth. Users' main expectations and necessities have been identifies through a well-designed questionnaire.

This end users questionnaire has been elaborated with the use of defined categories of respondents for each question (following police leaders' advice). It is planned:

- For general citizens main issues addressed are security protection approach and mechanisms followed in order to detect and fight against criminal activities, concerns about privacy, web-page facilities such as user-friendliness or intelligent support for queries, etc.
- For police officers as final users of the portal main questions deal with the capabilities expected from the portal and best practises regarding security mechanisms and compliance with regulations.
- System developers in charge of implementing technical solutions are also questioned as they are relevant actors for the design and development of the application.

Then, main questions have been organised in accordance with the following categories:

- Hardware: devices' properties used for detection, acquisition and the analysis of sound and image.
- Software – the part related to functionality of working systems.
- Security and privacy – the part currently related to access levels, law, and the possibility of allowing different procedures mentioned in other parts of the questionnaire.
- Events – the part related mainly to detection of threat, important events, recognition of a crime attempt, etc.

- Workflow – the part related mainly to messaging, reporting, and the flow of information

However, it is a long process that has not been finished yet. At this moment the security agents' needs are being addressed and the questionnaire results are being processed. After that, a questionnaire oriented to the citizens will be elaborated. The task 6.1, not finished yet, will ensure the INDECT Portal fulfils the end users' needs, meeting the WP targets.

In order to simplify the information processes in accordance with user requirements specific scenario types have been modelled describing the system profile.

The expected scenarios have to contain relevant information about the system and its time ordered functionality descriptions to fulfil the structured information need of the programmers. Apart from other factors, the level of success of a software system can be assessed with regard to two complementary types of requirements: The first type groups user-oriented requirements that concentrate on the functionalities provided to the system's user. The second type of requirements consists of context-oriented requirements. These can be system-related and human-related requirements.

8.4 Evaluation of Components Related to Objective 6.2

8.4.1 The initial evaluation of functional requirements coverage

According to the system requirements specified in D6.1, in result of the discussion with the police officers and analysis of the End Users Questionnaire, logical components that are responsible for realisation of those requirements were specified in D6.2 (see D6.2 Section 5.2). Definition of logical components was one of the crucial steps of the process of preliminary architecture definition.

In order to assure that requirements of all stakeholders were fulfilled, the architecture of the WP6 system has been presented as a set of different views (the business, logical, module, behavioural and deployment view). Each view addresses different aspect of the requirements. For example, in the logical view, *Secure Content and Access* subsystem is dedicated to handle needs corresponding to REQ-F400 *Content and access secure* requirements group); it consists of several components defined as corresponding to particular security-related and privacy-related tasks such as access control, authentication and authorisation or policy management. The deployment view shows how those elements will be deployed within the distributed servers' infrastructure.

The components responsible for realisation of a particular group of requirements related with WP6 are presented in the Table 1.

Table 1. Mapping between group of WP6 requirements and WP6 logical components.

		General Requirements (REQ-F100)	User manager (REQ-F200)	Content Management (REQ-F300)	Content and access secure (REQ-F400)	Ontological manager (REQ-F500)	GUI management (REQ-F600)	Video-call manager (REQ-F700)	DTMF manager (REQ-F800)	Message manager (REQ-F900)
1.	Portal Controller	+	+	+	+	+	+	+		+
2.	Web GUI	+	+	+	+	+	+	+		+
3.	GUI Plugin Manager	+				+	+			
4.	Semantic Mapper GUI						+			
5.	Shared Content Manager		+	+		+				+
6.	Content Handler			+	+					+
7.	Repository Manager		+	+	+	+		+		
8.	Workflow Manager		+	+		+		+		+
9.	Exporter			+	+					
10.	IVAS Composer			+			+	+	+	
11.	User Manager		+				+			
12.	Privilege Manager		+	+	+			+		+
13.	Context Sensing Engine		+	+						
14.	Profile Manager			+						+
15.	Activity Register		+	+	+					+
16.	Presence Manager		+							+
17.	Message Handler		+	+						+
18.	Message Box		+	+						
19.	Connection Manager						+			+
20.	Subscription Manager			+						+
21.	Access Controller		+	+	+					+
22.	Authentication/Authorisation Handler		+	+	+					+
23.	Policy Manager		+	+	+					+
24.	Cryptographic Key Manager		+		+					+
25.	Audit and Reporting Engine		+							
26.	Ontology Handler					+				
27.	Indexing Engine		+	+						
28.	Recommendation Engine		+							
29.	Data Mediator			+						
30.	Local Repository Handler		+	+						+
31.	Remote Repository Handler			+						
32.	IVAS Call Manager			+				+	+	
33.	IVAS Multimedia Manager			+				+	+	
34.	DTMF Handler			+					+	

The logical components correspond to the defined WP6 requirements at various scope, although it can be seen that most of them has to meet a higher number of requirements (i.e. uniform allocation of requirements have been quite successfully achieved). On the other hand, each requirement is fulfilled by at least one component what assures the full requirements coverage at the WP6 logical components' level.

8.4.2 The non-functional requirements validation

In current stage of the WP6 work, after defining the preliminary architecture, (see D6.2), some of the initial performance evaluations can be considered. The set of the most important non-functional requirements which were taken in to account is described below:

Security

It is one of the crucial aspect of the INDECT project. Security is understood as the ability to prevent and/or forbid access to the system by unauthorised users. Taking into account the importance of this aspect, a separate group of requirements was defined – *Content and access security* (D6.1 - REQ-F300). The essential elements of security include well-defined groups of end users and their privileges were defined in D6.2. Taking into consideration the presence of all the elements, it is much clearer that the INDECT Portal should offer not only a simple access control, but also more sophisticated access control methods. For example, it can provide actions such as subscription to a specified content, with access being confirmed by the end user with higher privileges (i.e., a police officer wants to analyse the document to which he/she does not have access, but can issue a request for it to his/her supervisor).

Performance

Considering the fact that the INDECT Portal will provide data to a wide group of users, the volume of data systems accessible to non-authenticated citizens may have a serious impact on the performance. The analysis of the preliminary portal architecture shows the presence of some potential bottlenecks and the components corresponding to these bottlenecks. One of such components is the Portal Controller, which is involved in almost all the users' online activities. In order to obtain a satisfactory level of performance, the INDECT Portal architecture has been defined in a way that enables the multiplication of components, thus potentially effectively coping with the architectural bottleneck problems.

Concurrency

This non-functional aspect has a direct relation to the system performance. However, as mentioned before, essential components, which has been identified as possible bottlenecks, should be implemented in a way enabling multi-access handling, and guaranteeing the data integrity, as well as avoiding potential deadlocks. This aspect is important because of the nature of the INDECT Portal: it handles a huge number of users that frequently

simultaneously access the same resources. For example, the portal will manage huge sets of documents that will be processed concurrently in different ways. For this reason, a special component (Shared Content Manager) has been defined - it will be responsible for handling all the simultaneous processes related to document management. Number of direct connections (interfaces) between components are shown below.

Component	Number of connected components
Local Repository Handler	19
Workflow Manager	18
Portal Controller	16
Repository Manager	12
Activity Register	8
Recommendation Engine	8
IVAS Composer	6
Presence Manager	6
IVAS Multimedia Manager	5
Context Sensing Engine	5
Access Controller	5
Semantic Mapper	5
Indexing Engine	5
IVAS Call Manager	4
Content Handler	4
Profile Manager	4
Message Handler	4
Remote Repository Handler	4
Semantic Mapper GUI	3
Exporter	3
User Manager	3
Privilege Manager	3
Message Box	3
Subscription Manager	3
Transformation Generator	3
Data Mediator	3
DTMF Handler	2
Shared Content Manager	2
Connection Manager	2
Authentication/Authorisation Handler	2
Ontology Handler	2
Web GUI	1
GUI Plugin Manager	1
Policy Manager	1
Cryptographic Key Manager	1
Audit and Reporting Engine	1

8.5 State of the Art Solutions Analysis

INDECT Portal will serve both police officers (equipped with PCs while working in the office and equipped with mobile phones while serving in the field) and citizens. Moreover it will provide end-users with a highly ergonomic (contextual, collaborative and personalized) single ‘point of access’ to multiple data sources. Given the heterogeneous information handled for full interaction with users, there will be an Advanced Interactive Video Streaming System

(IVAS) that will support communication and processing of multimedia streams by applying state-of-the-art multimedia streaming content processing and management methods.

The system will provide its global functionality by means of three layers that have been implemented as part of logical SOA-based architecture. Those are:

- Data acquisition and filtering layer (DAF) with basic functionalities related to WS Protocol Stack and the access and management of data stored in the servers used by WP6 Portal.
- Data correlation and semantic integration and decision support (DCDS) layer that provides intelligent processing of the collected data and a knowledge base to interface the information sources to the decision support system.
- Content management layer and data visualisation (VDM) layer that is responsible for the secure interaction with users through Web page or video call interfaces and the management of workspaces.

At current stage of WP6 work, several technologies and approaches to the system design have been identified. The Service Oriented Architecture (SOA) is considered as the basic design approach. Furthermore, two technologies have been chosen as the implementation platforms for the SOA approach: Web Services (WS) based on Simple Object Access Protocol (SOAP) and Representatial State Transfer (REST) web services. The Business Process Execution Language (BPEL) is regarded as a language useful to define WP6 workflows in terms of document management and decisions making (e.g., this related to crisis management). The interoperability will be achieved by using standard protocols (SOAP, HTTP, etc.) and languages such as the Extensible Markup Language (XML). Adaptability may be achieved by using the semantic technologies such as ontologies (e.g., PROTON) defined in Ontology Web Language (OWL) together with appropriate reasoning engines. The AJAX technology will be mainly used to create ergonomic user friendly graphic user interface.

Because the INDECT Portal architecture is designed on the basis of the SOA approach, it utilises standard protocols and enables comparatively easy (i.e., supported by the portal's 'intelligence') integration with new data sources and services. Therefore the INDECT Portal may be used to access data from already operating non-INDECT systems or from a system that will be developed in the future.

Usability of SOA approach

The WP6 Portal is focused on integrating functionalities provided by other WPs in one place. Taking into account the main market trends and relevant standards, as well as the fact that the system has a heterogeneous structure, the best approach to follow is the SOA (Service Oriented Architecture). Following SOA approach guarantees that the software provided by

other entities (other WPs) can be easily used and integrated, while not losing the ability to adapt to possible future changes. The advantages of the SOA approach, as the state of the art methodology for system integration, are as follows:

- SOA radically simplifies system integration, by allowing to bind systems dynamically. It is widely regarded as the state-of-the-art method for the design and implementation of complex IT systems based on combinations of universally interconnected and interdependent components called services.
- SOA is a technology, which is future-proof for at least the next 5 years.
- It is widely used on the world-wide IT market (it is supported by companies like Microsoft, IBM, Oracle, SAP, as well as by the leading open source projects including WSO2, FUSE, JBoss).
- It isolates user-centric application logic from the service implementation details (services are used as black boxes).
- It simplifies the usage of the existing (implemented) systems.
- One of its goals is to achieve increased interoperability (information exchange, reusability and composability of functions).
- It is platform independent.
- It complies with OASIS Standard 1.0 for WS_Security (an assessment of the required security mechanisms is done in D8.1 '*Specification of Requirements for Security and Confidentiality of the System*).

Heterogeneity of web services' interfaces

The main SOA implementations are using the so called "big web services" (i.e., the widely adopted web services based on SOAP, WSDL and UDDI), however, the REpresentational State Transfer (REST) is another important method for integration of self-contained (i.e., composable but at the same time providing interdependent functionalities) components of complex software architectures that has recently been gaining popularity (particularly among Internet companies). Because large part of WP6 Portal project is focused on document management functionalities, it is important that the REST approach may radically simplify the integration with at least some of the external INDECT components, especially those that are document-oriented rather than RPC-oriented. Reflecting the resource-centric view on web services (which is one of the foundational concepts of the REST methodology) in the INDECT Portal, the architecture design is especially important for the successful integration of the INDECT Portal with all the INDECT subsystems (in particular those developed in other WPs) which do not provide sophisticated programmatic Web interfaces. Building systems (developed outside of WP6) in the so-called RESTful way may help to reduce effort on cross-

WP integration, while still assuring an appropriate level of usability of document management functionalities (at the level close to this available when using SOAP-based “big web services”).

Semantic integration patterns

Experts from IBM (with a rich experience in application of the pragmatic SOA approach) have identified a few patterns of achieving semantic interoperability in SOA. Those patterns influence the key part of any semantically integrated system, in particular a portal, by defining the topology of distributed components composed into a semantic integration middleware. The INDECT system is a complex system used and managed by a single non-profit organisation, and at least some of the patterns presented below are relevant to this particular case.

The key patterns for achieving semantic interoperability in SOA may be roughly represented by the following cases:

In the case of point-to-point semantic integration, each data source has its own proprietary semantic model and semantic translation is performed in a distributed manner. The impact on other systems caused by a change of data definition in one of the integrated components is often unpredictable. On the other hand, many Enterprise Service Bus (ESB) projects still perform point-to-point semantic integration in SOA, as – when used selectively – it may ensure high performance and create a "fast track". It may be noticed that workflow and RPC-centric point-to-point communication model (that is very likely to be used in many INDECT systems integrated with the portal) may utilize business process modelling and development solutions, including tools for XML Schema mapping performed within BPEL systems.

In the case of hub-and-spoke semantic integration, each system operates on the basis of its own ‘local’ semantic ‘meaning’ mapped on a logical data model, which can be instantiated as a physical federated model or a canonical message model. Semantic interoperability is achieved at reduced redundancy and maintenance cost (thanks to avoiding point-to-point integration). Well-architected ESBs frequently use this pattern to map messages on a canonical message model and achieve semantic interoperability.

In the case of semantic interoperability based on Master Data Management (MDM), a single system connects heterogeneous information sources and produces a single version of the key information: a data instance, or metadata. As the MDM system is based on open standards, data may be truly reused as an organisation’s asset. Even when built separately from the existing systems (in order to reduce the drastic impact on businesses), MDM offers an option of legacy systems future migration to MDM.

The availability of industry information models (usually including XML messages and a message schema), also known as Domain Information Models (DIMs), depends on the

particular industry. On the other hand, some standard organisations tend to look at information from horizontal and cross-industry perspectives. For example, RosettaNet helps companies from multiple industries to provide RosettaNet Technical Dictionaries. DIMs are typically XML-based and are used to exchange messages in a business-to-business (B2B) environment. DIMs prompt a greater level of semantic interoperability and encourage asset reuse.

The widely referenced approach referred to as the Semantic Web assumes linking elements of the data model to a common ontology, thus semantically linking applications, even those operated by different organisational entities. According to the Semantic Web approach, the Resource Description Framework and the Web Ontology Language are used to allow data to be shared and reused on the Web.

An appropriately balanced trade-off of adopting different patterns of semantic interoperability should be the result of an analysis conducted at least from the perspective of the following issues:

- The strategy of the organisation (a company, or a non-profit organisation) for the next few years.
- Overall system dynamics (expected changes, especially related to semantic models).
- Planned initiatives to prompt greater asset reuse, minimize the maintenance cost and share the development cost.
- The regulatory environment of the IT group performing semantic integration.

It should be stressed that taking into account a pragmatic approach to semantic integration is especially important, because academic activities conducted in the area in the last few years (with the ambitious Semantic Web approach as the most popular research topic) correspond very weakly to the true market demands.

8.6 Progress beyond State-of-the-Art Solutions

Initial evaluation of WP6 components from the perspective of scientific and research advances shows that those of the INDECT Portal components, that are defined in D6.2 as means for achieving semantic interoperability, in particular enabling recommendation of semantic mapping (based not only on explicitly defined semantic models but also on behavioural models), represent an interesting, fairly new research topic, and a potentially rich source of innovative solutions.

One of the main problems that will be addressed by the WP6 Portal functionalities is the ability to correlate and integrate the information provided by other WPs. To achieve this goal, the INDECT Portal will be aimed at ‘understanding’ the data provided by other WPs. The variety of data types is caused by the diversity of the provided functionalities, e.g., metadata

describing pictures provided by WP5 may be closely related to the MPEG7 format, but those added to a text document found on the Internet (provided by WP3) may be based on the PROTON ontology. To effectively cope with the problem of data heterogeneity, the portal will include the SIX P2P component (Semantic Integration of XML Data in Peer to Peer Environment). It will support the construction and use of the common metadata format – the INDECT ontology (using the semantic analysis for that purpose), as well as online semantic transformation.

Semantic Mapper and Transformation Generator components will help to generate transformations (defined in XSLT) that are applied 'between' the WP6 data (represented in the form of WP6 XML Schema) and data formats used in other WPs (WPx XML Schema). This operation will be done during the development/maintenance phase once and will have to be redone after changing any of the XML Schemas involved in the WP6-WPx interactions. The Semantic Mapper GUI component will be an innovative graphical user interface suggesting to the user (the maintenance administrator) possible mappings between the terms in different XML Schemas. In order to prepare the mapping recommendations, the SIX P2P component will use a reasoning engine and a recommendation engine operating on Semantic Vector Space Model (SVSM), the INDECT ontology and additional results of the analysis of textual corpora. Data Mediator component will use generated XSLT documents to perform effective online translation between data formats used by the WP6 Portal metadata description and the data formats used by web services provided by other WPs.

The user-centric perspective (that is reflected by the initial INDECT scenarios defined in D6.1) implies that ‘understanding’ the semantics of the relations between the entities in the heterogeneous data offered by various systems from different INDECT WPs (rather than simple data availability) may be regarded as the key WP6 success factor. The component to be used within the INDECT WP6 architecture (with the aim of assuring ‘semantic compatibility’ of systems provided by different WPs both SOAP-based and REST-based) cannot (and should not) operate fully automatically. For this reason the INDECT SIX-P2P component will include semantic groupware system that will allow for manual refinement of both rules for online cross-WP XML data translation and the resulting XSLT documents.

Correspondences between diverse underlying XML schemas (specific to different INDECT WPx systems and obtained by means of the automatic XSD matching process driven by the INDECT ontology and enhanced by probabilistic reasoning and vector space similarity evaluation techniques) confirmed by the *integration maintainers* will not serve as the only basis for cross-WP XML data mediation. When relevant to some specified interface being under development, the automatically identified correspondences will be recommended to the INDECT system integrators as potential target rules for online cross-WP XML data translation. In order to make such recommendations useful (thus effectively reducing the

number of necessary human interventions and making the system more ergonomic), the level of relevancy of each rule to some specified case of XML data translation task will be automatically evaluated by means of probabilistic reasoning and vector space similarity evaluation techniques. Only after being manually approved by the *integration maintainers*, a can new rule be actually used for inter-WP online translation of XML data carried in SOAP messages or responses of RESTfull web services.

To the knowledge of WP6 partners, it is very likely that all the WP6 semantic integration components mentioned above will be fairly innovative solutions – well beyond the state of the art in the area of semantic web system integration.

It should be stressed that long-term, 'scientifically-oriented' research is not the only field of expected technological advance to be achieved by WP6. As far as advances in system integration is concerned, at least some features of WP6 components and the corresponding WP6 architectural solutions (both described in D6.2) may be regarded as reaching beyond the state-of-the-art in semantic web system integration. In particular, placing a component responsible for the offline task of semantic models alignment (performed mainly by means of the SIX P2P, Semantic Mapper and Transformation Generator components) at the core of the INDECT Portal architecture will enable the INDECT Portal to cope effectively with the dynamic nature of semantic models (used by experimental subsystems implemented in different WPs of INDECT).

From the general system integration perspective, the INDECT Portal can be seen as a system enabling innovative semantics-driven SOA-based integration of the portal with many different systems provided by other INDECT WPs (this view corresponds to the INDECT system integrators' perspective), at the same time being a semantic Web mashup, or smashup (this view corresponds to the portal end user's perspective). Effective semantic models alignment in heterogeneous service environment of dynamically changing semantic models should be achievable as a result of enabling appropriately quick and not complicated responses of the INDECT Portal semantic integration middleware to changes being introduced to the WPx semantic models (such changes that are very likely to be necessary throughout a few year-long project).

One of advances in system integration to be achieved by WP6 is represented by the WP6 semantic system integration approach assuming 'semantic combination' of SOAP-based "big web services" and semantic mushup-supporting RESTful web services. Usually, SOAP-based web services are regarded as basic components of an SOA – 'building blocks' being composed into 'larger' components offering more abstract (i.e., higher-layer) functionalities. However, following the SOA approach is not contradictory to the document-centric (in particular the 'RESTfull') view on web services. The INDECT Portal semantic integration

system will provide an unique value by enabling the semantic combination of SOAP-based "big web services" and semantic mushup-supporting RESTful web services – still staying in line with the basic principles of the SOA-based system design methodology. In the advent of migration of the leading Web applications design methodology from Web 2.0 into Web 3.0 vision, such a functionality represents a very important advantage of the INDECT Portal architecture over widely referenced approaches to semantic system integration (usually dealing with cases of much less heterogeneous data). At the same time, providing intelligent components enabling effective semantic system integration is probably one of the key INDECT WP6 challenges (a source of potentially highly innovative research results).

What is probably the most important means for effective integration of both RPC-oriented SOAP-based web services and resource-centric REST web services as components of a single system of the INDECT Portal (as seen from the perspective of WP6) is extensive reuse of features of the INDECT Portal system components responsible for semantic integration (first of all the ontology-based automatic semantic mediation features) achieved with regard to all the interoperating systems. Non-SOAP/WSDL services are widely powering Web 2.0 applications – experts see a clear value in providing semantic-based solutions to create powerful smashups where it is easier to integrate data and services on the client-serving system. It should be noted that, apart from service discovery (which is not of high importance in the case of the INDECT system), design of the INDECT Portal components will reflects semantics-related requirements of all the stages of the Semantic Web Services (SWS) life cycle: annotation, publication, discovery, data mediation, composition or configuration, orchestration, and execution.

It should be noted that in the field of business process semantic integration, a clear demonstration of how SWS can ease data-mediation (transforming the output of one task into the input of the subsequent task or tasks) still remains a challenge – one to be addressed by WP6.

8.7 Summary of Preliminary Results

After discussions with the police officers involved in the project and according to the first questionnaire results, the WP6 scenarios have been defined. On such basis it is possible to highlight how important is for INDECT end users to be provided with features like data access personalization, user-friendly interface and availability of access through different terminals. Moreover, it should be underlined that the INDECT PORTAL will enable ergonomic access to advanced functionalities provided by other WPs. Detailed scenarios have been developed satisfying the detected expectations and the foreseen WP6 goals. The analysis has been carried out not only from the end-users point of view but also from a technological point of view.

Thanks to the 'intelligence' of the mentioned WP6 components responsible for semantic integration and mediation, the INDECT Portal will not only be able to adapt to dynamically changing formal semantics of data provided by different INDECT WPs, but also to dynamically changing implicit semantics of data exchanged by these systems – what will be a clearly innovative feature of the mentioned WP6 components. As stressed by some widely cited authors, it is unrealistic to assume that semantic descriptions of services are fully correct and complete — that is, that they duplicate service functionality at the description level. More generally, targeting full completeness and correctness of data semantics representation is not reasonable anymore, as the underlying (implicit) data semantics is changing faster than any reasoning process can process it. Respectively, as far as the INDECT Portal architecture is concerned, reasoning technologies will mainly be applied to semi-automatic heterogeneous data mediation and will be based not only on formal semantic models, but also on models reflecting the 'open world' and the behavioural view on data semantics, e.g., the information retrieval-derived methods for the observation and modelling of semantic mapping utility. Such an approach represents a potential source of highly innovative results to be achieved in WP6 as a result of the work on the INDECT Portal.

Taking into consideration the fact that the INDECT Portal is based on the SOA approach and that some of the functionalities provided by other WPs may not be accessible online, the proposed architecture assumes the use of plug-ins. This simple solution makes the portal more flexible in terms of user-friendly interface design, and gives the possibility to use already implemented software solutions. WP6 Portal will integrate heterogeneous data sources, including different sources of streaming data (which can be both acquired and post-processed). The functions of the streaming-oriented IVAS subsystem are focussed on multimedia data handling. This aspect of the INTECT Portal functionality is also in the logical architecture where Data visualisation layer is responsible for the interaction with users through the video call interfaces (IVAS).

The WP6 Portal, as an electronic communication infrastructure, will create a secure environment facilitating both regulatory compliance and business improvement and considering the security and privacy concerns of the users. In other words, the security subsystem will be made strictly policy-driven in order to provide the capability of adaptation to different legal obligations. Moreover, the INDECT Portal will be made capable to create audit trails. The behaviour of the portal users will be monitored in order to assure that the portal is used according to its purpose and that no information security policies are being violated.

9 Work-Package 7

In near future, the best solution for extraction of information from signals to protect against terrorism, to manage crises, and to operate in situations of danger will consist of a combination of the advantages of artificial intelligence, biometrics (understood in a wide sense), communication systems, and computer networks together with the human experience. Thus, automatic recognition of criminal and/or dangerous situations including automatic recognition of offenders and/or criminal behaviour via parametric modelling and description should serve in the construction of intelligent systems for data processing and mining, which will form the basis of security systems.

The intelligent information system which supports observation, searching and detection for security of citizens must have following features:

− Applications with intelligent algorithms

− Devices, which are connected with applications

− An Integration platform which supports: the workflow for security issues with a graphical user interface. Transfer of data between different kind of applications with proper interfaces and data base system which manages the data which is to be shared among applications

9.1 Objective 7.1: Integration of Security Systems

Integration of security systems with emergent wireless communication systems and self-organizing computer networks in order to achieve their functionality and interoperability for extraction, processing, distribution, and supporting security information on citizens of urban environments is the key objective within the WP7. The developed services should include such tasks as monitoring of figureheads, detection of criminal behaviour, detection of threats, as well as automatic and intelligent notification of people and their protection.

9.2 Objective 7.2: Artificial Intelligence and Biometrics

Another objective is the artificial intelligence and biometrics, which is understood in a wide sense, i.e., not only as a collection of methods for identification/recognition of individual offenders but also for automatic detection of criminal behaviour of anonymous people. It should be used in order to protect against terrorism, to manage various crisis situations, and to tackle threats typical for a society living in a modern, highly technological and dense urban environment.

9.3 Objective 7.3 Automatic Detection of Threats

The last objective is automatic detection of threats and recognition of criminal behaviour of individuals. Modern pattern recognition algorithms, digital signal processing, artificial intelligence procedures, as well as methods for motion and features analysis will be taken advantage of for monitoring phenomena and behaviour in the environment. Thus, an approach for the automatic detection and prevention of situations with an increased probability of danger needs to be developed and put into practice.

9.4 Summary of the Preliminary Results

For the WP7, system architecture was defined. The WP7 integration platform can be used as a command and control system with typical functions. Communication with XML is the basic approach for the integration of external applications. XEP will be used for the exchange of XML files. System components have to be specified as well as interfaces to the external applications and main processes for security services and a functional specification. The implementation of the integration platform and integration of some applications has to be done. A detailed validation of the integration platform will be performed within the project.

10 Work-Package 8

The WP8 includes several objectives, including specification of requirements and solutions for secure data transfer, definition of rules and mechanisms for privacy protection, definition of procedures for confidential information access, development of a federated identity model, security and privacy for the overall system, and promotion and standardization.

10.1 Objective 8.1: Specification of Requirements and Solutions for Secure Data Transfer and Privacy Protection

Objective 8.1 assumes analysis of current security solutions applied in telecommunications networks for secure information transfer and in-depth research to verify security of cryptographic algorithms based on proposed S-boxes (Substitution boxes in cryptographic algorithms). The next task within this objective is analysis of properties of S-boxes, implementation of S-box generator and optimization and evaluation of a block cipher. Next, a proposal of new methods of concurrent error detection in hardware implementation of a cipher will be considered and suitable methods for secure transfer of data acquired in real-time in Node Station to the Central Station and then to police mobile terminals will be proposed. A very important task is security and authentication development for mobile ad-hoc terminals at the network layer. In particular a new secure ad-hoc routing protocol with multi-path capabilities will be designed in order to withstand both logical attacks (i.e. routing-based) as well as physical attacks (i.e. jamming, uncooperative MAC).

Requirements for overall security and confidentiality of stored and transferred information will be specified, including analysis of current security solutions applied for secure information transfer. Furthermore, differentiation of security and privacy levels will be specified in respect to confidentiality and sensitivity of the information. Definition of schemes and mechanism for control and monitoring of the access to the information stored and exchanged within INDECT system will be provided within the WP8. In particular, Federated ID technologies will be employed in order to ease authentication and authorization management of multiple IT systems across different organizations or departments of a single organization. Quantum cryptography methods for security and privacy assurance will be proposed and examined within this objective.

11 Work-Package 9

WP9 includes several objectives, including: test-bed for intelligent monitoring and automatic detection of threats in urban areas, test-bed on identification and observation of mobile objects in urban environment, test-bed on detection of threats and criminal activities in complex real/virtual environment, system for information extraction from Web and semi-structured data, test-bed for search engine for fast detection of persons and documents based on watermarking technology, test-bed for interactive multimedia applications gateway for intelligent observation system, test-bed for biometrics and intelligent methods for extraction of required information, test-bed for overall system control, security and privacy management, and - dissemination and exploitation.

11.1 Objective 9.1 Dissemination and Exploitation

The dissemination of the research results, the maintenance of permanent relationship with Standardization Bodies, and the continuous technology transfer to industry are the main objective within WP9. The project will be widely publicized and promoted within the research community by publications and participation in conferences and workshops. Periodical workshops for European Commission representatives (Project Officers, etc.) will be organized in Brussels. Active contributions will be provided to standardization bodies like ETSI and the IETF.

11.2 Objective 9.2: Test-Bed for System Components

The Node Station will be installed between months 35 and 47 in various points of city agglomerations and schools. During a cycle of autonomous processing its accuracy at threat detection and person identification will be verified. Its performance in the following scenarios is to be verified: Person identification for schools, Hallways monitoring (left luggage, fights, falls), Audio signals recognition (calling for help, direction estimation), Traffic monitoring. Working scenarios will be presented to End-Users within this objective.

Symmetric block ciphers require a high level of trust in their security. Evaluation of these ciphers is heuristic. Only those attacks which are known at the time can be taken into account. The most important attacks include linear cryptanalysis and differential crypt-analysis. They are applied individually to each cipher. They enable the evaluation of both the components, and the general structure of the newly constructed cipher. Unfortunately, in the case of large S-boxes and complex structure, such an evaluation requires huge amounts of computer time and memory.

Stream ciphers require consideration of some aspects distinct from those required in the design of block ciphers because they are used in such applications as on-line ciphering. However, the same kinds of components used in block ciphers are also used in stream ciphers, e.g. substitution block.

A new dedicated hash algorithm designed in answer to weaknesses of the MD/SHA family will be developed. Recently proposed attacks on well-known and widely used hash functions motivated a design of new hash functions. We propose a hash algorithm in which many elements can be parameterized. Changing the parameter of the function dynamically changes the way a hash value is computed. This approach makes the hash algorithm interesting not only because of its construction, but also because of enhanced security against attacks.

Evaluation of the vulnerability of new constructions to errors will be carried out by computer simulation. Both random errors and premeditated errors will be considered. The Active – HDL simulator will serve as a basic tool. Results will be used to strengthen the construction against errors and to improve utility and security.

Assuming sufficiently large sizes of S-boxes, evaluation of ciphers can be restricted to the case of linear cryptanalysis. The following three methods of block cipher evaluation are distinguished. In the first, exact method, the best nonzero linear approximation of a cipher is determined. In the second, rough method, the best non-zero linear approximation of a cipher is assumed to be a composition of the best nonzero linear approximations of a single iteration. In the third, intermediate method, it can be found that the best zero-nonzero approximation of a cipher that fulfils the so-called approximation conditions. For a cipher graph, G, of zero-nonzero approximations, the SP algorithm calculates the shortest path of a specified length in the graph, G. This path determines the best zero-nonzero approximation of the cipher that fulfils approximation conditions. The effectiveness of this approximation is used to evaluate the cipher.

New methods of error detection will be proposed. Special stress will be put on those methods requiring low overheads. Single and multiple faults, transient and permanent faults will be considered. The analysis of the effectiveness of these methods will be carried out using the Active-HDL simulator. Implementation of the proposed methods of error detection will improve the utility and the security of the cipher reducing the probability of the cipher being broken by exploiting errors.

The system provided to the user consists of different tools to process the information produced by different kind of verbal behaviour, i.e.: tools for information extraction from texts provided by citizens through emergency websites, high precision tools for finding highly specific information on the Internet, e.g. sectarian websites or recipes for homemade

explosives hidden among cookbook recipes, etc., precision tools for analysis of server content looking for particular information, e.g. e-mails which contain information similar to the information derived from some external source, Web server logs, etc., tools for filtering, classifying and profiling textual information. All the tools are based on different techniques of automatic information extraction from text. Tools developed by tasks of WP4 will serve as the basic components of the system for information extraction. Where text data sources are bolded and system modules are marked by tool ID. Systems of this type will be flexible enough to: extract information from texts related to different forms of verbal behaviour and obtained from multiple sources, i.e. Web, citizen provided and user selected or provided, perform multi-paradigm reasoning and automated operations on text information and knowledge, support the user in decision making. Components of the system will be tested by potential users during the course of the project:

Document Updates

Version[2]	Date[3]	Updates and Revision History[4]	Author
1	2009.12.20		N. Johanning
1.1	2009.12.29		N. Johanning, J. Danda, M.Sobczak, J. Derkacz, M. Maderska, A. Dziech
2	2010.01.15		N. Johanning, M. Sobczak, A. Figaj, M. J. M. Gil, J. Derkacz, J. Danda

[2] In form of "vYYYYMMDD"; Version number and edition should correspond to the actual document name conventions.
[3] In form of "DD/MM/YYYY"
[4] Attach as appendix document reviews when appropriate; describe also the current status of the document e.g. "released for internal review", "released for comments from partners"