This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 607881

Petty cRiminality diminution through sEarch and Analysis in multi-source video Capturing and archiving plaTform

P-REACT

Instrument:	Research and Innovation Action
Thematic Priority:	FP7- SEC - 2013.7.2-1
Grant Agreement:	607881

D2.2 - P-REACT Local Embedded framework and system on the cloud requirements

Deliverable Number	D2.2
Title	P-REACT Local Embedded framework and system on the cloud requirements
Version	4.6
Date	21/07/2015
Status	Final version
Dissemination Level	PU (Public)
Nature	Report

EC Distribution

Project Partners: Vicomtech-IK4 (VICOM); Kinesense (KS); Aditess (ADI); Future Intelligence (FINT); Center for Research and Technology Hellas (CERTH); Center for Security Studies (KEMEA); Societa Reti e Mobilita (SRM)

Contributors:

Main: Naiara Aginako [VICOM]

Others: Marcos Nieto [VICOM]
Peter Leškovský [VICOM]
Anargyros Sideris [FINT]
Giorgios Kioumourtzis [KEMEA]
Prokopios Drogkaris [KEMEA]
Georgios Stavropoulos [CERTH]
Mauro Borioni [SRM]
Alexandros Kyriakides [ADITESS]
Nikos Koutras [ADITESS]
Nectarios Efstathiou [ADITESS]
Mark Sugrue [KINESENSE]
Sarah Doyle [KINESENSE]

Every effort has been made to ensure that all statements and information contained herein are accurate; however the Partners accept no liability for any error or omission in the same.

Document Control

Version	Date	Author	Modifications
0.1	18/07/2014	Naiara Aginako	Initial draft
0.2	20/07/2014	Mark Sugrue (KS)	Added some sub headings
0.3	20/07/2014	Georgios Stavropoulos (CERTH)	Initial contributions to Sections 2 and 5
0.4	20/07/2014	Georgios Kioumourtzis (KEMEA)	Initial contributions to Section 2
0.5	01/08/2014	Anargyros Sideris (FINT)	Initial contributions to Sections 6, 6.1, 7, 7.1
0.6	08/08/2014	Mauro Borioni (SRM)	Initial contributions to Section 4.9
0.7	10/08/2014	Mauro Borioni (SRM)	Added contributions to Section 4.9
1.0	10/08/2014	Naiara Aginako (VICOM)	Initial contributions to Section 5 Consolidate first version
1.1	02/09/2014	Anargyros Sideris (FINT)	Updated contribution to Sections 5, 6, and 7
1.2	02/09/2014	Georgios Kioumourtzis (KEMEA)	Contributions to Section 2
1.3	08/09/2014	Prokopios Drogkaris (KEMEA)	Contributions to Section 1
2.0	08/09/2014	Naiara Aginako (VICOM)	Consolidate second version
2.1	01/10/2014	Georgios Stavropoulos (CERTH)	Updated contribution to Sections 4 and 6.2
2.2	02/10/2014	Anargyros Sideris (FINT)	Addition of some corrections to Sections 5.1 and 5.2
2.3	06/10/2014	Naiara Aginako (VICOM)	Updated contributions to Sections 1, 5 and 6.2.3
3.0	07/10/2014	Naiara Aginako (VICOM)	Final version
3.1	07/010/2014	Juan Arraiza (VICOM)	Final version revision
3.2	08/10/2014	Georgios Kioumourtzis	Final version revision
3.3	08/10/2014	Mauro Borioni (SRM)	Final version revision
4.0	09/10/2014	Naiara Aginako (VICOM)	Revised Final version
4.1	03/07/2015	Peter Leškovský (VICOM)	Major changes after rejection in Mid-Term Review. Added: section 5.1 on scenario description, section 6.3 on HW reqs for each scenario and section 8 on evaluation methodology and proposed performance measures
4.2	05/07/2015	Marcos Nieto (VICOM)	Revision and contributions in sections 5-8
4.3	09/07/2015	Sarah Doyle (KS)	Quality Review and Final Revision
4.4	17/07/2015	Peter Leškovský (VICOM)	Revision after review. Main changes in sections 5 and 8
4.5	20/07/2015	Juan Arraiza (VICOM)	Updates in sections 5 and 8 to harmonise with deliverables D1.2 and D2.3
4.6	21/07/2015	Georgios Stavropoulos (CERTH)	Minor updates with regard to Depth analytics in Section 8.2 and tables 10 and 21

Table of contents

Annexes

Tables

Figures

1. Overview

P-REACT's main goal is to design and develop a low cost surveillance platform that will detect Petty Crime incidents. The solution is comprised of two primary components, a local embedded platform and a cloud platform. The inclusion of intelligent video and audio analysis algorithms running on the embedded and cloud platform will create a pro-active, reliable and scalable solution to address petty crime incidents.

P-REACT's DoW describes this deliverable as: *D2.2) P-REACT Local Embedded framework and system on the cloud requirements [month 6].* The aim of this document is to give a detailed description of P-REACT's requirements both for the Local Embedded Framework and for the system on the cloud. Hardware, software, communication, security and other requirements are addressed here within. This document includes inputs from Task *T2.2 Security requirements for Data Integrity and Privacy* and Task *T2.4 System Capabilities Analysis and Specifications*.

This document includes the following sections:

- **Section 2 - European Legal Framework for Security and Privacy**: In this section, a description about the European Legal Framework for Security and Privacy is presented. The main goal of this European Framework is handling and protecting European citizen's personal data.
- **Section 3 - Social and Ethical Considerations**: A short introduction of the considerations that will be considered in P-REACT regarding the social implications of surveillance systems. The relevant aspects are presented.
- **Section 4 - Data Physical Protection**: As stated in the title, in this section, requirements regarding data physical protection are presented.
- **Section 5 - P-REACT platform**: A brief introduction to the overall architecture is presented in this section, in order to identify the communication and security requirements of the P-REACT platform. In addition, a description of the considered end use scenarios, with preliminary analytic and Hardware (HW) requirements, is presented.
- **Section 6 - Local Embedded System**: In this section, identified hardware and software requirements for local embedded framework are described. This section presents also the interfaces between the Sensor Manager and the analytics modules.
- **Section 7 - Cloud Based System**: Similar to in the previous section, hardware and software requirements for Cloud Based System are presented. Software requirements are described

separately for each of the main modules of the defined P-REACT architecture.

- **Section 8 – Evaluation methodology**: This section presents the methodology that will be applied for functional testing and quality assessment of the analysis modules, with metrics and statistical measures that will be used to quantify the performance of the analytical modules.
- **Section 9 - Conclusion**: Finally conclusions of the deliverable and the work done in tasks T2.2 and T2.4 is are described.

As it can be inferred from the description of the sections, Sections 2-4 are more related to Task *T2.2 Security requirements for Data Integrity and Privacy*. This deliverable describes what aspects have to be taken into account, and which are the considerations to tackle within P-REACT. The way these aspects are being implemented will be described in deliverable *D2.3 P-REACT Conceptual architecture including functional, technical, interoperability and large scale deployment specifications*.

Sections 5-7 are focused on the description of hardware, software, communication and security requirements, more related to activities accomplished in Task *T2.4 System Capabilities Analysis and Specifications*. For a better understanding of the presented requirements, a short description of P-REACT's architecture is given, which will be explained further in D2.3.

2. European Legal Framework for Security and Privacy

This section presents and describes the regulatory European Framework that will be applied to P-REACT project. As mentioned in the previous section, the completion of the actions accomplished within P-REACT in order to fulfill this European Framework will be described in deliverable D2.3. The aim of this section is focusing on the Legal Framework that will be followed in the project.

Personal data is collected and used in many aspects of everyday life and can be collected directly or indirectly and it can be used for different purposes than the ones initially appointed, European regulation framework handles and protects this data by overcoming potential discrepancies among national laws. Therefore, national level Legal Frameworks are not considered in this document due to the need of a common framework for all the European countries. This doesn't mean that national level framework directives will not be implemented; this issue will be tackled before the installation of P-REACT platform in the local area, adapted to the corresponding framework.

2.1. European Convention on Human Rights (ECHR)

The scope of the right to privacy, from the intimacy of the home, has been gradually extended by the European Court on Human Rights to portions of peoples' lives that are not necessarily '*intimate'* strictly speaking, and may be relevant to personal behaviours, attitudes, held outside personal homes and private premises. In fact, the scope of privacy is thus not limited by the non-intimate nature of personal information or acts concerned, nor by their public occurrence. Individuals enjoy a right to privacy even with regards to behaviours, attitudes and communications in public spaces like streets, shopping malls, airports or even at work. This means that recording, storage and use of information related to individuals in these places constitutes an invasion of their privacy that must, in order to be lawful, comply with the conditions set at the article 8 of the European Convention of Human Rights:

> *"There shall be no interference by a public authority with the exercise of this right except such as is in accordance with the law and is necessary in a democratic society in the interests of national security, public safety or the economic well-being of the country, for the prevention of disorder or crime, for the protection of health or morals, or for the protection of the rights and freedoms of others."*

2.2. Charter of Fundamental Rights of the European Union (CFREU)

Within Article 1 of the Charter of Fundamental Rights of the European Union is specified that human dignity is inviolable. It must be respected and protected. The principle of human dignity attests to the fundamental and guiding role occupied, in our western legal culture, by the ethical imperative of conceiving and dealing with human beings always as ends in themselves and never as means to an end. Moreover, Article 8 specifies that “Everyone has the right to the protection of personal data concerning him or her. Such data must be processed fairly for specified purposes and on the basis of the consent of the person concerned or some other legitimate basis laid down by law. Everyone has the right of access to data which has been collected concerning him or her, and the right to have it rectified.”

2.3. Directive 95/46/EC

The Protection of Private Data Directive 95/46/EC is a directive adopted by the European Union designed to protect the privacy and protection of all personal data collected for or about citizens of the EU, especially as it relates to processing, using, or exchanging such data. It includes all key elements from article 8 of the European Convention on Human Rights, which states its intention to respect the rights of privacy in personal and family life, as well as in the home and in personal correspondence. The Directive 95/46/EC was developed to harmonize national laws for personal data protection and movement of data, based on the existing national legislations of the EU Member States.

The Directive 95/46/EC is the reference text, at European level, on the protection of personal data that sets up a regulatory framework which seeks to strike a balance between a high level of protection for the privacy of individuals and the free movement of personal data within the European Union. To do so, the Directive sets strict limits on the collection and use of personal data and demands that each Member State sets up an independent national body responsible for the protection of these data.

To properly understand implications of the directive, it is important to describe the following terms used in the directive:

- **Anonymous data**: Any data that is rendered in such a way that the data subject is no longer identifiable either directly or indirectly, in particular by reference to an identification number or by one or more factors specific to his physical, physiological, mental, economic, cultural or social identity.
- **Data subject:** The data subject is an identifiable natural person whose personal data are collected, held or processed.
- **Data controller**: The data controller is the person or administrative entity (e.g. a General Director or a Head of Unit of the European Commission) that determines the purposes and means of the processing of personal data on behalf of an institution or body. In particular, the controller has the duties of ensuring the quality of data and, in the case of the EU institutions and bodies, of notifying the processing operation to the data protection officer (DPO). In addition, the data controller is also responsible for the security measures protecting the data. The controller is also the person or entity that receives a request from a data subject to

exercise his or her rights. The controller must cooperate with the DPO, and may consult him or her for an opinion on any data protection related question.

- **Personal Data**: Any information relating to an identified or identifiable natural person, referred to as *data subject*. An identifiable person is someone who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his or her physical, physiological, mental, economic, cultural or social identity.
- **Processor:** A processor is a natural or legal person, public authority, agency or any other body which processes personal data on behalf of the controller. The essential element is therefore that the processor only acts on behalf of the controller and thus only subject to the controller's instructions.
- **Processing (of personal data)**: Processing of personal data refers to any operation or set of operations which is performed upon personal data, whether or not by automatic means, such as collection, recording, organisation, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, blocking, erasure or destruction.
- **Sensitive data**: Sensitive data include data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, trade-union membership, and the processing of data concerning health or sex life.

2.4. Directive 97/66/EC

Directive 97/66/EC of the European Parliament and of the Council of 15 December 1997 concerns the processing of personal data and the protection of privacy in the telecommunications sector. It provides guidance for the harmonization of the provisions of the Member States required to ensure an equivalent level of protection of fundamental rights and freedoms, and in particular the right to privacy, with respect to the processing of personal data in the telecommunications sector and to ensure the free movement of such data and of telecommunications equipment and services in the Community. For P-REACT project purposes, the following articles are of particular interest.

- **Article 4, Security:** The provider of a publicly available telecommunications service must take appropriate technical and organizational measures to safeguard security of its services, if necessary in conjunction with the provider of the public telecommunications network with respect to network security. In case of a particular risk of a breach of the security of the network, the provider of a publicly available telecommunications service must inform the subscribers concerning such risk and any possible remedies, including the costs involved.

- **Article 5, Confidentiality of the communication:** Member States shall ensure via national regulations the confidentiality of communications by means of a public telecommunications network and publicly available telecommunications services (this does not apply for any legally authorized recording of communications in the course of lawful business practice for the purpose of providing evidence of a commercial transaction or of any other business communication).

The 1997 directive was replaced in 2002 by Directive 2002/58/EC which updated the aforementioned rules and principles.

2.5. Directive 2002/58/EC

Directive 2002/58/EC focuses on issues concerning the processing of personal data and the protection of privacy in the electronic communications sector. The following definitions apply:

- **Traffic Data** means any data processed for the purpose of the conveyance of a communication on an electronic communications network or for the billing thereof.
- **Location data** means any data processed in an electronic communications network, indicating the geographic position of the terminal equipment of a user of a publicly available electronic communications service.

More specifically, we mention the following areas related to the P-REACT project:

- **Processing Security**: Electronic communications services must be securely protected by their provider by (a) ensuring authorized only personal data access, (b) protecting personal data integrity and (c) ensuring the implementation of a security policy on the processing of personal data. In the case of a personal data breach, the provider must inform the person concerned, as well as the National Regulatory Authority (NRA).
- **Data retention:** The Directive determines that traffic data and location data must be erased or made anonymous when they are no longer required, except if the subscriber has given their consent. Regarding data retention, the Directive states that Member States may withdraw the protection of data only in the exceptional cases of criminal investigations or in order to safeguard national defence and public security. Such action may be taken only where it constitutes a "necessary, appropriate and proportionate measure within a democratic society".
- **Controls:** Member States must implement a system of penalties, in the case of data breach to the provisions of this Directive, and ensure that the national competent authorities have the necessary powers and resources to monitor and control compliance with the national provisions

2.6. Directive 2006/24/EC

The directive focuses on issues concerning the retention of data generated or processed in connection with the provision of publicly available electronic communications services or of public communications networks and amending. The EU Data Retention Directive 2006/24/EC in addition to the directive 2002/58/EC, defines that at member state level each EU country should have their own version of the Data Retention Directive embodied and incorporated into their national laws.

The data retention regulations will impact public communication providers (fixed, mobile telecoms, ISPs) that have communications data generated or processed on their networks or from using the services they provide. The regulations require traffic, location and subscriber data to be maintained for a minimum of 6 months up to 4 years. The regulations also outline four data security principles that should apply to retained data:

- **Security**: Data must have the same security levels and quality during their retention period.
- **Responsible Management:** Technical and organizational measures must protect against accidental or unlawful disclosure and data loss.
- **Accessibility**: Retained data must only be able to be accessed by authorized persons.
- **Destruction**: All data retained must be completely destroyed at the end of the retention period.
- **Transmission**: When data is requested by law enforcement the data must be able to be transmitted without undue delay.

2.7. Directive 2009/136/EC

Directive 2009/136/EC amends the Directive 2002/58/EC; First of all the focus enlarged in order to ensure not only the right to privacy but right to privacy and confidentiality, so as to stress that equal importance has to be devoted for ensuring that information is accessible only to those authorized to have access. Moreover, it corrects the definition of location data by including also the data processed in an electronic communications network or by an electronic communications service. Processing security will have to ensure that personal data can be accessed only by authorized personnel for legally authorized purposes; to protect personal data stored or transmitted against accidental or unlawful destruction, accidental loss or alteration, and unauthorized or unlawful storage, processing, access or disclosure.

2.8. Extent of Surveillance and Privacy in EU

In 2007, Privacy International[1] and the Electronic Privacy Information Clearinghouse (EPIC)[2] released a report entitled "Leading Surveillance Societies in the EU and the World 2007" [1]. This report assesses the state of privacy in several countries, including the EU and paints a grim picture. For each category, the following grading system was undertaken:

- **5** - no invasive policy or widespread practice/leading in best practice 4.1-5.0; Consistently upholds human rights standards; Improving Country has improved since last year.
- **4** - comprehensive efforts, protections, and safeguards for privacy 3.6-4.0; Significant protections and safeguards; Deteriorating Country has dropped by one category.
- **3** - some safeguards, relatively limited practice of surveillance 3.1-3.5; Adequate safeguards against abuse; Decaying Alarming rate of fall in protections.
- **2** - few safeguards, widespread practice of surveillance 2.6-3.0; Some safeguards but weakened protections.
- **1** - extensive surveillance/leading in bad practice.

	Constitutional protection	Statutory protection	Privacy Enforcement	Data-sharing	Visual surveillance	Communication interception	Communication Data Retention	Government Access to Data	Surveillance of Medical, Financial, and Movement	Border and trans-border issues	Leadership	Democratic safeguards
GREECE	4	3	4	-	3	1	-	3	3	-	4	3
ROMANIA	3	3	4	-	-	2	3	2	-	-	2	4
HUNGARY	4	4	4	3	1	1	4	3	2	3	1	4
SLOVENIA	4	4	4	3	4	2	1	2	2	-	2	3
PORTUGAL	4	4	3	2	2	2	-	-	3	-	2	4
LUXEMBOURG	2	3	3	2	-	2	3	-	4	-	1	4
GERMANY	4	4	4	4	2	2	1	3	4	2	1	4
ITALY	4	4	4	-	3	1	1	2	2	3	3	3
ESTONIA	3	3	4	-	-	2	-	3	3	-	2	3
BELGIUM	4	4	4	1	-	2	2	3	3	2	1	4
CZECH REP.	4	3	4	1	2	1	2	2	2	2	3	4
FINLAND	3	3	3	1	-	3	3	2	2	2	2	4
IRELAND	2	3	4	2	-	3	1	2	3	2	1	4
MALTA	2	4	3	-	-	2	-	2	-		2	2

[1] Privacy International - http://privacyinternational.org

[2] Electronic Privacy Information Clearinghouse (EPIC) - http://epic.org/

POLAND	3	4	3	3	2	1	1	2	2	-	3	3
SPAIN	3	4	4	-	2	1	2	2	1	-	1	4
AUSTRIA	2	3	2	1	2	2	4	2	3	2	1	4
CYPRUS	3	3	3	-	2	1	-	-	2	2	2	3
EU	3	2	3	2	-	-	1	2	3	2	2	3
LATVIA	3	2	2	2	-	2	-	2	2	2	2	3
NETHERLANDS	2	4	4	1	2	1	1	2	2	2	1	4
SLOVAKIA	4	3	3	-	-	2	1	1	2	-	2	2
SWEDEN	3	2	3	2	3	2	1	1	1	2	1	4
DENMARK	3	2	2	1	3	2	1	1	1	1	2	3
BULGARIA	3	2	3	-	-	1	2	2	2	-	2	2
LITHUANIA	3	3	2	-	1	1	3	-	-	-	2	3
FRANCE	3	2	3	1	2	2	1	1	2	1	1	4
UK	1	2	2	1	1	1	1	2	1	1	1	3
England & Wales	1	2	2	1	1	-	-	2	1	-	1	2
Scotland	1	2	2	3	2	-	-	3	2	-	3	4

Table 1 Privacy Ranking 2007 [1]

3. Social and Ethical Considerations

Social implications are neither imposed by some legal document nor are they referenced in relevant legislations. They are driven by the social acceptance of the wider public and represent the social opinion and impacts, focusing attention not on the individual, as this is more apparent in ethical considerations, but rather on the society, i.e. the interactions of the individuals. The inevitable and continuous changing face of CCTV technology suggests that surveillance is in a constant state of unrest, in terms of technical features, the public's reaction, its use and management by authorities, and the nature of security. Up to now there has been a great number of studies arguing that CCTV technology is an efficient and successful tool for reducing crime rates within targeted areas [1].

The use of CCTV technology is, in several occasions, raising ethical concerns related, directly or indirectly, to the lack of privacy protection, the repression of individual liberties for the 'greater good' and augmenting the feeling of insecurity. Such concerns have stimulated the rise of regulations aiming to protect individuals' rights and freedoms as well as regulate the use and output of information collected, processed and retained by such platforms. On the other hand, those who abide CCTV as an effective and successful method in the field of crime prevention, suggest that the presence of such systems in public spaces act as a deterrence to criminals or potential offenders. Therefore, innocent individuals should not be bothered by its presence. The cameras target offenders and thus offer no harm to the general public [2]. This mentality is widely used to convince the public that CCTV systems are used for a specific reason and do not impinge on issues of privacy or civil liberties. In fact it poses an ethical concern, which assumes that in general individuals are innocent and must give up some liberties for the 'greater good' [1]. Another ethical issue surrounding CCTV is its role in the increasing exclusion and discrimination of certain groups and individuals which may result in unfair targeting of groups and stigmatization [2].

P-REACT approach is based on the belief that there is no single direction to follow and that a more dynamic approach is required in order to consider the different criteria and determine how these are enforced. Within the following, we explore a set of challenging issues related to social and ethical acceptance and balance. These issues pertain:

- Surveillance, expanding to:
    - Holistic awareness, being realized by the integration of different sensors allowing information acquisition on multiple levels;
    - Wide coverage, implying the ability of a surveillance system to be applied to multiple locations where the petty crime incidents may occur.
- Internet and other ICT technologies.
- Automations, resulting to minimum human interventions (motion, cognition, creation, etc.).

An initial consideration regarding social implication of surveillance systems, and correspondingly the P-REACT platform, is related to the fact that most of these systems change the value and/or meaning of the body from a very private and personal entity to an exploitable thing, object of public use and object of study and analysis. A structured approach to identify and appraise the key criteria and concerns for societal acceptance and ethical matters will be reported to D1.6 which will become a Guideline for Societal Acceptance and Ethical Considerations for all relevant stakeholders. Prior to this, an overall analysis and report on project's privacy impact assessment will be performed (D1.5) and the aim of this report is to perform an analysis of how collected data is handled regarding the applicable legal and regulatory requirements, to identify any induced risks and evaluate protection mechanisms and processes.

4. Data Physical Protection

Nowadays, most of the data that is being transmitted over computer networks is transmitted as plain text and unencrypted, something that leaves it vulnerable to anyone with the knowhow to apprehend it. Since P-REACT system will deal with private data coming from cameras and microphones, and transmitting this data through the internet or local networks, it is imperative to ensure a high level of security for both the data itself, and the access to all parts of the system. In the following sections strategies and techniques for secure communication in the system's network, local and remote data transfer of the system and physical protection of them that can be utilized to strengthen P-REACT's security are presented. Also, methods for secure system access control are proposed and techniques for data storage and retention are analysed. Data protection requirements are presented in the following table, while a short analysis is presented in the following paragraphs.

Requirement	Title	Description
DPR_01	Data transfer	All data must be transmitted safely through the system
DPR_02	Multimedia transfer	Media from the cameras/microphones must be securely transmitted to the embedded system and/or cloud
DPR_03	Real-time life streaming transmission	The video and audio transmission shall be in real-time, in case an abnormal event has been detected.
DPR_04	Data storage and retrieval	Data should be stored safely to avoid apprehension.
DPR_055	Access to recordings	Recording must be secured and accessible only from the administrator, an authorised monitor and the owner of the shop.
DPR_06	Access Control	There must be a secured access system and only authorized users will have access to. The access control system must allow only a limited number of authentication attempts and also to provide protection in every communication.
DPR_07	Access to the system	Users can only access the system with secure authentication methods.
DPR_08	Access to the embedded platform	Only the administrator and the owner of the shop will have access to the physical embedded system.
DPR_09	Data storage	The system shall be able to protect storage, transmission and access of the recording data within the system.
DPR_10	Data retention	In order for the data to be safe from system errors, an efficient backup system should be provided.
DPR_11	Local data management	All the transmissions within the platform must be safe.
DPR_12	Private data	Private data must be available only to authorized users.
DPR_13	Remote data management	The transmission of the data from/to the system has to be secure and quick.
DPR_14	Portable equipment	All of the portable equipment must be secured and all the communications to the system must be encrypted.
DPR_15	System long time performance	The system shall be able to run for long periods of time without data corruption, slowdown, or servers needing to be rebooted

Requirement	Title	Description
DPR_16	Receive suspicious event notifications	Administrator should receive event notifications coming from the DERs when an abnormal behaviour occurs.
DPR_17	Database protection	All of the recorded data have to be stored to the database with cryptographic and other methods to be safe.
DPR_18	Physical protection of the embedded system	The Platform must be in a secure place and the storage must be protected from malicious attacks.
DPR_19	Physical protection of data	Sensitive data must be destroyed when they are no longer required.
DPR_20	Physical protection of equipment	Portable equipment must be safe, locked and without allowing the removal of computers and storage. Control room and the corresponding equipment should be monitored.
DPR_21	Availability of the embedded system	The embedded system must be permanently accessible and equipped with systems in order to avoid any failure of the system.
DPR_22	System reliability	The system should be able to come back from power failures
DPR_23	System Integration	The system shall be able to integrate any type of sensor, actuator, and devices from an initial list of devices to be defined in an early stage of the system analysis.
DPR_24	System Extensibility	The system shall be able to integrate any type of device in the Local hub data.
DPR_25	Data management	Data has to be interpreted before they can be utilized to ensure data integrity and validity of the data. Data management is also required not only to secure and maintain primary data, but also to process them efficiently.

Table 2 Data protection requirements

4.1. General Purpose Secure Communication

When transmitting private data over a computer network, whether local or the internet, use of a secure protocol is necessary. Although the most common transfer protocols like FTP or HTTP are easier to use, their insecure nature makes them inhibitory for the P-REACT's needs, even with the addition of authentication layers, since they have been proven to be easy prey for malicious actors. This forces the adoption of strong, cryptographically secure authentication methods for the transmission of sensitive data.

There are a notable number of these methods but not all of them are ideal for the P-REACT's needs.

Some of the most used solutions are Kerberos solution kFTP (Kerberos FTP) and GridFTP. The first presents a problem for smaller organizations and it has limited distribution and the second has expensive maintenance while it limits larger institution installation. Most effective transfer network protocols for P-REACT's requirements are the SSH protocol and the SSL/TLS protocol, a brief description of which follows.

4.1.1. SSH Protocol

The Secure Shell protocol is a cryptographic secure protocol designed to be easy use and inexpensive to implement. It offers remote login and many network services. It is easy to install and use, painless to configure, and trivial to administer.

The SSH protocol provides many safeguards that makes it effective for the project requirements. Firstly, all communications between the client and server are encrypted, it encrypts everything it sends and receives. Since each packet is encrypted using a key known only by the local and remote systems, any attempts to spoof the identity of either side of a communication do not work. As a result, a secure transmission of sensitive data that P-REACT project handles can be achieved. The client transmits its authentication information to the server using strong encryption making transmissions extremely difficult to decrypt and read from other malicious users. After an initial connection, the client can verify that it is connected to the same server it had connected too previously. In addition, the client can forward X11 applications from the server providing a secure means to use graphical applications over a network. Furthermore, it allows many different types of services to run on the top of it and the protocol's open architecture permit these services to run all at the same time without impeding each other.

For the protection of the integrity of an SSH communication, a cryptographic handshake is made for the client verification with the correct server. Also, the transport layer of the connection between the client and remote host is encrypted using a symmetric cipher and client has to authenticate itself to the server. The transport layer provides compression, speeding the transfer of information making the transmission of real time data of P-REACT's project easier. Consequently, the remote client interacts with the remote host over the encrypted connection. Finally, another advantage of SSH protocol is that it supports multiple opened channels. Each channel handles communication for different terminal sessions and for forwarded X11 sessions. As a result, different types of sessions do not affect one another and when a given session ends, channel can be closed without disrupting the primary SSH connection. Moreover, it allows flexibility in handling different types of remote connections without having to change the basic infrastructure of the protocol.

4.1.2. SSL/TLS Protocol

The SSL/TLS protocol is a secure network protocol that is used to encrypt confidential data sent over an insecure network. The SSL/TLS security protocol works between the application layer and the TCP/IP layer and supports multiple application layer protocols. It secures and then sends application data to the transport layer. Furthermore, it is divided in two layers. The first layer is the Handshake Protocol and the second layer is the Record Protocol Layer. The Handshake Protocol layer consists of three sub-protocols: the Handshake Protocol, the Change Cipher Spec Protocol, and the Alert protocol. Moreover, Handshake sub-protocol provides very important security functions that are necessary for P-REACT needs. Also, it performs a set of exchanges that starts from authentication, encryption, hash, and compression algorithms.

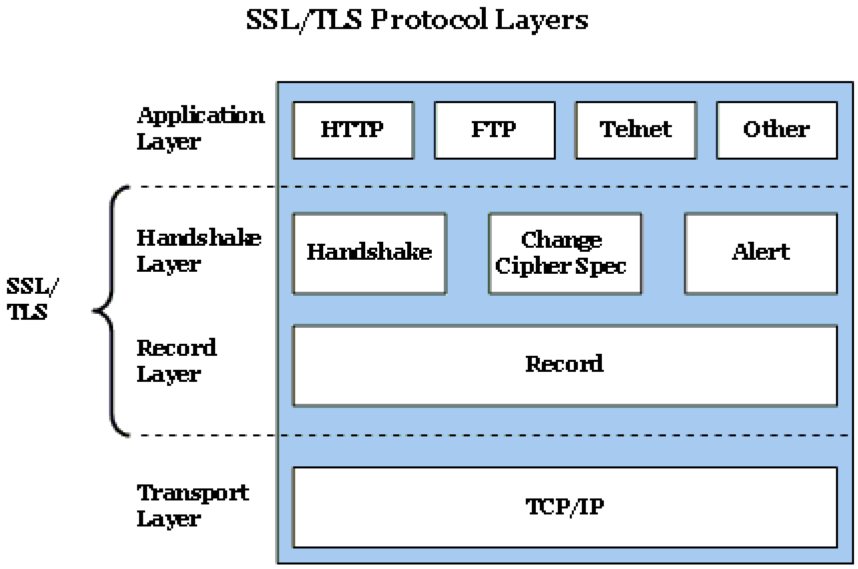


Figure 1 SSL/TLS Protocol Layers

SSL/TLS for authentication purposes uses an X.509 certificate that is a digital form of identification and contains a validity period, a public key, a serial number, and the digital signature of the issuer. In addition, there are two types of encryption and it uses both, symmetric key and asymmetric key. Also, SSL/TLS uses a public key named session key. This is used in symmetric algorithms to encrypt the bulk of the data, providing the benefit of asymmetric encryption for authentication with the faster, less intensive symmetric key encryption.

In conclusion, the main benefits of SSL/TLS protocol for the P-REACT project is the prevention of unwanted eavesdropping that provides and the integrity of data during transport that ensures with the use of hash algorithms

4.2. Secure Multimedia Communication

Besides generic data transfer requirements, P-REACT also will deal with data from cameras and microphones that need to be transferred on local or distributed networks with ease and efficiency. An example of this is the usage of an IP camera that needs to stream its data to the embedded system either via Ethernet or wireless. This will require the use of specific multimedia protocols for their secure transmission. There is a variety of strategies and technologies that address the requirements and implementation elements of secure multimedia communications. Most efficient strategy is the use of WebRTC and WebSocket protocols.

4.2.1. WebRTC

Web Real-Time Communication is a collection of standards and protocols that supports browser-to-browser applications for voice calling, video chat, and P2P file sharing without plugins. WebRTC turns real-time communication into a standard feature that any web application can have and transports its data over UDP. Furthermore, it prevents platform and device independence. It provides fully featured engines to the browser for the signal processing that can be applied to the real time video that the camera of the P-REACT's system records. It uses an algorithm for hiding the negative effects of network jitter and packet loss. Also the video engines optimize image quality, by picking the compression and codec settings, applying jitter and packet-loss concealment. All of the processing is performed by the browser. As a result, the cloud can receive the optimized media stream. WebRTC creates a MediaStream object that represents a real-time media stream and allows the application code to acquire data, manipulate individual tracks, and specify outputs. Summarizing, some benefits of the use of WebRTC in P-REACT project are the advanced voice and video quality that provides, the secure voice and video transmission because of the good encryption and the its adaptability to network conditions.

4.2.2. WebSocket

The WebSocket Protocol enables two-way communication over a TCP connection. It delivers communication between the client and server in both directions simultaneously More specifically, there is a persistent connection where client and server can start sending data any time. In addition it increases client and server communication efficiency and because it is an independent TCP-based protocol, it does not require HTTP tunneling. It reduces unnecessary network traffic and latency using full-duplex through a single connection. Furthermore, it provides streaming through proxies and firewalls, supporting simultaneously upstream and downstream communication. It is ideal for P-REACT system because it allows client to send data to server without needing to re-establish a connection. Client can send data any time while the connection remains open. Using WebSocket, P-REACT can ensure secure connections which are beneficial for authentication purposes. The open connection eliminates network overhead and allows server to send data continuously.

4.3. Virtual Private Networks

In the P-REACT system, the need for secure network is necessary in order to transmit sensitive personal data. A VPN can provide secure site-to-site connectivity and remote access. Virtual Private Network extends a private network across a public network. Also, with the use of a tunnelling protocol can achieve security procedures that involve encryption. VPN technologies have three types of security, trusted VPNs, secure VPNs and hybrid VPNs. Each one provides different level/options of security. Furthermore, a VPN technology does the IP encapsulation, the encryption and the authentication.

For more productivity enhancements and cost savings of P-REACT system, the SSL VPN (Secure Sockets Layer Virtual Private Network) can offer additional information security challenges. SSL VPN allows users to connect to the network through an authenticated pathway by encrypting all network traffic. More specifically, it is used to give remote users with access to web applications, client server applications and internet network connections. It consists of VPN devices that are used for the user and web browser connection. This connection is encrypted with the SSL protocol. An SSL VPN solution provides versatility and accessing resources from many locations. Finally, SSL Portal VPN and SSL Tunnel VPN are the two major types. With SSL Portal VPN type, a single connection to web site is provided so the end user can have access to multiple network services and with the SSL Tunnel VPN type, a secure access from web browser to multiple networks services is allowed.

4.4. Local and Remote Data Transfer and Management

For the most efficient local and remote data transfer and management of the system, the following plan is proposed.

About the embedded platform and cloud connection, a safe and encrypted Virtual Private Network (VPN) should be utilized, because of the good encryption and compression that it provides. A restricted network should be used, so that all clients will be able to connect to the cloud with a VPN connection. Since VPN is widely accepted as the safest way for secure communication, this approach is the best to be applied from all the security systems that it will be used.

Moreover, the embedded platform should always be connected with a VPN tunnel with the cloud, in order to avoid the possibility that an untrusted party will connect to the embedded platform. In the embedded system, in case a wireless network is used, it must be secret and cryptographic, and the camera must use a secure protocol. Also, the clip object that the embedded system will use to relay information to the cloud will be routed through VPN, and the data will be deleted from the embedded system only when the cloud response for the transfer is successful. This communication should be utilized by using JSON (JavaScript Object Notation) lightweight data-interchange format with SSH protocol for best security or with another SSL protocol type. In addition, all the multimedia clips can be transferred with WebRTC and Web sockets protocols.

Finally, periodically the embedded system must do a general check on the system and all the sensors, and if there is a problem it must report it to the cloud system.

4.5. Data Storage and Retention

Data storage and retention is an important issue for the successful operation of the system. In the embedded platform the clip is going to be saved until the cloud sends the signal that it has received it. Also, it will keep the data recorded for a short period of time for the audio and video analytics algorithms needs.

In the cloud, a secure database with the clips with a restore and backup system that cloud computing platforms like OpenShift provides should be used. Moreover, the database must not allow direct access to data and every local network must be protected through a firewall. The access to the local area network must be through a central access point. In order to avoid any kind of penetration or eavesdropping multiple cryptographic protocols and algorithms are preferred. In addition, server-side SSL identity protects from possible malicious attacks. Additionally, methods for encryption and identity control should be applied. Furthermore, with the use of SSL for TCP/IP protocol for the integrity of the data a secure communication is ensured. Finally, according to the European laws clips must be stored for at least six months.

4.6. Access Control

Since P-REACT system deals with personal and private data, it requires an extremely secure access control system for protecting them. In order to ensure the protection of data in accordance with national laws, data access must oblige with the following guidelines:

- Data can only be accessed by authorized users.
- A central access point will manage every login request using a secure protocol.

Access control must also offer flexible control over the user’s access rights because not only the access control of the system is important but also who is going to have access to the system and what rights they will have. More specifically, in the embedded system, access should only be granted to the system administrator and the owner of the shop, while for the cloud system only authorized personnel should have access.

Although traditional access control methods, such as a username and a password, can provide enough security, when dealing with sensitive personal data as the P-REACT does, a more sophisticated method should be utilized. Such a method could be the use of a multi-factor authentication method that will utilize a smart card or a token device. More secure methods, such as one time password sent over SMS or an NFC enabled smart phones used as a token could also increase the security of the system, but they would also increase the complexity and cost in both development and deployment.

In any case, the access control system must take into account the following precautions:

- Strong cryptographic protocols customized to system specificities are required.
- The access control system must only allow a limited number of authentication attempts.
- The system must protect against any malicious software that could compromise the security of the system.
- The system must protect every communication channel using a firewall and intrusion detection systems.
- All communication protocols that are used from the biometric device and the database should not have logical errors that could enable an attacker to violate its safety.
- Every single action must be recorded in order to detect a possible attack.
- The following penetration testing must be applied
    - Network level penetration testing;
    - Infrastructure level penetration testing;
    - Level penetration testing.

Finally, in order to enhance even more system's security, biometric authentication methods could be utilized such as fingerprint or retinal scans or face recognition technologies. This type of user authentication is out of the scope of the project, but it could be examined in case the consortium decides that the proposed methods do not provide sufficient security.

4.7. Physical Protection of Data

Except the protocols and security systems physical protection of data is also needed in order to ensuring their complete protection. Physical protection consists of a variety of measures in order to protect equipment and data from malicious attacks. Without any physical protection, other types of protection are not enough against attackers. Moreover, the detection of an attack may be sufficient to minimize its effect. Some measures for the physical protection of P-REACT system are:

- The Platform must be in a safe place and the storage must be encrypted, in case of stolen no one can get the data.

- Not allowed for anyone to connect personal laptops or any other computing device to the system’s network.
- Not allowed for anyone to add/modify hardware or software to the system without proper authorization.
- Monitors, printers and all hardware equipment must be placed away from windows and areas where unauthorized persons could have access.
- All sensitive information and media when it's no longer necessary must be destroyed.
- Portable equipment must be locked in a safe storage place overnight.
- Not allowed the removal of computers or storage media from the work area or facility without ensuring that the person removing it has authorization and a valid reason.
- Locks or cables must be provided in order to prevent theft at the equipment rooms.
- Video surveillance should be utilized to monitor the control room and the corresponding equipment.

4.8. Availability and Platform Restoration

The embedded platform is equated with a variety of features in case of power loss or the lack of internet connection. A powerful UPS (uninterruptible power supply should be used, in case of power failure in order to keep system’s platform available and online for more than 12 hours. In addition the platform could be equipped with an optional 3/4G connection, in order to assure that all the services will be online and permanently accessible. Finally, for any problems that cannot be evaluated and/or repaired remotely (e.g. problems with the internet connection), or in case of a mechanical damage/failure, the local operator (e.g. the owner of a shop) should contact the technical stuff in order to schedule a repair.

4.9. Protection Against Attacks

The video-surveillance systems have the purpose of protecting goods and people against illegal acts, but they may be subject themselves to external attacks, both from the physical point of view and cyber attacks. Then the companies equip themselves in order to adequately react to potential threats.

Beyond any theoretical and very technical consideration about the solutions offered by the market and the systems and tools that can potentially be implemented, stakeholders and potential end-users may provide more useful description about the protection systems used in real world. What follows is then a description of the current situation, described based on the information gathered during meetings with potential end-users.

As a preliminary consideration, accordingly with the experience of the stakeholders interviewed, attacks against various components of video-surveillance systems are actually very rare, considering both physical and cyber attacks. An exception is represented by act of vandalism against cameras placed on the field, which occur with higher frequency. Moreover, it should be considered that the level of response by the companies to potential external attacks is almost proportional to the value of the assets protected and the sensible information that could be damaged or stolen. The systems and procedures implemented in a bank, for example, are different from those implemented to control a bus depot or a store. The latter reflection is relevant, especially considering the defence systems against cyber attacks. It would be counterproductive to propose to a shop to implement a too high level solution, but even more expensive. It is decisive that the proposed solution is fitted with the actual value and importance of asset or information to be protected.

4.9.1. Physical Attacks

Physical attacks against video-surveillance systems can be different and could have different purposes. For example, the components positioned on the field (i.e. cameras, wireless, etc.) could be subjected to acts of vandalism without any further specific purpose. The same components could be damaged to turn off video-surveillance or at least reduce its operability to avoid being captured on camera during an illegal action. Or somebody could try to enter in premises where images are collected and stored, with the purpose of destroying potential evidence of a crime.

As per the feedback collected from stakeholders and potential end-users, physical attacks are infrequent, except as acts of camera vandalism. Indeed cameras may be removed, or just turned and diverted, to avoid coverage of a crime. For example, it was reported by stakeholders interviewed, that offenders succeeded in obstructing lens with paint using a sling shot to reach the highly positioned cameras. In any case, usually these types of attacks are perpetrated by general criminal and not cyber experts.

Generally, there are two methods to counteract physical camera attacks: prevention of potential attacks (before the attack is executed), and the detection of attacks (in real time, during the attack). For instance, cameras can be positioned in not accessible areas, or accessible only with extreme difficulty. Some cameras are equipped with software able to detect when the camera is diverted (e.g. when the camera is moved to face another field of view), or the lens is covered, or more frequently when the signal is lost for a pre-defined period. More sophisticated models are equipped with an internal gyroscope able to detect unexpected modifications of tilt angle or shaking. In these cases, the device sends an alert via SMS or e-mail to the control centre requesting for a reaction by the security team.

The backend solution and the storage of data can be placed in non-accessible rooms. Doors could be armoured and access could be possible restricted using biometrics and other means. Indeed video analytics could be used to detect unauthorised access. The door of the server cabinet itself could be alarmed etc.

4.9.2. Cyber Attacks

According to interviews with stakeholders and potential end-users, cyber attacks against video-surveillance systems are quite exceptional (actually, they did not report any of them). Despite this, they should be considered related to video-surveillance systems.

Before the widespread of DSL (Digital Subscriber Line) technology for the digital transmission of data, the majority of video-surveillance systems functioned as closed-loop systems. In a nutshell, in a closed-loop system the entire network is dedicated and internally wired, potentially without access to and from the outside. If set in such a way, the system would be immune from attack. However, such systems are often sold with a remote access to the Internet set with weak passwords by default. This could allow hackers to gain control of them. To make the system secure, it is necessary to disable remote access and equip the systems with stronger passwords - which is not always possible, given that often end-users do not even know this issue and they are not experts in the field. In any case, the main problem with closed circuit systems is the very high cost, especially, if they are intended to control a wide area, due to the high cost for cabling.

The recent diffusion of DSL technology today enables high performance transmission of data, through wider bandwidth. This allows file sending to the Internet even in the case of large data size such as with video. Generally, the main structure is built and owned by an external telecommunications company which provides the connection by subscription. This way a video-surveillance system using DSL technology allows the cost to be reduced when compared with a closed-loop system.

However, these systems are more open to cyber attacks. To overcome this risk, anti-intrusion and anti-virus software is used. Nevertheless, none of these software packages are able to guarantee an absolute and complete safety against cyber attack.

The most used prevention systems are firewalls and sniffers.

- A **firewall** is a passive component for perimeter defence of a computer network, which can also act as a link between two or more sections of the network, thus ensuring protection in terms of security of the network itself. The firewall filters all incoming and outgoing packets of data to and from a network or a computer, according to predefined rules that contribute to the safety of the network. A firewall can be a single computer (with at least two network adapters, one for input and one for output, and appropriate software), it can be a software included in a router or it can be implemented on a dedicated hardware apparatus. The main functionality is to essentially create a filter on the incoming and outgoing connections, raising the level of network security.

- **Sniffing**, however, is the activity of passive interception of data in transit in a computer network. Commonly "sniffing" is linked with unlawful practices (e.g. fraudulent interception of passwords or other sensitive information), but actually this activity can be carried out for legitimate purposes (e.g. analysis and the identification of communication problems or intrusion attempts).

Of course, within the IT field several systems exist with advanced technologies to protect computer networks. These are used by facilities with greater protection requirements (e.g. banks). However, based on the feedback received from the stakeholders interviewed, in real world corporate networks are usually protected only by ordinary firewall and sniffers. As already mentioned, the networks are usually provided with a level of protection commensurate with the value of the data and information that must be protected.

4.10. Auditing

As for many IT systems or applications, even a video-surveillance system should be subject to a security audit process.

A security audit is a technical assessment aimed to check if and at what level the system, with its processes and way of work, is compliant with the requirements established by the security policy in place (usually defined by means of dedicated Procedural Manual and Code of Practice). Such technical assessment is supposed to be systematic and measurable, it should be conducted by security auditors external to the system owners and operations staff, and it should be well-structured on the basis of a predefined check-list and terminate with a detailed audit report.

A security audit can be performed through staff personal interviews, security vulnerability tests, analysis of the way to access the system, control of operating settings and network shares, and other measures necessary to verify how effectively the security policy is being implemented in the system. Finally, even any physical possibility to access to the system and its components (e.g. on-site equipment, embedded system, back-end services, control room, etc.) should be considered.

The final purpose of a security audit is to provide a measurable way to examine how secure the system really is, and establish possible improvements to make it more secure.

The targets of a security audit depend on the system to be reviewed and on the condition it is working in, anyway some examples are the following:

- Passwords used to access the system and to use it: checking of the level of difficult to crack them.

- Systems and procedures to control who access to the system and data stored, also historical.
- Incidents logs: accuracy and completeness, also historical.
- Audit logs to register who accesses systems and data, also historical.
- Updating of the security tools and methodology, in accordance with the ones available in the market, checking if they are correctly patched to the last version.
- Backup media: how it is stored, who can access to it.
- Critical events, disasters, or black out: a recovery plan should be necessary.
- Data encryption: tools are appropriate, properly configured, and adequately updated?
- Performance measurements of the security measures and tools in place.
- Appropriateness of the procedure in use for registration of visitors to the system domain.
- Behaviour of operations staff: verify of compliance with the security police adopted (and if security requirements are even known).
- Compliance of the security settings (and security policy in general) adopted with the real functioning of the system itself and the practices being used by the operating staff.
- Compliance of the system with the updated data protection legal framework.
- Adequateness of signage in compliance with the legal requirements.

In some cases, “penetration tests” could be carried out, focusing on the search for security gaps in critical components of the system, such as firewalls or web servers.

4.11. Cloud Services and Emerged Potential Barriers

Very recently, cloud services have been proposed as an interesting solution for storage and data management. Cloud service providers offer storage availability to keep data collected (e.g. by a video-surveillance system), usually in pay per service modality. The P-REACT project intends to take advantage by this innovative modality to handle files and data collected by the video-surveillance system under development.

Although it does not represent a full solution against cyber attacks, it is more cost effective. However, according to stakeholders interviewed, this solution may present some problems to be handled. In particular:

- Psychological barriers and cultural acceptability of the system: IT departments and end-users are still somewhat sceptical about the use of cloud solutions due to data security and unfamiliarity issues.
- Reluctance to lose direct control of data storage and security issues and rely on unfamiliar third parties who are potentially located in other jurisdictions.
- Reluctance to being tied to a particular provider who is offering proprietary technology that restricts future changes or at least is costly to make changes.

5. P-REACT Platform

An overall high level architecture of P-REACT platform is presented in this section, in order to understand the proposed solution and the requirements exposed in the following sections. A deeper analysis of the architecture and its modules is documented in *D2.3 P-REACT Conceptual Architecture, including functional, technical and large-scale deployment specifications.*

As stated in the DoW, the P-REACT project proposes the design and development of a low-cost end-to-end platform that enhances the surveillance and the protection of small businesses against petty crimes and enables the easy exploitation of multi-source archiving system.

P-REACT platform is divided into two main platforms: the cloud system and the local embedded framework.

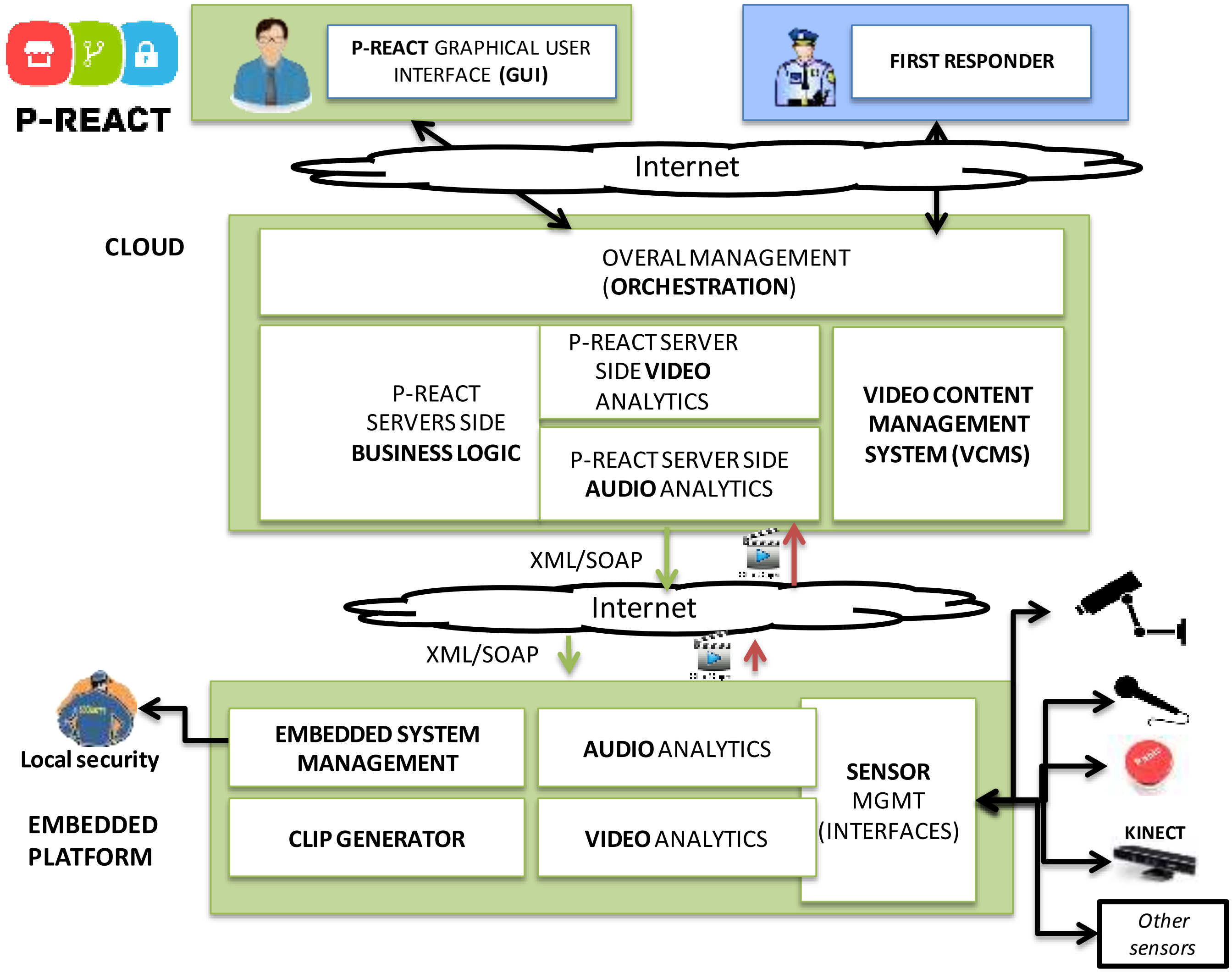

P-REACT
P-REACT GRAPHICAL USER INTERFACE (GUI)
FIRST RESPONDER
Internet
CLOUD
OVERAL MANAGEMENT (ORCHESTRATION)
P-REACT SERVERS SIDE BUSINESS LOGIC
P-REACT SERVER SIDE VIDEO ANALYTICS
P-REACT SERVER SIDE AUDIO ANALYTICS
VIDEO CONTENT MANAGEMENT SYSTEM (VCMS)
XML/SOAP
Internet
XML/SOAP
Local security
EMBEDDED PLATFORM
EMBEDDED SYSTEM MANAGEMENT
AUDIO ANALYTICS
CLIP GENERATOR
VIDEO ANALYTICS
SENSOR MGMT (INTERFACES)
KINECT
Other sensors

Figure 2 Simplified architecture of P-REACT platform

5.1. Communication Requirements

This section discusses the network requirements that are needed to be met towards achieving a smooth communication amongst the various components/entities of the P-REACT project, such as the sensors, embedded platform, cloud platform, and the internal/external users. In general the network equipment should provide the bandwidth needed to transfer quickly and efficiently all the information (e.g. Video Clips, OS images, control messages, etc.) needed to be exchanged amongst P-REACT's key entities/components, besides supporting the most widely used protocols, including IPv4, IPv6, TCP, UDP, HTTP, RTP and RTSP, both for wired and wireless transmissions. The following two sections present in more detail the communication requirements from the scope of the embedded system and the cloud platform respectively.

5.1.1. Embedded System's Communication Requirements

As Figure 2 presents, the embedded system needs to communicate with the sensors, the cloud and the local security personnel, each one having different bandwidth demands. More specifically, the network between the embedded platform and the sensors must provide the bandwidth needed not only to transfer the sensors' data to the embedded system but also all the various control and management messages stemming from the embedded system or even the cloud. In general, the various control and management messages require several Kbps for their transfer, but the sensors' data, especially Video, require much more. In more detail, several considerations must be taken into account for deriving a more accurate estimation of the sensors' bandwidth requirements including:

1. The number of sensors.
2. Whether sensors' data streaming will be continuous or based on triggering events.
3. For video sensors:
   a. Frames per second (FPS);
   b. Image resolution (VGA, 720p, 1080p, etc.);
   c. Video compression type (H264, MJPEG, MPEG-4, etc.);
   d. Scenery complexity (rich in action video, lighting conditions, etc.).

The following two tables taken from [4] and [5] respectively, provide some insight regarding the bandwidth needed to transfer efficiently various video resolutions.

Video codec	Resolution and aspect ratio	Maximum video payload bitrate (Kbps)	Minimum video payload bitrate (Kbps)
H.264	320x180 (16:9) 212x160 (4:3)	250	15
H.264/RTVideo	424x240 (16:9)) 320x240 (4:3	350	100
H.264	480x270 (16:9) 424x320 (4:3)	450	200
H.264/RTVideo	640x360 (16:9) 640x480 (4:3)	800	300
H.264	848x480 (16:9)	1500	400
H.264	960x540 (16:9)	2000	500
H.264/RTVideo	1280x720 (16:9)	2500	700
H.264	1920x1080 (16:9)	4000	1500
H.264/RTVideo	960x144 (20:3)	500	15

H.264	1280x192 (20:3)	1000	250
H.264	1920x288 (20:3)	2000	500

Table 3 Bandwidth rates for various H264 video resolutions

Video Resolution	Video rates range (Mbps)
QVGA	<0.6
WVGA	0.6-1.2
VGA	1.2-3
720p 2/3	2-3.5

Table 4 Video Rates range per Video Resolution

In general it is common for surveillance systems to have up to 10 cameras and to use a standard 100Mbps network switch. It is noted that if some of the sensors are wireless the switch must support also IEEE 802.11b/g/n technologies.

The communication requirements between the embedded system and the cloud consist on having the necessary bandwidth that will allow for: a) the efficient, error free and fast upload of the Clip objects to the cloud; b) the download of the system updates from the cloud; and c) the exchange of the various control and management messages between cloud and the embedded system itself. In general, the various control and management messages require several Kbps for their transfer, but the bandwidth needed to transfer the Clip objects and the system updates will be in the order of several hundreds of Kbps or even Mbps---it will be a matter of the Clip objects size, System updates magnitude and of any transfer time constraints. Taking into account that the Clip object, besides any metadata, will contain (or accompanied by) Video data and that the minimum resolution requirement for Video is VGA—that according to tables 1 and 2 requires a minimum bandwidth of 300Kbps (min Quality) and can reach up to 3 Mbps—it is anticipated that at least ADSL2 (or even VDSL) network connections should be exploited.

Finally, concerning the communication with the local security personnel, the bandwidth requirements will be in the order of several Kbps, assuming that the message will contain plain text. This demand will be easily satisfied from the embedded system's xDSL connection or even a local wireless network.

ID	Requirement	Justification
ECR_1	100 Mbps Ethernet switch	This switch can accommodate for the bandwidth needed to transfer the sensor's data and exchange the various control and management messages. It is noted that if wireless sensors are used the switch should supported by exploiting an IEEE 802.11b/g/n module.
ECR_2	ADSL2/VDSL network connections	It is anticipated that these networks can provide for the minimum upload speed needed to transfer efficiently the clip objects to the cloud, besides allowing the uninterrupted exchange of the various control messages between the cloud and the

ID	Requirement	Justification
		embedded system.

Table 5 Embedded system's communication requirements

5.1.2. Cloud's Communication Requirements

As Figure 2 presents, the cloud system needs to communicate with the sensors (via the embedded system), the embedded system, and its external users. More specifically, the cloud platform should have an egress network able to: a) receive all the data stemming from the embedded systems (mainly the Clip objects); b) send system updates and various control management messages to the embedded systems (or to the sensors via the corresponding embedded systems); c) support the amount of information needed to be exchanged between the Cloud System and its users (First responders/GUI users). It is not uncommon in production cloud environments the deployment of at least a 1 Gbps egress network, but as this is strongly related with the number of running embedded systems, it is anticipated that for P-REACT's needs, a 100 Mbps egress network would easily satisfy the bandwidth requirements.

Finally special care should be taken for the cloud's intra-network that should accommodate the fast transfer of OS images and the Video archives. For these purposes switches of 10 Gbps, usually, suffice, but for P-REACT's demonstration needs a 1 Gbps switch could bare, almost effortlessly, the anticipated network load.

ID	Requirement	Justification
CCR_1	Egress network of ~ 100 Mbps.	This bit rate covers easily the bandwidth demands of the running embedded systems and users described in P-REACT's use cases.
CCR_2	10 Gbps Ethernet switch	This equipment will provide the necessary capacity needed to transfer the OS Images and Video archives amongst the cloud nodes.

Table 6 Cloud's communication requirements

5.2. Security

P-REACT should take care that the communication between its entities is secured, especially taking into account the sensitive "nature" of the exchanged data. More specifically, the following communication channels should be shielded against eavesdropping and data tampering: Embedded System - Sensors, Embedded System - Local Security, Embedded System - Cloud, Cloud - Users. The most widespread solution towards securing communications are Virtual Private Networks (VPNs), which can be realised using IPSec, or TLS/SSL based technologies, such as OpenVPN and HTTPS. IPSec and OpenVPN are being exploited mainly for protecting all types of network traffic whereas HTTPS is being used for guarding HTTP based communication. Finally, when wireless connectivity is an option, special care is taken for deploying the wireless channels exploiting the IEEE 802.11i protocol suite (WPA/WPA2). Based on the above the following Table holds the requirements for having a secure communication over P-REACT's platform.

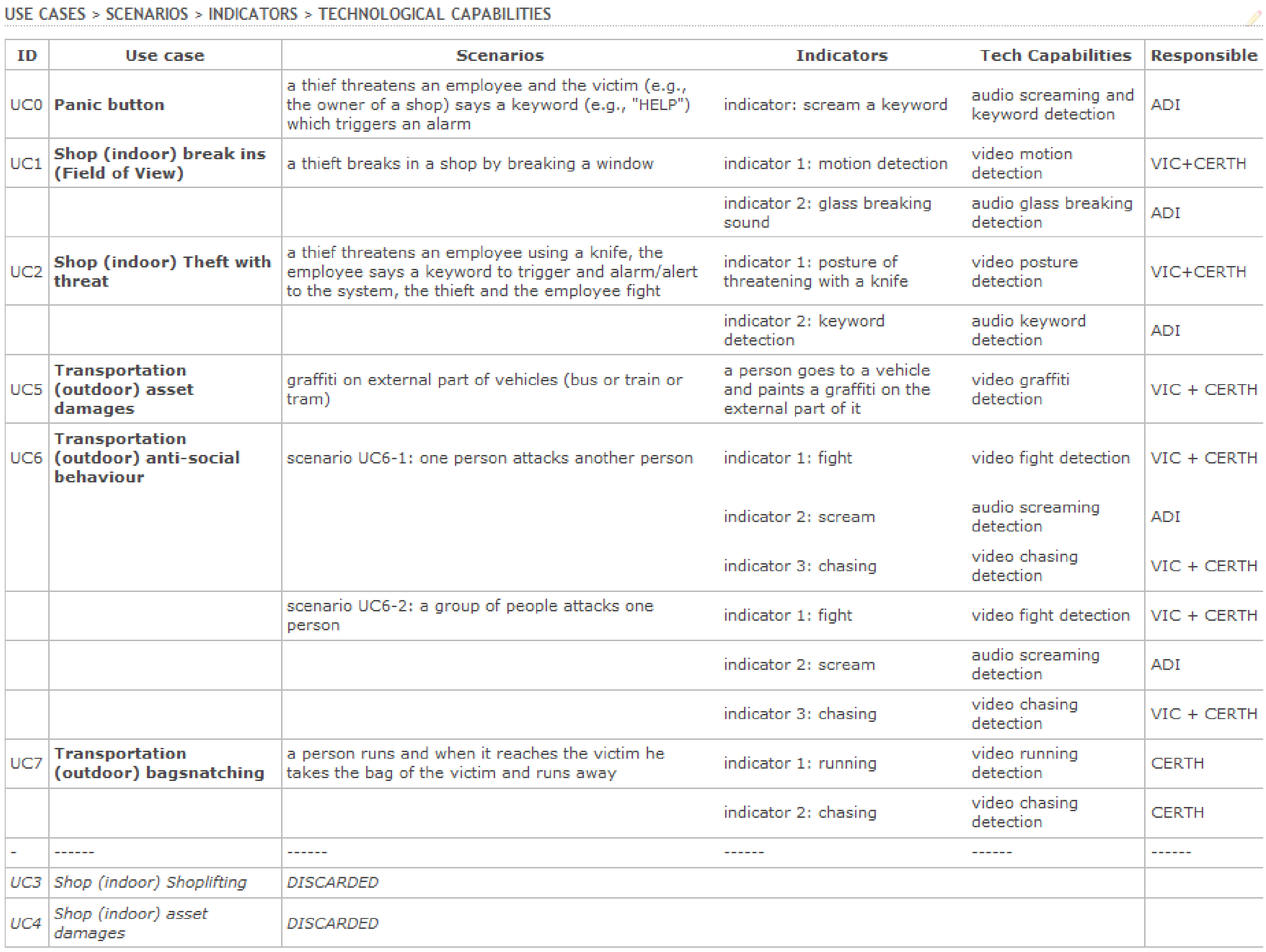
ID	Requirement	Justification
SCR_1	Edge Network equipment must support either IPsec or OpenVPN technologies.	Both of these technologies enable for the secure transfer of data over P-REACT's communication channels.
SCR_2	Web GUIs must be exposed via TLS/SSL (HTTPS)	HTTPS ensures the secure communication between the cloud and its users (e.g. GUI users)
SCR_3	Any Wireless Connectivity should be offered exploiting IEEE 802.11i (WPA/WPA2).	Wireless transmissions can be easily detected and monitored; therefore the wireless channel must be always secured.

Table 7 Communication's Security requirements

5.3. P-REACT's Prototype Scenarios

The P-REACT project will seek to address petty crimes situations as outlined in D2.1. A number of petty crime behaviours such as: antisocial behaviour, fight, vandalism, graffiti, small shop robbery, unauthorized access to tunnels or private premises, intrusion detection, theft of public structure elements, theft of bikes, etc. have been identified as key concerns to stakeholders. Development towards some of these crime types in certain scenarios will be addressed during the P-REACT project. Considering their significantly large occurrence in cities, these crimes often lead to major economic loss, or, considering a non-controlled escalation, even to serious injuries or trauma of involved persons. Their early detection is therefore of major importance.

After the mid-term review of the P-REACT project the use cases and scenarios to be developed were revisited. The table below summarises the list of use cases and scenarios that were decided to be finally developed and included in the P-REACT solution during the project. The table also shows the indicators and technological capabilities that were defined for each scenario, as well as which partner would be responsible for developing the different technological capabilities.



USE CASES > SCENARIOS > INDICATORS > TECHNOLOGICAL CAPABILITIES

ID	Use case	Scenarios	Indicators	Tech Capabilities	Responsible
UC0	**Panic button**	a thief threatens an employee and the victim (e.g., the owner of a shop) says a keyword (e.g., "HELP") which triggers an alarm	indicator: scream a keyword	audio screaming and keyword detection	ADI
UC1	**Shop (indoor) break ins (Field of View)**	a thieft breaks in a shop by breaking a window	indicator 1: motion detection	video motion detection	VIC+CERTH
			indicator 2: glass breaking sound	audio glass breaking detection	ADI
UC2	**Shop (indoor) Theft with threat**	a thief threatens an employee using a knife, the employee says a keyword to trigger and alarm/alert to the system, the thieft and the employee fight	indicator 1: posture of threatening with a knife	video posture detection	VIC+CERTH
			indicator 2: keyword detection	audio keyword detection	ADI
UC5	**Transportation (outdoor) asset damages**	graffiti on external part of vehicles (bus or train or tram)	a person goes to a vehicle and paints a graffiti on the external part of it	video graffiti detection	VIC + CERTH
UC6	**Transportation (outdoor) anti-social behaviour**	scenario UC6-1: one person attacks another person	indicator 1: fight	video fight detection	VIC + CERTH
			indicator 2: scream	audio screaming detection	ADI
			indicator 3: chasing	video chasing detection	VIC + CERTH
		scenario UC6-2: a group of people attacks one person	indicator 1: fight	video fight detection	VIC + CERTH
			indicator 2: scream	audio screaming detection	ADI
			indicator 3: chasing	video chasing detection	VIC + CERTH
UC7	**Transportation (outdoor) bagsnatching**	a person runs and when it reaches the victim he takes the bag of the victim and runs away	indicator 1: running	video running detection	CERTH
			indicator 2: chasing	video chasing detection	CERTH
-	------	------	------	------	------
UC3	*Shop (indoor) Shoplifting*	*DISCARDED*			
UC4	*Shop (indoor) asset damages*	*DISCARDED*			

NOTE: use cases UC3 and UC4 were discarded due to a lack of resources to tackle the challenges they imply on top of the selected use cases (UC0, UC1, UC2, UC5, UC6, and UC7).

5.3.1. Prototype’s Requirements and Functionalities

We consider that user requirements are translated into a set of functional specifications of the developed analysis system. The high level analysis system further consists of several modules, such as computer vision, time-series analysis or machine learning application, which provide a specific set of analytical functionalities.

A correct functionality of the analytical modules largely depends on the operation context. The operation context describes the variability of the environment in which the developed applications must operate, considering that they do not provide equal behaviour in different contextual conditions, which can be influenced by: the variability of the system installation, weather or lighting conditions, type or variability of public infrastructure (e.g. station, pedestrians' walk, outdoor street environment), presence and level of occupation of the monitored area with people and vehicles, etc. The specification of the operation context therefore largely influences the correct application and use of the entire petty crime analysis system.

In Table 8, Table 9, Table 10, Table 11, Table 12, and Table 13 we present the identified end user requirements and basic system functionalities for each of the considered prototype scenarios. Additionally, we provide the first estimate of the basic analytic functions of the system, which will be necessary for the implementation of the P-REACT's security system. Finally, we indicate the preferable operation context, which, however, will be finally determined by the context of the available training datasets and trial scenarios.

	UC0 - Panic button
End user requirements	1. Notify security of dangerous events. 2. Allow for remote live monitoring of the situation.
System functionality	1. Send alarm signal to remote server. 2. Start scene recording.
Analytic functions	Embedded system: 1. Audio analytics – recognition of key events. 2. Speech analysis - recognition of keywords.
Operating context	Indoor environment (small shops), public transport (taxi, bus). Camera pose and view set to capture the scene so that a future identification of the aggressor is possible. Regular ambienta noise and acoustic conditions. Single actor.

Table 8 Panic button scenario: end-user requirements and analytical functionalities

	UC1 - Shop (indoor) break ins (Field of View)
End user requirements	Detect when a thief breaks in a shop after breaking a window when the shop is closed
System functionality	1. Detect glass breaking and motion and then send alarm signal to remote server. 2. Start scene recording. and monitoring (life video streaming).
Analytic functions	Embedded system: 1. Audio analysis – glass breaking detection.

	2. Video analysis – video motion detection.
Operating context	Indoor environment (small shops). Camera including micro pose and view set to cover as much of the shop as possible, including the window(s). Low ambient noise and acoustic conditions.

Table 9 Shop robbery (indoor): break ins (Field of View)

	UC2 - Shop (indoor) Theft with threat
End user requirements	Detect when a thief threatens an employee using a knife, the employee says a keyword to trigger and alarm/alert to the system, the thief and the employee fight.
System functionality	1. Detect the pose of a person holding a knife. 2. Detect the employee shouting a keyword (e.g., “HELP”). 3. Detect fighting.
Analytic functions	1. Audio analytics to detect screaming. 2. Audio analytics to detect keyword (e.g., “HELP”). 3. Depth and/or Video analytics to detect fighting people.
Operating context	Indoor environment (small shops). Depth camera, video camera and micro covering the cashier’s area. Regular ambient noise and acoustic conditions.

Table 10 Shop robbery (indoor): Theft with threat

	UC5 - Transportation (outdoor) asset damages
End user requirements	Detect when a person breaks in the depot where the vehicles are parked and paints graffiti on external part of vehicles (bus or train or tram).
System functionality	1. Detect the presence of a person close to a vehicle. 2. Detect graffiti being painted on a vehicle.
Analytic functions	1. Video analytics to detect people. 2. Video analytics to detect graffiti being painted.
Operating context	Outdoor environment (vehicle depot) Camera pose and view set to cover as much of the depot as possible, including the vehicles. Calibrated camera to set the area of interest or a prohibited area. Regular ambient noise and acoustic conditions.

Table 11 Transportation (outdoor) asset damages

	UC6 - Transportation (outdoor) anti-social behaviour
End user requirements	Scenario UC6-1: one person attacks another person: 1. One person attacks another person and they start fighting. 2. One or both of the involved persons scream. 3. The victim tries to escape and runs and the attacker chases the victim. Scenario UC6-2: a group of people attacks one person: 1. A group of people attack one person and they start fighting. 2. One or more of the involved persons scream. 3. The victim tries to escape and runs and the attackers chase the victim.
System functionality	Scenario UC6-1: one person attacks another person: 1. Detect two people fighting. 2. Detect screaming. 3. Detect one person chasing another one. Scenario UC6-2: a group of people attacks one person: 4. Detect a group of people fighting with one person. 5. Detect screaming. 6. Detect a group of people chasing one person.
Analytic functions	Scenario UC6-1: one person attacks another person: 1. Video analytics to detect two people fighting. 2. Audio analytics to detect screaming. 3. Video analytics to detect one person chasing another one. Scenario UC6-2: a group of people attacks one person: 4. Video analytics to detect a group of people fighting with one person. 5. Audio analytics to detect screaming. 6. Video analytics to detect a group of people chasing one person.
Operating context	Outdoor environment (e.g., a bus stop). Camera with microphone pose and view set to cover as much of the spot as possible. Regular ambient noise and acoustic conditions.

Table 12 Transportation (outdoor) anti-social behaviour

	UC7 - Transportation (outdoor) bag-snatching
End user requirements	Detect when a thief runs and when he/she reaches a victim, he/she takes the bag of the victim and runs away.
System functionality	1. Detect the presence of a person running and getting close to another person. 2. Detect when the victim (or some other person) chases the theft that is running away.
Analytic functions	1. Video analytics to detect running. 2. Video analytics to detect chasing.
Operating context	Outdoor environment (e.g., a bus stop). Camera with microphone pose and view set to cover as much of the spot as possible. Regular ambient noise and acoustic conditions.

Table 13 Transportation (outdoor) bag-snatching

6. Local Embedded System

This section discusses the Hardware and Software requirements of the embedded system (platform and sensors) that will be installed at the SMEs and/or public Transportation stations for detecting petty crime incidents, informing local authorities to intervene, and sending “scene” material to the P-REACT’s cloud system, where, after further analysis, this material could be stored and used as evidence to the court.

6.1. HW Requirements

According to DoW, the embedded system should be a low cost platform able to manage data from multiple sensors; to perform lightweight Video and Audio analysis; to forward Sensors data along with Analysis metadata (as Clips objects) to the cloud; to inform local authorities upon petty crime detection; and to handle cloud’s requests for Sensors’ activation/deactivation, analytics update, clip generation, etc. Therefore, the hardware that it will be used should satisfy the processing, storage, I/O connectivity and networking requirements needed for the smooth execution of the above operations, and at the same time meet the cost constraints. However, Depth analytics require processing power which for the time being cannot be offered from low cost embedded systems. As a result P-REACT will develop two embedded systems, one low cost that it will use data from conventional RGB Cameras, and one more advanced in processing capabilities—but more expensive—that will be used for analysing Depth Data coming from depth sensors. The following two tables contain the hardware requirements for these two models, and are a result of preliminary tests of running basic audio, video and depth analysis functionalities on a dedicated computational platform.

ID	Requirement	Justification
ELR_1	CPU >= 1GHz	Preliminary tests and previous experience indicate that this processing capacity is enough to handle the lightweight Video and Audio analytics, clip generation, cloud communications and sensors management.
ELR_2	RAM >= 512MB, DDR3	Preliminary tests and previous experience indicate that 512MB is the minimum capacity that could handle memory needs for the lightweight Video and Audio analytics, clip generation, cloud communications and sensors management.
ELR_3	Storage of 8 GB	It is anticipated that OS and additional libraries will take 3-4 GB of disk size. It is assumed that the remaining 4 GB is the minimum storage space needed to temporarily store analysis meta-data, clip objects, etc.
ELR_4	OS should be Linux, Android compatible	Open Source solutions should be preferred for meeting the low cost constraints, leaving aside their anyway good performance on embedded systems.
ELR_5	USB, at least 2 slots	It should support sensors that support only this connection type (mainly camera)
ELR_6	Other I/O ports such as HDMI, GPIO	It is reasonable to anticipate for future connection of the embedded system to displays or the integration of sensor modules in the embedded system's board.
ELR_7	Network should support 100Mbps Ethernet and should be able to use wifi, if needed.	It is necessary for communicating with the cloud system and also to support IP based sensors (mainly cameras)

Table 14 Low cost embedded platform's HW minimum requirements

ID	Requirement	Justification
EAR_1	CPU Dual core - 2GHz	According to Depth analysis experts this is the minimum processing requirements to run efficiently the Depth analytics.
EAR_2	GPU with Support of computational capabilities (e.g. CUDA, OpenCL, OpenGL ES v2.0)	This is needed for supporting the deployment of more efficient analytics algorithms at the embedded system level.
EAR_3	RAM with 4GB DDR3	It handles memory needs for the lightweight Depth, Video and Audio analytics, clip generation, cloud communications and sensors management.
EAR_4	Storage of 100 GB	It is anticipated that OS and additional libraries will take 3-10 GB of disk size. The rest could be used to temporarily store a big amount of analysis meta-data, clip objects, etc.
EAR_5	OS Windows 8.1, Linux compatible	The new Depth sensors are usually compatible only with 'Windows and at later time an open source driver is provided.
EAR_6	USB3/2	The new Sensors may be provided only with USB3 support, while older ones could be not compatible with USB3 ports.
EAR_7	Network 100Mbps Ethernet	It is necessary for communicating with the cloud system and also to support IP based sensors (mainly cameras)

Table 15 Advanced platform's HW minimum requirements

As far as it concerns the sensors, their capabilities should meet, at least, the minimum input requirements the Video, Audio, and Depth analytics algorithms set for being able to successfully identify petty crime events. In addition, they should meet the low cost constraints imposed by the project's objectives and should be compatible with the IO ports and OS of the embedded system. Taking all this into account, the following tables present sensors' requirements.

ID	Requirement	Justification
CSR_1	Resolution at least 640 x 480 (VGA)	According to Video analysis experts this is the minimum video resolution needed for Video analytics.
CSR_2	Night/Day switch (IR)	It is needed for providing petty crime detection during night.
CSR_3	Framerate 25-30 fps	According to Video analysis experts this is the frame rate needed for Video analytics.
CSR_4	MJPEG, H.264 or any other standard video format	According to Video analysis experts standard video formats are compatible with the vision libraries used for the analysis.
CSR_5	Should support Network (IP), or USB connectivity or to be able to integrated using GPIO in the embedded system.	The embedded system will provide, network (IP) and USB ports. A possibility for integrating the camera module to the embedded system is left open.
CSR_6	Linux compatible	The RGB cameras will be connected to the low cost embedded system that runs Linux.

Table 16 RGB Camera's minimum requirements

ID	Requirement	Justification
DSR_1	Operating distance ~7m	According to Depth analysis experts this is the typical operational range of depth sensors and is suitable for covering indoor places.
DSR_2	Accuracy ~5cm	According to Depth analysis experts this accuracy is enough for the used algorithms.
DSR_3	Should support Network (IP), or USB connectivity or to be able to integrated using GPIO in the embedded system.	The embedded system will provide, network (IP) and USB ports. A possibility for integrating the camera module to the embedded system is left open.
DSR_4	MS Window and/or Linux compatibility	The Depth Sensors will be first deployed in Microsoft systems but later when an open source driver provided it may be also deployed in a Linux system and therefore reduce the cost.

Table 17 Depth Sensor's minimum requirements

ID	Requirement	Justification
ASR_1	Codec Lossless PCM	According to Audio analysis experts this codec is the preferable input to their algorithms.
ASR_2	Channels Mono	According to Audio analysis experts one channel is enough.
ASR_3	Sampling Rate 16 KHz	According to Audio analysis experts this codec is the preferable input to their

ID	Requirement	Justification
		algorithms.
ASR_4	Bits per sample 16	According to Audio analysis experts this codec is the preferable input to their algorithms.

Table 18 Audio Sensor's requirements

6.2. HW Requirements of the Prototype Scenarios

In the following we present the mapping of suggested and preferable HW requirements to the system's functionalities for each of the prototype scenarios.

UC0 – Panic button

Component	Requirements	Functional requirements
Embedded platform	Low cost (see Table 14) with wireless or mobile internet.	Low computational power requirements.
Audio	Close-range (0.5m - 3m) sensitivity, mono channel, 16 KHz sampling rate, lossless PCM Codec, 16 bits per sample.	Recognition of screaming and keywords.
Video	Close-range (0.5m - 3m) application, minimum 640x480 (VGA) resolution, night/day IR switch, 25-30 fps frame rate, standard video format (MJPEG, H.264, …).	Recording of the scene.
Depth Sensor	None.	

Table 19 HW requirements of the Panic Button scenario.

Comment: Although a panic button as a security event trigger is the simplest solution, as it starts recording of video and audio for remote analysis, a possible technical barrier is the necessary bandwidth to handle the data upload to the cloud. This applies especially in the case of mobile transportation platforms.

UC1 – Shop (indoor) break in (Field of View)

Component	Requirements	Functional requirements
Embedded platform	Low cost (see Table 14) with wireless or mobile internet.	Low computational power requirements.
Audio	Close-range (0.5m - 3m) sensitivity, mono channel, 16KHz sampling rate, lossless PCM Codec, 16 bits per sample.	Recognition of breaking glass.
Video	Close-range (0.5m - 3m) application, minimum 640x480 (VGA) resolution, night/day IR switch, 25-30 fps frame rate, standard video format (MJPEG, H.264, …).	Video motion detection, recording of the scene.

Table 20 HW requirements of the UC1 – Shop (indoor) break-in scenario.

Comment: It is considered that audio algorithms could be used in real time (no audio recording) to detect a potential event which would trigger the camera images being recorded and streamed to the control room for incident verifications.

UC2 – Shop (indoor) Theft with threat

Component	Requirements	Functional requirements
Embedded platform	Advanced platform (see Table 15).	High processing requirements for person detection and tracking.
Audio	Mid-range (3m - 10m) sensitivity, mono channel, 16 KHz sampling rate, lossless PCM Codec, 16 bits per sample.	Recognition of screaming people, recognition of keywords and panic words.
Video	• Close-range (0.5m - 3m) for cash desk and entrance monitoring in indoor environments. • Mid-range (3m - 10m)	Close-range for detection of the entering security area event. Mid-range for action recognition (fighting).

	application for general shop surveillance. Minimum 640x480 (VGA) resolution, 25-30 fps framerate, standard video format (MJPEG, H.264, …).	
Depth Sensor	Mid-range (3m – 10m) application, Kinect V2, depth images (424x512).	Activity detection

Table 21 HW requirements of the UC2 – Shop (indoor) break-in scenario.

Comment: Depth-analysis plays a key role for this use case since it is the only sensors which may provide reliable information about the pose of the different persons in the scene, allowing the execution of action recognition algorithms.

UC5 – Transportation (outdoor) asset damages

Component	Requirements	Functional requirements
Embedded platform	Advanced platform (see Table 15), wireless or mobile connection for public transportation vehicles.	High processing requirements for person detection and tracking.
Audio	None	
Video	Far-range (10m - 30m) application for public areas monitoring. Minimum 640x480 (VGA) resolution, 25-30 fps framerate, standard video format (MJPEG, H.264, …).	Long-term background analysis to evaluate the persistence of new background (graffiti).
Depth Sensor	None	

Table 22 HW requirements of the UC5 – Transportation (outdoor) asset damages.

UC6 – Transportation (outdoor) anti-social behaviour

Component	Requirements	Functional requirements
Embedded platform	Advanced platform (see Table 15),	High processing requirements for

	wireless or mobile connection for public transportation vehicles.	person detection and tracking.
Audio	• Mid-range (3m - 10m) sensitivity, mono channel, 16 KHz sampling rate, lossless PCM Codec, 16 bits per sample. • Far range (10m - 30m), high quality omni-directional audio on platforms, 44KHz sampling rate, lossless PCM Codec, 16 bits per sample.	Volume and noise analysis, recognition of screaming people. Recognition of keywords and panic words in case of high quality audio.
Video	• Mid-range (3m - 10m) application for close transportation platform and inside vehicle monitoring. • Far-range (10m - 30m) application for public transportation platform monitoring. Minimum 640x480 (VGA) resolution, 25-30 fps framerate, standard video format (MJPEG, H.264, …).	Mid-range monitoring for action recognition. Far-range for group and trajectory analysis, related relations group vs. individual and related events.
Depth Sensor	Mid-range (3m – 10m) application, Kinect V2, depth images (424x512).	Activity detection

Table 23 HW requirements of the UC6 – Transportation (outdoor) anti-social behaviour.

UC7 – Transportation (outdoor) bag-snatching

Component	Requirements	Functional requirements

Embedded platform	Advanced platform (see Table 15), wireless or mobile connection for public transportation vehicles.	High processing requirements for person detection and tracking.
Audio	None	
Video	• Mid-range (3m - 10m) application for close transportation platform and inside vehicle monitoring. • Far-range (10m - 30m) application for public transportation platform monitoring. Minimum 640x480 (VGA) resolution, 25-30 fps framerate, standard video format (MJPEG, H.264, …).	Mid-range monitoring for action recognition. Far-range for group and trajectory analysis, related relations group vs. individual and related events.
Depth Sensor	Mid-range (3m – 10m) application, Kinect V2, depth images (424x512).	Pose recognition, action classification.

Table 24 HW requirements of the UC7 – Transportation (outdoor) bag snatching.

6.3. SW Requirements

As mentioned in the previous section, the two flavours of the P-REACT's embedded systems will be (a) a low-cost system capable of running light video and audio analytics algorithms and (b) a more advanced system capable of utilizing depth sensors along with audio. The following sections present the software requirements of both the embedded system flavours along with a brief description of each.

6.3.1. Embedded System - Audio Analytics

In this subsection the requirements that the audio analytics module located in the embedded system are listed. These requirements must be fulfilled in order to ensure the correct integration of the audio analytics module.

1. The **audio analytics algorithm** will have access to an audio stream, so that it can work in real-time. This access will be provided by the Sensor Manager module that is the responsible of capturing the

sensor's data and streaming it to the modules that are requiring it.

2. The **audio stream** will be lossless PCM. The encoding will be Little Endian, 16 bits per sample, Signed Integers (two's complement) and sample rate will preferably be 16 KHz.

6.3.2. Embedded System - Video Analytics

In this subsection the requirements that the video analytics module located in the embedded system are listed. These requirements must be fulfilled in order to ensure the correct integration of the video analytics module taking into account the limitations of this analysis module due to the processing capacity of embedded systems.

1. C++ will be used for the video analytic modules implementation.

2. The video analytics algorithm module will have access to the camera's data via an intermediate module (Sensors Manager) that will utilise OpenCV (or other libraries) APIs.

3. BSD or LGPL licenses of 3rd Party dependencies for all SW components.

7. Cloud Based System

This section presents the Hardware (HW) and Software (SW) requirements needed to deploy the cloud platform that will host the Video Content Management System (VCMS) system along with the advanced Video/Audio analytics, the business logic and the orchestrator responsible for the overall management of the cloud based system.

7.1. HW Requirements

The cloud platform, being the "heart" and "mind" of P-REACT's overall system, should be able to manage all the installed embedded devices and sensors; to receive, analyse and archive a significant amount of data stemming from the sensors (mainly) and/or citizens' end devices; and to inform authorities for detected crime events. All these require substantial processing, storage and networking resources, therefore the deployed hardware should take advantage of processors supporting full virtualization, high speed networking equipment and enough storage capacity. Taking all this into account, the following tables present the cloud's hardware requirements.

ID	Requirement	Justification
CR_1	CPU should be Multi-core with support for full virtualization, i.e. processors using	The processing power should be able to support the heavy analysis of the data stemming from multiple embedded systems, the concurrent execution of meta-analysis requests and VCMS operation.

ID	Requirement	Justification
	technologies such as AMD-V or Intel VT.	
CR_2	RAM should be at least 32 GB	The memory capacity should be enough to support the heavy analysis of the data stemming from multiple embedded systems, the concurrent execution of meta-analysis requests and VCMS operation.
CR_3	Storage should be at least 1 TB, with RAID support	The storage space should cover the needs for storing archive content (VCMS). RAID technology should be used for hard disk failure and data resiliency.
CR_4	Network should be at least 1Gbps Ethernet	It should accommodate the necessary bandwidth needed to receive with no delay all the data stemming from the embedded systems. Extra precaution is needed in the case of VM image transfers across the cloud nodes.
CR_5	Cloud Framework should be compatible with open source solutions	There are several popular and well documented open source solutions, supported from many vendors, promoting the use of open APIs, thus enabling for interoperability with other platforms and services.
CR_6	Hypervisor should be compatible with native (Type 1) and hosted (Type 2) solutions	Being flexible to support both approaches.

Table 25 Cloud system hardware minimum requirements

7.2. SW Requirements

Besides the aforementioned hardware requirements, the cloud system's software should be able to handle input from various sources (i.e. different cameras, microphones etc.) and be able to execute a number of advance video and audio analytics both in real time and in batch mode. In the following subsections, Video Content Management System, Orchestration and Linking of Software Components and Business Logic Component are described. These modules are the core of the Cloud System and therefore their software requirements demand a deeper explanation.

7.2.1. Video Content Management System

The Video Content Management System (VCMS) will be used in order to store clip objects coming from the embedded systems as well as video and audio analytics algorithm objects that can be used on either the cloud system or on the embedded ones. All objects that will be stored should include metadata that will be used for indexing and fast retrieval. Taking into account that the clips can be provided in various formats, the need for real-time processing, and the required ability to remotely configure the embedded systems, the following table represents the VCMS's software requirements:

ID	Requirement	Justification
VCMSR_1	Video Transcoding	Ability to easily and seamlessly transcode video clips from one format to another.

ID	Requirement	Justification
VCMSR_2	Open Source Software	Use of open source software and open standards as much as possible. An exception to this might be certain video codecs which are not available as open source.
VCMSR_3	Video Classification	The P-REACT system will be used for classifying video clips. Classification data will be stored as metadata. The classification will be carried out both manually by humans with the aid of GUIs, and automatically by intelligent classification algorithms.
VCMSR_4	Data Encryption	The data transmitted and stored should have the option to be encrypted, and the confidentiality and integrity of the data should be protected. Also, an access rights policy and mechanism must be provided.
VCMSR_5	Distributed Data Storage	The option to store the data in various physical locations should be available. Legal regulations may dictate that certain video clips cannot be stored in certain locations. The physical location will depend on the cloud infrastructure and how it is deployed (e.g. public cloud, private cloud).
VCMSR_6	Real time & Batch mode operation	The P-REACT system will work in both a real-time mode and in batch mode. The VCMS should have a way to accommodate for this. ○ Real-time mode will be used for cases where a quick response is necessary for an incoming clip object. For example, certain clip objects might issue an alert, because they reflect an emergency situation which needs to be addressed as quickly as possible. For this to happen, the incoming clip objects need to be processed and classified in real-time in order to allow a timely response (within a few minutes) to the emergency situation. ○ In batch mode the clip objects are processed and analysed in large batches. For example, a batch mode process could be one which searches for a video clip in the archive in which humans are present at a specific location during a specific time range in the past.
VCMSR_7	System Scalability	The VCMS needs to be scalable and highly interoperable. It has to be able to handle huge amounts of data, many files, high bandwidth, without degrading performance as the numbers become large. It should integrate with existing cloud and big data technologies. All data should be available for export in standard formats.
VCMSR_8	System Configuration	We need full access to the source code. We need to have the ability and the right to make changes and additions to the source code as needed in the future.
VCMSR_9	High-Level Programming Language	Preferred programming languages are Java, Python, C, and C++.
VCMSR_10	Remote configuration of the Embedded system	The VCMS will also store video/audio analytics algorithms which will be used to dynamically configure the embedded systems. These objects can be in the same form as the video clip objects, but with different metadata. A possible solution is to store generic "data objects" in the VCMS, with a metadata tag which indicates if the object is a video clip object or an analytics algorithm object.

Table 26 Video Content Management System Requirements

7.2.2. Orchestration and Linking of Software Components

The orchestration and Linking Software Components are the ones that link all the other P-REACT components together. Communication between the various components uses web services that allow different components to run on different systems even different operating systems. The following table presents the requirements of these components, taking into account the various P-REACT components that need to be interconnected.

ID	Requirement	Justification
OLSCR_1	Communication between different systems/OS's	Web services should be used to enable communication between various systems and operating systems. SOAP or REST are proposed, with the latter being easier to use.
OLSCR_2	Logging	The system must keep a centralized detailed log of any actions performed
OLSCR_3	Queuing	In order to handle large numbers of service request a queuing system must be utilized. A simple FIFO queue will not suffice, since some request will be more urgent than others (e.g. video analytics vs. software update)
OLSCR_4	Authentication/Authorization	Both service requests and operator requests must be authorized and authenticated in order to ensure the integrity and credibility of the system
OLSCR_5	Clip object manipulation	Clip objects are sent from the embedded system to the VCMS and are also requested from the analytics modules and the operators through the GUI. The "Glue" components must handle this information exchange seamlessly.
OLSCR_6	Analytics algorithm management	These components should also decide which of the available analytics algorithm should be called for each clip object processed.

Table 27 Orchestration and Linking of Software Components requirements

7.2.3. Business Logic Component

The Business Logic Component is the component which manages the cameras, VCMS and interfaces with the human operators via the GUI. Its main functionalities are (1) alert raising, (2) Interaction & control and (3) real time monitoring. Taking these into account, the following requirements are gathered:

ID	Requirement	Justification
BLCR_1	Sensor management/configuration	The "Brain" component should be able to switch sensors on or off and change their parameters (e.g. brightness, resolution etc.)
BLCR_2	Video/audio analytics configuration	It is necessary in order to optimize performance. This includes configuration of existing algorithms and even upload of new ones.
BLCR_3	Embedded system configuration/management	It should be able to configure the software of the embedded system and/or upload updates, perform maintenance etc.
BLCR_4	Raise alert	It should make the final decisions on which of the clips are "important" enough to raise an alarm and, if necessary, activate neighbouring sensors.
BLCR_5	Real time monitoring	It should be able to display on the GUI clips on demand, either live or stored ones, and provide the operator the ability to validate them.

Table 28 Business Logic Component requirements

8. Evaluation Methodology

The methodology presented herein addresses the need to evaluate the performance, accuracy, capabilities and limitations of the functionalities of the P-REACT tasks. The evaluation will be made at two levels: (i) individual analysis modules, and (ii) end-to-end use case scenarios.

While in the first case (i) the focus is to evaluate the different functional characteristics of the developed components and applications, from their behaviour at the basic implementation level (see analytic specifications defined in Section 5.3.1), in the second case (ii) an evaluation of a combination of the developed modules, which results in successful detection of the specific use case event, will be made.

In the following, we describe the general quality evaluation methodology (Section 8.1) defining the considered analysis objects, the performance evaluation measures, the commonly used, statistical, quality indicators and end-to-end time performance measures. Then, a description of the end-to-end evaluation of the use case scenarios, considering a combination of the results of the analysis modules used to monitor that scenario is presented (Section 8.2). Finally, a list of public available datasets, which could potentially be used for the evaluation of the P-REACT system, is given (Section 8.3).

8.1. Quality Assessment Methodology

The quality assessment of the P-REACT system suggests a **three stages evaluation** plan, out of which only the first two stages will be held during the runtime of the P-REACT project. The last stage, being dependent on the final installation at the end-user side, is to be processed after the deployment of the system and is to be used to fine-tune the parameters of each module as well as set up the most appropriate response logic and the business model.

Evaluation stages:

1. **Lab testing –** assessing the analytical functionalities at the module level (i.e. evaluation level (i)), based on carefully prepared datasets related to each analytic function and obtained within an controlled environment. Public and internally generated datasets will be used.
2. **Field testing –** assessing the analytical and system functionalities (i.e. evaluation levels (i) and (ii)) in real-world scenarios corresponding to the defined use cases. Datasets corresponding to the defined use case scenarios, obtained within a semi-controlled environment, will be used.
3. **End user testing –** testing the integrated system functionalities after the final P-REACT system deployment. In addition to the performance measures, the whole system is to be evaluated based on an assessment of **operating availability**. In this context, statistics indicating the percentage of time within which the overall system was operational and running above reliability thresholds are to be collected and evaluated.

As a pre-requisite for each quality test, the following functional, implementation related and operational details must be established:

- **Functionality**: A set of functional specifications or functionalities of the tested module, i.e.

extended and concrete functionalities of the preliminary ones defined in Section 5.3.1.

- **Parameters**: A set of intrinsic and extrinsic parameters of the tested module which govern its behaviour and should thus be included and analysed during the evaluation process.
- **Datasets**: Ground truth datasets on which the evaluation will be conducted. Note that during each evaluation stage, different evaluation datasets must be captured and prepared.

8.1.1. Quality Assessment Analysis Objects

We propose that the basic analysis units in the assessment methodology are the **object** and the **event**. Objects have spatial or spatio-temporal character. The objects represent entities that are tangible, measurable and not subject to interpretation or subjectivity. Events describe actions or semantic properties of one or more objects, having temporal or spatio-temporal component. Events are descriptions that may be subject to some subjectivity.

We consider the following **four types of features**, which are to be recognized by the P-REACT system:

1. **Spatial objects** – represent one or more magnitudes of an entity at a given, indefinite, time. They are specified by the Region of Interest (ROI) in the spacial domain of the input data (i.e. image, in the context of video processing). The spatial component of the object can therefore be represented by geometric entities such as points, circles, polygons, etc.
2. **Spatio-temporal objects** – represent a spatial object at specific instants of time. In the video analysis context we consider it to be a set of spatial objects, of the same entity, associated to one or more frames of the video sequence.
3. **Temporal events –** represent actions, events or relations observable at a certain time interval. It may include links to objects that form part of the action or property defined by the considered event. All audio events of our system will be considered as temporal events, reflecting the recording of mono channel audio streams.
4. **Spatio-temporal events** – represent events in which objects with spatial characteristics are involved. The spatial characteristics of the considered objects form the spatial component of the event itself, at each instance of the input sequence considered.

In general, two **comparison measures** evaluating the relation between the spatial objects and the temporal events on the ground truth dataset will be employed:

- **ROI overlap** – considering that each spatial object is represented by its spatial occurrence, a ratio of the ROI overlap is to be used for the localisation quality measure. In case of spatial description by points, relation to object size defining the acceptable distance between the annotation points is to be considered.

- **Time overlap –** to assess the correct determination of temporal occurrence of events or objects, the ratio of the intersection of annotated time intervals and the union of the detected intervals is to be considered.

8.1.2. Quality Assessment Evaluation Measures

In the following, we present the performance measures and their related performance statistics which will be used for the quality assessment of the applied analysis tasks at the module level (considering the applied classifiers, detectors and recognitions) as well as at the scenario level (considering the recognition of the use case specific events). In case of object detections or event recognitions, we evaluate the posterior classification of the detections or recognitions into correct, wrong and missed ones (a binary classification with one positive class and unknown, but large, number of negative cases).

Performance measures:

To assess the quality of object or event detection and classification, we calculate the **True Positive** (TP), i.e. correct detection / recognition, **False Positive** (FP), i.e. wrong detection / recognition, **False Negative** (FN), i.e. missed detection / recognition, and **True Negative** (TN), i.e. correct rejection , responses of the classification or detection tasks.

In general, a **correct** (TP) **detection of spatial object or temporal event** is considered if more than 50% overlap (ROI or time) of the matched detection with the ground truth provided annotation is detected. Different thresholds may be applied in order to weight the detection importance of certain objects or events.

For the evaluation of **spatio-temporal** objects or events detection related to a video sequence input, we employ the following three strategies:

- **Frame level** – in this case the temporal information of the appearance is ignored and only the spatial correlation (i.e. the overlap) between the ground truth and the detections at each frame is taken into account. As such, a single object appearing in a sequence of 1000 frames can generate 1000 TP if correctly detected at each frame, 1000 FN if not, and as many FP as unmatched detections.

- **Object level** – evaluates the appearance of the object at each instance of the sequence, throughout the entire sequence. A compromise in the exact estimation of the start and the end of the detection interval is given by considering a boundary interval **Δt** around the start and the end of the evaluated detections. As in the previous example, an object appearing in 1000 frames will result in 1 TP if correctly detected in all the frames, allowing only for a slight variance of the start and the end of the detection interval; in 1 FN if not, and as many FP as are unmatched detections.

- **Sparse spatio-temporal level** – in this case a ground truth object is associated with a detection (i.e. TP) if there is sufficient temporal and spatial similarity. We adopt a one-to-many data association procedure which allows associating a single, e.g. long-lasting, ground truth object with a number of sparse, brief, detected objects. In Figure 3, four cases of performance assessment are presented. Case A represents a typical situation in which an object is detected sparsely as different tracks, with short miss-detection periods. An overlap threshold is set up to define how permissive is the evaluation with these holes. An extreme but still acceptable case of correct detection is case B, where the overlap is above the set threshold. Finally, not valid detections are exemplified in cases C and D, where lower event overlap has been detected. In case D, in addition, the first detected event has been considered as unmatched, due to its extension beyond the set border margin **Δt**.

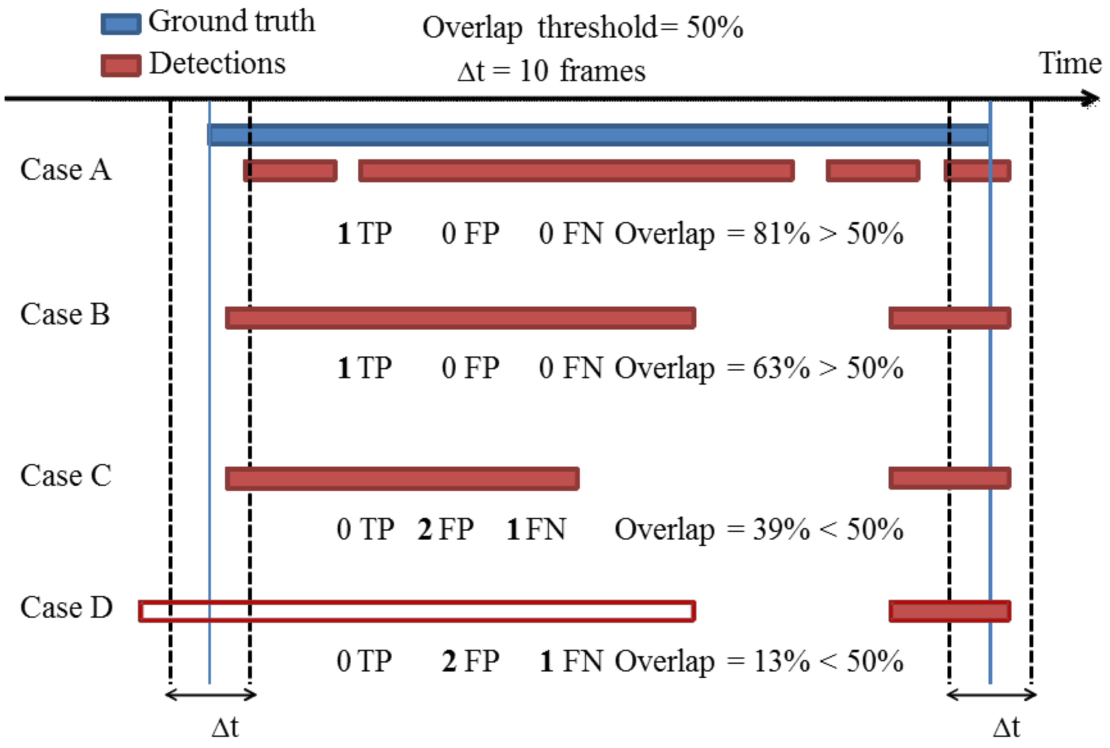

Ground truth
Detections
Overlap threshold= 50%
Δt = 10 frames
Time
Case A
1 TP 0 FP 0 FN Overlap = 81% > 50%
Case B
1 TP 0 FP 0 FN Overlap = 63% > 50%
Case C
0 TP 2 FP 1 FN Overlap = 39% < 50%
Case D
0 TP 2 FP 1 FN Overlap = 13% < 50%
Δt
Δt

Figure 3 Sparse one-to-many temporal data association between ground truth and detections.

For the evaluation of **temporal events** or object detections, similar strategies will be used:

- **Frame level** – considering that the event detection must be correct throughout the entire recorded sequence (i.e. at all frames).
- **Object level** – evaluating only the correct detection of the event within the sequence, i.e.

evaluating the occurrence of the event.

- **Sparse temporal level –** implementing one-to-many data association procedure in the correspondence assessment between the ground truth and the detected events, allowing for certain level of faulty or missing detections (defined by the overlap threshold). A boundary interval of **Δt** is added to compensate for exact start and end of the event detection. An illustrative example is given in Figure 3.

Applying the aforementioned detection and recognition measures, quality assessment of the analytical systems can be done based on a variety of statistical performance indicators.

Performance assessment statistics:

- **Generation of confusion matrices:** specifying the misclassification among several classes. In the binary classification case, the confusion matrix is represented by the TP, FP, FN and TN classification responses with respect to the target class.
- **Classification performance** in terms of:
  - *Precision* – specifying a fraction of correct object classification among all objects assigned to that class;
  - *Recall* – specifying the fraction of correctly classified objects;
  - *Error* – specifying the probability of erroneous classification (equals to 1-recall);
  - *F-measure* – a harmonic mean of the precision and recall, providing a single measure of performance of the test for the positive class;
  - *Sensitivity* – specifying the performance for objects of a certain target class;
  - *Specificity* – specifying the performance for all objects outside the target class.
- ROC (*Receiver Operator Characteristic Curve*) and AUC (*Area Under the Curve*) metric: The ROC represents a graph illustrating the performance of a binary classifier with respect to the discrimination threshold. It can be used for setting the final parameters of the system, considering a trade-off between cost of wrong responses and performance of correct classification or detection; The AUC represents the probability that the classifier will assign a higher positive classification score to a randomly chosen positive example than to a randomly chosen negative example.

8.1.1.Time Response Performance Assessment

The time responses of each component can be used to evaluate the end-to-end performance of the overall P-REACT system. Sequence diagrams can be used to identify how the response time of the system is computed. The following diagrams (see Figure 4 and Figure 5) illustrate the defined behaviour of the different modules that the P-REACT system launch at the embedded-side and cloud-side subsystems, respectively (more details about the modular architecture and the cloud-side dynamic model can be found in “D2.3 P-REACT Conceptual Architecture”, section 6.3.2).

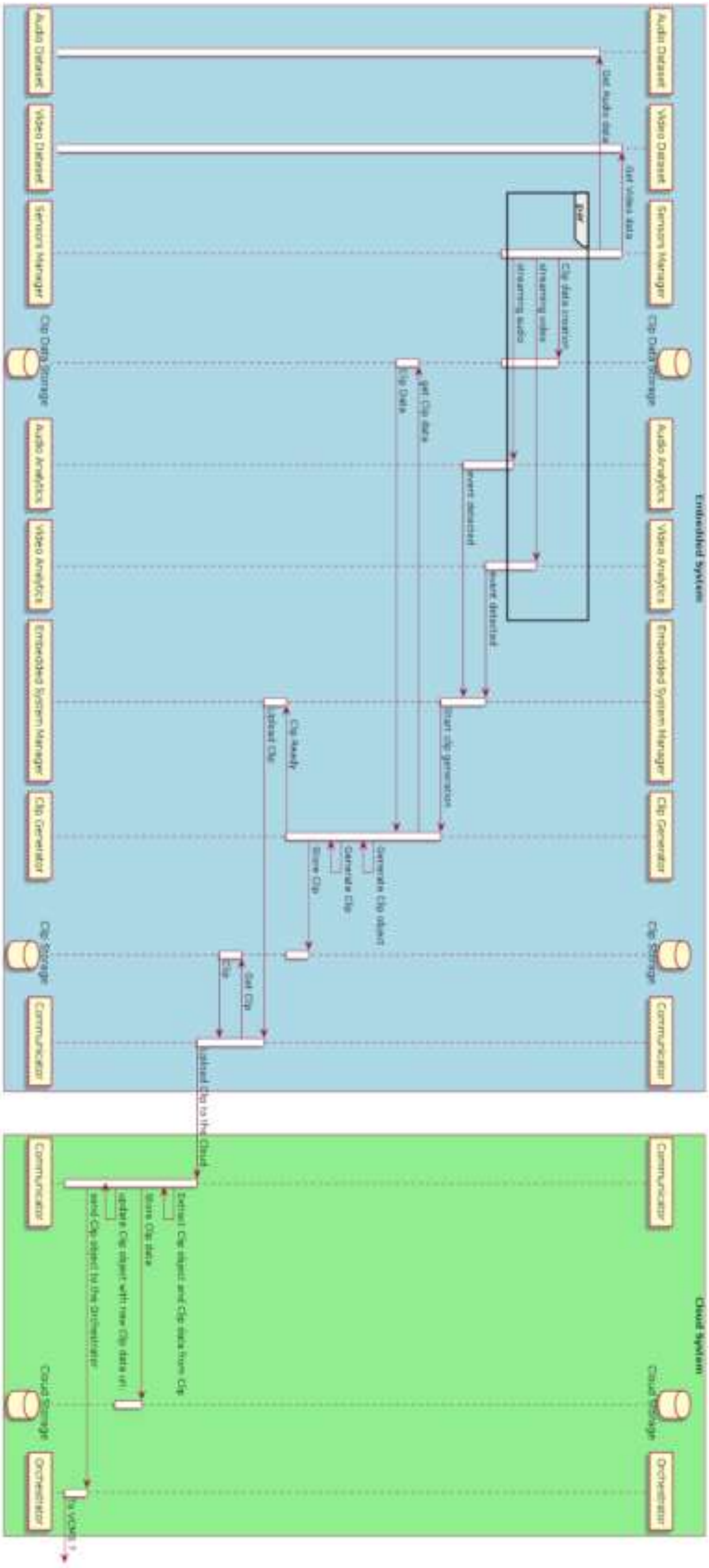


Figure 4 Detailed sequence diagram of the embedded-side components, and the connection with the cloud-side subsystem.

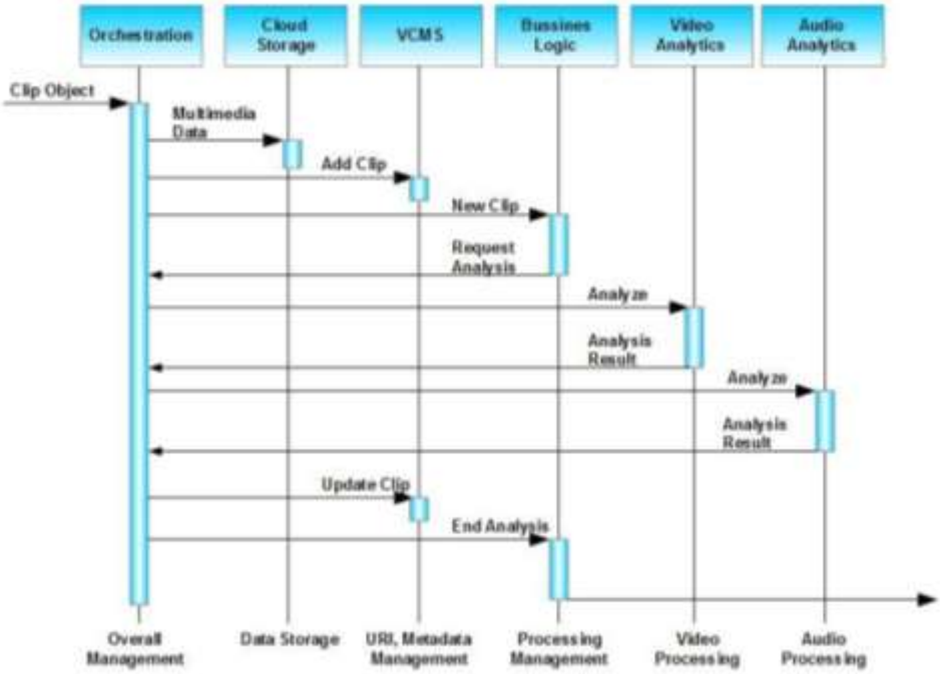


Figure 5 Detailed sequence diagram of the cloud-side subsystem.

The response time can be, therefore, measured on the basis of the different analysis and transmission times that happen inside and between modules. In particular, the embedded-side analytics are designed to work in real-time. A general interpretation of this term is to provide a result in time, just when it is required. In the context of video and audio analysis it is translated to the ability of processing the input data streams with a sufficient frequency, ideally equal to the input frequency (i.e. all the input information is processed at the frequency it comes, being usually around 25-30 Hz for video processing).

The response of the cloud-side subsystem can be analysed analogously, though real-time operation is not a requirement since the information is not streamed and lost, but stored in cloud-storage databases. Preferably, the response time of the cloud-side should be a one that permits a human operator to receive alerts from the P-REACT system in time to take the required actions (e.g. calling to first responders).

The tests on time response measures will be carried out individually for the embedded-side and cloud-side, considering the different functionalities they might implement. A general end-to-end time response for specific use cases are to be defined during the first trials.

8.2. End-to-end Use Case Scenarios Evaluation

The P-REACT project combines various modalities and analytical functions for the final detection of each use case security event. This allows for setting different weights for each of the basic analysis modules in order to tune the detection ratio of the overall P-REACT system and to favour or penalize e.g. false alarms. Moreover, it also allows for setting different alarm triggering thresholds based on the performance and individual processing time requirements of each module.

For the evaluation of the use case scenarios, one has to consider the modular and multi-layer architecture of the entire system. At the bottom level, the basic analytical functions, considering e.g. person detection and tracking, pose recognition, etc., are implemented. These are then combined in order to obtain the detection of higher level end-user scenario indicators, like chasing, fighting etc., as shown in Section 5.3.1. Finally, the business logic implements the combination of several key indicator detectors, which may be based on the analysis of different sensing modalities (e.g. audio, video or depth) as well as different combination of the end-user scenario indicators. This last level is where the concrete user requirements and installation possibilities are to be considered.

For the final deployment and evaluation of the P-REACT system, an evaluation report is to be prepared where the performance of the basic analytical modules is evaluated under several contextual conditions and considering a variety of internal parameters. Based on this report, the most suitable combination of analytical modules and the corresponding alarm triggering thresholds for their combination, which are to be applied within the recognition business logic, can be assessed. Several factors influence the final selection of the individual module parameters and the combination thresholds: a) individual detection performance; b) individual time performance; c) end-user preferences and requirements on the detection response (e.g. favouring false negatives instead of false positives); d) end user requirements on the end-to-end response time; e) operational context; and e) end user installation possibilities and preferences.

NOTE: At this moment of the project development it is not possible to anticipate any of the final detection rates of the analysis modules under real-world conditions, nor do we have any realistic expectations from the end-users on the performance of the entire system.

As an example we describe the evaluation methodology for the *UC2 – Shop (indoor) Theft with threat* end user scenario:

> The UC2 scenario considers the combination of video/depth activity detection and audio keyword detection as the key indicators of the final theft with threat event detection. The evaluation methodology described in Section 8.1 will be applied separately for the video, depth and the audio analysis modules. The lab-testing or field testing datasets can be used for the evaluation of the modules. Then, the application context of the scenario is determined based on the ground truth dataset of the UC2 scenario at hand. Based on the performance statistics (time and detection) for the considered context, the best parameters and combination thresholds of

the business logic will be set up. Finally, the end-to-end evaluation of the UC2 scenario will be again evaluated based on the methodology described in Section 8.1, but this time considering the output of the business logic module, triggering the final alarm.

8.3. Datasets

In many cases, the usage of real data will not be possible (e.g. vandalism or antisocial behaviour), or limited to a reduced material. For the evaluation of the algorithms, the use of **existing datasets** will be a must. Such datasets will provide the necessary **training material** to develop the algorithms, evaluate their performance with **ground truth**, and eventually, **compare it with competing solutions**.

The main challenge is to select and use correctly the datasets, to make the obtained results extensible to similar or analogous situations (e.g. a dataset for graffiti detection only in day time can be used to evaluate the performance of the system in day time, not in night time, which remains unknown until a specific dataset is used).

A preliminary list of datasets for video analytics can be found in the following table:

Name	Interesting for P-REACT	Crimes	Link	Details
BEHAVE Interactions Test Case Scenarios	Group behavior and interaction scenes	Yes (Fighting, Chasing, etc.)	http://groups.inf.ed.ac.uk/vision/BEHAVEDATA/INTERACTIONS/	This dataset comprises two vides of an outdoor environment with actors performing interaction activities. Annotated with Viper
PETS 2014	Fighting outdoor scene	Yes	http://www.cvg.rdg.ac.uk/PETS2014/	This dataset consists on several cameras attached to a parked large vehicle, around which the action happens. There are also surveillance cameras monitoring the whole scene.
CAVIAR Test Case Scenarios	Indoor scenarios of a mall, with examples of events in a shop.	Yes (Fighting, Running, Chasing, Abandoned bags)	http://groups.inf.ed.ac.uk/vision/CAVIAR/CAVIARDATA1/	Several cameras, calibrated.
UT (University of Texas)	Some (fake) fighting scenes	Somehow	http://cvrc.ece.utexas.edu/SDHA2010/Human_Interaction.html	Slightly high view of outdoor scenes where actors play human actions including some fighting.
PETS 2007	Crowded scenes in stations	Yes (Loitering, Attended luggage removal)	http://www.cvg.rdg.ac.uk/PETS2007/data.html	very complex sequences, really challenging because of the presence of multiple persons
ICDP (International Conference	Ground truth of human actions including	Somehow	http://www.icdp-conf.org/	Restricted access upon participation in the conference

on Imaging for Crime Detection and Prevention)	fighting, kicking, etc.			
OTCBVS (Dataset 06: Terravic Weapon IR Database)	Weapon detection and weapon discharge using thermal imagery	Somehow	http://www.vcipl.okstate.edu/otcbvs/bench/Data/06/download.html	5 collections
LIRIS (Human activities dataset)	Indoor depth dataset, with human actions	Somehow (Unsuccessfully try to enter a room)	http://liris.cnrs.fr/voir/activities-dataset/download.html	Low perspective scenes, like taken by a robot
Stony Brook University	Two-persons interaction with Punching, Kicking, etc., depth information	Somehow (Fighting)	http://www.cs.stonybrook.edu/~kyun/research/kinect_interaction/index.html	
VISOR (parking areas, human action)	Some urban/outdoor scenes, and human action sequences	No	http://www.openvisor.org/video_categories.asp	Not all videos are annotated
PETS 2006 Benchmark Data	Left-luggage, multi-camera	No	http://www.cvg.rdg.ac.uk/PETS2006/data.html	Multisensor sequences in stations
CVAP Recognition of human actions	Isolated human action recognition	No	http://www.nada.kth.se/cvap/actions/	Can be used to train classifiers that detect patterns of human actions

Table 29 Video analytics data sets for human activities and interactions.

The usage of datasets is subject to the use case that needs to be evaluated: graffiti detection, motion detection, unauthorised access, fighting, chasing, etc.

The creation of datasets must be handled carefully due to the large amount of effort required to prepare, acquire, classify and annotate the material. This task is typically very time consuming and costly, and only in the case of an absolute need of it should be executed by the project (e.g. if there is a major lack of datasets for a critical scenario). For instance, after observing the existing, available datasets from the scientific community, it seems necessary to create a customized dataset for graffiti detection and also for vandalism analysis once the use cases scenarios are detailed enough.

9. Conclusions

In this document we have described the initial requirements for the P-REACT system. These requirements are the basis for the definition of the technical and functional specifications of the system, above which the solution is being deployed.

From Sections 2 to 4, the aspects that have to be taken into account regarding Data Integrity and Privacy in P-REACT are highlighted. The guidelines provided outline different solutions that can be achieved, highlighting relevant aspects that must be fulfilled from the ethical and social points of view. Deliverable D2.3 will document the committed actions to ensure the aforementioned data integrity and privacy requirements are fully covered.

As previously mentioned, P-REACT platform implementation will be based on the European Legal Framework, even though during the installation of the system the corresponding national legal framework will be studied to ensure that they are also covered.

Instead, Sections 5 to 7 are more focused on the description of the system requirements of the P-REACT's platform, at the level of the local, embedded, and the cloud components. Note that for a detailed description of the P-REACT's architecture, one should consult the deliverable D2.3, "P-REACT Conceptual Architecture".

An initial description of the platform and the considered application scenarios has been given, with a mapping to the considered functional, analytical and hardware requirements. A detailed description of the software, hardware, communication and security requirements can be found in this document. The presented requirements follow the initial analysis of the user specifications and may be subject to slight changes during the system implementation and deployment, following the iterative development methodology of the entire platform. Further details are expected to be given at the time of testing the system (i.e. lab and field testing).

Finally, the Section 8 presents a methodology for functional testing and quality assessment of the analysis modules. We propose three stages of quality evaluation with necessary formalisation of the pre-requisites of each quality test, considering its functional, implementation and operational requirements. At last, the metrics that will be used to quantify the performance of the analytical modules, in the context of the selected scenarios, have been given.

ANNEX I. GLOSSARY AND ACRONYMS

The table below shows the most significant acronyms used and/or cited to prepare this document:

Term	Definition / Description
ADSL	Asymmetric digital subscriber line
AMD	Advanced Micro Devices
API	Application Programming Interface
CFREU	Charter of Fundamental Rights of the European Union
CUDA	CUDA® is a parallel computing platform and programming model invented by NVIDIA.
DDR	Double data rate
DSL	Digital Subscriber Line
ECHR	European Convention on Human Rights
FIFO	First In First Out
FPS	Frames per second
FTP	File Transfer Protocol
GPIO	General-purpose input/output
GridFTP	Grid FTP
H.264	MPEG-4 Part 10, Advanced Video Coding (MPEG-4 AVC)
HDMI	High-Definition Multimedia Interface
HTTP	Hypertext Transfer Protocol
HTTPS	Secure HTTP
ICT	Information and Communication Technologies
IPSec	Internet Protocol Security (IPsec) is a protocol suite for securing Internet Protocol (IP) communications by authenticating and encrypting each IP packet of a communication session.
IR	Infrared
kFTP	Kerberos FTP
MJEPG	Motion JPEG
OpenCL	Open Computing Language (OpenCL) is a framework for writing programs that execute across heterogeneous platforms
OpenGL	OpenGL (Open Graphics Library) is a cross-language, multi-platform application programming interface (API) for rendering 2D and 3D vector graphics.
OpenVPN	Open VPN
PCM	Pulse-code modulation
RAID	Redundant array of independent disks

Term	Definition / Description
RAM	Random-access memory
REST	Representational state transfer
RTP	Real-time Transport Protocol
RTSP	Real Time Streaming Protocol
SOAP	Simple Object Access protocol
SSH	Secure Shell is a cryptographic network protocol for secure data communication
UDP	User Datagram Protocol
UML	Unified Modeling Language
UPS	Uninterruptible Power Supply
USB	Universal Serial Bus
VDSL	Very high bit-rate Digital Subscriber Line
VGA	Video Graphics Array
VPN	Virtual Private Network

Table 30 Glossary and Acronyms

ANNEX II. REFERENCES

The table below shows the most significant references used and/or cited to prepare this document:

Reference	Source
[1]	Privacy International - National Privacy Ranking 2007 - Leading Surveillance Societies Around the World, 2007, https://www.privacyinternational.org/sites/privacyinternational.org/files/file-downloads/phrcomp_sort_0.pdf , Accessed on 09/2014
[2]	V. Carli, Assessing CCTV as an effective safety and management tool for crime-solving, prevention and reduction, International Centre for the Prevention of Crime, Montreal, 2008
[3]	C. Norris, G. Armstrong, The maximum surveillance society : the rise of CCTV, Oxford Berg, 2010
[4]	Network bandwidth requirements for media traffic in Lync Server 2013, Available from: http://technet.microsoft.com/en-us/library/jj688118.aspx, Accessed on 08/2014
[5]	Video Quality Optimization • Multi-rate Video Encoding, Available from: http://www.envivio.com/files/white-papers/WP_MultiResEncoding-highres.pdf, Accessed on 08/2014.